



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

OPEN ACCESS

Edited by:

Jiajun Zhao,
Shandong Provincial Hospital, China

Reviewed by:

Ya-Xiong Tao,
Auburn University, United States
Pierre De Meyts,
Université Catholique de Louvain,
Belgium

*Correspondence:

Balamurugan Dhayalan
bdhayal@iu.edu
Michael A. Weiss
weissma@iu.edu

†Present address:

Alexander N. Zaykov,
Novo-Nordisk Research Center,
Seattle, WA, United States

‡These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Diabetes: Molecular Mechanisms,
a section of the journal
Frontiers in Endocrinology

Received: 23 November 2021

Accepted: 17 January 2022

Published: 01 March 2022

Citation:

Dhayalan B, Glidden MD, Zaykov AN,
Chen Y-S, Yang Y, Phillips NB,
Ismail-Beigi F, Jarosinski MA,
DiMarchi RD and Weiss MA
(2022) Peptide Model of the
Mutant Proinsulin Syndrome. I.
Design and Clinical Correlation.
Front. Endocrinol. 13:821069.
doi: 10.3389/fendo.2022.821069

Balamurugan Dhayalan^{1*†}, Michael D. Glidden^{2,3,4†}, Alexander N. Zaykov^{5†},
Yen-Shan Chen¹, Yanwu Yang¹, Nelson B. Phillips², Faramarz Ismail-Beigi^{2,3,4},
Mark A. Jarosinski¹, Richard D. DiMarchi⁵ and Michael A. Weiss^{1*}

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 (Leu^{B15}→Pro, Leu^{A16}→Pro, and Phe^{B24}→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 Diabetes 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 Lys^{B28} 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., Phe^{B24} 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 Pro^{B28} 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 His^{B10}→Asp (native α -helix α_1 spanning residues B9-B19) and Thr^{A8}→Glu (native α -helix α_2 ; A1-A8) (58, 59). Representative MIDY mutations (Leu^{B15}→Pro, Leu^{A16}→Pro, and Phe^{B24}→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 Ser^{B24} [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 ¹H-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

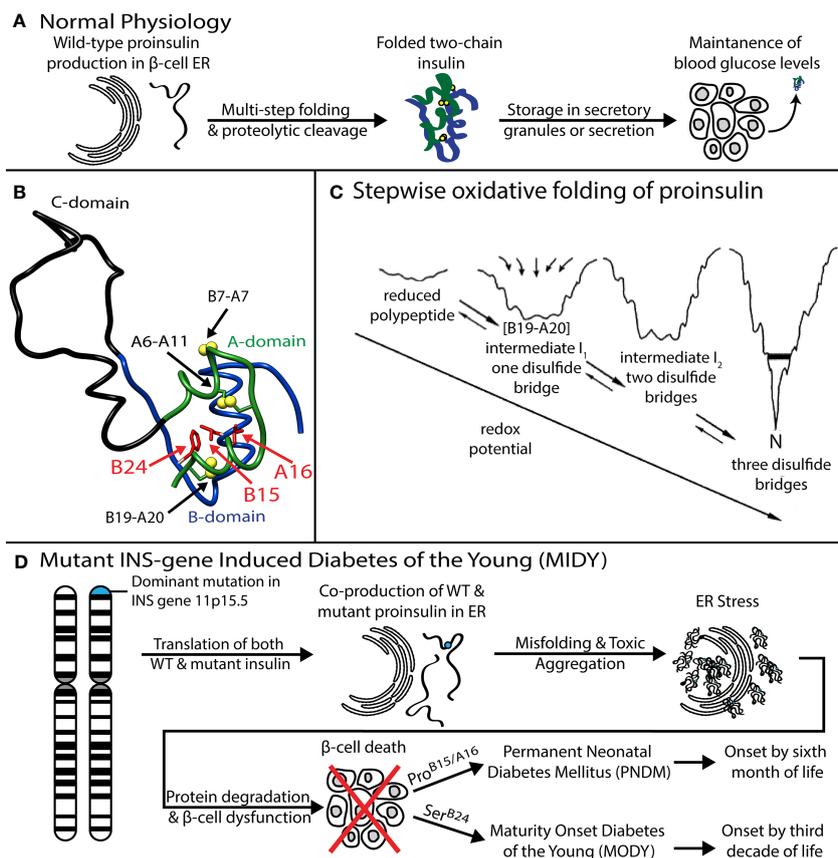


FIGURE 1 | Genetic etiology, pathophysiology, and structural basis of Mutant Proinsulin Syndrome. **(A)** Overview of normal proinsulin production and maintenance of glucose homeostasis. **(B)** Solution structure of WT proinsulin (7) showing the A- (green), B- (blue), and C- (black) domains and three native disulfide bridges (yellow spheres). Wild-type residues Phe^{B24}, Leu^{B15} and Leu^{A16}, which are examined in this study as sites of mutation, are shown as red side chains. **(C)** Folding of proinsulin proceeds via sequential disulfide linkage steps from the fully unfolded state through one (1SS), two (2SS), and three disulfide bonded (3SS) conformational ensembles before reaching the native state (N). Chemical trapping experiments showed that the formation of B19-A20 disulfide as the predominant key first step in the hierarchical disulfide pathway (8–10). Reprinted with permission from Hua QX, et al. *Biochemistry*. 2001;40:12299-311. Copyright (2001) American Chemical Society. **(D)** Dominant mutations in the *INS* gene causing co-production of both WT and mutant proinsulin, leading to ER stress, β -cell dysfunction, and, ultimately, diabetes. Clinically identified mutations Ser^{B24}, Pro^{B15} and Pro^{A16} present with a spectrum of disease severity and age of diabetes onset.

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 (Cys^{B19} and Cys^{A20}; 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 (25). 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.

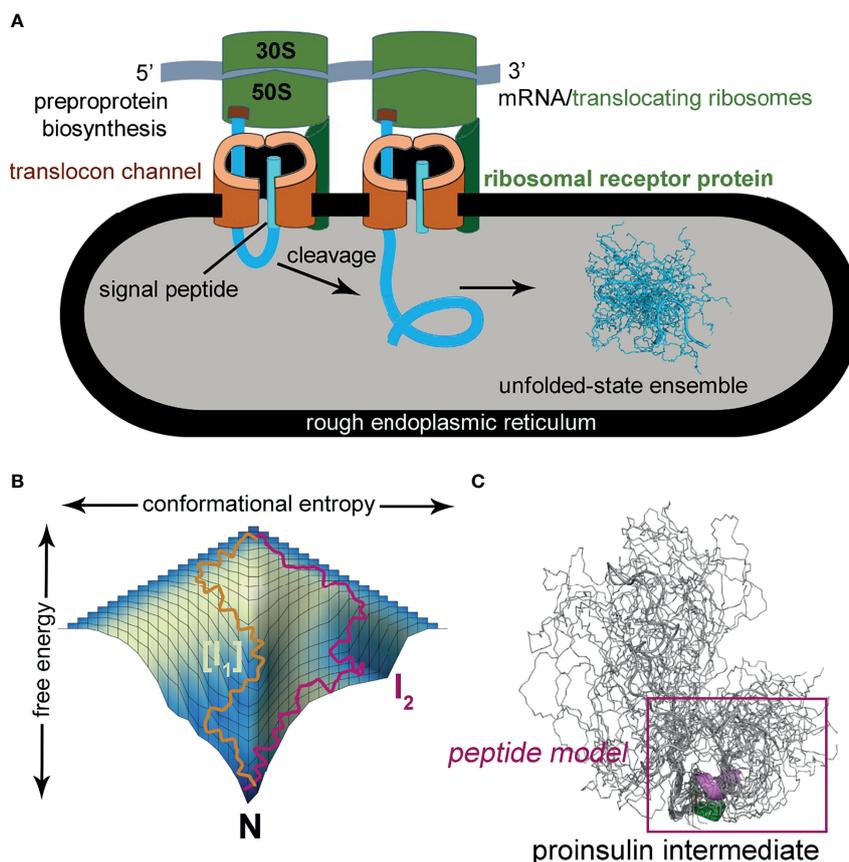


FIGURE 2 | Peptide models of protein-folding intermediates probe critical biosynthetic steps. **(A)** Preproprotein biosynthesis in rough endoplasmic reticulum: nascent polypeptide chain (blue ribbon) enters ER (gray) with cleavage of signal peptide (cyan ribbon) via translocon channel (brown). Protein folding may in part be co-translational or begin from an unfolded-state ensemble (blue chains at right); this model was calculated by Xplor-NIH software and ensemble generated by PyMOL (ribbon and sticks). The cartoon was otherwise adapted from (56). **(B)** Energy folding landscape. Yellow and magenta lines indicate alternative paths to the native state (N): I_1 , set of unobserved intermediate species; I_2 , observed intermediate that can accumulate in a kinetic trap. Panel adapted from reference (57). **(C)** Conformational ensemble of a one-disulfide proinsulin intermediate containing cystine B19-A20. A minimal model (magenta box) was designed based on a 49-residue peptide (50) as described in the main text. The present CD studies suggest formation of nascent α -helices in the central B domain and C-terminal A domain (dark green and magenta, respectively).

solid-phase fluorenylmethyloxycarbonyl (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 [afgbde], cycle-#aa+2 [ffbd], 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. α -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 verification 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 dissolved in 20 mM glycine 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. Peptide purity 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 8.5) at 12°C water bath for 24 h. After analytical rp-HPLC indicated two-chain conversion (typically 60–80%), the reaction mixture was acidified and purified on semi-preparative rp-HPLC. Fractions containing clean protein were pooled, lyophilized and masses were confirmed by LC-MS.

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, ¹⁹F, 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. Automated macros were utilized that acquired full far-UV CD spectra (255–190 nm) in 2°C steps from 4°C to 40°C (plus 25°C and 37°C) with a wavelength resolution of 0.5 nm and 30 sec. photocount averaging time. Following this, 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 obtained 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, $\langle [\theta]_{222\pm 1nm} \rangle$, 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.

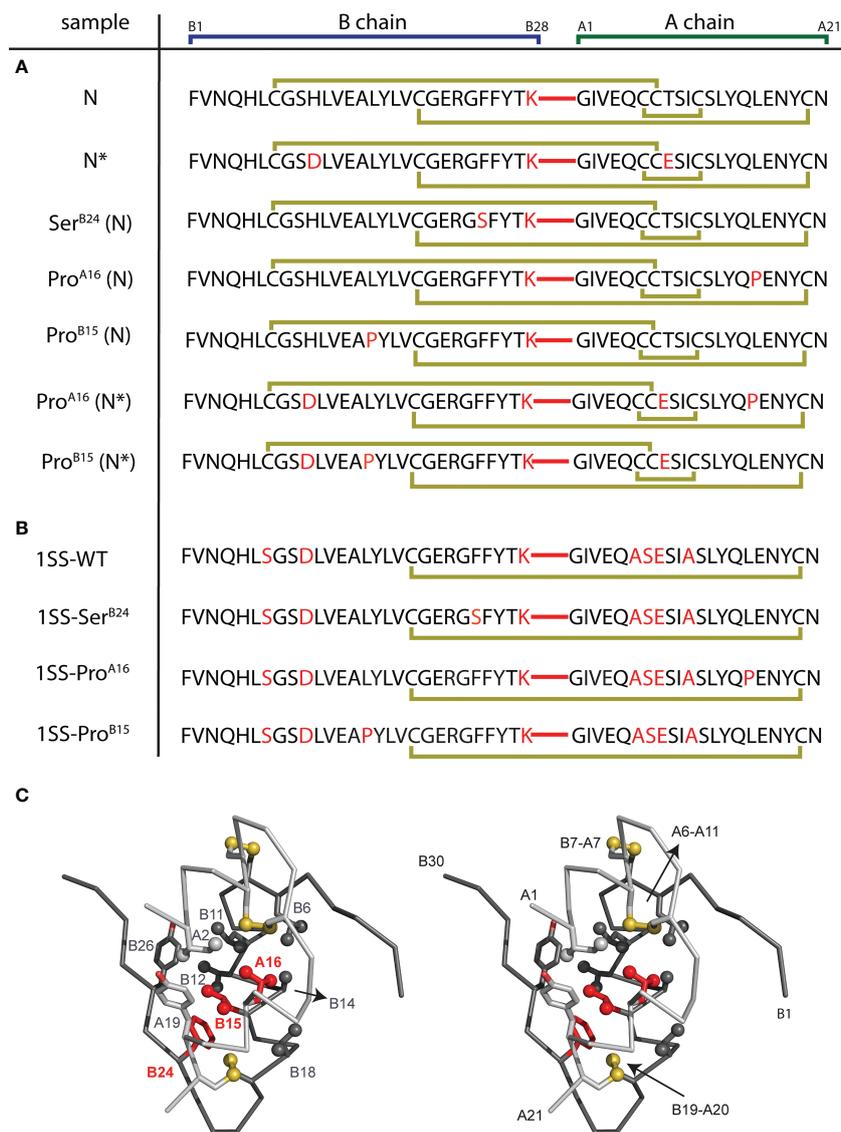


FIGURE 3 | DesDi variant protein sequences. **(A)** Sample names and sequences of three disulfide bond containing analogs (N and N*). **(B)** Sample names and sequences of one disulfide-bond-containing analogs ([B19-A20]; abbreviated as 1SS). Yellow lines show disulfide linkages. A red line connecting the C-terminal B-domain and N-terminal A-domain sequences signifies the presence of a peptide bond between residues B28 and A1. All 1SS samples and N* have additional mutations Glu^{A8} and Asp^{B10} to enhance solubility at neutral pH (e.g., to enable the high protein concentrations needed for NMR spectroscopy). Residues in red highlight mutations to the WT insulin sequence. **(C)** Stereo view showing ribbon model of an insulin monomer [as extracted from the T₆ Zn hexamer; PDB entry 4INS (60)]. Sites of clinical mutation investigated here [B15, B24 and A16; highlighted in red] and neighboring residues are shown as sticks. Sulfur atoms in disulfide bridges are shown as gold spheres, and selected side-chain methyl groups as spheres (one-third van der Waals radius). A- and B chains are otherwise shown in light and dark gray, respectively.

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 ISS 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, $[\Theta]$, were calculated by dividing raw ellipticities by the corrected protein concentration ($[C]_{cal} = [C]_{KP}/M_G$) 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):

$$\Theta(c) = \frac{\Theta_A + \Theta_B e^{(-\Delta G - mc)/RT}}{1 + e^{(-\Delta G - mc)/RT}} \quad [1]$$

where ΔG is the Gibbs free energy of unfolding, C is guanidine concentration, R is the ideal gas law constant, T is temperature, and Θ_A and Θ_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

ISS-peptides (50 μ g, 45 μ M final concentration) were treated with 25 mM reduced glutathione (GSH) and 5 mM oxidized glutathione (GSSH) in 200 μ L 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 QuikChangeTM (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 μ L/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 ISS-DesDi variants; selected distance restraints were extracted from NOESY spectrum of an engineered insulin monomer (23). For the parent ISS DesDi

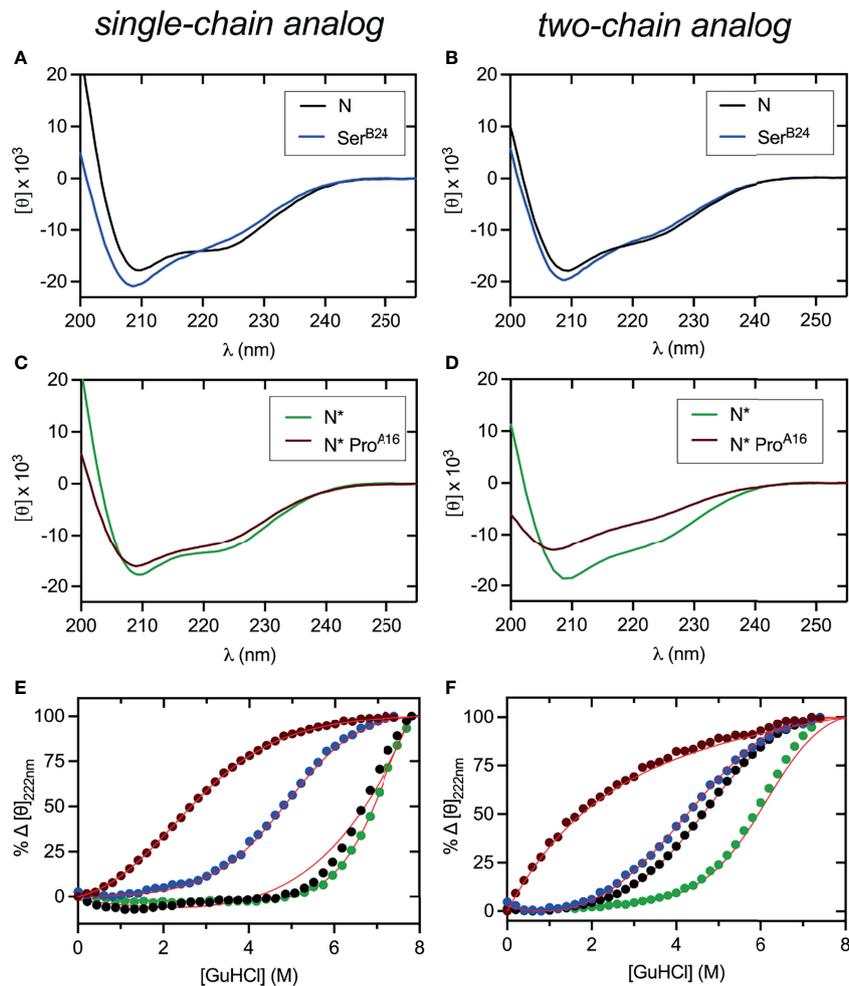


FIGURE 4 | Far UV-CD and stability studies of three disulfide containing DesDi variants. **(A, B)** CD wavelength scans of samples (N) acquired at 25°C. **(C, D)** CD wavelength scans of N*-samples acquired at 25°C. In panels **(A–D)**, left hand side represent single-chain data and right-hand side two-chain data. **(E, F)** Comparative CD-guanidine titrations of all three disulfide samples (solid lines are fits). Color legend: N (black), N* (green), Ser^{B24} (blue) and N* Pro^{A16} (maroon). ΔG_u values obtained by titration curve fitting are presented in **Table 1**. All titrations were performed at 25°C, except for N (37°C), Ser^{B24} (37°C) and N* (50°C) single-chain samples where higher temperature was used to enhance their unfolding transition.

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 [Asp^{B10}, Glu^{A8}]-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 Ser^{B24} (least perturbed relative to the N parent peptide) \gg Pro^{A16} = Pro^{B15} (no folded product detected). Introduction of N* substitutions Asp^{B10} and Glu^{A8} rescued inefficient but detectable folding of Pro^{A16}-N* but not Pro^{B15}-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 Pro^{A16} and Pro^{B15}. 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

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

native state peptide domain	isolated yield (mg) ^a	ΔG_u (kcal/mol) ^b	<i>m</i> value(kcal/mol/M)	C_{mid} (M)
single-chain (N) ^c	21.7 mg	>4 ^c	N.D. ^c	N.D. ^c
single-chain (N*) ^c	21.2 mg	7.6 ± 0.4 ^c	0.99 ± 0.02	7.68 ± 0.1
Ser ^{B24} single-chain (N) ^c	12.0 mg	3.4 ± 0.1 ^c	0.69 ± 0.02	4.93 ± 0.1
Pro ^{A16} single-chain (N) ^d	0	–	–	–
Pro ^{A16} single-chain (N*) ^d	2.1 mg	0.9 ± 0.1	0.51 ± 0.02	1.76 ± 0.1
Pro ^{B15} single-chain (N) ^e	0	–	–	–
Pro ^{B15} single-chain (N*) ^e	0	–	–	–

^aIsolated yield represents amounts after *rp*-HPLC purification from a folding reaction of 100 mg reduced polypeptide.

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

^cCD-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 ($R^2 = 0.9994$), but analysis of the N titration curve could only place a lower limit on thermodynamic stability of 4 kcal/mol.

^dNo product obtained. Folding was nonetheless rescued in part by Asp^{B10} and Glu^{A8} substitutions (N*).

^eNo isolable product obtained. In the case of Pro^{B15}, inclusion of stabilizing substitutions Asp^{B10} and Glu^{A8} (N*) did not rescue foldability; only disulfide isomers were observed.

the Lys^{B28}-Gly^{A1} peptide bond (right-hand panels in **Figure 4**).¹ MIDY mutations Ser^{B24} and Pro^{A16} are each associated with reduced α -helix content (**Table 2**).² The extent of perturbation was more marked in the case of Pro^{A16}, 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_u ; **Table S1**, column 3). In accordance with their relative susceptibilities to guanidine-induced unfolding, apparent by qualitative inspection of the denaturation data, Pro^{A16} is more profoundly destabilizing than is Ser^{B24}. 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 Pro^{A16} N* analog did not exhibit a cooperative

unfolding transition, and so its stability could not be estimated by this method. The stabilities of corresponding Pro^{B15} native state analogs could not be assessed due to absence of folded product. Functional studies of two-chain versions of Ser^{B24}-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 Pro^{A16}-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-UV spectra at 4, 25 and 37°C (**Figures 5A–C**) suggest rank order N* (most structured) > N > insulin *lispro* >> parent ISS model > ISS-Ser^{B24} > ISS-Pro^{A16} > ISS Pro^{B15} (least structured). These inferences 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

¹Normalized CD spectra are unaffected by changes in peptide concentration in the 20–90 μ M range.

²CD -based estimates of secondary structure content are in good agreement with published values based on SELCON-3 deconvolution (77).

TABLE 2 | CD-derived secondary-structure contents^a.

sample Name	α -helix	β -sheet	disordered
A)			
insulin <i>lispro</i>	38 – 40%	9 – 10%	32 – 33%
wild-type insulin ^b	47 – 56%	3 – 10%	24 – 29%
single-chain (N)	49 – 64%	2 – 10%	24 – 27%
single-chain (N*)	49 – 58%	4 – 10%	24 – 26%
Ser ^{B24} single-chain (N)	55 – 61%	3 – 7%	27 – 28%
Pro ^{A16} single-chain (N*)	43 – 48%	11 – 14%	27 – 28%
two-chain DesDi (N)	47 – 53%	10 – 12%	26 – 28%
two-chain DesDi (N*)	56 – 61%	7 – 10%	24 – 26%
two-chain Ser ^{B24} (N)	50 – 58%	5 – 6%	27 – 29%
two-chain Pro ^{A16} (N*)	25 – 26%	22 – 23%	31 – 32%
B)			
1SS-WT	27 – 30%	18 – 20%	30 – 31%
1SS-Ser ^{B24}	19 – 20%	26 – 27%	31 – 32%
1SS-Pro ^{A16}	12 – 14%	32 – 33%	30 – 31%
1SS-Pro ^{B15}	10 – 11%	29 – 32%	32 – 34%

^aTotal percent α -helix, β -sheet, and disordered coil were obtained from spectra acquired at discreet temperatures of 4°C, 25°C, and 37°C using the SELCON-3 algorithm (66–68). Estimated percentages are presented as the minimum to maximum content calculated across the three sampled temperatures.

^bEstimates 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).

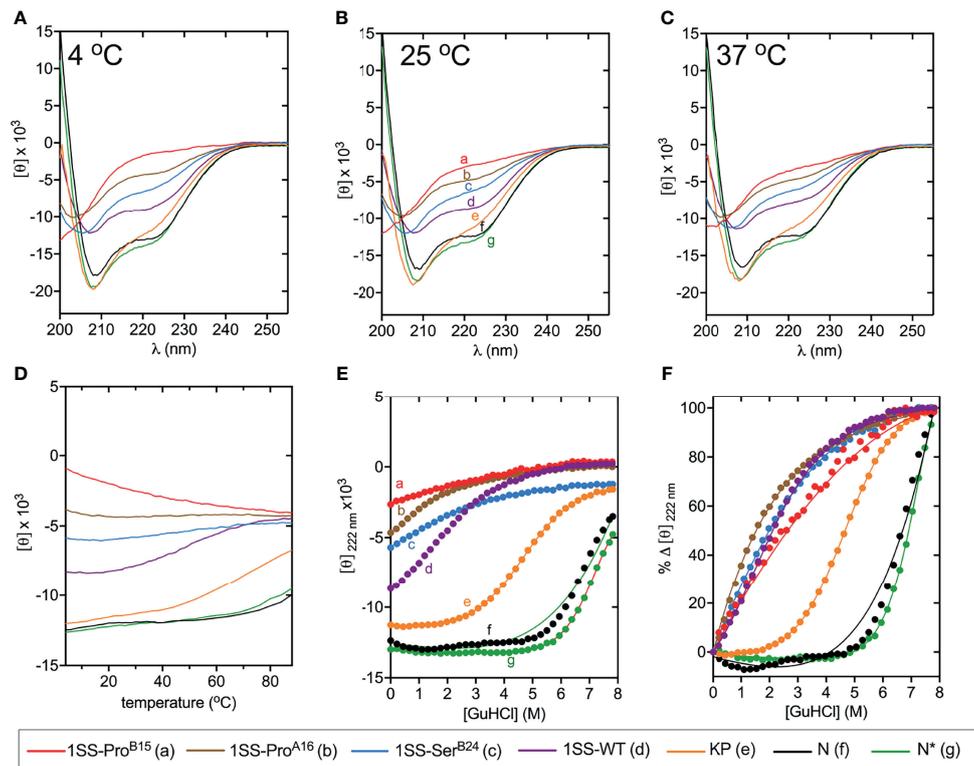


FIGURE 5 | Folding and thermodynamic stability of 1SS variants and controls. **(A–C)** CD wavelength scans of all 1SS variants, N, and N* acquired at **(A)** 4°C, **(B)** 25°C, and **(C)** 37°C. The weak temperature dependence of each secondary structure is visualized with **(D)** a plot of $\langle[\theta]_{222\pm 1\text{nm}}\rangle$ versus temperature. **(E, F)** Comparative CD-guanidine titrations of insulin *lispro* and all 1SS samples (solid lines are fits). Color legend: insulin *lispro* (orange), N* (green), N (black), 1SS-WT (purple), 1SS-Ser^{B24} (blue), 1SS-Pro^{A16} (brown) and 1SS-Pro^{B15} (red). ΔG_u values obtained by titration curve fitting are presented in **Table 1** and **Table S1**. To enhance their unfolding transition, N and N* titrations were performed at 37°C and 50°C, respectively; all other titrations were performed at 25°C.

temperatures³ (**Figure 5D**) and possibly a cooperative guanidine-denaturation transition with $\Delta G_u < 1$ 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 α -helical structure occurs in the B domain (residues B9-B19; helix α_1 in insulin and proinsulin) and A domain (A12-A18; helix α_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 cysteine 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 ¹H-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 ¹³C_α and ¹H_α 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

³For single-chain N and N* parent domains, their limited temperature dependence of $\langle[\theta]_{222\pm 1\text{nm}}\rangle$ in the range 4–40°C correlates with enhanced folding efficiency and increased thermodynamic stability.

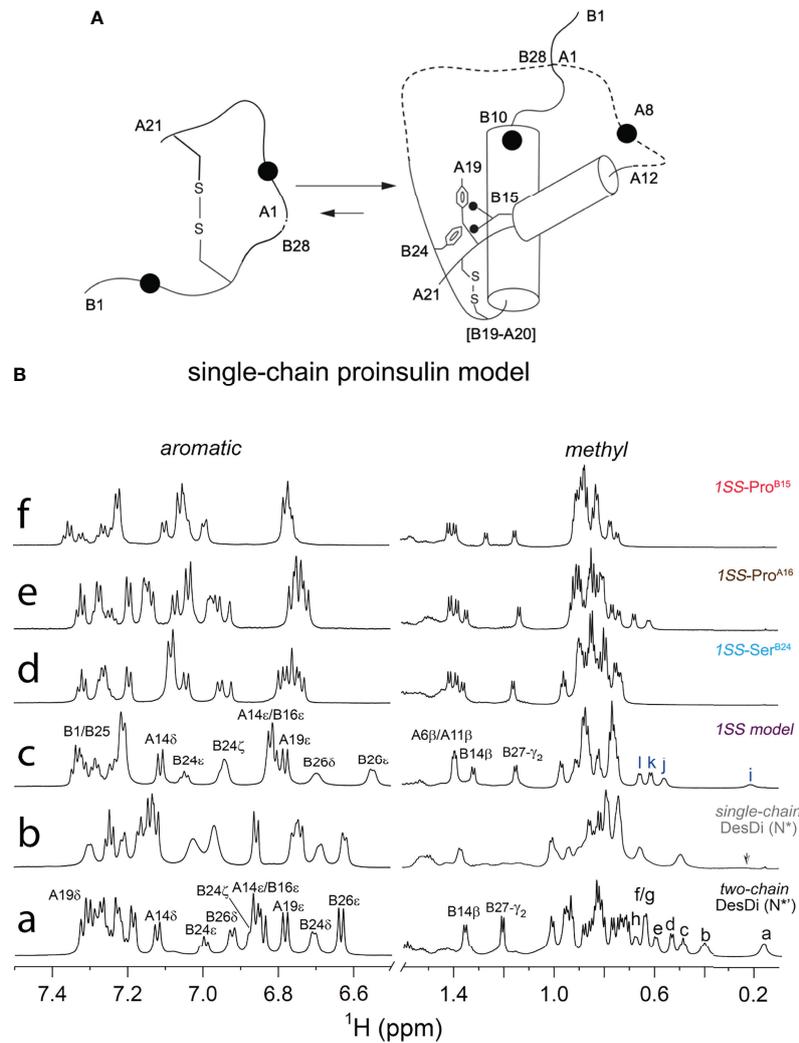


FIGURE 6 | Folding pathway of proinsulin model and NMR spectroscopy. **(A)** Model of proinsulin's specific folding nucleus based on single-chain 1SS-DesDi analog. Initial pairing of B19-A20 disulfide in combination with the formation of two-helices (B9-B19 and A12-A20) are considered to be central events in the formation of folding nucleus. Dotted lines represent disordered regions. Larger circles represent the Asp^{B10} and Glu^{A8} substitutions. Key residues are highlighted by their sequence position [panel modified from reference (51)] **(B)** Stack plot of 1D ¹H-NMR spectra of DesDi insulin analogs: aromatic (*left panel*) and methyl (*right panel*) region. (a) two-chain DesDi (N^{*}). Selected resonance assignments are as indicated in the top of signals. The well-resolved methyl resonances are indicated as a: Leu^{B15} δ₁-CH₃; b: Ile^{A2} δ₁-CH₃; c: Ile^{A10} δ₁-CH₃; d: Leu^{B15} δ₂-CH₃; e: Ile^{A2} γ₂-CH₃; f/g: Ile^{A2} γ₂-CH₃ and Leu^{B11} δ₁-CH₃; h: Leu^{B11} δ₂-CH₃; (b) single-chain DesDi (N^{*}). The arrow indicates broadening signal of Leu^{B15} δ₁-CH₃. (c) single-chain DesDi model with one disulfide bond at [B19-A20]. Cys^{A6} and Cys^{A11} were replaced by alanine, and Cys^{A7} and Cys^{B7} were mutated by serine. The 1SS model exhibits a spectral property of native-like insulin as observed in the aromatic region and in the upfield-shifted methyl region (*far right*). Selected resonance assignments are as indicated in the top of signals. The well-resolved methyl resonances are indicated as i: Leu^{B15} δ₁-CH₃; j: Leu^{B11} δ₁-CH₃; k: Leu^{B15} δ₂-CH₃ and l: Leu^{B11} δ₂-CH₃. (d) single-chain 1SS-Ser^{B24} analog. (e) single-chain 1SS-Pro^{A16} analog and (f) single-chain 1SS-Pro^{B15} analog. Aromatic and methyl resonances of signature residues for three 1SS variants shifted to downfield, as well as exhibited a reduction in chemical-shift dispersion. Spectra were acquired at pD 7.4 (direct meter reading) at 35°C in D₂O.

redox stability parent 1SS model (most stable) > 1SS-Ser^{B24} > 1SS-Pro^{A16} >> 1SS-Pro^{B15} (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 α-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 α-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 β-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

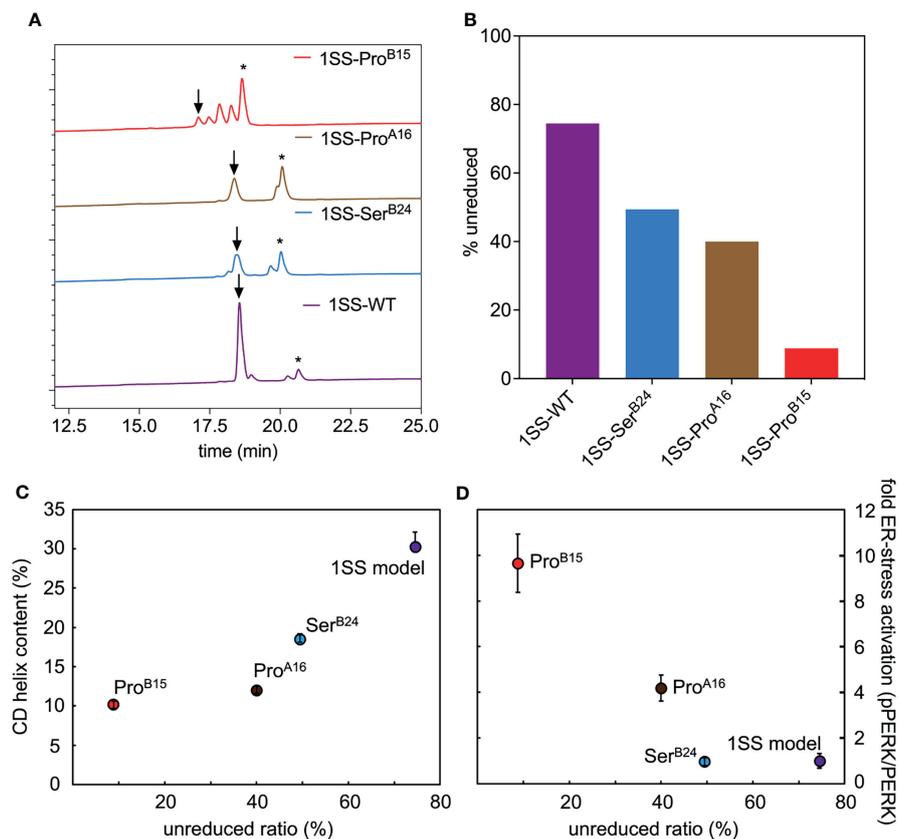


FIGURE 7 | Determination of equilibrium ratios of 1SS analogs in redox buffer containing glutathione reduced (GSH) and glutathione oxidized (GSSH) in 5:1 ratio. **(A)** Reverse-phase HPLC profiles of the equilibrium reaction mixtures after 3 h at 25 °C. Downward arrow indicates the remaining amount of polypeptide (in oxidized state) and asterisk (*) indicates the fully reduced polypeptide. Glutathione adducts were seen intermediate of the oxidized and reduced peaks. **(B)** Bar graph showing percentage remaining unreduced 1SS-polypeptides in the GSH-GSSH redox equilibrium buffer. **(C, D)** Correlations between the stability of 1SS-peptides in redox buffer with other biophysical and biological parameters reflects the same trend as the clinical severity of these mutants. Plots showing the amount of remaining unreduced 1SS-peptides correlates with their CD α -helical content **(C)** and ER-stress levels in a cell-based assay **(D)**.

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 Pro^{B15} (highest ER stress) > Pro^{A16} > Ser^{B24} > 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 Cys^{B19} and Cys^{A20}, an early step in biosynthesis (8, 24). We exploited this platform to investigate three clinical mutations, two with neonatal onset [Pro^{B15} and Pro^{A16} (82, 83)] and one with onset in early adulthood [Ser^{B24} (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.

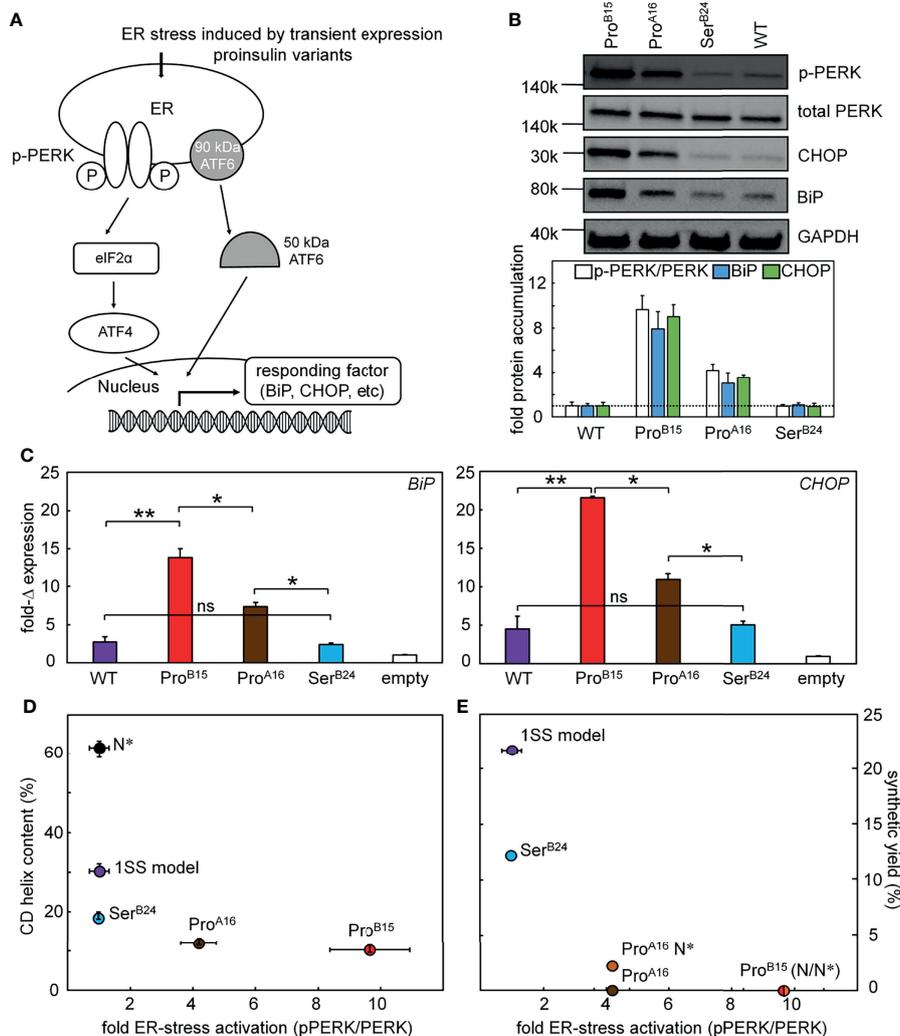


FIGURE 8 | Proinsulin mutants induce ER stress. **(A)** Schematic picture shows that the induced ER stress would be regulated by the expression of different proinsulin variants. Proinsulin variants give rise to accumulation of unfolded protein in the ER resulting in the phosphorylation of PERK and induction of ISR. As common markers, CHOP is a downstream response of the ISR and BiP is a chaperon, which would exhibit increased expression. Western blot was applied to test the protein levels of p-PERK and the accumulations of BiP and CHOP in panel **(B)** The transcription responses of these two ER stress markers were monitored by qPCR assay in panel **(C)**. **(B)** Western-blot assays probing the ER stress markers including p-PERK/PERK, CHOP, and BiP. A total of 3 blots provided ER stress markers: (i) p-PERK alone (due to a specific protocol requirement for antibody), (ii) PERK and GAPDH and (iii) BiP and CHOP. Samples were obtained from the same lysate. WT proinsulin induces low ER stress. Representative gel images (above) shows the WB results and the histogram (bottom) presents the quantification of the WB signals ($n=3$, biological replicates). **(C)** Real-time qPCR assay probing the transcription responses of *BIP* and *CHOP* genes induced by WT (purple) and variants of proinsulin. Gene markers for ER stress were significantly activated by the expressions of Pro^{B15} (red) and Pro^{A16} (brown) variants but not the Ser^{B24} (blue) proinsulin. Endogenous extent of GAPDH mRNA is applied as internal control to normalize the quantitative analysis of PCR gene expression. Asterisks (*) and (**) indicate p -value < 0.05 and < 0.01 . The “ns” indicates p -value > 0.05 (ns - not significant). **(D, E)** Plots show the relationship between cellular ER stress activation and biophysics phenomena. **(D)** The plot shows the relationship between extents of fold of activated ER stress associated with different proinsulin variants (X-axis; using the representative ER stress marker pPERK/PERK ratio) and their percentage of helix contents evaluated by CD (Y-axis). **(E)** Plot showing the result of activated cellular ER stress correlating to the synthetic yield of these analogs in a three-disulfide containing DesDi model. N* indicates Asp^{B10} and Glu^{A8} substitutions that are used to rescue the folding in those cases.

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 Lys^{B28}-Gly^{A1} peptide bond (cleavable by protease Lys-C) enables

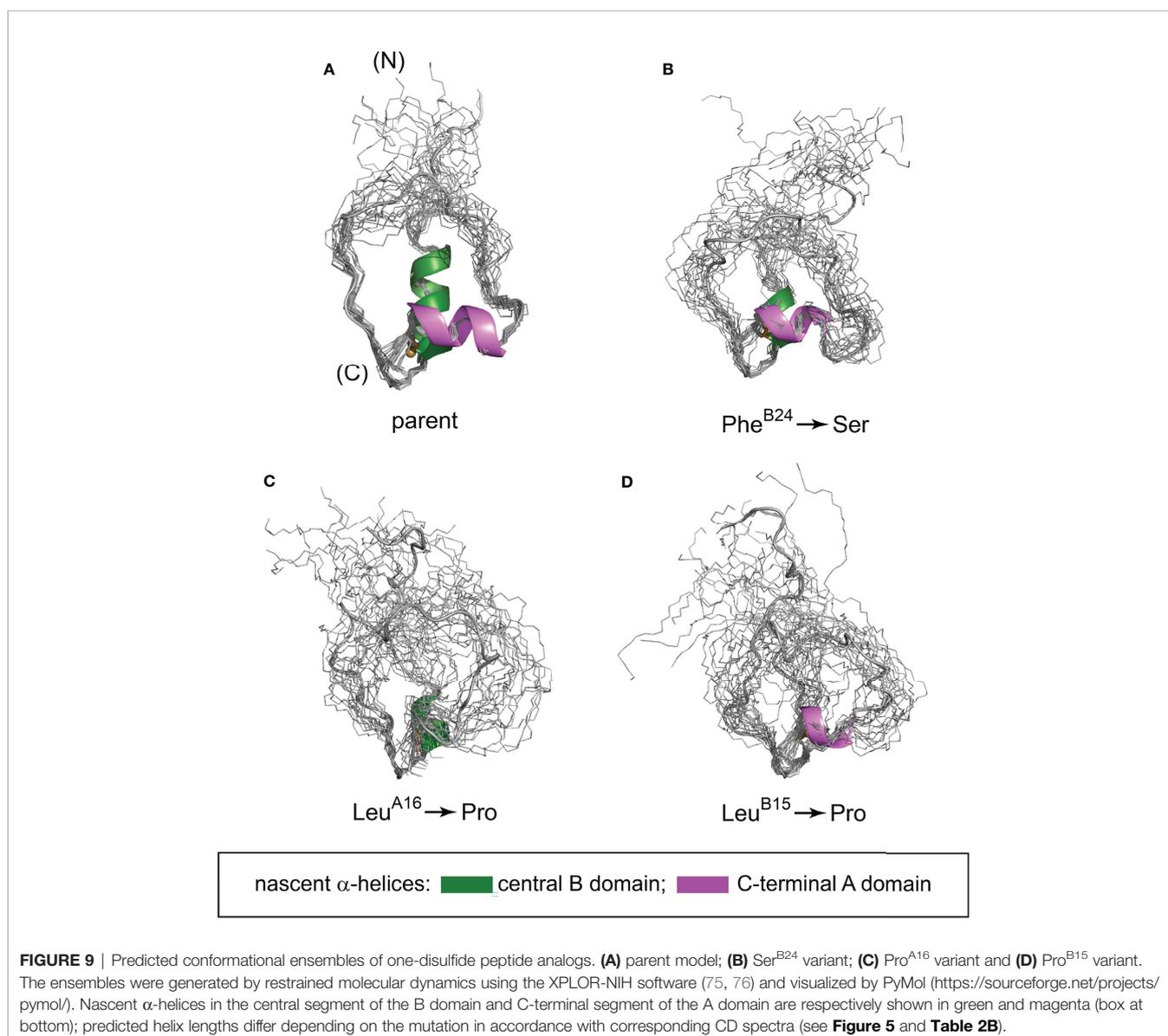
productive folding even of mutant insulins otherwise refractory to classical insulin chain combination (50), presumably by constraining the orientation of A- and B-chain residues to stabilize a specific folding nucleus. We obtained a one-disulfide model through pairwise substitutions of cystines B7-A7 (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 His^{B10}→Asp (58) and Thr^{A8}→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 Pro^{A16} analog, albeit in small yield. Pro^{B15} blocks folding of DesDi even in the presence of Asp^{B10} and Glu^{A8}. 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 > Ser^{B24} >> Pro^{A16} > Pro^{B15} (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^{A16}) 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 -helices 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 Supplement 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 were 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)

REFERENCES

- 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
- 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
- Aguilar-Bryan L, Bryan J. Neonatal Diabetes Mellitus. *Endocr Rev* (2008) 29(3):265–91. doi: 10.1210/er.2007-0029
- Edghill EL, Flanagan SE, Patch A-M, Boustred 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: 10.2337/db07-1405
- Ron D. Proteotoxicity in the Endoplasmic Reticulum: Lessons From the Akita Diabetic Mouse. *J Clin Invest* (2002) 109:443–5. doi: 10.1172/JCI0215020
- 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: 10.1177/153537020322801018
- 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
- 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
- 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/bi001613r
- Yan H, Guo ZY, Gong XW, Xi D, Feng YM. A Peptide Model of Insulin Folding Intermediate With One Disulfide. *Protein Sci* (2003) 12:768–75. doi: 10.1110/ps.0237203
- 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.0013333
- 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
- Weiss MA. Proinsulin and the Genetics of Diabetes Mellitus. *J Biol Chem* (2009) 284(29):19159–63. doi: 10.1074/jbc.R109.009936

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.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2022.821069/full#supplementary-material>

- 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/db09-1091
- 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
- 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
- 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/dom.13378
- Stoy J, De Franco E, Ye H, Park S-Y, Bell GI, 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
- 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
- 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
- Blundell T, Humbel R. Hormone Families: Pancreatic Hormones and Homologous Growth Factors. *Nature* (1980) 287(5785):781–7. doi: 10.1038/287781a0
- Nakagawa SH, Zhao M, Hua Q-x, Hu S-Q, Wan Z-I, 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
- 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.2010908117
- 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
- Oas TG, Kim PS. A Peptide Model of a Protein Folding Intermediate. *Nature* (1988) 336(6194):42–8. doi: 10.1038/336042a0
- 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.
- Baldwin RL. Early Days of Studying the Mechanism of Protein Folding. *Protein Folding Handbook* (2005) 1:3–21. doi: 10.1002/9783527619498.ch1

28. Hartl FU, Hayer-Hartl M. Converging Concepts of Protein Folding *In Vitro* and *In Vivo*. *Nat Struct Mol Biol* (2009) 16(6):574. doi: 10.1038/nsmb.1591
29. Pederson T. The Ribosome: A Structural Biology Triumph Offering New Horizons. *FASEB J* (2019) 33(4):4655–56. doi: 10.1096/fj.190401ufm
30. Cassaignau AM, Cabrita LD, Christodoulou J. How Does the Ribosome Fold the Proteome? *Annu Rev Biochem* (2020) 89:389–415. doi: 10.1146/annurev-biochem-062917-012226
31. Onuchic JN, Nymeyer H, Garcia AE, Chahine J, Succi ND. The Energy Landscape Theory of Protein Folding: Insights Into Folding Mechanisms and Scenarios. *Adv Protein Chem* (2000) 53:87–152. doi: 10.1016/S0065-3233(00)53003-4
32. Baker D. A Surprising Simplicity to Protein Folding. *Nature* (2000) 405(6782):39–42. doi: 10.1038/35011000
33. Baldwin RL. The Nature of Protein Folding Pathways: The Classical Versus the New View. *J Biomol NMR* (1995) 5(2):103–9. doi: 10.1007/BF00208801
34. Englander SW, Mayne L. The Nature of Protein Folding Pathways. *Proc Natl Acad Sci USA* (2014) 111(45):15873–80. doi: 10.1073/pnas.1411798111
35. Arunagiri A, Haataja L, Cunningham CN, Shrestha N, Tsai B, Qi L, et al. Misfolded Proinsulin in the Endoplasmic Reticulum During Development of Beta Cell Failure in Diabetes. *Ann NY Acad Sci* (2018) 1418(1):5. doi: 10.1111/nyas.13531
36. Ghosh R, Colon-Negron K, Papa FR. Endoplasmic Reticulum Stress, Degeneration of Pancreatic Islet β -Cells, and Therapeutic Modulation of the Unfolded Protein Response in Diabetes. *Mol Metab* (2019) 27:S60–8. doi: 10.1016/j.molmet.2019.06.012
37. Gemmer M, Förster F. A Clearer Picture of the ER Translocon Complex. *J Cell Sci* (2020) 133(3):jcs231340. doi: 10.1242/jcs.231340
38. Fass D, Thorpe C. Chemistry and Enzymology of Disulfide Cross-Linking in Proteins. *Chem Rev* (2018) 118(3):1169–98. doi: 10.1021/acs.chemrev.7b00123
39. Hwang J, Qi L. Quality Control in the Endoplasmic Reticulum: Crosstalk Between ERAD and UPR Pathways. *Trends Biochem Sci* (2018) 43(8):593–605. doi: 10.1016/j.tibs.2018.06.005
40. Ushioda R, Nagata K. Redox-Mediated Regulatory Mechanisms of Endoplasmic Reticulum Homeostasis. *Cold Spring Harb Perspect Biol* (2019) 11(5):a033910. doi: 10.1101/cshperspect.a033910
41. Phillips BP, Gomez-Navarro N, Miller EA. Protein Quality Control in the Endoplasmic Reticulum. *Curr Opin Cell Biol* (2020) 65:96–102. doi: 10.1016/j.ccb.2020.04.002
42. Robinson PJ, Bulleid NJ. Mechanisms of Disulfide Bond Formation in Nascent Polypeptides Entering the Secretory Pathway. *Cells* (2020) 9(9):1994. doi: 10.3390/cells9091994
43. Guo Z-Y, Qiao Z-S, Feng Y-M. The *In Vitro* Oxidative Folding of the Insulin Superfamily. *Antioxid Redox Signal* (2008) 10(1):127–40. doi: 10.1089/ars.2007.1860
44. Lazaridis T, Karplus M. "New View" of Protein Folding Reconciled With the Old Through Multiple Unfolding Simulations. *Science* (1997) 278(5345):1928–31. doi: 10.1126/science.278.5345.1928
45. Brooks CL, Grubele M, Onuchic JN, Wolynes PG. Chemical Physics of Protein Folding. *Proc Natl Acad Sci USA* (1998) 95(19):11037–8. doi: 10.1073/pnas.95.19.11037
46. Kim PS, Baldwin RL. Intermediates in the Folding Reactions of Small Proteins. *Annu Rev Biochem* (1990) 59(1):631–60. doi: 10.1146/annurev.bi.59.070190.003215
47. Creighton TE. Protein Folding Coupled to Disulphide Bond Formation. *Biol Chem Hoppe Seyler* (1997) 378(8):731–44. doi: 10.1515/bchm.1997.378.8.731
48. Xiao S, Bi Y, Shan B, Raleigh DP. Analysis of Core Packing in a Cooperatively Folded Miniature Protein: The Ultrafast Folding Villin Headpiece Helical Subdomain. *Biochemistry* (2009) 48(21):4607–16. doi: 10.1021/bi8021763
49. Clarke J, Fersht AR. Engineered Disulfide Bonds as Probes of the Folding Pathway of Barnase: Increasing the Stability of Proteins Against the Rate of Denaturation. *Biochemistry* (1993) 32(16):4322–9. doi: 10.1021/bi00067a022
50. Zaykov AN, Mayer JP, Gelfanov VM, DiMarchi RD. Chemical Synthesis of Insulin Analogs Through a Novel Precursor. *ACS Chem Biol* (2014) 9(3):683–91. doi: 10.1021/cb400792s
51. Hua QX, Mayer J, Jia W, Zhang J, Weiss MA. The Folding Nucleus of the Insulin Superfamily: A Flexible Peptide Model Foreshadows the Native State. *J Biol Chem* (2006) 281:28131–42. doi: 10.1074/jbc.M602616200
52. Narhi LO, Hua QX, Arakawa T, Fox GM, Tsai L, Rosenfeld R, et al. Role of Native Disulfide Bonds in the Structure and Activity of Insulin-Like Growth Factor 1: Genetic Models of Protein-Folding Intermediates. *Biochemistry* (1993) 32:5214–21. doi: 10.1021/bi00070a033
53. Weiss MA, Hua QX, Jia W, Chu YC, Wang RY, Katsoyannis PG. Hierarchical Protein "Un-Design": Insulin's Intrachain Disulfide Bridge Tethers a Recognition α -Helix. *Biochemistry* (2000) 39:15429–40. doi: 10.1021/bi001905s
54. Hua QX, Nakagawa SH, Jia W, Hu SQ, Chu YC, Katsoyannis PG, et al. Hierarchical Protein Folding: Asymmetric Unfolding of an Insulin Analogue Lacking the A7-B7 Interchain Disulfide Bridge. *Biochemistry* (2001) 40:12299–311. doi: 10.1021/bi011021o
55. Jia XY, Guo ZY, Wang Y, Xu Y, Duan SS, Feng YM. Peptide Models of Four Possible Insulin Folding Intermediates With Two Disulfides. *Protein Sci* (2003) 12:2412–9. doi: 10.1110/ps.0389303
56. Neuhof A, Rolls MM, Jungnickel B, Kalies K-U, Rapoport TA. Binding of Signal Recognition Particle Gives Ribosome/Nascent Chain Complexes a Competitive Advantage in Endoplasmic Reticulum Membrane Interaction. *Mol Biol Cell* (1998) 9(1):103–15. doi: 10.1091/mbc.9.1.103
57. Waudby CA, Dobson CM, Christodoulou J. Nature and Regulation of Protein Folding on the Ribosome. *Trends Biochem Sci* (2019) 44(11):914–26. doi: 10.1016/j.tibs.2019.06.008
58. Kaarsholm NC, Norris K, Jorgensen RJ, Mikkelsen J, Ludvigsen S, Olsen OH, et al. Engineering Stability of the Insulin Monomer Fold With Application to Structure-Activity Relationships. *Biochemistry* (1993) 32:10773–8. doi: 10.1021/bi00091a031
59. Weiss MA, Hua Q-X, Jia W, Nakagawa SH, Chu Y-C, Hu S-Q, et al. Activities of Monomeric Insulin Analogs at Position A8 Are Uncorrelated With Their Thermodynamic Stabilities. *J Biol Chem* (2001) 276(43):40018–24. doi: 10.1074/jbc.M104634200
60. Baker EN, Blundell TL, Cutfield JF, Dodson EJ, Dodson GG, Hodgkin DMC, et al. The Structure of 2Zn Pig Insulin Crystals at 1.5 Å Resolution. *Philos Trans R Soc Lond B Biol Sci* (1988) 319(1195):369–456. doi: 10.1098/rstb.1988.0058
61. Shoelson S, Polonsky K, Zeidler A, Rubenstein A, Tager H. Human Insulin B24 (Phe—Ser). Secretion and Metabolic Clearance of the Abnormal Insulin in Man and in a Dog Model. *J Clin Invest* (1984) 73(5):1351–8. doi: 10.1172/JCI11338
62. Hua QX, Shoelson SE, Inouye K, Weiss MA. Paradoxical Structure and Function in a Mutant Human Insulin Associated With Diabetes Mellitus. *Proc Natl Acad Sci USA* (1993) 90(2):582–6. doi: 10.1073/pnas.90.2.582
63. Pandeyarajan V, Smith BJ, Phillips NB, Whittaker L, Cox GP, Wickramasinghe N, et al. Aromatic Anchor at an Invariant Hormone-Receptor Interface: Function of Insulin Residue B24 With Application to Protein Design. *J Biol Chem* (2014) 289(50):34709–27. doi: 10.1074/jbc.M114.608562
64. Yang Y, Glidden MD, Dhayalan B, Zaykov A, Chen Y-S, Wickramasinghe NP, et al. Peptide Model of the Mutant Proinsulin Syndrome. II. Nascent Structure and Biological Implications. *Front Endocrinol* (2022) 13:821091. doi: 10.3389/fendo.2022.821091
65. Glidden MD, Aldabbagh K, Phillips NB, Carr K, Chen Y-S, Whittaker J, et al. An Ultra-Stable Single-Chain Insulin Analog Resists Thermal Inactivation and Exhibits Biological Signaling Duration Equivalent to the Native Protein. *J Biol Chem* (2018) 293(1):47–68. doi: 10.1074/jbc.M117.808626
66. Sreerama N, Woody RW. A Self-Consistent Method for the Analysis of Protein Secondary Structure From Circular Dichroism. *Anal Biochem* (1993) 209(1):32–44. doi: 10.1006/abio.1993.1079
67. Sreerama N, Woody RW. Poly (Pro) II Helices in Globular Proteins: Identification and Circular Dichroic Analysis. *Biochemistry* (1994) 33(33):10022–5. doi: 10.1021/bi00199a028
68. Johnson WC. Analyzing Protein Circular Dichroism Spectra for Accurate Secondary Structures. *Proteins: Struct Funct Bioinform* (1999) 35(3):307–12. doi: 10.1002/(SICI)1097-0134(19990515)35:3<307::AID-PROT4>3.0.CO;2-3
69. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. *Proteomics Protoc Handbook* (2005) 571–607. doi: 10.1385/1-59259-890-0:571
70. Sosnick TR, Fang X, Shelton VM. Application of Circular Dichroism to Study RNA Folding Transitions. *Methods Enzymol* (2000) 317:393–409. doi: 10.1016/S0076-6879(00)17026-0

71. Pace CN, Shaw KL. Linear Extrapolation Method of Analyzing Solvent Denaturation Curves. *Proteins: Struct Funct Bioinform* (2000) 41(S4):1–7. doi: 10.1002/1097-0134(2000)41:4+<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)00108-8
 73. Yan W, Frank CL, Korth MJ, Sopher BL, Novoa I, Ron D, et al. Control of PERK Eif2 α 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* (1988) 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 Nucl Magn Reson Spectrosc* (2006) 48(1):47–62. doi: 10.1016/j.pnmrs.2005.10.001
 77. Hua Q-x, Weiss MA. Mechanism of Insulin Fibrillation: The Structure of Insulin Under Amyloidogenic Conditions Resembles a Protein-Folding Intermediate. *J Biol Chem* (2004) 279(20):21449–60. doi: 10.1074/jbc.M314141200
 78. Shoelson 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(1):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 NKX2-2 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.