Peptide Model of the Mutant Proinsulin Syndrome. I. Design and Clinical Correlation

Balamurugan Dhayalan1*, Michael D. Glidden2,3,4†, Alexander N. Zaykov5†, Yen-Shan Chen1, Yanwu Yang1, Nelson B. Phillips2, Faramarz Ismail-Beigi2,3,4, Mark A. Jarosinski1, Richard D. DiMarchi5 and Michael A. Weiss1*

1 Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, United States, 2 Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, OH, United States, 3 Department of Physiology & Biophysics, Case Western Reserve University School of Medicine, Cleveland, OH, United States, 4 Department of Medicine, Case Western Reserve University School of Medicine, Cleveland, OH, United States, 5 Department of Chemistry, Indiana University, Bloomington, IN, United States

The mutant proinsulin syndrome is a monogenic cause of diabetes mellitus due to toxic misfolding of insulin’s biosynthetic precursor. Also designated mutant INS-gene induced diabetes of the young (MIDY), this syndrome defines molecular determinants of foldability in the endoplasmic reticulum (ER) of β-cells. Here, we describe a peptide model of a key proinsulin folding intermediate and variants containing representative clinical mutations; the latter perturb invariant core sites in native proinsulin (LeuB15→Pro, LeuA16→Pro, and PheB24→Ser). The studies exploited a 49-residue single-chain synthetic precursor (designated DesDi), previously shown to optimize in vitro efficiency of disulfide pairing. Parent and variant peptides contain a single disulfide bridge (cystine B19-A20) to provide a model of proinsulin’s first oxidative folding intermediate. The peptides were characterized by circular dichroism and redox stability in relation to effects of the mutations on (a) in vitro foldability of the corresponding insulin analogs and (b) ER stress induced in cell culture on expression of the corresponding variant proinsulins. Striking correlations were observed between peptide biophysical properties, degree of ER stress and age of diabetes onset (neonatal or adolescent). Our findings suggest that age of onset reflects the extent to which nascent structure is destabilized in proinsulin’s putative folding nucleus. We envisage that such peptide models will enable high-resolution structural studies of key folding determinants and in turn permit molecular dissection of phenotype-genotype relationships in this monogenic diabetes syndrome. Our companion study (next article in this issue) employs two-dimensional heteronuclear NMR spectroscopy to define site-specific perturbations in the variant peptides.

Keywords: monogenic diabetes, endoplasmic reticular stress, peptide chemistry, protein folding, oxidative folding intermediate
INTRODUCTION

The mutant proinsulin syndrome (MPS) is a monogenic cause of diabetes mellitus (DM) presenting at a broad range of ages: onset can occur either in the neonatal period, childhood, adolescence or early adulthood (1–4). Characterization of this syndrome and related mouse models (5, 6) has established the paradigm that DM may arise as a proteotoxic disorder of insulin biosynthesis (Figure 1A). Also designated Mutant INS-gene Induced Diabetestes of the Young (MIDY) (11), MPS thus pertains to patients traditionally classified, on the basis of age at presentation, as either Permanent Neonatal Diabetes Mellitus (PNDM) or Maturity Onset Diabetes of the Young (MODY) (12). This phenotypic spectrum may reflect polygenic differences in β-cell biology (13) or intrinsic mutation-dependent biophysical properties of the variant proinsulins (11, 14, 15).

MIDY mutations are ordinarily dominant and associated with misfolding of the proinsulin variant in the endoplasmic reticulum (ER), leading to β-cell dysfunction and eventual death [(16); for reviews, see (17, 18)]. Whereas the majority of MIDY mutations introduce or remove a Cys residue—in either case leading to an odd number of thiol groups and hence risk of aberrant intermolecular disulfide pairing (19)—PNDM- and MODY phenotypes are also associated with non-cysteine-related mutations (20). The latter amino-acid substitutions generally occur at sites conserved among vertebrate insulins [and in most cases also shared by vertebrate insulin-like growth factors (IGFs) (21)]. These mutational “hot spots” define the structural framework of native insulin (Figure 1B) and are of mechanistic interest as putative molecular determinants of protein folding efficiency (20, 22, 23). The present study, based on the oxidative refolding pathway of native proinsulin (Figure 1C), describes the design and chemical synthesis of a single-chain peptide model of a key on-pathway proinsulin folding intermediate (8, 24). This model enables comparative biophysical studies of representative MIDY “hot-spot” mutations with neonatal or delayed disease onset (Figure 1D). First introduced in studies of bovine pancreatic trypsin inhibitor (25), peptide models of protein-folding intermediates have provided a general approach toward dissecting critical molecular interactions guiding the conformational search of a nascent polypeptide (see Box 1 and Figure 2) (31, 32).

Abbreviations: MIDY, mutant INS-induced diabetes of the young; MODY, maturity-onset diabetes of the young; MPS, mutant proinsulin syndrome; PNDM, permanent neonatal diabetes mellitus; DesDi, single-chain insulin analog insulin containing LysB28 and lacking residues B29 and B30; N, DesDi insulin template with all three native disulfide bonds; N*, DesDi insulin control with solubility enhancements; [B19-A20], DesDi variant based on N* retaining only the B19-A20 disulfide linkage; ER, endoplasmic reticulum; SCI, single-chain insulin; CD, circular dichroism; NMR, nuclear magnetic resonance; rp-HPLC, reverse-phase high-performance liquid chromatography; LC-MS, liquid chromatography-mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization - time of flight; and IGF, insulin-like growth factor. Insulin residues are denoted by residue type (in standard three-letter code) followed by the chain and position (e.g., PheB9 designates a phenylalanine at the 24th position of the B chain).

The present peptide model contains a single disulfide bridge, internal cystine B19-A20, the first partial fold to accumulate in chemical-trapping studies of proinsulin (or homologous IGF-I) refolding in vitro (8, 9, 24) (see Supplemental Discussion regarding effects of pH in such refolding assays). Its framework derives from a foreshortened single-chain synthetic precursor of insulin optimized for efficiency of disulfide pairing (50). Designated “DesDi”, the parent analog comprises 49 residues (Figure 3A): B-chain residues B1-B28 followed by A-chain residues A1-A21. Substitution of ProB28 by Lys enables enzymatic cleavage to liberate an active two-chain insulin analog (50). Our one-disulfide model contains pairwise substitution of cystine B7-A7 by Ser and pairwise substitution of cystine A6-A11 by Ala (Figure 3B). Segmental α-helical propensity and solubility were augmented by substitutions HisB10→Asp (native α-helix α1 spanning residues B9-B19) and ThrA8→Glu (native α-helix α2: A1-A8) (58, 59). Representative MIDY mutations (LeuB15→Pro, LeuA16→Pro, and PheB24→Ser) were introduced into the parent DesDi framework with six cysteines (“native state”) and the peptide model (“1SS”). These amino-acid substitutions were chosen based on phenotype and structural interest: the proline variants (each neonatal in onset) are predicted to perturb nascent α-helical folding and native packing of the hydrophobic core (20, 60) whereas SerB24 [with onset in adolescence or early adulthood (61)] perturbs the “aromatic anchor” of the B-chain β-strand (B24-B28) (60, 62, 63). Each of these conserved side chains contributes to core packing near internal cystine B19-A20 (Figure 3C). A related two-chain one-disulfide model of the homologous IGF-I folding nucleus has previously been described (51); the corresponding three side chains (IGF-I residues Leu14, Phe23 and Leu57) were observed to participate in its molten native-like structure at low temperatures.

The present study (the first of two in this issue) describes the synthesis and respective foldabilities of the above set of native state and 1SS 49-residue peptides. Circular dichroism (CD) is employed to probe α-helix contents, thermal unfolding profiles and thermodynamic stabilities. Characterization of such chemical and biophysical properties were extended through cell-biological assays of ER stress induced in human cells on transient expression of the corresponding mutant proinsulins. Striking correlations were observed among foldability, nascent α-helical folding and extent of ER stress in human cell culture. Our companion study further interrogates the nascent structures of these peptide models and their mutational perturbation by 1H and 13C NMR spectroscopy [(64); following article in this issue]. Together, our findings establish a general platform for biophysical studies of a subset of MIDY mutations in relation to molecular mechanisms of proinsulin biosynthesis in pancreatic β-cells.

MATERIALS AND METHODS

Automated DesDi Peptide Synthesis

Peptides were synthesized either with an ABI 433A Peptide Synthesizer (Applied Biosystems) or Tribute 2-Channel peptide synthesizer (Gyros Protein technologies) using a preprogrammed
A Normal Physiology  
Wild-type proinsulin production in β-cell ER  

[Diagram showing normal physiology]  

B C-domain  
A-domain  
A6-A11  
B7-A7  

[Diagram showing C-domain]  

B24  
B19-A20  

[Diagram showing B-domain]  

D Mutant INS-gene Induced Diabetes of the Young (MIDY)  

[Diagram showing MIDY]  

BOX 1 Peptide models of protein-folding intermediates.

Studies of the mutant proinsulin syndrome (17) have built on general principles of cell biology and protein chemistry established over the past sixty years (26–28). This deep interdisciplinary background highlights mechanisms underlying the biosynthesis of disulfide-stabilized secretory proteins from the scale of organelles and macromolecular complexes (29, 30) to the molecular biophysics of a nascent polypeptide chain’s conformational search (31–34).

Ribosomal translation at the outer surface of the rough ER is coupled to cleavage of the signal peptide and associated translocation into the ER [Figure 2A; for review, see (37)]. The latter environment provides chaperones and oxidative machinery for disulfide bond formation, rearrangement, and quality control (38–42). These general processes pertain to β-cell physiology and dysfunction in DM (17, 35, 36). Although initial steps of protein folding can in some cases be co-translational (30), the nascent proinsulin chain is likely to form an unfolded-state ensemble (at right in Figure 2A) to enable initial pairing of two cysteines distant in the sequence (CysB19 and CysA20; residues 43 and 109 in preproinsulin) (9, 24, 43). The present study has exploited a subset of diabetes-associated mutations to investigate such long-range pairing.

Analysis of atomic-scale events in the in vitro refolding of polypeptide chains and their computational simulation (44, 45) provide insight into the challenges faced in cellular protein biosynthesis (27, 46). Folding is visualized as proceeding through funnel-shaped free-energy landscapes (Figure 2B), in general via multiple trajectories (e.g., yellow or magenta lines) (31). Dissection of globular proteins into peptide models, pioneered by Oas and Kim in 1988 (25) has been broadly influential in enabling key steps to be identified (47). Applications have been described to both oxidative folding intermediates (25), the rapid autonomous folding of disulfide-free subdomains (48) and fragments containing engineered disulfide bridges (49). Use of peptide models may circumvent the usual cooperativity of globular protein folding, which can obscure discrete steps (29). The latter perspective has been reinforced by studies of intact proteins by native-state amide-proton exchange kinetics (34).

The present studies have exploited a peptide model of a key one-disulfide proinsulin folding intermediate (simulated ensemble in Figure 2C); its features favor formation of on-pathway nascent structure [see also Figure 9 and Discussion (50)]. Peptide design builds on prior studies of insulin-related polypeptides lacking specific disulfide bridges (10, 51–55). To our knowledge, this is the first investigation of clinical mutations in a peptide model of a proinsulin folding intermediate.
solid-phase fluorenlymethyloxycarbonyl (Fmoc) protocol designed for standard 0.1 mmol scale syntheses. ABI protocols consist of the following modules: cycle-1 [d], cycle-2 [aibde], cycle-3 with number of repetitions equal #aa-1 [afgbd], cycle #aa+2 | [ffbdc], where #aa is a number of amino acids in the sequence. Automated couplings utilized diisopropylcarbodiimide (DIC)/6-Cl-hydroxybenzotriazole (6-Cl-HOBt) in N-Methyl pyrrolidone (NMP) whereas Fmoc deprotections used 20% piperidine in NMP. a-carboxyl-protected Asp was used in place of Asn in all syntheses of DesDi analogs to accommodate the use of ChemMatrix® Rink-Amide resin (loading = 0.46 mmol/g). The Tribute peptide synthesizer used heating protocols: couplings were done for 6 min at 60°C except for Cys/His (2 min at 25°C, then 5 min at 60°C) and Arg (20 min at 25°C, then 5 min at 60°C); deprotection was done twice (30 sec at 50°C, then 3 min at 50°C). Reagent conditions were otherwise similar to ABI protocols except that DMF was used as solvent and choice of the resin was H-Asn(Trt)-HMPB-ChemMatrix® resin. Peptides were cleaved with TFA cocktail (2.5% vol/vol of each: β-mercaptoethanol, triisopropylsilane, anisole, and water) followed by ether precipitation.

**Folding and Purification of N and N* DesDi Analogs**

Crude peptides from ether precipitation were dissolved in glycine buffer (20 mM glycine and 2 mM cysteine hydrochloride, pH 10.5) to a final peptide concentration of 0.1 mM. The pH of this solution was readjusted to 10.5 to account for traces of residual TFA present in lyophilized peptides. This solution was then stirred while open to air at 4°C until reaction completion (usually overnight). After monitoring of the folding reaction by analytical HPLC to determine the extent of conversion, the pH of the solution was then lowered to ~2.0 with 5N HCl to neutralize the folding reaction. Folded peptide was then purified by preparative rp-HPLC using Waters 2545 Quaternary pumping system equipped with FlexInject. Chromatographic separations were
performed on a C4 Proto (20x250 mm) 300 Å, 10 μm, Higgins Analytical Inc. column, using 25-50% solvent B (0.1% TFA in acetonitrile) in solvent A (0.1% TFA in H₂O) over 35 minutes (min) at a flow rate of 20 ml/min with detection by UV absorption at 215 nm. Fractions containing clean peptide were pooled and lyophilized. Purity of the materials was confirmed by LCQ Advantage Ion Trap Mass Spectrometer System coupled to an Agilent 1100 Series HPLC system. Masses were obtained by online electrospray mass spectrometry. MS data shown were collected across the entire principal UV-absorbing peak in each chromatogram; LC-MS retention times and mass measurements are given in Table S1.

**Folding and Purification of [B19-A20]-SS DesDi Analogs**

Purification of single-disulfide analogs was performed in a two-step process. First, the crude peptide was fully dissolved in pH 11 buffer in presence of excess DTT. The pH was then lowered to 8.0 before purification by semi-preparative rp-HPLC under alkaline conditions using 25 mM ammonium bicarbonate buffer (pH 8.0) and acetonitrile (60 min gradient of 20→50%) as eluents on a TRIART C18, 250 x 10mm 5μm, 120Å column. Fractions containing linear, fully reduced peptides were pooled and lyophilized. Folding was performed either using room air oxidation as described for N and N* analogs or utilizing cysteine-cystine redox pair. For the latter, peptides were oxidized by acidic RP-HPLC conditions using acetic acid, 50% acetonitrile elution gradient as described above. Collected fractions were pooled and lyophilized. When necessary, an additional purification using acidic RP-HPLC conditions was performed. Purity of the materials was confirmed by analytical LC-MS as described for N and N* DesDi analogs above (LC-MS retention times and mass measurements are given in Table S1).

**Two-Chain DesDi Conversion**

In cases of three disulfide containing DesDi analogs, single-chain insulins were converted to two-chain version by Lys-specific enzyme. In a typical experiment, single-chain DesDi analog was treated with Endo Lys-C enzyme (65) in 25 mM Tris base, 100 mM urea buffer (pH 10.5) at a final 0.1 mM peptide concentration followed by addition of 1:1 cystine/cysteine (2 mM each). Folding was allowed to proceed for 1 hour (hr), followed by purification by rp-HPLC using the same 20→50% acetonitrile elution gradient as described above. Collected fractions were pooled and lyophilized. When necessary, an additional purification using acidic RP-HPLC conditions was performed. Purity of the materials was confirmed by analytical LC-MS as described for N and N* DesDi analogs above (LC-MS retention times and mass measurements are given in Table S1).

**Purification of Insulin Analogs**

Wild-type human insulin and insulin lispro were purified from U-100 pharmaceutical formulations of Humulin® and Humalog® (Eli Lilly and Co.), respectively, using preparative rp-HPLC (C4 10μm 250x20mm Proto 300 Column; Higgins Analytical, Inc.) utilizing Buffer A (0.1% TFA in H₂O) and a 10-min elution gradient of 20→70% Buffer B (0.1% TFA in acetonitrile). Following lyophilization of the collected protein fraction, purity was verified using analytical rp-HPLC (C4 5μm 250x4.6mm Proto 300 Column; Higgins Analytical, Inc.) with a 35-min elution gradient of 25→50% Buffer B; molar mass was verified with an Applied Biosystems 4700 Proteomics Analyzer utilizing MALDI-TOF in reflector mode; chromatographic retention times and mass measurements for these clinical analogs are given in Table S1.

**Nuclear Magnetic Resonance (NMR) Spectroscopy**

All spectra were acquired at a protein concentration of ~0.2 mM in 100% D₂O (pD 7.4) at 35°C with a Bruker AVANCE 700-MHz spectrometer, as described (23). All chemical shifts were calibrated in parts per million (ppm) relative to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as an internal standard. The spectrometer was equipped with ¹H, ¹³C, ¹⁵N quadruple resonance cryoprobe.

**CD Spectropolarimetry**

Far ultraviolet (255-190 nm) CD spectra were obtained at high signal-to-noise for WT insulin, insulin lispro, and all [B19-A20], N and N* DesDi peptides (summarized in Figure 4) using a CD spectropolarimeter (Aviv-400 or Jasco J-1500) equipped with temperature control and an automated titration unit. Samples were prepared at a concentration of 20-70 μM protein in degassed potassium phosphate (10 mM KH₂PO₄/K₂HPO₄ with 50 mM KCl), brought to pH 7.4 with KOH, and placed in a parafilm-sealed 1-mm pathlength quartz cuvette. Utilizing this approach, spectra were acquired from 4-88°C in 4°C steps. To reduce acquisition time at high temperatures, ellipticity measurements made above 40°C included only wavelength sets of 254(± 1), 222(± 1) and 208(± 1) nm using a 0.5-nm wavelength resolution and 30-sec. detector averaging time. Buffer-only CD spectra were obtained using degassed buffer with no protein at temperatures of 4, 25 and 37°C using a 0.5-nm wavelength resolution and 90 sec. averaging time. The linear temperature dependence at all wavelengths of buffer-only spectra allowed interpolation and extrapolation of buffer-only spectra at any temperature. Extrapolated reference spectra were then subtracted from all CD spectra acquired at the same temperature. Estimates of secondary-structure content were then determined from normalized spectra acquired at 4, 25 and 37°C using the SELCON-3 algorithm packaged with the CDPro spectral analysis software (66–68).

The temperature dependence of protein folds was assessed by plotting the average of molar ellipticity values at wavelengths straddling the α-helix sensitive wavelength of 222 nm (i.e., 221-223 nm with 0.5-nm resolution) against temperature. The quantity resulting from averaging wavelengths 221-223 nm, \(\langle \theta \rangle_{222±1nm} \), enhances signal-to-noise while ensuring that any difference in ellipticity that may be observed between sequential temperature steps is not confounded by random error.

---

**Figure 4**

**Mutant Proinsulin Syndrome I**
Quantitation of Samples and Normalization of CD Spectra

CD samples were quantitated via reference-subtracted UV-Vis spectra acquired in 10 mM potassium phosphate (pH 7.4) with a UV-Vis spectrometer (Aviv Biomedical Inc., Lakewood, NJ) and a 3-mm quartz cuvette. Protein concentrations in potassium phosphate, $[C]_{KPi}$, were calculated using absorbance at $\lambda=280$ nm and estimated extinction coefficients predicted by the online ExPASy ProtParam tool (69), which estimates molar absorptivity using the amino-acid sequence and cysteine sulfur oxidation states of each peptide. Calibration of these estimated extinction coefficients, which do not anticipate the effects of intramolecular dipole-dipole interactions in a folded protein, was achieved by quantifying $[C]_{GuCl}$, the concentration of the same amount of protein in a buffer containing 8M guanidine hydrochloride, 10 mM potassium phosphate (pH 7.4) and 50 mM KCl; and then calculating a unique correction factor $M_G = [C]_{KPi}/[C]_{GuCl}$ for each sample. As DesDi protein samples containing all three
native disulfide bridges do not fully unfold in 8M guanidine, extinction coefficients for N, N* and WT insulin were calibrated by determining amino-acid compositions of the samples from three cycles of N-terminal Edman degradation sequencing using an Applied Biosystems Procise 494 Sequencer. This approach was validated by the fact that calibration via UV-Vis spectroscopy with potassium phosphate buffer containing guanidine hydrochloride yielded the same correction factors for wildtype insulin. Percent error in the estimated extinction coefficient of WT insulin and insulin lispro was ~4% whereas that of N* and N was 9% and 24%, respectively. The predicted extinction coefficients of all 1SS peptides, which are mostly unfolded in zero denaturant (see Figure 4), had <0.5% deviation from the experimental value. CD spectra reporting molar ellipticity per residue, [θ], were calculated by dividing raw ellipticities by the corrected protein concentration ([C]_cal = [C]_KP/M_C) and the number of amino acids, N, for each protein.

CD-Monitored Guanidine-Induced Unfolding Studies

Thermodynamic stabilities of all insulin peptides in 10 mM potassium phosphate buffer (pH 7.4) at 25°C were determined using guanidine hydrochloride titrations monitored by CD at the α-helix-sensitive wavelength 222 nm as described (23). Using non-linear least squares regression, plots of ellipticity vs. guanidine concentration were fit to a two-state unfolding model (70):

\[
\theta(c) = \theta_A + \frac{\theta_B e^{-\Delta G/m_C}/RT}{1 + e^{-\Delta G/m_C}/RT}
\]

where \(\Delta G\) is the Gibbs free energy of unfolding, \(C\) is guanidine concentration, \(R\) is the ideal gas law constant, \(T\) is temperature, and \(\theta_A\) and \(\theta_B\) are baseline ellipticity values reflecting the folded and unfolded state. Baseline ellipticities were calculated via simultaneous fitting of linear equations \(\theta_A(c) = \theta_A + m_Ac\) and \(\theta_B(c) = \theta_B + m_Bc\) as described (71).

GSH-GSSH Assay

1SS-peptides (50 µg, 45 µM final concentration) were treated with 25 mM reduced glutathione (GSH) and 5 mM oxidized glutathione (GSSH) in 200 µL buffer (20 mM sodium phosphate [pH 8.2] and 100 mM NaCl) and allowed to attain equilibrium at 25°C. After 3 hrs, an aliquot was acidified with 6 M guanidine hydrochloride in 0.1% trifluoroacetic acid and analyzed by analytical rp-HPLC using 5-65% solvent B (0.1% TFA in acetonitrile) in solvent A (0.1% TFA in water) over 23 min.

Proinsulin Constructs

Plasmids expressing full-length human proinsulin or variants were constructed by polymerase chain reaction (PCR). Mutations in proinsulin were introduced using QuikChange™ (Stratagene). Constructions were verified by DNA sequencing.

Mammalian Cell Culture and ER Stress Assays

Human embryonic kidney 293T cells were purchased from American Type Culture Collection and cultured in Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin as recommended. Transfections were performed using Lipofectamine 3000 as described by the vendor (Invitrogen). Transfected HEK 293T cells were subjected to the Bio-Rad one-step real-time qPCR protocol. Readouts were provided by the up-regulation of ER stress markers CHOP and BiP. The gene expression values were normalized by the expression of the gene encoding glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as internal control. The mRNA abundances were measured in triplicate. In Western blot assay probing ER stress markers (72, 73), after 24 hr post transient transfection, cells were lysed by RIPA buffer (Cell Signaling Technology; CST). Protein concentrations in lysates were measured by BCA assay (Thermo) and subjected to 4-20% SDS-PAGE and WB using anti-pPERK. Anti-PERK, anti-BiP and anti-CHOP antibodies (CST) at a dilution ratio of 1:1000; GAPDH provided a loading control.

Rat Experiments

Animals were maintained in accredited facility of Case Western Reserve University School of Medicine. All procedures were approved by the Institutional Animal Care and Use Committee (IACUC) office of the University. Animal care and use was monitored by the University’s Veterinary Services.

Measurement of the Glucose-Lowering Effect of Insulins in Diabetic Rats

Male Lewis rats (average body mass of ~300 g) were rendered diabetic by streptozotocin (STZ) as described (23). Insulin analogs were dissolved in Lilly® Diluent buffer with the specified dose and injected in 100 µL/300 g rat. Lispro insulin (KP) was diluted as appropriate in Lilly® Diluent buffer. Control rats received the appropriate volume of the Lilly buffer. For intravenous (IV) injection, rats were anesthetized in a chamber for 5 min using a mixture of 5% isoflurane and 95% oxygen. Following cleaning of the tail, rats were injected while under anesthesia using the lateral tail vein. For subcutaneous (SQ) experiments, rats were injected under the skin into the soft tissue in the posterior aspect of the neck. Following injection, blood glucose was measured using a small drop of blood obtained from the clipped tip of the rat’s tail using a clinical glucometer (EasyMax® V Glucose Meter, Oak Tree Health, Las Vegas, NV). Blood-glucose concentrations were measured at time t=0, and every 10 min for the first hr, every 20 min for the second hr, every 30 min for the third hr, and then each hr for the rest of the experiment.

Molecular Modeling

Structural ensembles were calculated by simulated annealing using XPLOR-NIH (74–76). A model of the one-disulfide proinsulin intermediate (containing cystine B19-A20; see Box 1) was generated using distance restraints pertaining to residues A16-A21 and B15-B26 as observed in an engineered proinsulin monomer (7). A similar modeling protocol was employed to generate ensembles for 1SS-DesDi variants; selected distance restraints were extracted from NOESY spectrum of an engineered insulin monomer (23). For the parent 1SS DesDi
model, helix-related distance restraints were corroborated by NMR [see companion study (64)]. To allow for protein flexibility in these partial folds, upper bounds on long-range distance restraints were increased by 3 Å relative to NMR-derived bounds obtained in prior studies of insulin and proinsulin (7, 23).

RESULTS

Eleven 49-residue peptides were prepared by solid-phase peptide chemistry (Figure 3); the red segments represent a peptide bond between residues B28 and A1. Seven peptides contained insulin’s canonical six cysteines with intended disulfide pairing indicated in gold (Figure 3A). As expected (50), peptide N and N* (DesDi and [AspB10, GluA8]-DesDi, respectively) underwent oxidative folding with high efficiency to yield a single predominant product (Table 1). DesDi’s foldability was decreased or blocked by the MIDY substitutions in order SerB24 (least perturbed relative to the N parent peptide) >> ProA16 = ProB15 (no folded product detected). Introduction of N* substitutions AspB10 and GluA8 rescued inefficient but detectable folding of ProA16-N* but not ProB15. N* (Table 1). Four 1SS model peptides were also synthesized (Figure 3B). Because these contain only two cysteines, disulfide pairing was efficient in each case, including in the presence of ProA16 and ProB15. Reverse-phase HPLC retention times and molecular masses are given in Table S1. Two-chain versions of N/N* analogs and the parent 1SS model peptide were obtained following enzymatic cleavage with Lys-C protease (50). Analytical rp-HPLC chromatograms and LC-MS profiles are provided as Figures S1–S15.

CD studies were conducted of the native-state analogs as single chains (left-hand panels in Figure 4) and on cleavage of...
the Lys₁₂₈-Gly₁³₁ peptide bond (right-hand panels in Figure 4).¹ MIDY mutations SerB₂₄ and ProA₁₆ are each associated with reduced α-helix content (Table 2).² The extent of perturbation was more marked in the case of ProA₁₆, especially in the two-chain context (Figure 4D and Table 2, row 8 versus row 10). Thermodynamic stabilities were inferred from CD-detected guanidine denaturation studies (Figures 4E, F). Application of a two-state model provided estimates of free energies of unfolding (ΔG°), Table S1, column 3). In accordance with their relative susceptibilities to guanidine-induced unfolding, apparent by qualitative inspection of the denaturation data, ProA₁₆ is more profoundly destabilizing than is SerB₂₄. In each case imposition of the B₂₈-A₁ peptide bond enhances stability [which may rationalize its utility as a vehicle for oxidative folding (50)]. The two-chain ProA₁₆ N* analog did not exhibit a cooperative unfolding transition, and so its stability could not be estimated by this method. The stabilities of corresponding ProB₁₅ native state analogs could not be assessed due to absence of folded product. Functional studies of two-chain versions of SerB₂₄-N demonstrated reduced but substantial activity in a rodent model of DM (Figure S16) in accordance with past studies (23, 61, 78). The two-chain derivative of ProA₁₆-N* was inactive as was the two-chain derivative of the ISS parent peptide.

CD studies of ISS analogs are shown in Figure 5 in relation to the parent ISS model (labeled d in panel B), insulin lispro (labeled e in panel B) and native-state domains N and N* (respectively labeled f and g in panel B); for clarity, a color code is shown at bottom. Qualitative inspection of the far-U assay for 4, 25 and 37°C (Figures 5A–C) suggest rank order N* (most structured) > N > insulin lispro >> parent ISS model > 1SS-SerB₂₄ > 1SS-ProA₁₆ > ISS ProB₁₅ (least structured). These differences are in accordance with CD deconvolution (Tables 2A, B) and two-state thermodynamic modeling (Table 1 and Table S1). Of the ISS analogs, only the parent model peptide exhibits, to a small extent, a cooperative thermal unfolding curve at low

### Table 1 | Folding yields of DesDi synthetic precursors and thermodynamic stabilities.

| native state peptide domain | isolated yield (mg) | ΔG°(kcal/mol) | m value(kcal/mol/M) | C_mol(M) |
|-----------------------------|---------------------|---------------|---------------------|---------|
| single-chain (N)⁰           | 21.7 mg             | >4¹           | N.D.⁵              | N.D.⁵   |
| single-chain (N')⁶          | 21.7 mg             | 7.8 ± 0.4¹    | 0.99 ± 0.02         | 7.68 ± 0.1 |
| SerB₂₄ single-chain (N)⁶    | 12.0 mg             | 3.4 ± 0.1¹    | 0.69 ± 0.02         | 4.93 ± 0.1 |
| ProA₁₆ single-chain (N)⁶    | 0                   | –             | –                   | –       |
| ProA₁₆ single-chain (N*)⁶   | 2.1 mg              | 0.9 ± 0.1     | 0.51 ± 0.02         | 1.76 ± 0.1 |
| ProB₁₅ single-chain (N)⁶    | 0                   | –             | –                   | –       |
| ProB₁₅ single-chain (N*)⁶   | 0                   | –             | –                   | –       |

¹Isolated yield represents amounts after rp-HPLC purification from a folding reaction of 100 mg reduced polypeptide.

²ΔG° values provided were obtained from curve fitting of CD-guanidine titrations to a two-state unfolding transition model as described in Methods (23).

³CD-guanidine titrations performed at 37°C for N and 50°C N* (see Figure 4) yielded partially folded peptides at the maximum guanidine concentration. Fitting of the N* titration was successful (R² = 0.9994), but analysis of the N titration curve could only place a lower limit on thermodynamic stability of 4 kcal/mol.

⁴No isolable product obtained. Folding was nonetheless rescued in part by AspB₁₀ and GluA₈ substitutions (N*).

⁵No isolable product obtained. In the case of ProB₁₅, inclusion of stabilizing substitutions AspB₁₀ and GluA₈ (N*) did not rescue foldability: only disulfide isomers were observed.

### Table 2 | CD-derived secondary-structure contents.

| sample Name | α-helix | β-sheet | disordered |
|-------------|---------|---------|-----------|
| **A** | | | |
| insulin lispro | 38 – 40% | 9 – 10% | 32 – 33% |
| wild-type insulin⁷ | 47 – 56% | 3 – 10% | 24 – 29% |
| single-chain (N) | 49 – 64% | 2 – 10% | 24 – 27% |
| single-chain (N')⁶ | 49 – 58% | 4 – 10% | 24 – 26% |
| SerB₂₄ single-chain (N)⁶ | 55 – 61% | 3 – 7% | 27 – 28% |
| ProA₁₆ single-chain (N')⁶ | 43 – 48% | 11 – 14% | 27 – 28% |
| two-chain DesDi (N)⁶ | 47 – 53% | 10 – 12% | 26 – 28% |
| two-chain DesDi (N')⁶ | 56 – 61% | 7 – 10% | 24 – 26% |
| two-chain SerB₂₄ (N)⁶ | 50 – 58% | 5 – 6% | 27 – 29% |
| two-chain ProA₁₆ (N')⁶ | 25 – 26% | 22 – 23% | 31 – 32% |
| **B** | | | |
| 1SS-WT | 27 – 30% | 18 – 20% | 30 – 31% |
| 1SS-SerB₂₄ | 19 – 20% | 26 – 27% | 31 – 32% |
| 1SS-ProA₁₆ | 12 – 14% | 32 – 33% | 30 – 31% |
| 1SS-ProB₁₅ | 10 – 11% | 29 – 32% | 32 – 34% |

¹Total percent α-helix, β-sheet, and disordered coil were obtained from spectra acquired at discrete temperatures of 4°C, 25°C, and 37°C using the SELCON-3 algorithm (60–62). Estimated percentages are presented as the minimum to maximum content calculated across the three sampled temperatures.

²Estimates were calculated from WT insulin spectra to confirm that SELCON-3 processing of our own CD data gives values that match those published in the literature (77).
temperatures\(^3\) (Figure 5D) and possibly a cooperative guanidine-denaturation transition with \(\Delta G_u < 1\ \text{kcal/mole}\) (Figures 5E, F and Table S1).

We imagine that the 1SS peptides exist as a conformational equilibrium between a disulfide-tethered random coil (at left in Figure 6A) and a collapsed conformation in which nascent \(\alpha\)-helical structure occurs in the B domain (residues B9-B19; helix \(\alpha_t\) in insulin and proinsulin) and A domain (A12-A18; helix \(\alpha_3\) in insulin and proinsulin). Diffusion-collision of these nascent helices creates a molten proto-core engaging (in the parent model) Leu\(_{B11}\), Val\(_{B12}\), Leu\(_{B15}\), Val\(_{B18}\), Phe\(_{B24}\), Tyr\(_{B26}\), Leu\(_{A16}\) and Tyr\(_{A19}\) in the neighborhood of internal cystine B19-A20. This scheme envisions that this molten-core functions as proinsulin’s specific folding nucleus (51) and is destabilized by the MIDY mutations but to different extents. Evidence supporting this hypothesis was provided by one-dimensional \(^1\text{H}-\text{NMR}\) spectroscopy (Figure 6B). Whereas native state analogs (single-chain N\(^*\) and two-chain N\(^*\)) exhibit marked chemical-shift dispersion as expected of native globular domains, such dispersion is attenuated among the 1SS analogs in order N\(^*\) > parent 1SS model > 1SS-Pro\(_{A16}\) > 1SS-Pro\(_{B15}\). Preservation of chemical-shift dispersion in the proto-core of the parent 1SS peptide is remarkable, as evident by the upfield chemical shifts of aromatic resonances (Phe\(_{B24}\), Tyr\(_{B26}\) and Tyr\(_{A19}\)) and aliphatic resonances (Leu\(_{B11}\) and Leu\(_{B15}\)). Qualitative interpretation of the NMR spectrum of Ser\(_{B24}\) native state and 1SS analogs was confounded by the absence of the diamagnetic ring-current field of Phe\(_{B24}\), a major source of chemical-shift dispersion in native insulin (79). These NMR features are investigated further by two-dimensional heteronuclear NMR in our companion study with focus on \(^{13}\text{C}_\alpha\) and \(^1\text{H}_\alpha\) secondary NMR shifts (64).

The relative stabilities of the 1SS proto-cores, although inaccessible to guanidine denaturation studies (above), were instead probed by resistance to reduction at equilibrium in a defined redox buffer (25 mM reduced glutathione and 5 mM oxidized glutathione). Initial solutions contained only the disulfide-constrained peptides and were allowed to come to equilibrium as monitored by serial rp-HPLC chromatograms; representative steady-state chromatograms are shown in Figure 7A. Quantitation of the surviving disulfide-constrained elution peaks (arrows in Figure 7A) indicates a rank order of

\(^3\)For single-chain N and N\(^*\) parent domains, their limited temperature dependence of in the range 4-40°C correlates with enhanced folding efficiency and increased thermodynamic stability.
redox stability parent 1SS model (most stable) > 1SS-SerB24 > 1SS-ProA16 >> 1SS-ProB15 (least stable; histogram in Figure 7B). This trend is in accordance with effects of these mutations on native state DesDi folding yields (Table 1) and relative $\alpha$-helix contents of the 1SS models (Figure 7C). Encouraged by this coherence, we speculate that the relative native state folding yields mirror the efficiency of initial closure of cystine B19-A20, in turn dependent on diffusion-collision of the proposed $\alpha$-helical proto-core.

In an effort to connect the above chemical and biophysical properties to cell biology—and ultimately to the pathophysiology of the MPS in patients—we undertook studies of ER stress induced by transient expression of wild-type or mutant proinsulins in a human kidney-derived embryonic cell line readily grown in culture and amenable to transient transfection (HEK 293T cells). Although not related to $\beta$-cell lineages, Arvan and colleagues have shown the utility of these cells in studies of proinsulin biosynthesis (Figure 8A) (11). ER stress was probed through Western-blot...
studies of the pPERK/PERK ratio, induction of ER chaperone BiP (a member of the HSP70 family), and ER-stress-responsive transcription factor CHOP (Figure 8B). Changes in these markers (relative to the wild-type proinsulin baseline; horizontal dashed line in Figure 8B) are shown in Figures 8C (BiP on left and CHOP on right). The rank order of ER stress was ProB15 (highest ER stress) > ProA16 > SerB24 > WT > empty vector control. This pattern parallels the sensitivity of the 1SS models to reduction in a defined redox buffer (Figure 7D), 1SS α-helix contents (Figure 8D) and native-state DesDi folding yields (Figure 8E). Such extensive correlations provide evidence that the biophysical and biochemical properties of DesDi 1SS and native-state peptides relate to the pathophysiology of proinsulin biosynthesis in the ER of a human cell.

DISCUSSION

How proteins fold and misfold define key problems at the intersection of biophysics, cell biology and medicine (80). The mutant proinsulin syndrome highlights the importance of foldability in the process of insulin biosynthesis (81). Mutational impairment of native disulfide pairing in the ER of pancreatic β-cells leads to ER stress, β-cell dysfunction and eventual death (17, 18). This syndrome ordinarily exhibits genetic dominance, implying that misfolding of a variant proinsulin impairs bystander biosynthesis of wild-type proinsulin (19). This monogenic diabetes syndrome thus illuminates structural determinants of specific disulfide pairing (20) and folding efficiency as an implicit evolutionary constraint (23). The present study sought to develop a peptide model of a one-disulfide proinsulin folding intermediate as a general platform for studies of a mechanistic subclass of MIDY mutations: those that impair the nascent conformational search leading to initial pairing of CysB15 and CysA20, an early step in biosynthesis (8, 24). We exploited this platform to investigate three clinical mutations, two with neonatal onset [ProB15 and ProA16 (82, 83)] and one with onset in early adulthood [SerB24 (11, 61, 62)]. The present study extends the use of peptide models of protein-folding intermediates (see Box 1 (25)) to investigate the molecular pathogenesis of a monogenic syndrome of toxic protein misfolding.
Our peptide model is a 49-residue “mini-proinsulin” based on the DesDi framework as developed by DiMarchi and colleagues to optimize the efficiency of disulfide pairing in an enzyme-cleavable synthetic precursor (50). This framework contains B-chain residues B1-B28 followed by A-chain residues A1-A21. Its LysB28-GlyA1 peptide bond (cleavable by protease Lys-C) enables productive folding even of mutant insulins otherwise refractory to classical insulin chain combination (50), presumably by constraining the orientation of A- and B-chain residues to stabilize a specific folding nucleus. We obtained a one-disulfide model through pairwise substitutions of cystines B7-A7 (by serine) and A6-A11 (by alanine). Choice of Ser or Ala was
determined by solvent exposure of these disulfide bridges in native insulin (60). Nascent structure in the 1SS model, presumably stabilized by the B28-A1 peptide bond, was further favored by helicogenic substitutions HisB10→Asp (58) and ThrA8→Glu (59). Their additional negative charges were also intended to enhance solubility and mitigate the propensity of partial folds to aggregate via exposed nonpolar surfaces. In native state DesDi analogs introduction of these acidic side chains rescues the folding of a ProA16 analog, albeit in small yield. ProB15 blocks folding of DesDi even in the presence of AspB10 and GluA8. Insight into even such deleterious mutations can nonetheless be obtained through studies of corresponding 1SS peptide models.

We have characterized the above three clinical mutations in both native state and 1SS contexts. Physico-chemical properties included oxidative folding yield, nascent structure (including CD-defined α-helix contents), stability to chemical denaturation, and stability to reduction under defined redox conditions. A consistent trend in rank order of perturbations was observed: wild-type > SerB24 >> ProA16 > ProB15 (grossly perturbed). Although CD provides only a low-resolution structural probe, the wealth of prior information (including NMR studies of insulin, proinsulin and homologous growth factors) enables construction of molecular models (Figure 9). Intended as working hypotheses, these models highlight the following predicted features:

i. **Parent ensemble.** The overall conformation is globular as a partial fold, stabilized by the confluence of a central B-domain α-helix (green ribbon in Figure 9A) and C-terminal A-domain α-helix (magenta ribbon). The mini-core contains a native-like cluster of nonpolar and aliphatic
side chains (Leu_B11, Val_B12, Leu_B15, Val_B18 and Leu_B16) near cystine B19-A20, extended by packing of a nascent C-terminal B-domain β-strand (Phe_B24 and Tyr_B26).

ii. Ser_B24 ensemble. Substitution of Phe_B24 by Ser would replace aromatic packing against Val_B12, Leu_B15 and Cys_B19 by a small hydrophilic side chain. Our model posits substantial retention of the parent α-helix but with increased flexibility in the N-terminal portion of the B-chain α-helix due to loss of stabilizing Phe_B24-Val_B12 and Phe_B24-Leu_B15 as well as transmitted destabilization of corresponding long-range Tyr_B26 contacts.

iii. Pro_A16 ensemble. Proline has low intrinsic helical propensity (84). Accordingly, substitution of Leu_A16 by Pro would be expected to destabilize the nascent A-domain α-helix and perturb its packing against the B domain. Our model posits substantial retention of the parent B-domain super-secondary structure but with increased flexibility in these elements and as broadly transmitted in the globule.

iv. Pro_B15 ensemble. Substitution of Leu_B15 by Pro would likewise be predicted to destabilize the nascent B-domain α-helix and also introduce multiple long-range perturbations, both to the C-terminal B-domain β-strand (Phe_B24 and Tyr_B26) and within the mini-core. Our model posits N-terminal shortening of both B- and A-domain α-helices in accordance with CD spectra (Figure 5 and Table 2B). These segmental and long-range perturbations would be associated with a global enhancement of conformational fluctuations.

Together, our experimental findings and molecular modeling are in accordance with genotype-phenotype relationships as Ser_B24 (only mildly perturbed in the peptide model) is associated with adult-onset disease (MODY) whereas the more severe Pro substitutions are associated with neonatal onset (PNDM). A direct connection between chemistry and biology was further suggested by correlations between our synthetic and biophysical findings and assays of ER stress induced by the corresponding proinsulin variants in a human cell line. Such coherence among diverse probes validates the present peptide model as a general platform for mechanism-based studies of MIDY mutations mapping near cystine B19-A20. In our companion study in this issue (64), we describe more detailed two-dimensional NMR studies in an effort to deepen the biophysical characterization of this platform. These findings validate major features of the parent model depicted in Figure 9A.

Pairing of Cys_B19 and Cys_A20 represents only a first step in a complex choreography of disulfide pairing leading to the native state (Figure 1C). Indeed, many MIDY mutations map outside of the present folding nucleus as exemplified by substitutions at positions B5, B8 and A4 (18, 20). We envisage that in the future peptide-based strategies can be extended to two-disulfide models that encompass such additional MIDY mutations. Together, reductionist approaches promise to dissect molecular events that underlie aberrant disulfide pairing in a monogenic disease of toxic protein misfolding (17, 18). The evolution of wildtype insulin at the edge of foldability (23) suggests that such studies may inform baseline mechanisms of β-cell ER stress in the natural history of non-syndromic Type 2 DM. Structural lessons of the mutant proinsulin syndrome (20) thus promise to uncover a new layer of understanding in deciphering the informational content of insulin sequences (60). This layer, although of only fleeting importance in biosynthesis and hidden once the native state is reached, may nonetheless underlie the phenomenon of β-cell "exhaustion" in Type 2 DM.

Peptide models can facilitate analysis of protein folding by reducing to a minimum the complexity of a globular protein architecture. In favorable cases such simplification can enable critical determinants of folding efficiency to be dissected. We nonetheless caution that such models can be a double-edged sword: the same simplification can lead other structural contributions to foldability (or kinetic obstacles) to be overlooked. In the present case omission of proinsulin’s C domain is likely to introduce offsetting advantages and disadvantages. On the one hand, 1SS DesDi exhibits a surprising richness of structure amenable to high-resolution NMR study (64). On the other hand, insertion of the long and flexible C domain, which impairs the in vitro refolding efficiency of single-chain insulin analogs, may allow a broader set of native and non-native disulfide bridges to be formed (8), some as off-pathway kinetic traps (43, 85, 86). Further, the present model pertains to only a subset of clinical mutations; other models may be required to investigate mutations distant from cystine B19-A20. These caveats notwithstanding, the complex biophysical chemistry of disulfide pairing in proinsulin, considered in its entirety, poses a foundational problem in protein science, central to the pathogenesis of β-cell dysfunction as a pandemic disease of civilization.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The animal study was reviewed and approved by Institutional Animal Care and Use Committee.

AUTHOR CONTRIBUTIONS

Chemical peptide syntheses were performed by BD, AZ, MJ, and RD. CD studies were performed by BD, MG, and NP. ER stress assays were conducted by Y-SC. NMR studies were performed and interpreted by YY, and MW. Figures were prepared by BD, MG, AZ, Y-SC, and MW. Supplemental rat studies were overseen by NP and FI-B. The Supplemental was prepared by BD, AZ, YY, and MW. All authors contributed to editing the manuscript with first draft prepared by BD, MG, and MW. Overall experimental design and oversight were provided by MW. All authors contributed to the article and approved the submitted version.
FUNDING
This work was supported in part by grants to MW from the National Institutes of Health (R01 DK040949 and R01 DK069764). AZ was supported in part by Calibrum, LLC. NP was supported in part by the American Diabetes Association (grant no. Grants 7-13-IN-31 and 1-08-RA-149). MG was a Pre-doctoral Fellow of the National Institutes of Health (Medical Scientist Training Program 5T32GM007250-38 and Fellowship 1F30DK104618-01).

ACKNOWLEDGMENTS
We thank Prof. T. Gerken (CWRU) for N-terminal peptide sequencing, members of the Ismail-Beigi laboratory (A. Tustan, K. Carr, S. Poordian, R. Grabowski and K. Mackenzie; CWRU) for assistance with supplemental rat studies, and Dr. J.P. Mayer (Indiana University) for advice on peptide synthesis. We also thank Prof. P. Arvan (University of Michigan), M. Liu (Tianjin Research Institute of Endocrinology, Tianjin, China) and R. Wek (IUSM) for advice regarding ER stress assays. We dedicate this article to the memory of our late colleagues, Prof. D.F. Steiner (University of Chicago) and Prof. G.G. Dodson (University of York, U.K.), to whom the senior author (MW) is grateful for their advice and collaboration in the early stages of this work.

REFERENCES
1. Stoy J, Edghill EL, Flanagan SE, Ye H, Paz VP, Pluzhnikov A, et al. Insulin Gene Mutations as a Cause of Permanent Neonatal Diabetes. *Proc Natl Acad Sci USA* (2007) 104(38):15040–4. doi: 10.1073/pnas.0707291104
2. Colombo C, Porzio O, Liu M, Massa O, Vasta M, Salardi S, et al. Seven Mutations in the Human Insulin Gene Linked to Permanent Neonatal/Infancy-Onset Diabetes Mellitus. *J Clin Invest* (2008) 118(6):2148–56. doi: 10.1172/JCI33777
3. Aguilar-Bryan L, Bryan J. Neonatal Diabetes Mellitus. *Endocr Rev* (2008) 29(3):265–91. doi: 10.1210/er.2007-0029
4. Edghill EL, Flanagan SE, Patch A-M, Bousted C, Parrish A, Shields B, et al. Insulin Mutation Screening in 1,044 Patients With Diabetes: Mutations in the INS Gene Are a Common Cause of Neonatal Diabetes But a Rare Cause of Diabetes Diagnosed in Childhood or Adulthood. *Diabetes* (2008) 57(4):1034–42. doi: 10.23736/s0012-1878.07-00150-7
5. Ron D. Proteotoxicity in the Endoplasmic Reticulum: Lessons From the Akita Diabetic Mouse. *J Clin Invest* (2002) 109:443–5. doi: 10.1172/JCI0215020
6. Araki E, Oyadomari S, Mori M. Impact of Endoplasmic Reticulum Stress Pathway on Pancreatic β-Cells and Diabetes Mellitus. *Exp Biol Med* (2003) 228(10):1213–7. doi: 10.17157/153537020322801018
7. Yang Y, Hua QX, Liu J, Shimizu EH, Choquette MH, Mackin RB, et al. Solution Structure of Proinsulin: Connecting Domain Flexibility and Prohormone Processing. *J Biol Chem* (2010) 285:7847–51. doi: 10.1074/jbc.C109.084921
8. Qiao Z-S, Min C-Y, Hua Q-X, Weiss MA, Feng Y-M. *In Vitro* Refolding of Human Proinsulin Kinetic Intermediates, Putative Disulfide-Forming Pathway, Folding Initiation Site, and Potential Role of C-Peptide in Folding Process. *J Biol Chem* (2003) 278(20):17800–9. doi: 10.1074/jbc.M300906200
9. Qiao ZS, Guo ZY, Feng YM. Putative Disulfide-Forming Pathway of Porcine Insulin Precursor During Its Refolding *In Vitro*. *Biochemistry* (2001) 40:2662–8. doi: 10.1021/bi010613r
10. Yan H, Guo ZY, Gong XW, Xi D, Feng YM. A Peptide Model of Proinsulin Folding Intermediate With One Disulfide. *Protein Sci* (2003) 12:768–75. doi: 10.1002/pro.2037203
11. Liu M, Haataja L, Wright J, Wickramasinghe NP, Hua Q-X, Phillips NF, et al. Mutant INS-Gene Induced Diabetes of Youth: Proinsulin Cysteine Residues Impose Dominant-Negative Inhibition on Wild-Type Proinsulin Transport. *PloS One* (2010) 5(10):e13333. doi: 10.1371/journal.pone.001333
12. Liu M, Hodish I, Haataja L, Lara-Lemus R, Rajpal G, Wright J, et al. Proinsulin Misfolding and Diabetes: Mutant INS Gene-Induced Diabetes of Youth. *Trends Endocrinol Metab* (2010) 21(11):652–9. doi: 10.1016/j.tem.2010.07.001
13. Weiss MA. Proinsulin and the Genetics of Diabetes Mellitus. *J Biol Chem* (2009) 284(29):19159–63. doi: 10.1074/jbc.R109.009936
14. Meur G, Simon A, Harun N, Virally M, Dechaume A, Bonnefond A, et al. Insulin Gene Mutations Resulting in Early-Onset Diabetes: Marked Differences in Clinical Presentation, Metabolic Status, and Pathogenic Effect Through Endoplasmic Reticulum Retention. *Diabetes* (2010) 59(3):653–61. doi: 10.2337/db09-0077
15. Weiss MA. Diabetes Mellitus Due to the Toxic Misfolding of Proinsulin Variants. *FEBS Lett* (2013) 587(13):1942–50. doi: 10.1016/j.febslet.2013.04.044
16. Liu M, Hodish I, Rhodes CJ, Arvan P. Proinsulin Maturation, Misfolding, and Proteotoxicity. *Proc Natl Acad Sci USA* (2007) 104(40):15841–6. doi: 10.1073/pnas.0702697104
17. Liu M, Weiss MA, Arunagiri A, Yong J, Rege N, Sun J, et al. Biosynthesis, Structure, and Folding of the Insulin Precursor Protein. *Diabetes Obes Metab* (2018) 20:28–50. doi: 10.1111/dob.13378
18. Stoy J, De Franco E, Ye H, Park S-Y, Bell GJ, Hattersley AT. In Celebration of a Century With Insulin–Update of Insulin Gene Mutations in Diabetes. *Mol Metab* (2021) 52:101280. doi: 10.1016/j.molmet.2021.101280
19. Haataja L, Snapp E, Wright J, Liu M, Hardy AB, Wheeler MB, et al. Proinsulin Intermolecular Interactions During Secretory Trafficking in Pancreatic β Cells. *J Biol Chem* (2013) 288(3):1896–906. doi: 10.1074/jbc.M112.420018
20. Dhayalan B, Chatterjee D, Chen Y-S, Weiss MA. Structural Lessons From the Mutant Proinsulin Syndrome. *Front Endocrinol* (2021) 12:754693. doi: 10.3389/fendo.2021.754693
21. Blundell T, Humblet R. Hormone Families: Pancreatic Hormones and Homologous Growth Factors. *Nature* (1980) 287(5785):781–7. doi: 10.1038/287781a0
22. Nakagawa SH, Zhao M, Hua Q-x, Hu S-Q, Wan Z-l, Jia W, et al. Chiral Mutagenesis of Insulin. Foldability and Function Are Inversely Regulated by a Stereospecific Switch in the B Chain. *Biochemistry* (2005) 44(13):4984–99. doi: 10.1021/bi048025o
23. Rege NK, Liu M, Yang Y, Dhayalan B, Wickramasinghe NP, Chen Y-S, et al. Evolution of Insulin at the Edge of Foldability and Its Medical Implications. *Proc Natl Acad Sci USA* (2020) 117(47):29618–28. doi: 10.1073/pnas.2009891117
24. Miller JA, Narhi LO, Hua QX, Rosenfeld R, Arakawa T, Rohde M, et al. Oxidative Refolding of Insulin-Like Growth Factor 1 Yields Two Products of Similar Thermodynamic Stability: A Bifurcating Protein-Folding Pathway. *Biochemistry* (1993) 32:5203–13. doi: 10.1021/bi00070a032
25. Oas TG, Kim PS. A Peptide Model of a Protein Refolding Intermediate. *Nature* (1988) 336(6194):42–8. doi: 10.1038/336042a0
26. Rheinberger H-J. *A History of Protein Biosynthesis and Ribosome Research. Protein Synthesis and Ribosome Structure*. Weinheim, Germany: Wiley-VCH Verlag (2004). p. 1–51.
27. Baldwin RL. Early Days of Studying the Mechanism of Protein Folding. *Protein Folding Handbook* (2005) 1:3–21. doi: 10.1002/9783527619498.ch1
71. Pace CN, Shaw KL. Linear Extrapolation Method of Analyzing Solvent Denaturation Curves. *Proteins: Struct Funct Bioinform* (2000) 41(54):1–7. doi: 10.1002/1097-0134(2000)41:54<1::AID-PROT10>3.0.CO;2-2

72. Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M, et al. Regulated Translation Initiation Controls Stress-Induced Gene Expression in Mammalian Cells. *Mol Cell* (2000) 6(5):1099–108. doi: 10.1016/S1097-2765(00)01088-8

73. Yan W, Frank CL, Korth MJ, Sopher BL, Novoa I, Ron D, et al. Control of PERK eIF2α Kinase Activity by the Endoplasmic Reticulum Stress-Induced Molecular Chaperone P58IPK. *Proc Natl Acad Sci USA* (2002) 99(25):15920–5. doi: 10.1073/pnas.252341799

74. Nilges M, Clore GM, Gronenborn AM. Determination of Three-Dimensional Structures of Proteins From Interproton Distance Data by Hybrid Distance Geometry-Dynamical Simulated Annealing Calculations. *FEBS Lett* (1998) 229(2):317–24. doi: 10.1016/0014-5793(88)81148-7

75. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR Molecular Structure Determination Package. *J Magn Reson* (2003) 160(1):65–73. doi: 10.1016/S1090-7807(02)00014-9

76. Schwieters CD, Kuszewski JJ, Clore GM. Using Xplor–NIH for NMR Molecular Structure Determination. *Prog Nuc Mag Reson Spectros* (2006) 48(1):47–62. doi: 10.1016/j.pnmrs.2005.10.001

77. Hua QX, Weiss MA. Mechanism of Insulin Fibrillation: The Structure of a Protein in a Kinetic Trap. *J Biol Chem* (2004) 279(20):21449–60. doi: 10.1074/jbc.M314141200

78. Shoelison SE, Lu ZX, Parlautan L, Lynch CS, Weiss MA. Mutations at the Dimer, Hexamer, and Receptor-Binding, Surfaces of Insulin Independently Affect Insulin-Insulin and Insulin-Receptor Interactions. *Biochemistry* (1992) 31:1757–67. doi: 10.1021/bi00121a025

79. Jacoby E, Hua QX, Stern AS, Frank BH, Weiss MA. Structure and Dynamics of a Protein Assembly. 1H-NMR Studies of the 36 kDa R6 Insulin Hexamer. *J Mol Biol* (1996) 258(2):136–57. doi: 10.1006/jmbi.1996.0239

80. Dobson CM. Protein Folding and Misfolding. *Nature* (2003) 426(6968):884–90. doi: 10.1038/nature02261

81. Dodson G, Steiner D. The Role of Assembly in Insulin’s Biosynthesis. *Curr Opin Struct Biol* (1998) 8(2):189–94. doi: 10.1016/S0959-440X(98)80037-7

82. Flanagan SE, De Franco E, Allen HL, Zerah M, Abdul-Rasoul MM, Edge JA, et al. Analysis of Transcription Factors Key for Mouse Pancreatic Development Establishes NFKX2-2 and MNX1 Mutations as Causes of Neonatal Diabetes in Man. *Cell Metab* (2014) 19(1):146–54. doi: 10.1016/j.cmet.2013.11.021

83. Ortolani F, Piccinno E, Grasso V, Papadia F, Panzeca R, Cortese C, et al. Diabetes Associated With Dominant Insulin Gene Mutations: Outcome of 24-Month, Sensor-Augmented Insulin Pump Treatment. *Acta Diabetol* (2016) 53(3):499–501. doi: 10.1007/s00592-015-0793-1

84. Chakrabarty A, Kortemme T, Baldwin RL. Helix Propensities of the Amino Acids Measured in Alanine-Based Peptides Without Helix-Stabilizing Side-Chain Interactions. *Protein Sci* (1994) 3:843–52. doi: 10.1002/pro.5560030514

85. Hua QX, Gozani SN, Chance RE, Hoffmann JA, Frank BH, Weiss MA. Structure of a Protein in a Kinetic Trap. Nat Struct Biol* (1995) 2:129–38. doi: 10.1038/nsb0295-129

86. Hua QX, Jia W, Frank BH, Phillips NB, Weiss MA. A Protein Caught in a Kinetic Trap: Structures and Stabilities of Insulin Disulfide Isomers. *Biochemistry* (2002) 41:14700–15. doi: 10.1021/bi0202981

**Conflict of Interest:** This study received partial funding from Calibrium, LLC. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication. All authors declare no other competing interests.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Dhayalan, Glidden, Zaykov, Chen, Yang, Phillips, Ismail-Beigi, Jarosinski, DiMarchi and Weiss. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.