

# Incorporation of Aliphatic Proline Residues into Recombinantly Produced Insulin

Stephanie L. Breunig, Janine C. Quijano, Cecile Donohue, Amy Henrickson, Borries Demeler, Hsun Teresa Ku, and David A. Tirrell\*




Cite This: *ACS Chem. Biol.* 2023, 18, 2574–2581



Read Online

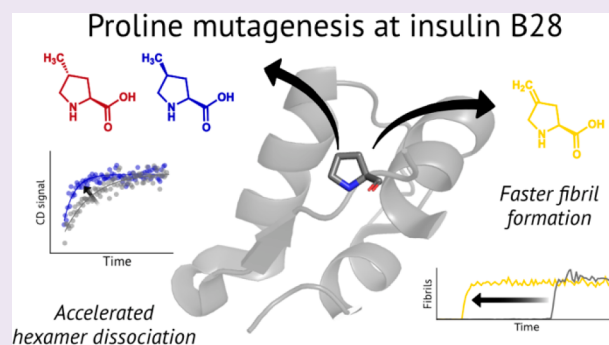
ACCESS |

 Metrics & More

 Article Recommendations

 Supporting Information

**ABSTRACT:** Analogs of proline can be used to expand the chemical space about the residue while maintaining its uniquely restricted conformational space. Here, we demonstrate the incorporation of 4*R*-methylproline, 4*S*-methylproline, and 4-methyleneproline into recombinant insulin expressed in *Escherichia coli*. These modified proline residues, introduced at position B28, change the biophysical properties of insulin: Incorporation of 4-methyleneproline at B28 accelerates fibril formation, while 4-methylation speeds dissociation from the pharmaceutically formulated hexamer. This work expands the scope of proline analogs amenable to incorporation into recombinant proteins and demonstrates how noncanonical amino acid mutagenesis can be used to engineer the therapeutically relevant properties of protein drugs.



## INTRODUCTION

Proline is unique among the canonical amino acids: the cyclic pyrrolidine side chain restricts the conformational space accessible to the residue. Replacing proline with any of the proteinogenic amino acids through standard mutagenesis approaches necessarily grants greater conformational freedom. Alternatively, noncanonical proline (ncPro) residues expand the chemical space about proline while maintaining a pyrrolidine (or pyrrolidine-like) side chain. Because the conformational preferences of many proline analogs are known,<sup>1,2</sup> ncPro residues are useful as probes of conformational effects on protein behavior. Proline analogs have enabled investigators to elucidate the importance of a key proline *cis-trans* isomerization event in 5HT<sub>3</sub> receptor opening,<sup>3</sup> modify the properties of elastin-like proteins,<sup>4–6</sup> determine the molecular origins of collagen stability,<sup>1,7</sup> and probe the role of *cis-trans* isomerization in  $\beta$ 2-microglobulin fibrillation.<sup>8,9</sup> Although Chatterjee and co-workers have reported progress in the development of orthogonal prolyl-tRNA synthetase/tRNA pairs,<sup>10</sup> to date, only residue-specific (rather than site-specific) incorporation approaches have been successful in introducing proline analogs into recombinant proteins.<sup>11</sup> Using *E. coli* as an expression host, these approaches have allowed efficient incorporation of various 3- and 4-functionalized proline residues and those with modified ring sizes and compositions.<sup>12</sup>

Insulin (Figure 1a) is a 5.8 kDa peptide hormone normally released from the pancreatic  $\beta$  cells in response to elevated levels of blood glucose. Its binding to the insulin receptor

induces intracellular responses that ultimately lower blood glucose concentrations.<sup>13</sup> Diabetes mellitus is a result of dysfunctional insulin signaling, either through an impaired ability to produce and secrete insulin (type 1) or through insulin resistance (type 2). Subcutaneous injection of exogenous insulin is a common strategy in diabetes treatment, especially for individuals with type 1 diabetes. The recombinant production of insulin<sup>14</sup> has significantly advanced diabetes treatment by enabling both its production at scale and the creation of modified insulins with desirable properties through standard mutagenesis and chemical modification approaches.<sup>15–18</sup>

To mimic the insulin profile of a healthy pancreas, two broad classes of insulin analogs have been developed: long-acting (or basal) insulins and fast-acting insulins (FAIs).<sup>19</sup> Long-acting variants recapitulate the lower levels of insulin secretion that maintain metabolism in an anabolic state. Conversely, FAIs aim to mimic the transient increases in insulin concentration stimulated by elevated blood glucose after a meal. Typical insulin replacement therapy relies on a combination of regular basal insulin treatments and FAI injections before meals.

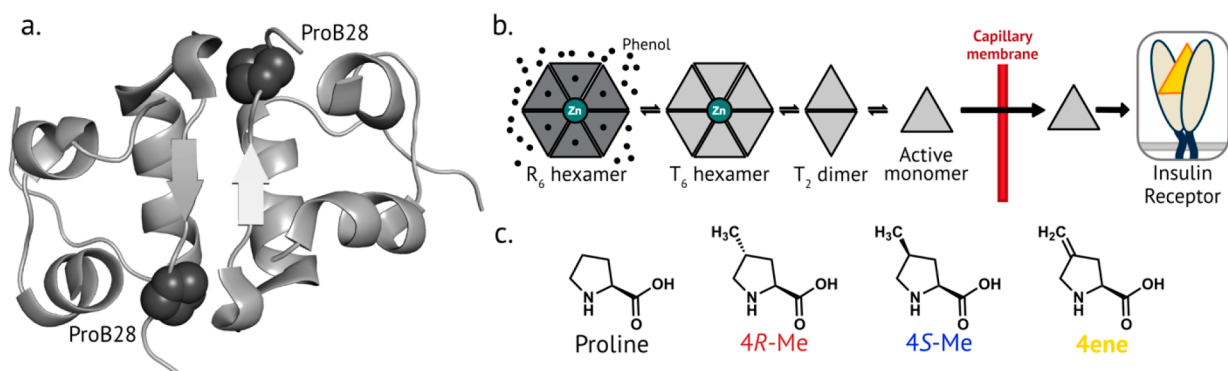
**Received:** September 12, 2023

**Revised:** October 24, 2023

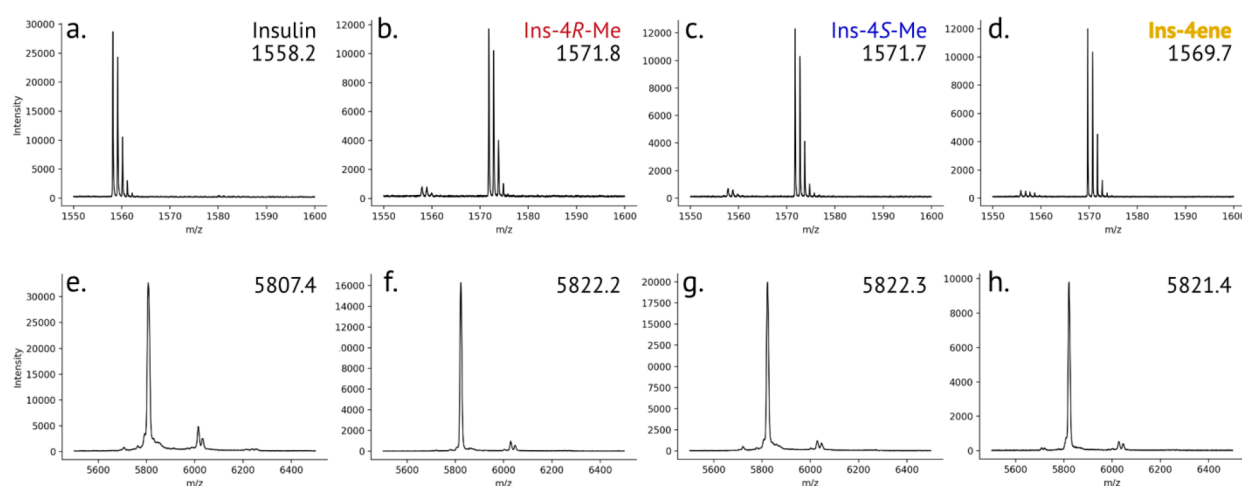
**Accepted:** October 27, 2023

**Published:** November 14, 2023





**Figure 1.** Proline mutagenesis at position B28 of human insulin. a. Crystal structure of insulin (PDB 1MSO), highlighting ProB28 located at the dimer interface. b. Simplified scheme of insulin dissociation after injection. Insulin exists as a hexamer in the R state in the presence of zinc and phenolic ligands such as in the pharmaceutical formulation. After injection, insulin dissociates into lower-order oligomeric species that can diffuse more easily across the capillary membrane, enter the bloodstream, and bind to the insulin receptor. c. The structure of proline and the aliphatic proline analogs used in this study.



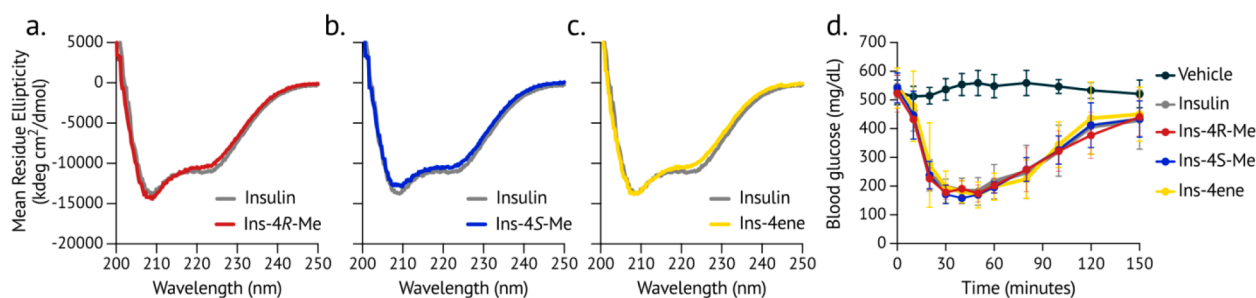
**Figure 2.** Mass spectrometry of insulin variants. a–d. Characterization of proline analog incorporation. The solubilized inclusion body (containing proinsulin) after expression in medium supplemented with proline (a), 4R-Me (b), 4S-Me (c), or 4ene (d) was digested with Glu-C and analyzed by MALDI-TOF MS. The peptide that contains position B28 of mature insulin is  $^{50}\text{RGFFYT}\underline{\text{P}}\text{KTRRE}$  (expected  $m/z = 1557.8$ ). e–h. MALDI-TOF characterization of mature and purified insulin variants: human insulin (e), ins-4R-Me (f), ins-4S-Me (g), and Ins-4ene (h). The peaks at  $m/z \sim 6050$  correspond to adducts of the sinapic acid matrix.

Many FDA-approved insulin variants have been engineered by altering the amino acid sequence in ways that cause pronounced pharmacokinetic effects.<sup>20</sup> Notably, insulin aspart<sup>17,21</sup> (NovoLog, marketed by Novo Nordisk) and insulin lispro<sup>15,22</sup> (Humalog, Eli Lilly) both involve changes to ProB28, a key residue<sup>15</sup> near the C-terminus of the B-chain (Figure 1a). Insulin aspart is achieved by the single point mutation ProB28Asp, while insulin lispro contains an inversion of ProB28 and LysB29; both changes destabilize insulin oligomers. Because the rate-limiting step for insulin absorption into the bloodstream is dissociation of oligomer to monomer<sup>23</sup> (Figure 1b), these changes accelerate insulin's onset of action. Insulins are also prone to chemical and physical denaturation,<sup>24–26</sup> processes slowed by the formation of protective oligomers. As a result, insulin production and distribution require a cold chain,<sup>27</sup> and maintaining protein stability in continuous subcutaneous insulin infusion pumps is challenging.<sup>28</sup>

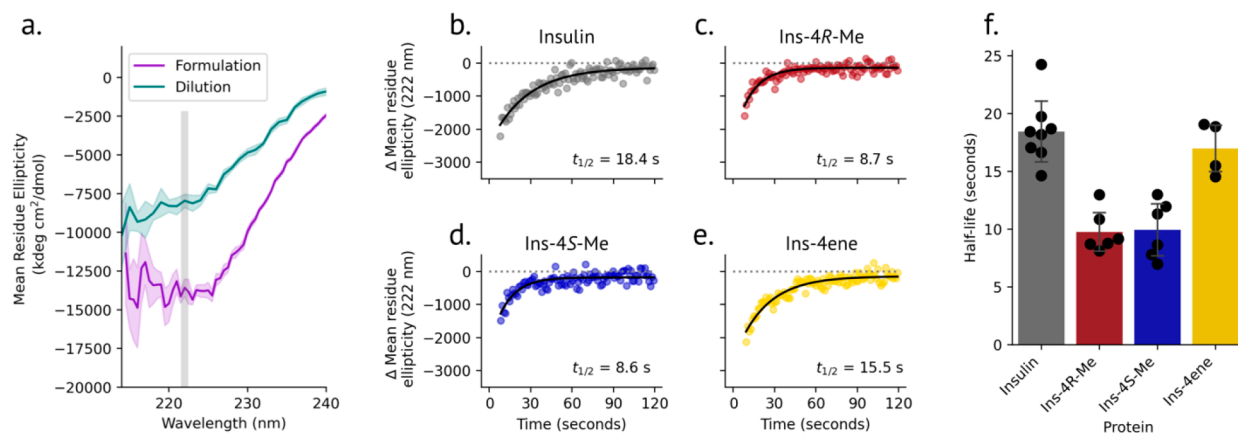
Intrigued by the role that ProB28 plays in insulin biophysics, we recently introduced a series of ncPro residues at position B28 of human insulin.<sup>29,30</sup> Because mature insulin contains one

proline, a residue-specific replacement approach results in site-specific proline replacement without the need for an orthogonal aminoacyl-tRNA synthetase/tRNA pair. These efforts, which focused on proline analogs known to incorporate well into recombinant proteins in *E. coli*,<sup>5</sup> illustrated how proline mutagenesis of insulin can be used to tune its biophysical characteristics.<sup>29,30</sup>

Here, we demonstrate the efficient incorporation of three new aliphatic proline residues (4R-methylproline, 4R-Me; 4S-methylproline, 4S-Me; and 4-methyleneproline, 4ene; Figure 1c) at position B28 of recombinantly produced insulin; these insulin variants will be referred to as ins-4R-Me, ins-4S-Me, and ins-4ene, respectively. We find that these modifications alter insulin behavior: replacement of ProB28 with 4ene speeds fibril formation, while 4-methylation accelerates hexamer dissociation without affecting stability against physical denaturation. This work expands the range of proline analogs that can be incorporated into recombinant proteins in *E. coli*. It also demonstrates how small molecular changes introduced through noncanonical amino acid mutagenesis can be used to



**Figure 3.** Circular dichroism spectroscopy and bioactivity of insulin variants. a-c. Far-UV circular dichroism spectra of insulin and insulin variants ( $60 \mu\text{M}$  in  $100 \text{ mM}$  phosphate buffer,  $\text{pH } 8.0$ ). The spectrum of each insulin variant is overlaid with that of human insulin (gray). d. Insulins were injected subcutaneously into diabetic mice, and blood glucose was measured over time after injection.



**Figure 4.** Hexamer dissociation kinetics of insulin variants. a. Equilibrium CD spectra of insulin before and after dilution. To measure the dissociation kinetics, the decrease in negative ellipticity at  $222 \text{ nm}$  was monitored over time after dilution. b-e. Representative dissociation plots for insulin (b), ins-4R-Me (c), ins-4S-Me (d), and ins-4ene (e). Each dilution experiment was fit to a monoexponential function, and the half-life for each displayed replicate is indicated. f. Summary of dissociation half-life values.

probe and engineer the therapeutically relevant properties of protein drugs.

## RESULTS AND DISCUSSION

**Aliphatic Proline Residues Are Accepted by the *E. coli* Translational Machinery.** To identify an expanded set of ncPro residues accepted by the *E. coli* translational machinery, we expressed proinsulin (a precursor to insulin) under conditions that favor ncPro incorporation.<sup>12</sup> We monitored ncPro replacement by proinsulin expression and mass spectrometry and noted a range of incorporation efficiencies for the 15 commercially available proline analogs tested (Figure S1 and Table S1). Notably, the aliphatic proline residues 4R-Me, 4S-Me, and 4ene led to high levels of proinsulin expression and good ( $\sim 90\%$ ) incorporation efficiencies under the optimized conditions (Figure 2a-d, Table S2, Table S3). We replaced ProB28 of human insulin with each of these three aliphatic proline analogs (Figure 2e-h) and determined the resulting effects on insulin behavior.

**Proline Analogs Do Not Affect Insulin Secondary Structure or Bioactivity.** The secondary structure of each insulin variant was assessed by circular dichroism (CD) spectroscopy (Figure 3a-c; Table S4). The CD spectrum of insulin is sensitive to the state of oligomerization of the protein: for monomeric insulins, the ratio of negative ellipticities at  $208$  and  $222 \text{ nm}$  is increased compared to that of the insulin dimer.<sup>31</sup> The far-UV CD spectrum for each variant closely matched that of human insulin at  $60 \mu\text{M}$ ,

suggesting that proline replacement has not significantly perturbed the secondary structure or dimer formation under these conditions.

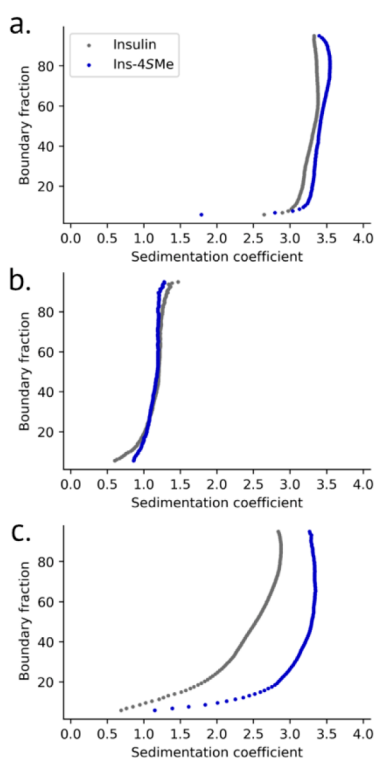
To validate biological activity *in vivo*, insulins were formulated with zinc and phenolic ligands and injected subcutaneously into diabetic mice; blood glucose was monitored over  $2.5 \text{ h}$ . Rodent models can assess insulin activity but cannot distinguish differences in time to onset of action for human insulin and fast-acting analogs.<sup>32</sup> Because the C-terminus of the B-chain does not interact with the insulin receptor,<sup>33,34</sup> we did not expect modification of ProB28 to affect bioactivity. As anticipated, all of the insulin variants reduced the blood glucose in diabetic mice (Figure 3d).

**4-Methylation of ProB28 Speeds Hexamer Dissociation.** Increased negative ellipticity at  $222 \text{ nm}$  in the CD spectrum of insulin is a signature of oligomerization.<sup>31,35</sup> Adopting a method reported by Gast and co-workers,<sup>36</sup> we measured the rate of dissociation of insulin variants to the monomer state. Insulins were formulated under conditions that mimic the pharmaceutical formulation ( $600 \mu\text{M}$  insulin,  $25 \text{ mM}$  *m*-cresol,  $250 \mu\text{M}$   $\text{ZnCl}_2$ ) and favor the  $\text{R}_6$  hexamer state.<sup>37</sup> Dissociation was monitored by tracking the mean residue ellipticity at  $222 \text{ nm}$  over time after 150-fold dilution into ligand-free buffer (Figure 4a). CD spectra after dilution are distinct from that of chemically denatured insulin (Figure S3, Figure S4), confirming that changes in the CD signal are not a result of denaturation. While ins-4ene dissociated at a rate similar to that of human insulin ( $t_{1/2} = 17.0 \pm 2.3$  and  $18.4 \pm 2.8 \text{ s}$ , respectively), dissociation of ins-4R-Me ( $t_{1/2} = 9.8 \pm$

1.8 s) and ins-4S-Me ( $t_{1/2} = 9.9 \pm 2.5$  s) was accelerated under these conditions (Figure 4b-f, Table S4).

It seems likely that steric effects play a role in accelerating the dissociation of ins-4R-Me and ins-4S-Me. The carbon atoms of 4R- and 4S-methyl substituents installed at ProB28 in published structures of the R<sub>6</sub> insulin hexamer are in close proximity (2.2 and 2.3 Å, respectively) to backbone carbonyl oxygen atoms of the adjacent monomer (Figure S5a); these distances are smaller than the sum of the van der Waals radii of the respective atoms (1.70 Å for carbon and 1.52 Å for oxygen).<sup>38</sup> Similar interactions are observed in the T<sub>6</sub> hexamer (Figure S5b). The resulting steric clashes may act to destabilize the hexamer and accelerate dissociation.

We examined the oligomerization of insulin and ins-4S-Me in more detail by sedimentation velocity analytical ultracentrifugation. At 300 μM in the presence of zinc and *m*-cresol, the sedimentation coefficients of both insulin (3.3S) and ins-4S-Me (3.4S) correspond to the hexamer<sup>39</sup> (Figure 5a, Table



**Figure 5.** Diffusion-corrected integral sedimentation coefficient distributions. Insulin and ins-4S-Me were formulated under the following conditions: a. 300 μM insulin, 12.5 mM *m*-cresol, and 125 μM Zn. b. 4 μM insulin, 167 μM *m*-cresol, 1.67 μM Zn. c. 300 μM insulin. Samples described in panels a and c were measured using interference optics due to high absorbance from insulin and *m*-cresol; those in panel b used absorbance at 225 nm. All samples were in 25 mM tris buffer, pH 8.0.

S5). After dilution to 4 μM, both insulins behave as monomers (Figure 5b, Table S5),<sup>40</sup> further validating that the dilution kinetics measured above describe hexamer dissociation.

Interestingly, we found that at these higher concentrations, 4S-methylation of ProB28 affects insulin oligomerization in the absence of ligands. Compared to the broad sedimentation profile of insulin, which indicates the presence of multiple oligomers between monomer and hexamer,<sup>41</sup> the sedimentation of ins-4S-Me (3.1S) is nearly unchanged when zinc and *m*-

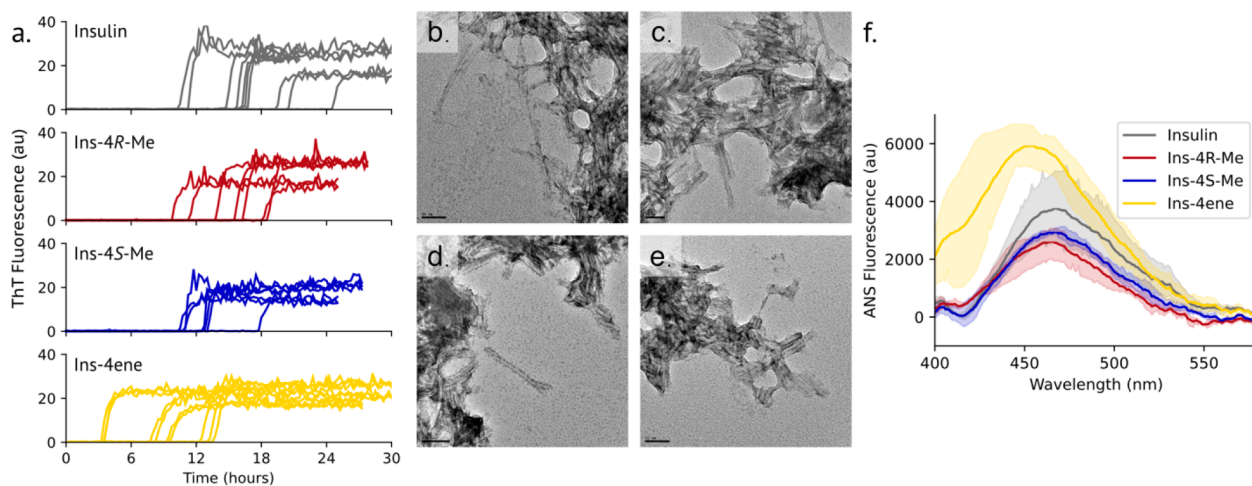
cresol are omitted (Figure 5c). These results are counter-intuitive: when formulated with ligands, ins-4S-Me dissociates more rapidly upon dilution relative to insulin (Figure 4f). However, at 300 μM and in the absence of ligands, 4S-methylation instead stabilizes the hexamer. We note that insulin behavior depends upon its formulation,<sup>24,39</sup> for instance, the crystal structure of the insulin hexamer in the presence of zinc and *m*-cresol<sup>37</sup> is distinct from that under ligand-free conditions.<sup>42</sup> Perhaps the effect of 4S-methylation on insulin oligomerization is similarly context-dependent.

**4ene at Position B28 Accelerates Fibril Formation.** To assess stability against aggregation, insulins were subjected to vigorous shaking at 37 °C, and fibril formation was monitored over time with the fibril-specific dye thioflavin T (ThT; Figure 6a; Table S4). Under these conditions, the lag time to fibril formation for insulin was  $16.6 \pm 4.1$  h. Introduction of either 4R-Me or 4S-Me at position B28 did not significantly affect the lag time ( $15.0 \pm 3.5$  and  $12.5 \pm 2.5$  h, respectively). However, ins-4ene ( $8.2 \pm 4.0$  h) formed fibrils more rapidly than the other insulins examined here. All aggregates were fibrillar in nature when assessed by transmission electron microscopy (TEM). Compared to the micrometer-long fibrils described in previous reports,<sup>43–46</sup> those observed here were relatively short (tens to hundreds of nanometers). Under these conditions, ins-4ene formed fibrils shorter than those of insulin, ins-4R-Me, or ins-4S-Me (Figure 6b-e).

We sought to better understand the molecular mechanism of the decreased stability of ins-4ene. An early step in the mechanism of insulin fibril formation is thought to be detachment of the C-terminus of the insulin B-chain from the core of the molecule, exposing hydrophobic residues.<sup>46</sup> We probed insulin disorder with the dye 8-anilino-1-naphthalene-sulfonic acid (ANS).<sup>47</sup> Compared to the other insulins measured, ins-4ene exhibits a blue-shift in the emission maximum and an increase in fluorescence intensity upon labeling with ANS (Figure 6f; Table S4), suggesting an increased exposure of buried hydrophobic residues. We analyzed DMSO-solubilized ins-4ene fibrils by mass spectrometry but did not observe chemical modification of the exocyclic alkene (Figure S6).

## CONCLUSION

In this work, we have expanded the set of ncPro residues that can be incorporated into proteins in *E. coli* and demonstrated their utility in modifying the properties of recombinantly produced insulin. We found that the proline analogs 4R-Me, 4S-Me, and 4ene could be incorporated with high efficiencies into recombinant proinsulin. As this work was coming to completion, the incorporation of 4R-Me into recombinant thioredoxin was reported.<sup>48</sup> In that report, incorporation of the 4S diastereomer could not be detected, perhaps because it interferes with thioredoxin folding and renders the protein unstable. In the work reported here, proinsulin is purified from the inclusion body fraction; therefore, any influence of ncPro replacement on protein stability during expression should be reduced. We were also able to detect low to modest levels of incorporation of other ncPro residues into proinsulin, such as 3-hydroxy analogs, 4-oxoproline, and the diazirine-containing ncPro photo-proline (Figure S1, Table S1). Engineering the prolyl-tRNA synthetase or other components of the *E. coli* translational machinery might lead to improved incorporation of these analogs.



**Figure 6.** Fibrillation of insulin variants. a. Insulin variants ( $60 \mu\text{M}$  in  $100 \text{ mM}$  phosphate buffer,  $\text{pH } 8.0$ ) were incubated at  $37^\circ\text{C}$  with vigorous shaking, and fibril formation was monitored by ThT fluorescence over time. b–e. Representative TEM images of insulin (b; scale bar:  $50 \mu\text{m}$ ), ins-4R-Me (c,  $50 \mu\text{m}$ ), ins-4S-Me (d,  $100 \mu\text{m}$ ), and ins-4ene (e,  $50 \mu\text{m}$ ) aggregates. f. ANS emission spectra of insulin variants ( $1 \mu\text{M}$  insulin variant labeled with  $5 \mu\text{M}$  ANS in  $100 \text{ mM}$  phosphate buffer,  $\text{pH } 8.0$ ).

In addition to expanding the side-chain volume and hydrophobic surface area that can be introduced into proteins at proline sites, the three ncPro residues studied here complement other translationally active proline analogs that have been used to probe the roles of proline conformation in determining the biophysical behavior of proteins. For instance, 4R-Me and 4S-Me have opposing conformational preferences relative to their more commonly used 4-fluoroproline counterparts<sup>1,2,49</sup> and together have been used to validate the stereoelectronic origin of collagen stability.<sup>7,49</sup> The amplitude of the pyrrolidine ring pucker of 4ene is expected to be attenuated compared to proline (Figure S7) and may approach that of the planar 3,4-dehydroproline.<sup>50</sup> Furthermore, the olefin present in 4ene might be used as a chemical handle for protein modification at proline residues.<sup>51,52</sup>

This work also demonstrates the extent to which small molecular changes can affect the protein behavior. For instance, introducing an exocyclic alkene at B28 of insulin shortened the lag time of fibril formation. This finding, together with increased solvent-exposed hydrophobic surface area of ins-4ene as evidenced by its ANS emission spectrum, supports current models for the mechanism of insulin fibrillation.<sup>46</sup> Furthermore, installing a simple methyl group at ProB28 of human insulin significantly accelerates dissociation from the insulin hexamer, behavior associated with achieving a rapid onset-of-action for monomeric FAIs.<sup>23</sup> Insulin fibrillation is thought to proceed from the monomer state, and conditions that disfavor insulin oligomerization often promote physical denaturation.<sup>46</sup> Notably, in the cases of Ins-4R-Me and Ins-4S-Me, rapid dissociation was achieved without accelerating fibril formation. Future work is needed to understand more fully the effect of 4-methylation on insulin oligomerization and to determine whether the faster dissociation observed for ins-4R-Me and ins-4S-Me *in vitro* results in faster onset-of-action *in vivo*.

## METHODS

**Detailed Methods Are Provided in the Supporting Information (SI).** *Proinsulin Expression, Refolding, and Purification.* Proinsulins containing proline analogs were expressed in the *E. coli* strain CAG18515/pQE80\_H27R-PI\_proS. This is a proline auxotrophic strain which carries a plasmid for prolyl-tRNA synthetase

(*proS*) overexpression and inducible expression of proinsulin. Briefly, the culture was grown in Andrew's Magical Medium,<sup>53</sup> a defined medium that contains  $50 \text{ mg L}^{-1}$  proline, until midlog phase. Cells were then washed and resuspended in a medium that lacks proline. The proline analog ( $0.5\text{--}1.0 \text{ mM}$ ) was added along with  $0.5 \text{ M}$  NaCl to induce osmotic stress conditions that increase cellular uptake of the proline analog. Proinsulin was expressed overnight with the addition of  $1 \text{ mM}$  IPTG. Specific expression conditions (*proS* variant, ncPro concentration) for each insulin variant are described in Table S3.

Proinsulin was isolated from the inclusion body fraction, solubilized ( $3 \text{ M}$  urea and  $10 \text{ mM}$  cysteine), and refolded ( $10 \text{ mM}$  CAPS,  $\text{pH } 10.7$ ,  $4^\circ\text{C}$ ) over the course of 2–3 days. Purified, mature insulin was obtained after digestion with trypsin and carboxypeptidase B and purification by reversed-phase high-performance liquid chromatography (HPLC). Fractions were analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE), and analytical HPLC to verify sample quality and ensure  $\geq 95\%$  purity for all downstream analyses. Yields for ncPro-containing insulins ranged from  $23$  to  $53 \text{ mg L}^{-1}$ , and mature insulin yields were approximately  $15\text{--}50\%$  of the proinsulin starting material. Insulin was expressed in Terrific Broth (TB) and similarly refolded and purified.

*Reduction of Blood Glucose in Diabetic Mice.* Adult (8 week old) male C57BL/6J mice were ordered from Jackson Laboratory (Bar Harbor, ME) and injected intraperitoneally with freshly prepared streptozotocin at  $65 \text{ mg kg}^{-1} \text{ day}^{-1}$  for 3 consecutive days. Insulin-dependent diabetes was confirmed by detection of high glucose levels ( $200\text{--}600 \text{ mg dL}^{-1}$ ) as measured with a glucometer. Insulin analogs were diluted to  $100 \mu\text{g mL}^{-1}$  in formulation buffer ( $1.6 \text{ mg mL}^{-1}$  m-cresol,  $0.65 \text{ mg mL}^{-1}$  phenol,  $3.8 \text{ mg mL}^{-1}$  sodium phosphate  $\text{pH } 7.4$ ,  $16 \text{ mg mL}^{-1}$  glycerol,  $0.8 \mu\text{g mL}^{-1}$   $\text{ZnCl}_2$ ) and injected ( $35 \mu\text{g kg}^{-1}$ ) subcutaneously at the scruff. Blood glucose was measured over 2.5 h from blood sampled from the lateral tail vein.

*Kinetics of Hexamer Dissociation.* Insulin samples were dialyzed overnight against  $28.6 \text{ mM}$  tris buffer,  $\text{pH } 8.0$ , and then formulated as follows:  $600 \mu\text{M}$  insulin,  $250 \mu\text{M}$   $\text{ZnCl}_2$ ,  $25 \text{ mM}$  m-cresol,  $25 \text{ mM}$  tris buffer,  $\text{pH } 8$ . To a stirred buffer solution containing  $2.98 \text{ mL}$  of  $25 \text{ mM}$  tris,  $\text{pH } 8.0$  in a  $10 \text{ mm}$  quartz cuvette was injected  $20 \mu\text{L}$  of the insulin formulation ( $150\text{-fold}$  dilution). Ellipticity was monitored at  $222 \text{ nm}$  over  $120 \text{ s}$  at  $25^\circ\text{C}$ . A typical run led to a rapid drop in the CD signal as mixing occurred ( $\sim 5 \text{ s}$ ) and then a gradual rise to an equilibrium ellipticity representative of an insulin monomer. Data preceding the time point with the greatest negative ellipticity (representing the mixing time) were omitted from further analysis. Runs were discarded if the maximum change in mean residue

ellipticity from equilibrium did not exceed  $750 \text{ deg cm}^2 \text{ dmol}^{-1}$ , which indicated poor mixing. The remaining data were fit to a monoexponential function using Scipy (Python). The data presented here are from at least two separate insulin HPLC fractions measured on two different days.

**Analytical Ultracentrifugation.** Insulins were dialyzed against 28.6 mM tris buffer, pH 8.0 and formulated at 300  $\mu\text{M}$  insulin, 12.5 mM *m*-cresol, and 125  $\mu\text{M}$   $\text{ZnCl}_2$ , in 25 mM tris buffer. Ligand-free insulins (300  $\mu\text{M}$ ) and diluted insulins (4  $\mu\text{M}$  insulin, 167  $\mu\text{M}$  *m*-cresol, and 1.7  $\mu\text{M}$   $\text{ZnCl}_2$ ) were also prepared from the same dialysis samples. Diluted insulin conditions were identical to those after dilution in the CD dissociation kinetics experiments and were incubated at RT for at least one hour after dilution prior to analysis.

Velocity sedimentation experiments were performed at the Canadian Center for Hydrodynamics at the University of Lethbridge. 300  $\mu\text{M}$  insulin samples were measured by interference optics due to the high absorbance from the protein and *m*-cresol. Diluted samples (4  $\mu\text{M}$  insulin) were measured using absorbance optics at 225 nm. All data were analyzed with UltraScan III version 4.0 release 6606.<sup>54</sup> Velocity data were initially fitted with the two-dimensional spectrum analysis<sup>55</sup> to determine meniscus position and time- and radially invariant noise. Subsequent noise-corrected data were analyzed by the enhanced van Holde-Weischet analysis<sup>56</sup> to generate diffusion-corrected integral sedimentation coefficient distributions.

**Fibrillation Lag Time.** Insulin samples (60  $\mu\text{M}$  in 100 mM sodium phosphate, pH 8.0) were centrifuged at 22,000 g for 1 h at 4 °C prior to the addition of 1  $\mu\text{M}$  thioflavin T (ThT). Each sample was shaken continuously in a 96-well, black, clear bottom plate at 37 °C, and fluorescence readings were recorded every 15 min (444 nm excitation, 485 nm emission). Fibrillation runs were performed on at least two separate HPLC fractions, each in triplicate or quadruplicate, on at least two different days. The growth phase of each fibrillation replicate was fit to a linear function, and fibrillation lag times were reported as the x-intercept of this fit. Fibril samples were stored at 4 °C until analysis by TEM.

**ANS Fluorescence.** Insulins (1  $\mu\text{M}$ ) were mixed with 5  $\mu\text{M}$  ANS in 100 mM phosphate buffer, pH 8.0. Fluorescence emission spectra were measured in 1 cm quartz cuvettes at ambient temperature (350 nm excitation). Measurements for each insulin were performed in triplicate from three separate HPLC fractions. Spectra were smoothed before plotting and determining the emission maxima.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscchembio.3c00561>.

Extended experimental details, materials, and methods; nucleotide and amino acid sequences; incorporation efficiencies of 15 proline analogs; mass spectrometry characterization of proline analog incorporation, mature insulin variants, and solubilized insulin fibrils; representative purity assessment of insulin variants by SDS-PAGE; proinsulin expression conditions and yields; summaries of insulin variant characterization data; equilibrium circular dichroism spectra; models of insulin variant hexamers; and conformational analysis of proline analogs (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

David A. Tirrell – Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States; [orcid.org/0000-0003-3175-4596](https://orcid.org/0000-0003-3175-4596); Email: [tirrell@caltech.edu](mailto:tirrell@caltech.edu)

## Authors

Stephanie L. Breunig – Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States; [orcid.org/0000-0002-8665-6363](https://orcid.org/0000-0002-8665-6363)

Janine C. Quijano – Department of Translational Research and Cellular Therapeutics, Diabetes and Metabolism Research Institute, Beckman Institute City of Hope, Duarte, California 91010, United States

Cecile Donohue – Department of Translational Research and Cellular Therapeutics, Diabetes and Metabolism Research Institute, Beckman Institute City of Hope, Duarte, California 91010, United States

Amy Henrickson – Department of Chemistry and Biochemistry, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada; [orcid.org/0000-0003-3266-5202](https://orcid.org/0000-0003-3266-5202)

Borries Demeler – Department of Chemistry and Biochemistry, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada; Department of Chemistry and Biochemistry, University of Montana, Missoula, Montana 59801, United States; [orcid.org/0000-0002-2414-9518](https://orcid.org/0000-0002-2414-9518)

Hsun Teresa Ku – Department of Translational Research and Cellular Therapeutics, Diabetes and Metabolism Research Institute, Beckman Institute City of Hope, Duarte, California 91010, United States; Irell & Manella Graduate School of Biological Science, City of Hope, Duarte, California 91010, United States; [orcid.org/0000-0001-5662-9893](https://orcid.org/0000-0001-5662-9893)

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acscchembio.3c00561>

## Notes

The authors declare the following competing financial interest(s): D.T. is an inventor on U.S. patents that describe the use of non-canonical amino acids in the engineering of insulin and other proteins.

## ■ ACKNOWLEDGMENTS

We sincerely thank C. Paavola, M. Akers, J. Moyers, and A. Mahdavi for advice on proinsulin refolding. A. Lopez assisted with cloning *proS* mutants, and A. Chapman performed pilot ncPro incorporation studies. Circular dichroism spectroscopy was performed at the Beckman Institute Laser Resource Center at Caltech with assistance from J. Winkler. Electron microscopy was performed at the Beckman Institute Resource Center for Transmission Electron Microscopy at Caltech with input from S. Chen. Mass spectrometry was performed at the Caltech Multiuser Mass Spectrometry Laboratory in the Division of Chemistry and Chemical Engineering with the help of M. Shahgholi and at the Proteome Exploration Laboratory at Caltech with assistance from B. Quan. This work was supported by grants from the National Institutes of Health (1R01GM134013-01 and 1R01GM120600), the Juvenile Diabetes Research Foundation (3-IND-2015-118-I-X), the Canada 150 Research Chairs program (C150-2017-00015), the Canada Foundation for Innovation (CFI-37589), and the Canadian Natural Science and Engineering Research Council (DG-RGPIN-2019-05637). UltraScan supercomputer calculations were supported through NSF/XSEDE grant TGMCB070039N and University of Texas grant TG457201. S.L.B. was supported by a National Science Foundation Graduate Research Fellowship under grant number 1745301.

The Canadian Natural Science and Engineering Research Council supports A.H. through a scholarship grant.

## REFERENCES

- (1) Shoulders, M. D.; Raines, R. T. Collagen Structure and Stability. *Annu. Rev. Biochem.* **2009**, *78*, 929–958.
- (2) Ganguly, H. K.; Basu, G. Conformational Landscape of Substituted Prolines. *Biophys Rev* **2020**, *12*, 25–39.
- (3) Lummis, S. C. R.; Beene, D. L.; Lee, L. W.; Lester, H. A.; Broadhurst, R. W.; Dougherty, D. A. *Cis-Trans* Isomerization at a Proline Opens the Pore of a Neurotransmitter-Gated Ion Channel. *Nature* **2005**, *438*, 248–252.
- (4) Kim, W.; Hardcastle, K. I.; Conticello, V. P. Fluoroproline Flip-Flop: Regiochemical Reversal of a Stereoelectronic Effect on Peptide and Protein Structures. *Angew. Chem., Int. Ed.* **2006**, *45*, 8141–8145.
- (5) Kim, W.; George, A.; Evans, M.; Conticello, V. P. Cotranslational Incorporation of a Structurally Diverse Series of Proline Analogues in an *Escherichia Coli* Expression System. *ChemBioChem* **2004**, *5* (7), 928–936.
- (6) Kim, W.; McMillan, R. A.; Snyder, J. P.; Conticello, V. P. A Stereoelectronic Effect on Turn Formation Due to Proline Substitution in Elastin-Mimetic Polypeptides. *J. Am. Chem. Soc.* **2005**, *127* (51), 18121–18132.
- (7) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. Code for Collagen's Stability Deciphered. *Nature* **1998**, *392*, 666–667.
- (8) Torbeev, V. Y.; Hilvert, D. Both the *Cis-Trans* Equilibrium and Isomerization Dynamics of a Single Proline Amide Modulate B2-Microglobulin Amyloid Assembly. *Proc Natl Acad Sci U S A* **2013**, *110* (50), 20051–20056.
- (9) Torbeev, V.; Ebert, M. O.; Dolenc, J.; Hilvert, D. Substitution of Proline32 by  $\alpha$ -Methylproline Preorganizes B2-Microglobulin for Oligomerization but Not for Aggregation into Amyloids. *J. Am. Chem. Soc.* **2015**, *137* (7), 2524–2535.
- (10) Chatterjee, A.; Xiao, H.; Schultz, P. G. Evolution of Multiple, Mutually Orthogonal Prolyl-tRNA Synthetase/tRNA Pairs for Unnatural Amino Acid Mutagenesis in *Escherichia Coli*. *Proc. Natl. Acad. Sci. USA* **2012**, *109* (37), 14841–14846.
- (11) Andrews, J.; Gan, Q.; Fan, C. “Not-so-Popular” Orthogonal Pairs in Genetic Code Expansion. *Protein Sci.* **2023**, *32* (2), e4559.
- (12) Breunig, S. L.; Tirrell, D. A. Incorporation of Proline Analogs into Recombinant Proteins Expressed in *Escherichia Coli*. *Methods Enzymol* **2021**, *656*, 545–571.
- (13) Hauesler, R. A.; McGraw, T. E.; Accili, D. Biochemical and Cellular Properties of Insulin Receptor Signalling. *Nat Rev Mol Cell Biol* **2018**, *19*, 31–44.
- (14) Goeddel, D. V.; Kleid, D. G.; Bolivar, F.; Heyneker, H. L.; Yansura, D. G.; Crea, R.; Hirose, T.; Kraszewski, A.; Itakura, K.; Riggs, A. D. Expression in *Escherichia Coli* of Chemically Synthesized Genes for Human Insulin. *Proc. Natl. Acad. Sci. USA* **1979**, *76* (1), 106–110.
- (15) Brems, D. N.; Alter, L. A.; Beckage, M. J.; Chance, R. E.; Dimarchi, R. D.; Green, L. K.; Long, H. B.; Pekar, A. H.; Shields, J. E.; Frank, B. H. Altering the Association Properties of Insulin by Amino Acid Replacement. *Protein Eng.* **1992**, *5* (6), 527–533.
- (16) Rakatzi, I.; Ramrath, S.; Ledwig, D.; Dransfeld, O.; Bartels, T.; Seipke, G.; Eckel, J. A Novel Insulin Analog With Unique Properties: LysB3, GluB29 Insulin Induces Prominent Activation of Insulin Receptor Substrate 2, but Marginal Phosphorylation of Insulin Receptor Substrate 1. *Diabetes* **2003**, *52*, 2227–2238.
- (17) Brange, J.; Ribel, U.; Hansen, J. F.; Dodson, G.; Hanse, M. T.; Havelund, S.; Melberg, S. G.; Norris, F.; Norris, K.; Snel, L.; Sørensen, A. R.; Voigt, H. O. Monomeric Insulins Obtained by Protein Engineering and Their Medical Implications. *Nature* **1988**, *333*, 679–682.
- (18) Lepore, M.; Pampanelli, S.; Fanelli, C.; Porcellati, F.; Bartocci, L.; Di Vincenzo, A.; Cordonì, C.; Costa, E.; Brunetti, P.; Bolli, G. B. Pharmacokinetics and Pharmacodynamics of Subcutaneous Injection of Long-Acting Human Insulin Analog Glargine, NPH Insulin, and Ultralente Human Insulin and Continuous Subcutaneous Infusion of Insulin Lispro. *Diabetes* **2000**, *49*, 2142–2148.
- (19) Mathieu, C.; Gillard, P.; Benhalima, K. Insulin Analogues in Type 1 Diabetes Mellitus: Getting Better All the Time. *Nat Rev Endocrinol* **2017**, *13*, 385–399.
- (20) Bolli, G. B.; Di Marchi, R. D.; Park, G. D.; Pramming, S.; Koivisto, V. A. Insulin Analogues and Their Potential in the Management of Diabetes Mellitus. *Diabetologia* **1999**, *42* (10), 1151–1167.
- (21) Owens, D.; Vora, J. Insulin Aspart: A Review. *Expert Opin. Drug Metab. Toxicol.* **2006**, *2* (5), 793–804.
- (22) Holleman, F.; Hoekstra, J. B. L. Insulin Lispro. *N Engl J Med* **1997**, *337* (3), 176–183.
- (23) Bakaysa, D. L.; Radziuk, J.; Havel, H. A.; Brader, M. L.; Li, S.; Dodd, S. W.; Beals, J. M.; Pekar, A. H.; Brems, D. N. Physicochemical Basis for the Rapid Time-Action of LysB28ProB29-Insulin: Dissociation of a Protein-Ligand Complex. *Protein Sci.* **1996**, *5* (12), 2521–2531.
- (24) Brange, J.; Langkjaer, L. Insulin Structure and Stability. In *Stability and Characterization of Protein and Peptide Drugs: Case Histories*; Wang, Y. J., Pearlman, R., Eds.; Plenum Press: New York, 1993; pp 315–350, DOI: 10.1007/978-1-4899-1236-7\_11.
- (25) Brange, J.; Havelund, S.; Hougaard, P. Chemical Stability of Insulin. 2. Formation of Higher Molecular Weight Transformation Products During Storage of Pharmaceutical Preparations. *Pharm. Res.* **1992**, *9* (6), 727–734.
- (26) Brange, J.; Langkjær, L.; Havelund, S.; Volund, A. Chemical Stability of Insulin. 1. Hydrolytic Degradation During Storage of Pharmaceutical Preparations. *Pharm. Res.* **1992**, *9* (6), 715–726.
- (27) Weiss, M. A. Design of Ultra-Stable Insulin Analogues for the Developing World. *Journal of Health Specialties* **2013**, *1* (2), 59–70.
- (28) Kerr, D.; Morton, J.; Whately-Smith, C.; Everett, J.; Begley, J. P. Laboratory-Based Non-Clinical Comparison of Occlusion Rates Using Three Rapid-Acting Insulin Analogs in Continuous Subcutaneous Insulin Infusion Catheters Using Low Flow Rates. *J Diabetes Sci Technol* **2008**, *2* (3), 450–455.
- (29) Fang, K. Y.; Lieblich, S. A.; Tirrell, D. A. Replacement of ProB28 by Pipecolic Acid Protects Insulin against Fibrillation and Slows Hexamer Dissociation. *J Polym Sci A Polym Chem* **2019**, *57* (3), 264–267.
- (30) Lieblich, S. A.; Fang, K. Y.; Cahn, J. K. B.; Rawson, J.; LeBon, J.; Teresa Ku, H.; Tirrell, D. A. 4S-Hydroxylation of Insulin at ProB28 Accelerates Hexamer Dissociation and Delays Fibrillation. *J. Am. Chem. Soc.* **2017**, *139*, 8384–8387.
- (31) Pocker, Y.; Biswas, S. B. Conformational Dynamics of Insulin in Solution. Circular Dichroic Studies. *Biochemistry* **1980**, *19* (22), 5043–5049.
- (32) Plum, A.; Agersø, H.; Andersen, L. Pharmacokinetics of the Rapid-Acting Insulin Analog, Insulin Aspart, in Rats, Dogs, Pigs, and Pharmacodynamics of Insulin Aspart in Pigs. *Drug Metab. Dispos.* **2000**, *28*, 155–160.
- (33) Scapin, G.; Dandey, V. P.; Zhang, Z.; Prosise, W.; Hruza, A.; Kelly, T.; Mayhood, T.; Strickland, C.; Potter, C. S.; Carragher, B. Structure of the Insulin Receptor-Insulin Complex by Single-Particle Cryo-EM Analysis. *Nature* **2018**, *556* (7699), 122–125.
- (34) Menting, J. G.; Whittaker, J.; Margets, M. B.; Whittaker, L. J.; Kong, G. K.-W.; Smith, B. J.; Watson, C. J.; Žáková, L.; Kletvíková, E.; Jiracek, J.; Chan, S. J.; Steiner, D. F.; Dodson, G. G.; Brzozowski, A. M.; Weiss, M. A.; Ward, C. W.; Lawrence, M. C. How Insulin Engages Its Primary Binding Site on the Insulin Receptor. *Nature* **2013**, *493*, 241–245.
- (35) Goldman, J.; Carpenter, F. H. Zinc Binding, Circular Dichroism, and Equilibrium Sedimentation Studies on Insulin (Bovine) and Several of Its Derivatives. *Biochemistry* **1974**, *13* (22), 4566–4574.
- (36) Gast, K.; Schüler, A.; Wolff, M.; Thalhammer, A.; Berchtold, H.; Nagel, N.; Lenherr, G.; Hauck, G.; Seckler, R. Rapid-Acting and Human Insulins: Hexamer Dissociation Kinetics upon Dilution of the Pharmaceutical Formulation. *Pharm. Res.* **2017**, *34* (11), 2270–2286.

- (37) Smith, G. D.; Ciszak, E.; Magrum, L. A.; Pangborn, W. A.; Blessing, R. H. R6 Hexameric Insulin Complexed with M-Cresol or Resorcinol. *Acta Crystallogr.* **2000**, *56* (12), 1541–1548.
- (38) Bondi, A. Van Der Waals Volumes and Radii. *J. Phys. Chem.* **1964**, *68* (3), 441–451.
- (39) Richards, J. P.; Stickelmeyer, M. P.; Flora, D. B.; Chance, R. E.; Frank, B. H.; DeFelippis, M. R. Self-Association Properties of Monomeric Insulin Analogs under Formulation Conditions. *Pharm. Res.* **1998**, *15* (9), 1434–1441.
- (40) Bakaysa, D. L.; Radziuk, J.; Havel, H. A.; Brader, M. L.; Li, S.; Dodd, S. W.; Beals, J. M.; Pekar, A. H.; Brems, D. N. Physicochemical Basis for the Rapid Time-Action of LysB28ProB29-Insulin: Dissociation of a Protein-Ligand Complex. *Protein Sci.* **1996**, *5*, 2521–2531.
- (41) Pekar, A. H.; Frank, B. H. Conformation of Proinsulin. A Comparison of Insulin and Proinsulin Self-Association at Neutral PH. *Biochemistry* **1972**, *11* (22), 4013–4016.
- (42) Mandal, K.; Dhayalan, B.; Avital-Shmilovici, M.; Tokmakoff, A.; Kent, S. B. H. Crystallization of Enantiomerically Pure Proteins from Quasi-Racemic Mixtures: Structure Determination by X-Ray Diffraction of Isotope-Labeled Ester Insulin and Human Insulin. *ChemBioChem* **2016**, *17* (5), 421–425.
- (43) Jansen, R.; Dzwolak, W.; Winter, R. Amyloidogenic Self-Assembly of Insulin Aggregates Probed by High Resolution Atomic Force Microscopy. *Biophys. J.* **2005**, *88*, 1344–1353.
- (44) Nielsen, L.; Frokjaer, S.; Brange, J.; Uversky, V. N.; Fink, A. L. Probing the Mechanism of Insulin Fibril Formation with Insulin Mutants. *Biochemistry* **2001**, *40* (28), 8397–8409.
- (45) Jiménez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R.; Caspar, D. L. D.; et al. The Protofilament Structure of Insulin Amyloid Fibrils. *Proc. Natl. Acad. Sci. USA* **2002**, *99* (14), 9196–9201.
- (46) Akbarian, M.; Yousefi, R.; Farjadian, F.; Uversky, V. N. Insulin Fibrillation: Toward Strategies for Attenuating the Process. *Chemical Communications* **2020**, *56* (77), 11354–11373.
- (47) Semisotnov, G. V.; Rodionova, N. A.; Razgulyaev, O. I.; Uversky, V. N.; Gripas', A. F.; Gilmanshin, R. I. Study of the "Molten Globule" Intermediate State in Protein Folding by a Hydrophobic Fluorescent Probe. *Biopolymers* **1991**, *31* (1), 119–128.
- (48) Caporale, A.; Loughlin, J. O.; Ortin, Y.; Rubini, M. A Convenient Synthetic Route to (2S,4S)-Methylproline and Its Exploration for Protein Engineering of Thioredoxin. *Org. Biomol. Chem.* **2022**, *20*, 6324–6328.
- (49) Shoulders, M. D.; Hodges, J. A.; Raines, R. T. Reciprocity of Steric and Stereoelectronic Effects in the Collagen Triple Helix. *J. Am. Chem. Soc.* **2006**, *128*, 8112–8113.
- (50) Kang, Y. K.; Park, H. S. Conformational Preferences and Cis-Trans Isomerization of L-3,4-Dehydroproline Residue. *Peptide Science* **2009**, *92* (5), 387–398.
- (51) Loosli, S.; Foletti, C.; Pappmeyer, M.; Zürich, E. Synthesis of 4-(Arylmethyl)Proline Derivatives. *Synlett* **2019**, *30*, 508–510.
- (52) Deming, T. J.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. Biosynthetic Incorporation and Chemical Modification of Alkene Functionality in Genetically Engineered Polymers. *Journal of Macromolecular Science - Pure and Applied Chemistry* **1997**, *34* (10), 2143–2150.
- (53) He, W.; Fu, L.; Li, G.; Andrew Jones, J.; Linhardt, R. J.; Koffas, M. Production of Chondroitin in Metabolically Engineered *E. Coli*. *Metab. Eng.* **2015**, *27*, 92–100.
- (54) Demeler, B.; Gorbet, G. E. Analytical Ultracentrifugation Data Analysis with Ultrascan-III. *Analytical Ultracentrifugation: Instrumentation, Software, and Applications* **2016**, 119–143.
- (55) Brookes, E.; Cao, W.; Demeler, B. A Two-Dimensional Spectrum Analysis for Sedimentation Velocity Experiments of Mixtures with Heterogeneity in Molecular Weight and Shape. *European Biophysics Journal* **2010**, *39* (3), 405–414.
- (56) Demeler, B.; Van Holde, K. E. Sedimentation Velocity Analysis of Highly Heterogeneous Systems. *Anal. Biochem.* **2004**, *335* (2), 279–288.