Chapter III

4*S*-FLUORINATION AT PROB29 OF INSULIN LISPRO SLOWS FIBRIL FORMATION

3.1 Contributions

Chad D. Paavola, Michael P. Akers, Julie S. Moyers, and Alborz Mahdavi provided advice on proinsulin refolding. Alex M. Chapman prepared lispro variants used for mouse experiments. Janine Quijano and Jeanne LeBon performed mouse experiments, with input from Hsun Teresa Ku. Analytical ultracentrifugation was performed by Maduni Ranasinghe at the Canadian Center for Hydrodynamics at the University of Lethbridge, with input from Borries Demeler. Jay Winkler, Songye Chen, Scott Virgil, and Mona Shahgholi assisted with circular dichroism spectroscopy, transmission electron microscopy, chromatography, and mass spectrometry, respectively. Stephanie L. Breunig prepared all other insulin samples and performed all other experiments. Alex M. Chapman and David A. Tirrell contributed to the writing of this text.

3.2 Abstract

Recombinant insulin is a life-saving therapeutic for millions of patients affected by diabetes mellitus. Standard mutagenesis approaches have led to insulin variants that enable improved control of blood glucose: for instance, the fast-acting insulin lispro contains two point mutations that destabilize dimer formation and expedite absorption. However, insulins undergo irreversible denaturation upon storage, a process accelerated for the insulin monomer. Here we replace ProB29 of insulin lispro with 4*R*-fluoroproline, 4*S*-

fluoroproline, and 4,4-difloroproline. For most properties measured, fluorination at B29 does not change the behavior of lispro: all of the resulting variants reduce blood glucose in diabetic mice, sediment as monomers by analytical ultracentrifugation, exhibit similar secondary structure as measured by circular dichroism, and rapidly dissociate from the zinc and resorcinol-bound hexamer upon dilution. Though we measured similar lag times to fibril formation for most variants, we find that 4*S*-fluorination of ProB29 delays fibril formation. These results demonstrate how subtle molecular changes enabled by non-canonical amino acid mutagenesis can improve the stability of protein therapeutics.

3.3 Main text

Insulin, a 5.8 kDa peptide hormone composed of two disulfide-linked chains (A and B), is a widely-used, essential therapeutic for individuals with diabetes mellitus. Insulin is normally secreted from pancreatic β-cells in response to elevated blood glucose; its binding to the insulin receptor induces a variety of intracellular responses that lead to decreased blood glucose concentration.[1] Diabetes, which affects more than 400 million adults world-wide,[2] arises from dysfunction of insulin signaling, either from loss of insulin secretion (type 1), or from development of insulin resistance (type 2). Advances in insulin technology over the last century, notably the recombinant production of insulin,[3] have dramatically reduced the mortality rate associated with diabetes. Individuals with type 1 diabetes typically rely on insulin replacement therapy to compensate for loss of pancreatic function, while those with type 2 often supplement lifestyle changes with oral medications that enhance insulin signaling, and in severe cases, receive exogenous insulin.

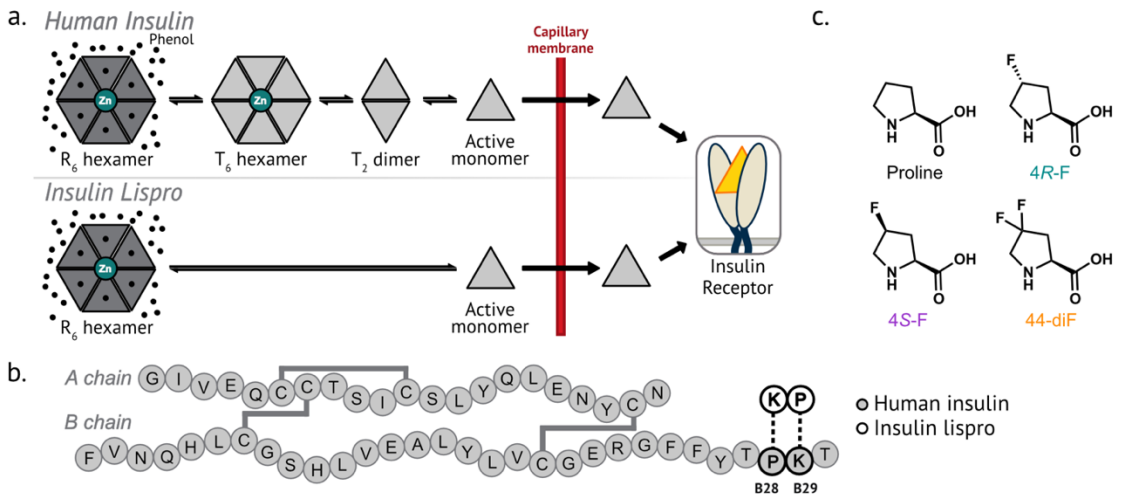

Figure 3.1. Proline mutagenesis at position B29 of insulin lispro. a. Simplified representation of the oligomerization states of human insulin and insulin lispro. Disrupted oligomerization speeds the release of lispro into the bloodstream. **b**. The amino acid sequences of human insulin and insulin lispro. **c**. The structures of proline and the 4-fluorinated proline analogs used in this study.

In its pharmaceutical formulations, insulin exists as a zinc- and phenol-bound hexamer (referred to as the R_6 state).[4] Upon subcutaneous injection, insulin hexamers dissociate through lower-order oligomeric species before crossing the capillary membrane in the active monomeric state[5,6] (Figure 3.1a). Because dissociation to the monomer is rate-limiting for insulin absorption into the bloodstream, disrupting insulin association also accelerates its onset of action.[7] Introducing point mutations that destabilize oligomer formation has resulted in a suite of FDA-approved fast-acting insulins (FAIs).[8–10] FAIs are rapidly released into the bloodstream shortly after injection,[11] resembling the transient increases in insulin concentration stimulated by elevated blood glucose (i.e., after a meal). Typical insulin replacement therapies combine regular treatments of basal (long-acting) insulin variants with injections of FAIs before mealtime to approximate the insulin-action

profile of a healthy pancreas.[11] However, FAIs tend to be more prone to chemical and physical denaturation, since the protective effects of oligomerization are interrupted.[12]

Insulin lispro (Humalog™, Eli Lilly), the first commercially-available FAI, disfavors monomer association through the inversion of two residues near the C-terminus of the insulin B-chain, ProB28 and LysB29 (Figure 3.1b).[8] Removal of the constrained pyrrolidine ring of proline at position B28 grants additional conformational degrees of freedom and removes hydrophobic packing interactions at the dimer interface, disrupting the homodimer interaction.[8] Consequently, insulin lispro hexamers dissociate to bioactive monomers considerably faster than human insulin, and are thus absorbed more rapidly.[7] However, the standard mutagenesis approach employed in the discovery of insulin lispro[8] (and other FDA-approved insulin variants[9,10,13]) is limited by the chemical functionalities present in the 20 canonical amino acids, and restricted with respect to the chemical space that can be explored. In particular, no other canonical amino acid can recapitulate the unique structural and conformational properties of proline.[14] In contrast, protein engineering strategies that draw upon non-canonical amino acids (ncAAs) afford access to chemistries beyond those found in natural or conventionally-engineered proteins.[15]

Non-canonical proline (ncPro) residues provide chemical diversity while maintaining proline's cyclic structure. Many ncPro residues exhibit conformational biases different from those of proline,[14] so can be used to interrogate importance of proline conformation in protein behavior. ncPro residues with 4*R* electron-withdrawing groups (such as 4*R*-fluoroproline, 4*R*-F; Figure 3.1c) promote a C^{γ}-*exo* ring pucker due to a gauche effect,

which in turn enforces the *trans* amide isomer. Conversely, ncPro analogs with 4*S*-electron withdrawing groups (such as 4*S*-fluoroproline, 4*S*-F) favor the C^{γ}-*endo* pucker and the *cis* amide isomer, compared to canonical proline.[14] The position of the amide *cis*-*trans* equilibrium preferences of 4,4-difluoroproline (44-diF) is similar to that of canonical proline.[16] Isomerization of the 4-fluorinated prolines is accelerated relative to proline, as an inductive effect reduces the bond order of the preceding amide.[16] Proline analogs have been used to determine the stereoelectronic origin of collagen stability,[17] and have helped identify a key *cis*-*trans* isomerization event in the opening of the 5-HT_3 receptor.[18]

The consequences of replacing proline with fluorinated analogs depend on protein context. Replacing proline with ncPro residues in barstar,[16] ubiquitin,[19] and the trp cage miniprotein[20] stabilized the protein if the ncPro conformational preferences matched those present in proline; proline analogs with opposed preferences destabilized the protein. In the context of thioredoxin, the conformational preferences of the analog made little difference; replacing a conserved *cis* proline residue with either diastereomer of 4-fluoroproline led to similar effects on protein stability, despite the opposed conformational preferences of the fluoroprolines.[21]

A key proline residue at position 32 of β2 microglobulin (β2m) was replaced by the fluoroprolines 4*R*-F, 4*S*-F, and 44-diF,[22] and by α-methylproline, which strongly favors the *trans* isomer.[23] Pro32 adopts the *cis* form in the natively folded protein, and the *trans* conformer in the amyloid fibril, suggesting that *cis*-*trans* isomerization is a key step in β2m fibrillation. Consistent with this expectation, ncPro *cis*-*trans* preferences affected soluble

β2m dynamics, oligomerization, and stability against thermal and chemical denaturation. However, the diverse fibrillation behaviors and fibril morphologies of the β2m variants suggested that fibril formation is more complex than simple association of *trans*-Pro32 β2m monomers.[22,23]

We recently demonstrated that introduction of ncPro residues at position B28 in human insulin can be used to tune its biophysical properties,[24,25] including results presented in Chapter II of this thesis. These ncPro mutagenesis experiments on human insulin prompted us to ask whether similar changes in the biophysical properties of insulin lispro might be achieved. To explore this question, we replaced ProB29 in insulin lispro with 4-fluorinated proline analogs (Figure 3.1c); these lispro variants will be referred to as KP-4*R*-F, KP-4*S*-F, and KP-44-diF.

Lispro variants were expressed in *Escherichia coli* as the corresponding proinsulin (a precursor to insulin lispro) under conditions that favor ncPro incorporation.[26] Briefly, the proline auxotrophic strain CAG18515 overexpressing the prolyl-tRNA synthetase was grown in M9 medium supplemented with all 20 amino acids until late log phase. Cells were washed and resuspended in a medium lacking proline, at which point high concentrations (0.3-0.5 M) of sodium chloride (to facilitate proline uptake), and the proline analog were added. Proline replacement was assessed by MALDI-TOF mass spectrometry of a proinsulin peptide fragment (Figure 3.2a-d), and of mature insulin (Figure 3.2e-h). Incorporation of all fluorinated proline analogs exceeded 90% (Table 3.S1).

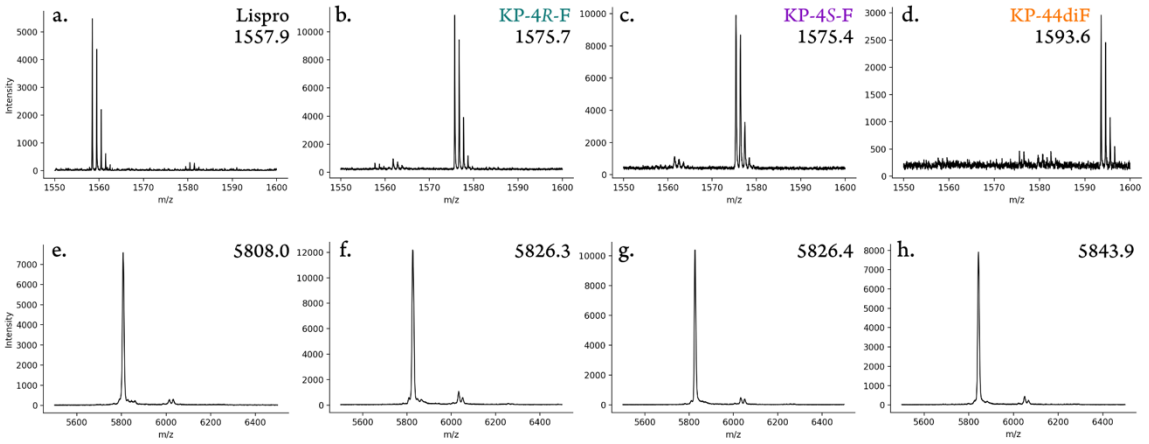

Figure 3.2. Mass spectrometry characterization of lispro variants. a-d. Characterization of proline analog incorporation. The solubilized inclusion body (containing proinsulin) after expression in the presence of proline (a), 4*R*-F (b), 4*S*-F (c), 44-diF (d) was digested with Glu-C and analyzed by MALDI-TOF MS. The peptide that contains position B29 of mature lispro is ^{50}RGFFYTK**P**TRRE (expected m/z = 1557.8). **e-h**. MALDI-TOF characterization of mature and purified insulin variants: insulin lispro (e), KP-4*R*-F (f), KP-4*S*-F (g), and KP-44diF (h). The larger molecular weight peaks (~6050 Da) present in these spectra correspond to adducts of the sinapic acid matrix.

To verify the ability of each ncPro-containing insulin lispro to reduce blood glucose *in vivo*, insulins were injected subcutaneously into diabetic mice, and blood glucose was monitored over the course of 2.5 h. Mouse models allow determination of activity but not time to onset of action.[27] Because the B-chain C-terminus does not participate in binding to the insulin receptor,[28] we did not expect substitution at position B29 to affect bioactivity. Indeed, all lispro variants retained activity (Figure 3.3a).

The secondary structure of each insulin lispro variant was probed by circular dichroism spectroscopy (Figure 3.3b-e). Compared to human insulin, which exists as a dimer at 60 μM, the monomeric insulin lispro exhibits increased negative ellipticity at 208 nm, and

decreased negative ellipticity at 222 nm, consistent with previous reports.[8] The CD spectra of all fluorinated lispro variants were similar to that of lispro, suggesting that fluorination does not alter the secondary structure of lispro. Analytical ultracentrifugation confirmed this preference for the monomer state across all lispro variants (Figure 3.3f, Table 3.S3).

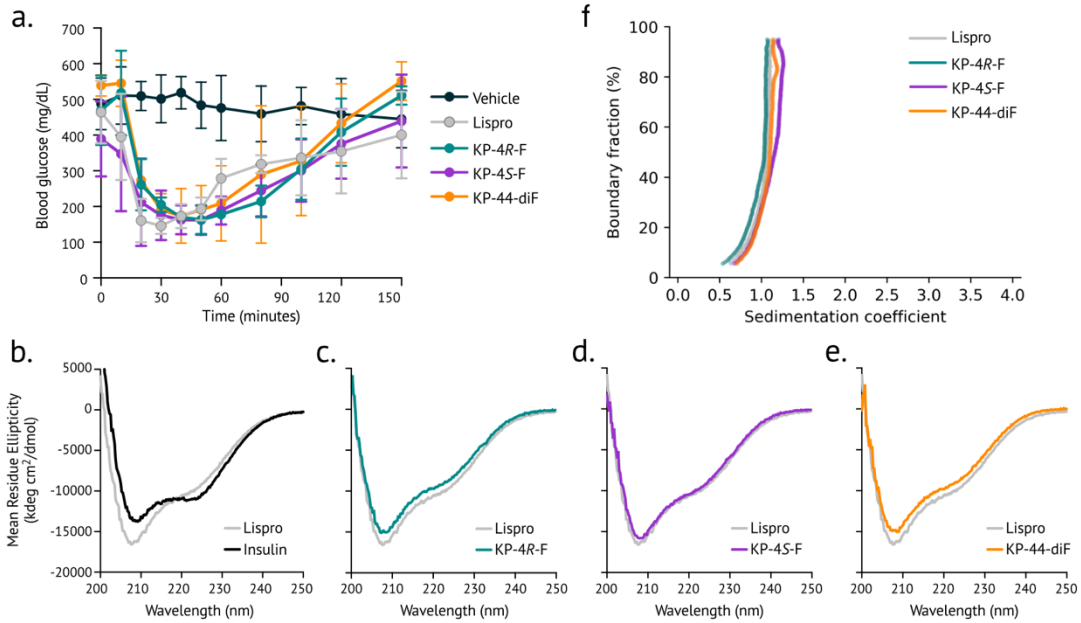

Figure 3.3. Bioactivity, circular dichroism spectroscopy, and sedimentation of lispro variants. **a**. Changes in blood glucose concentrations over time after diabetic mice were injected subcutaneously with lispro variants. **b-e**. Far-UV circular dichroism spectra (60 µM lispro in 10 mM phosphate buffer, pH 8.0). The spectrum of lispro is overlaid with that of human insulin (b.), KP-4R-F (c.), KP-4S-F (d.), and KP-44diF. **f**. Sedimentation velocity of lispro variants (125 µM) in 100 mM phosphate, pH 8.0.

We measured the half-life ($t_{1/2}$) for dissociation of the hexamer form of each insulin variant by monitoring the decrease in negative ellipticity at 222 nm over time after dilution (Figure 3.4a). We formulated insulin and lispro variants under conditions that mimic the pharmaceutical formulation, using resorcinol as the phenolic ligand to slow dissociation.[29]

Under these conditions (600 μM insulin, 25 mM resorcinol, 250 μM $ZnCl_2$), insulin lispro dissociated significantly more rapidly than human insulin ($t_{1/2}$ = 11.3 ± 3.8 and 30.2 ± 2.8 s, respectively), as anticipated. Accurately measuring dissociation behavior faster than lispro is expected to be challenging due to the resolution of this experiment, though comparisons to human insulin can still be made. We found that all fluorinated lispro variants dissociated similarly to proline-containing lispro (Figure 3.4b-g, Table 3.S3).

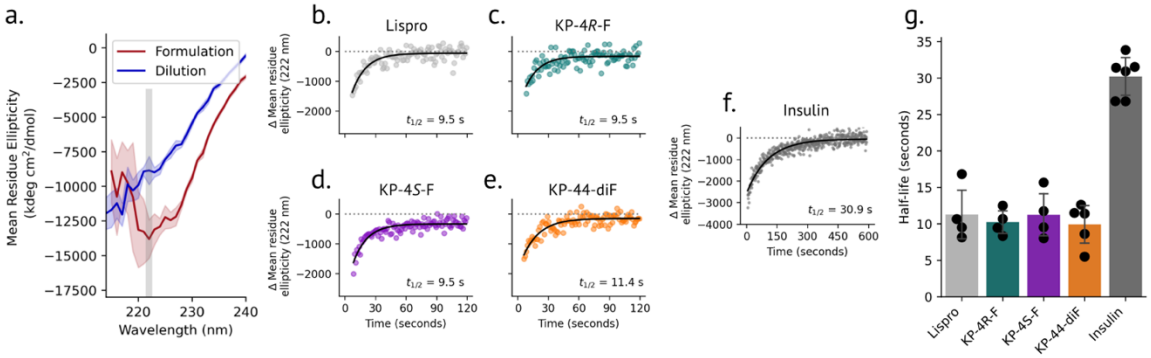

Figure 3.4. Hexamer dissociation kinetics of lispro variants. **a**. Equilibrium CD spectra of lispro before and after dilution. To measure dissociation kinetics, the decrease in negative ellipticity at 222 nm was monitored over time after dilution. **b-f**. Representative dissociation kinetics measurements for lispro (b), KP-4R-F (c), KP-4S-F (d), KP-44diF (e), and insulin (f). Note the extended x-axis for insulin in panel f. **g**. Summary of dissociation half-life values.

Each insulin lispro variant was subjected to continuous, vigorous shaking at 37°C to assess its stability against physical denaturation; fibrillation was monitored with the dye thioflavin T (ThT), (Figure 3.5a). Under these conditions, the fibrillation lag time of insulin lispro (10.9 ± 2.2 h) was shortened compared to that of human insulin (16.7 ± 4.1 h; see Chapter II), consistent with the notion that fibrillation proceeds from the monomer state.[30] Replacing ProB29 with 4*R*-F (10.3 ± 2.4 h) or 44-diF (9.0 ± 1.3 h) did not significantly change the fibrillation lag time. Notably, KP-4*S*-F was stabilized against fibril formation

(17.9 ± 0.8 h). We note that 4*S*-F favors the *endo* ring pucker (proline and the other ncPro variants are predominantly *exo*) and more strongly prefers the *cis* isomer than the other proline analogs examined in this study.[14,31] Perhaps these conformational preferences at the C-terminus of the B-chain inhibit the formation of nuclei that precede fibril formation. TEM analysis of the lispro fibrils revealed similar morphology among samples (Figure 3.5b-e).

We probed the disorder of each lispro variant using the dye 8-anilino-1-naphthalenesulfonic acid (ANS), which exhibits a blue-shift in its emission maximum upon binding to hydrophobic patches of proteins indicative of a molten globule state.[32] The emission spectra of all lispro variants were similar upon addition of ANS (Figure 3.5f), suggesting that the observed differences in fibrillation propensity were not due to variations in protein disorder.

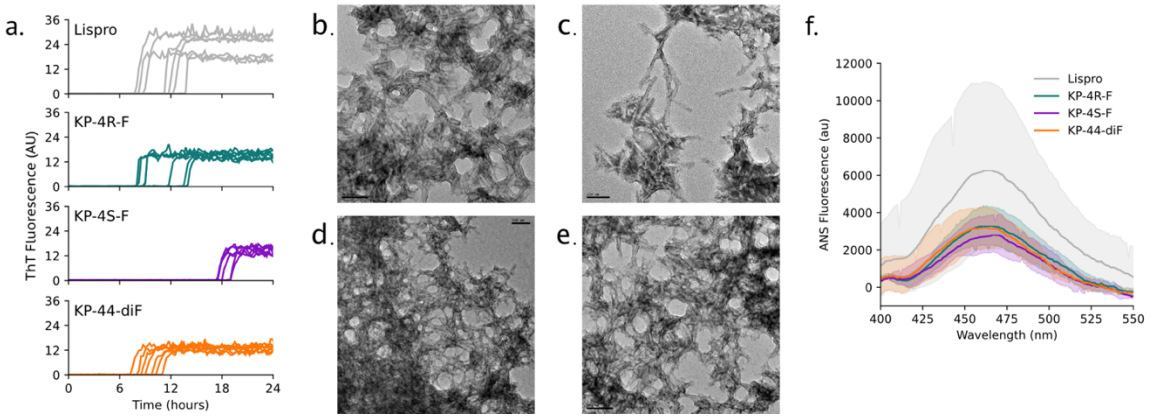

Figure 3.5. Fibrillation of lispro variants. **a**. Lispro variants (60 µM in 100 mM phosphate buffer, pH 8.0) were incubated at 37°C with vigorous shaking; fibril formation was monitored by ThT fluorescence. **b-e**. TEM images of lispro (b), KP-4R-F (c), KP-4S-F (d), and KP-44-diF (e) aggregates. **f**. ANS emission spectra of lispro variants (1 µM lispro variant labeled with 5 µM ANS in 100 mM phosphate buffer, pH 8.0).

In conclusion, we report the synthesis of insulin lispro variants fluorinated at ProB29. While most properties of lispro variants remain unchanged, we find that 4*S*-fluorination of ProB29 in insulin lispro stabilizes the protein against physical denaturation. Approaches to protect insulin against denaturation often rely upon promoting oligomerization.[24,33] However, here we observe stabilization of a monomeric FAI without changes in association behavior. Similar stabilization from the monomer state was observed for an insulin lispro variant that contains a 3-iodo substituent at TyrB26.[34] In both cases, the mechanism for stabilization is not fully established. Protein behavior is complex: ncPro substitution can impart long-range conformational effects,[23] and the precise molecular mechanism of insulin fibrillation is not fully understood.[30]

Regardless of the mechanism, the discovery of a monomeric insulin with an extended lag time to fibril formation (KP-4*S*-F) is relevant to the design of improved insulin therapies. Current insulin formulations must be stored at 4°C and have limited shelf-lives. The inherent sensitivity of insulin to chemical and physical denaturation is especially problematic for long-term storage in continuous subcutaneous insulin infusion (CSII) pumps,[35] and in efforts to provide insulin-replacement therapies to individuals without regular access to refrigeration.[36]

The results of this work highlight the ability of non-canonical amino acid mutagenesis to modulate pharmaceutically-relevant properties of therapeutic proteins. ncPro mutagenesis allows for modification of insulin lispro at position B29 without losing the unique conformational restrictions characteristic of proline, and the fluorinated insulin lispro

variants introduced here can be produced through straightforward bacterial expression. Future work is needed to more thoroughly elucidate the molecular origins for the stabilization of KP-4*S*-F.

3.4 Materials and methods

3.4.1 Chemicals

All chemicals were purchased from MilliporeSigma unless otherwise indicated. 4*R*-fluoroproline (4*R*-F) and 4*S*-fluoroproline (4*S*-F) were purchased from Bachem. 4,4-difluoroproline (44-diF) was purchased from Synthonix. All proline analogs were used as received.

3.4.2 Enzymes

Gibson Assembly enzymes were purchased as the Repliqa HiFi assembly mix from Quantabio. Trypsin was purchased from MilliporeSigma. Carboxypeptidase B was purchased from Worthington Biochemical. Glu-C peptidase was purchased from Promega.

3.4.3 Strains and plasmids

The proline-auxotrophic *E. coli* strain CAG18515 was obtained from the Coli Genetic Stock Center (CGSC) at Yale University. Strain DH10B was used for all cloning applications.

The plasmid pQE80_H27R-PI-KP_proS contains an IPTG-inducible proinsulin-lispro gene and an *E. coli* prolyl-tRNA synthetase gene controlled by its endogenous promoter. Proinsulin is translationally fused to an N-terminal leader peptide (H27R) that increases

expression yields,[37] and a 10x-his tag to facilitate proinsulin enrichment after refolding. The ProB28-LysB29 inversion present in insulin lispro was installed in plasmid pQE80_H27R-PI_proS (see Chapter II). We used a two-part Gibson Assembly approach with the two sets of primers described below. The two overlap regions were at the site of mutation, and within the ampicillin resistance gene selection marker. Correct installation of the desired mutation was verified by Sanger sequencing.

3.4.4 Primers

DNA oligos were purchased from Integrated DNA Technologies (IDT). Nucleotides responsible for the proline-lysine inversion are underlined.

Gibson assembly fragment 1:

SLB3099_GA-fwd: CGTGGTTTCTTCTACACGAAACCGACCCGCCGTGAAG

AmpR-GA_rev: GACAGTAAGAGAATTATGCAGTG

Gibson assembly fragment 2:

AmpR-GA_fwd: GCAGTGTTATCACTCATGG

SLB3099_GA-rev: CAGCTTCACGGCGGGTCGGTTTCGTGTAGAAGAAACCACG

3.4.5 Nucleotide and amino acid sequences

H27R-PI-KP: The mutations responsible for the proline-lysine inversion are underlined.

ATGACAATGATCACTAATTCACCCGAGATTTCCCACCATCATCATCATCATCA
CCACCACCATCAGTTGATCTCGGAGGCCCGTTTTGTGAACCAGCACCTGTGCG
GTAGCCACCTGGTGGAAGCTCTGTACCTGGTTTGCGGTGAGCGTGGTTTCTTC
TACACGAAACCGACCCGCCGTGAAGCTGAAGATCTGCAGGTGGGCCAGGTA

GAACTGGGCGGTGGTCCGGGTGCCGGCTCTCTGCAACCGCTGGCACTGGAAG
GTTCCCTGCAAGCGCGTGGTATCGTAGAGCAGTGCTGTACTTCTATCTGCTCC
CTGTACCAGCTGGAGAACTACTGTAATTAA

The sequence of the H27R leader peptide is underlined, proinsulin is in **bold**. The A-chain and B-chain in mature insulin are colored **red** and **blue**, respectively.

MTMITNSPEISHHHHHHHHHHQLISEAR**FVNQHLCGSHLVEALYLVCGERGFF**
YTKPTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQARGIVEQCCTSICS
LYQLENYCN

proS: the endogenous *proS* promoter is underlined, the coding sequence is in UPPERCASE.

attcacgccctctcttttgacatttcttttgcactggtaaactaaatcacttttttttgtcccaggctcgccttgagcctgttctaccttcc
aactggaaccgtaacaacATGCGTACTAGCCAATACCTGCTCTCCACTCTCAAGGAGAC
ACCTGCCGACGCCGAGGTGATCAGCCATCAGCTGATGCTGCGCGCCGGGATG
ATCCGCAAGCTGGCCTCCGGGTTATATACCTGGCTGCCGACCGGCGTGCGCG
TTCTGAAAAAAGTCGAAAACATCGTGCGTGAAGAGATGAACAACGCCGGTGC
GATCGAGGTGTCGATGCCGGTGGTTCAGCCAGCCGATTTGTGGCAAGAGAGT
GGTCGTTGGGAACAGTACGGTCCGGAACTGCTGCGTTTTGTTGACCGTGGCG
AGCGTCCGTTCGTACTCGGCCCAACTCATGAAGAAGTTATCACTGACCTGATT
CGTAACGAGCTTAGCTCTTACAAACAGCTGCCGCTGAACTTCTATCAGATCCA
GACCAAGTTCCGCGACGAAGTGCGTCCGCGTTTCGGCGTCATGCGTTCCCGC
GAATTCCTGATGAAAGATGCTTACTCTTTCCATACTTCTCAGGAATCCCTGCA
GGAAACCTACGATGCAATGTATGCGGCCTACAGCAAAATCTTCAGCCGCATG
GGGCTGGATTTCCGCGCCGTACAAGCCGACACCGGTTCTATCGGCGGCAGCG
CCTCTCACGAATTCCAGGTGCTGGCGCAGAGCGGTGAAGACGATGTGGTCTT
CTCCGACACCTCTGACTATGCAGCGAACATTGAACTGGCAGAAGCTATCGCG
CCGAAAGAACCGCGCGCTGCTGCTACCCAGGAAATGACGCTGGTTGATACGC
CGAACGCGAAAACCATCGCGGAACTGGTTGAACAGTTCAATCTGCCGATTGA

GAAAACGGTTAAGACTCTGCTGGTTAAAGCGGTTGAAGGCAGCAGCTTCCCG
CAGGTTGCGCTGCTGGTGCGCGGTGATCACGAGCTGAACGAAGTTAAAGCAG
AAAAACTGCCGCAGGTTGCAAGCCCGCTGACTTTCGCGACCGAAGAAGAAAT
TCGTGCCGTGGTTAAAGCCGGTCCGGGTTCACTGGGTCCGGTAAACATGCCG
ATTCCGGTGGTGATTGACCGTACCGTTGCGGCGATGAGTGATTTCGCTGCTGG
TGCTAACATCGATGGTAAACACTACTTCGGCATCAACTGGGATCGCGATGTC
GCTACCCCGGAAGTTGCAGATATCCGTAACGTGGTGGCTGGCGATCCAAGCC
CGGATGGCCAGGGTAGGCTGCTGATCAAACGTGGTATCGAAGTTGGTCACAT
CTTCCAGCTGGGTACCAAGTACTCCGAAGCACTGAAAGCCTCCGTACAGGGT
GAAGATGGCCGTAACCAAATCCTGACGATGGGTTGCTACGGTATCGGGGTAA
CGCGTGTGGTAGCTGCGGCGATTGAGCAGAACTACGACGAACGAGGCATCGT
ATGGCCTGACGCTATCGCGCCGTTCCAGGTGGCGATTCTGCCGATGAACATG
CACAAATCCTTCCGCGTACAAGAGCTTGCTGAGAAACTGTACAGCGAACTGC
GTGCACAAGGTATCGAAGTGCTGCTGGATGACCGCAAAGAGCGTCCGGGCGT
GATGTTTGCTGATATGGAACTGATCGGTATTCCGCACACTATTGTGCTGGGCG
ACCGTAACCTCGACAACGACGATATCGAATATAAATATCGTCGCAACGGCGA
GAAACAGTTAATTAAGACTGGTGACATCGTCGAATATCTGGTGAAACAGATT
AAAGGCTGA

MRTSQYLLSTLKETPADAEVISHQLMLRAGMIRKLASGLYTWLPTGVRVLKKVE
NIVREEMNNAGAIEVSMPVVQPADLWQESGRWEQYGPELLRFVDRGERPFVLG
PTHEEVITDLIRNELSSYKQLPLNFYQIQTKFRDEVRPRFGVMRSREFLMKDAYSF
HTSQESLQETYDAMYAAYSKIFSRMGLDFRAVQADTGSIGGSASHEFQVLAQSG
EDDVVFSDTSDYAANIELAEAIAPKEPRAAATQEMTLVDTPNAKTIAELVEQFNL
PIEKTVKTLLVKAVEGSSFPQVALLVRGDHELNEVKAEKLPQVASPLTFATEEEIR
AVVKAGPGSLGPVNMPIPVVIDRTVAAMSDFAAGANIDGKHYFGINWDRDVAT
PEVADIRNVVAGDPSPDGQGRLLIKRGIEVGHIFQLGTKYSEALKASVQGEDGRN
QILTMGCYGIGVTRVVAAAIEQNYDERGIVWPDAIAPFQVAILPMNMHKSFRVQ
ELAEKLYSELRAQGIEVLLDDRKERPGVMFADMELIGIPHTIVLGDRNLDNDDIE
YKYRRNGEKQLIKTGDIVEYLVKQIKG

3.4.6 Proinsulin expression

A single colony of *E. coli* strain CAG18515 harboring plasmid pEQ80_H27R-PI-KP_proS was used to inoculate 70 mL of Luria Bertani (LB) medium containing ampicillin, and the culture was grown overnight at 37°C to stationary phase. The overnight culture was used to inoculate 5 L (as 4x1.25 L cultures) of 1x Andrew's Magical Medium (AMM),[38] a defined medium containing all 20 proteinogenic amino acids, in 2.8 L Fernbach flasks. The composition of AMM was the following: 3.60 g L^{-1} glucose, 3.5 g L^{-1} KH_2PO_4, 6.56 g L^{-1} $K_2HPO_4 \bullet 3H_2O$, 3.5 g L^{-1} $(NH_4)_2HPO_4$, 8.37 g L^{-1} MOPS, 0.72 g L^{-1} tricine, 2.92 g L^{-1} NaCl, 0.51 g L^{-1} NH_4Cl, 0.26 g L^{-1} $MgCl_2 \bullet 7H_2O$, 50 mg L^{-1} K_2SO_4, 0.246 mg L^{-1} $MgSO_4 \bullet 7H_2O$, 12.3 mg L^{-1} $CaCl_2$,$\bullet 2H_2O$, 2.8 mg L^{-1} $FeSO_4 \bullet 7H_2O$, 0.5 mg L^{-1} thiamine, 24 μg L^{-1} boric acid, 1 μg L^{-1} trace metals (Cu^{2+}, Mn^{2+}, Zn^{2+}, MoO_4^{2-}), and 50 mg L^{-1} each amino acid.

When growth reached mid-exponential phase (OD_{600} ~0.8), the culture was subjected to a medium shift: cells were pelleted via centrifugation (5 kg, 5 min, 4°C) and washed twice with ice-cold 0.9% NaCl. Washed cells were resuspended in 1 L of 1.25x AMM –Pro, a 1.25x concentrated form of AMM that omits proline. Cells were incubated for 30 min at 37°C to deplete residual proline, after which 250 mL of a solution containing 2.5 mM ncPro and 1.5 M NaCl was added (0.5 mM ncPro and 0.3 M NaCl working concentrations). For the incorporation of 44-diF, the concentration of NaCl was 2.5 M (0.5 M working concentration). After 30 min of incubation at 37°C to allow for ncPro uptake, proinsulin expression was induced by the addition of 1 mM IPTG. Cultures were incubated overnight

at 37°C, after which cells were harvested via centrifugation and stored at -80°C until further processing.

Proline-containing proinsulins were expressed in rich medium: proinsulin-lispro was expressed using strain CAG18515 harboring plasmid pQE80-H27R-PI-KP_proS in 6 L (as 6 x 1.0 L cultures) of Terrific Broth (TB). 1 mM IPTG was added at mid-log phase (OD_{600} ~0.8) to induce proinsulin expression. Cultures were incubated at 37°C for 3 h, after which cells were harvested via centrifugation and stored at -80°C until further processing.

3.4.7 Proinsulin refolding

Cell pellets were warmed from -80°C to room temperature and resuspended in 5 mL IB buffer (50 mM tris, 100 mM NaCl, 1 mM EDTA, pH 8.0) per gram cell pellet. 1 mg L^{-1} lysozyme and 1 mM PMSF were added, and the slurry sat on ice for 30 min before cells were lysed via sonication. The lysate was centrifuged (14 kg, 30 min, 4°C) and the soluble fraction was discarded. The pellet was washed twice with IB buffer + 1% Triton X-100, once with IB buffer, and once with water; this final step required centrifugation for 45 min. The washed inclusion body pellet was resuspended in a minimal amount of water, and the mass of proinsulin in the inclusion body pellet was estimated by SDS-PAGE.

In preparation for proinsulin refolding, the inclusion body was resuspended in 3 M urea and 10 mM cysteine in water, such that the proinsulin concentration was 1 mg proinsulin per L total slurry. To dissolve proinsulin, the pH was adjusted to 12 and sample stirred for 1 h at room temperature. At this stage, ncPro incorporation was assessed by MALDI-TOF, which is described in section 3.4.10 below. The solubilized proinsulin solution was diluted

ten-fold into refold buffer (10 mM CAPS, pH 10.6) that had been pre-cooled to 4°C. The pH of the refold solution was adjusted to 10.7 and the sample stored at 4°C; care was taken to ensure that the solution pH remained between 10.6 and 10.8 throughout the refolding process. Proinsulin refolding progress was monitored by reverse-phase HPLC, and usually reached completion within 50 h.

Proinsulin was enriched from the refold solution after adjusting the pH to 8.0 and incubating the sample overnight with Ni-NTA resin and 10 mM imidazole. The resin was washed with wash buffer (25 mM imidazole in PBS, pH 8.0), and proinsulin was eluted with elution buffer (250 mM imidazole in PBS, pH 8.0). Fractions containing proinsulin were combined and extensively dialyzed against 10 mM sodium phosphate, pH 8.0.

3.4.8 Lispro maturation and purification

Refolded and dialyzed proinsulin was warmed to 37°C and digested with trypsin (20 U mL^{-1}) and carboxypeptidase-B (10 U mL^{-1}) at 37°C for 2.5 h to remove the N-terminal tag and C chain. Digestion was halted by adjusting the pH to ~3 with 6 N HCl.

Lispro variants were immediately purified after proteolysis by reverse-phase HPLC on a C_4 column (Penomenex Jupiter 5 μm particle size, 300 Å pore size, 250x10 mm) using 0.1% TFA in water (solvent A) and 0.1% TFA in acetonitrile (solvent B) as mobile phases. A gradient of 25-32% solvent B was applied over 65 min, and fractions containing lispro were collected. Samples for purity analysis were removed at this stage; the remaining portion of the fraction was lyophilized. Each lispro fraction was analyzed by analytical reverse-phase HPLC, MALDI-TOF MS (Figure 3.2e-h), and SDS-PAGE (Figure 3.S1) to

verify sample quality and ensure ≥95% purity for all downstream analyses. Lyophilized powders were stored at -20°C until further use.

3.4.9 Preparation of lispro variants for mouse experiments

We used an earlier version of this protocol to express lispro variants used in mouse studies.[24,25] Briefly, the hexahistidine-tagged version of proinsulin (lacking the H27R leader peptide) was expressed in M9 medium using the strain CAG18515 harboring plasmid pQE80PI-KP_proS. The medium shift described above was used to incorporate the ncPro residue of interest, and proinsulin was expressed for 2.5 h. After cell lysis, proinsulin was solubilized with 8 M urea and isolated from the washed inclusion body fraction by Ni-NTA under denaturing conditions. Refolding was achieved after oxidative sulfitolysis by gentle agitation at 12°C overnight in 0.5 mM BME, 0.3 M urea, 50 mM glycine, pH 12. After dialysis, lyophilization, and proteolysis by trypsin and carboxypeptidase B, mature lispro variants were purified by reverse-phase HPLC.

During the course of this work, we found that the refolding protocol described in section 3.4.7 resulted in an approximately seven-fold increase in refolding yield, and higher purity samples.

3.4.10 MALDI-TOF MS

To assess levels of incorporation of ncPros into the corresponding proinsulins, samples were subjected to Glu-C digestion, which results in a peptide fragment containing ProB29 (^{50}RGFFYTK**P**TRRE). A 20 µL aliquot of the solubilized proinsulin-containing inclusion body fraction was subjected to cysteine reduction (5 mM DTT, 55°C for 20 min) and

alkylation (15 mM iodoacetimide, RT for 15 min in the dark), prior to 10-fold dilution into 100 mM NH_4HCO_3, pH 8.0 (100 µL final volume). Digestion was started with addition of 0.6 µL Glu-C (0.5 µg μL^{-1} in ddH_20) at 37°C for 2.5 h. The digestion reaction was quenched by adding 10 µL of 5% TFA. Peptides were desalted using ZipTip C_{18} columns (MilliporeSigma) according to the manufacturer's protocol. Desalted peptides (in 50% acetonitrile, 0.1% TFA) were diluted 3:1 into the matrix solution (α-cyanohydroxycinnamic acid in 50% ACN, 0.1% TFA) and analyzed by MALDI-TOF MS. Analog incorporation was calculated by comparing the area under the curve (AUC) of the ncPro form of the peptide (m/z = 1576 for 4*R*-F and 4*S*-F, and 1595 for 44-diF) with the AUC of the canonical proline peptide (m/z = 1558).

HPLC-purified insulin and lispros were analyzed as full-length, mature proteins. Aliquots directly from HPLC purification (~30% ACN, 0.1% TFA) were mixed 1:1 with matrix solution (sinapic acid in 30% ACN, 0.1% TFA) before analysis by MALDI-TOF MS.

3.4.11 Reduction of blood glucose in diabetic mice

NODscid mice (NOD.CB17-Prkdcscid/J) were obtained from Jax Mice (Bar Harbor, ME). Mice were maintained under specific pathogen-free conditions, and experiments were conducted according to procedures approved by the Institutional Animal Care and Use Committee (IACUC) at the City of Hope. Adult (8-12 week old) male NODscid mice were injected intraperitoneally (45 mg kg^{-1} day^{-1} for 3 days) with freshly prepared streptozotocin (STZ) in 50 mM citrate buffer, pH 4.5 to induce diabetes. Diabetes was confirmed 3 weeks after the last dose of STZ by detection high glucose levels (200-600 mg dL^{-1}) as measured

by a glucomonitor (Freestyle, Abbott Diabetes Care, Alameda, CA) in blood sampled from the lateral tail vein. Lispro analogs were diluted to 100 μg mL^{-1} in formulation buffer (1.6 mg mL^{-1} *m*-cresol, 0.65 mg mL^{-1} phenol, 3.8 mg mL^{-1} sodium phosphate pH 7.4, 16 mg mL^{-1} glycerol, 0.8 μg mL^{-1} $ZnCl_2$). Lispro analogs were injected (35 μg kg^{-1}) subcutaneously at the scruff and blood glucose was measured at 0, 10, 20, 30, 40, 50, 60, 80, 100, 120, and 150 min.

3.4.12 Circular dichroism spectroscopy

Equilibrium measurements: The circular dichroism spectra of lispro samples (60 μM in 100 mM sodium phosphate, pH 8.0) were measured at 25°C in 1 mm quartz cuvettes on an Aviv Model 430 Circular Dichroism Spectrophotometer using a step size of 0.5 nm and averaging time of 1 s. A reference buffer spectrum was subtracted from each sample spectrum.

Kinetic measurements: Insulin samples in 100 mM sodium phosphate buffer pH 8.0 were dialyzed overnight against 28.6 mM tris buffer, pH 8.0 (Slide-A-Lyzer dialysis cassettes, 3.5 kDa MWCO, ThermoFisher. Insulins were formulated as the following: 600 μM insulin, 250 μM $ZnCl_2$, 25 mM resorcinol, 25 mM tris buffer, pH 8. To a stirred buffer solution containing 2.98 mL of 25 mM tris, pH 8.0 in a 10 mm quartz cuvette was injected 20 μL of the insulin formulation (150-fold dilution). Ellipticity was monitored at 222 nm over 120 s (1 s kinetic interval, 0.5 s time constant) at 25°C. A typical run led to a rapid drop in CD signal as mixing occurred (~5 s), then a gradual rise to an equilibrium ellipticity representative of an insulin monomer. Data preceding the timepoint 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 deg cm^2 $dmol^{-1}$, which was indicative of poor mixing. The remaining data were fit to a mono-exponential function using Scipy (Python); data presented here are from at least two separate HPLC fractions, measured on two different days.

An equilibrium spectrum for each protein was obtained after dilution; all spectra were indicative of an insulin monomer. The CD spectrum of lispro under pre-dilution formulation conditions was obtained using a 0.1 mm quartz cuvette. In each case, a blank spectrum containing all buffers and ligands was subtracted from the sample spectrum.

3.4.13 Fibrillation

Lispro samples (60 μM 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 μM thioflavin T (ThT). Each lispro (200 μL) was added to a 96-well, black, clear bottom plate (Greiner Bio-One) and sealed. Samples were shaken continuously at 960 rpm on a Varioskan multimode plate reader 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, and on 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.

3.4.14 Transmission electron microscopy

Lispro fibrils were centrifuged (5 kg, 1 min), then washed twice and resuspended in ddH_2O. Fibrils were stained with 2% uranyl acetate on a 300-mesh formvar/carbon coated copper grid (Electron Microscopy Sciences) and imaged on a Tecnai T12 LaB6 120 eV transmission electron microscope.

3.4.15 ANS fluorescence

1 μM lispro variant was mixed with 5 μM ANS in 100 mM phosphate buffer, pH 8.0. Fluorescence emission spectra were measured in 1 cm quartz cuvettes at ambient temperature using a PTI QuantaMaster fluorescence spectrofluorometer. A 350 nm excitation wavelength and scan rate of 2 nm s^{-1} were used. Measurements for each variant were performed in triplicate from three separate HPLC fractions.

3.4.16 Analytical ultracentrifugation

Lispro variants were formulated at 125 μM in 100 mM phosphate buffer, pH 8.0. Velocity sedimentation experiments were performed at the Canadian Center for Hydrodynamics by Maduni Ranashinghe, with input from Borries Demeler. Data were analyzed with UltraScan III version 4.0 release 6606.[39] Sedimentation coefficients were determined from the enhanced van Holde-Weischhet[40] analysis to generate diffusion-corrected sedimentation coefficient distributions.

3.5 Supplementary figures and tables

Figure 3.S1. Purity of lispro samples assessed by SDS-PAGE. HPLC-purified lispro (a), KP-4*R*-F (b), KP-4*S*-F (c), and KP-44diF were analyzed by SDS-PAGE to validate purity; shown are representative lanes corresponding to individual HPLC fractions.

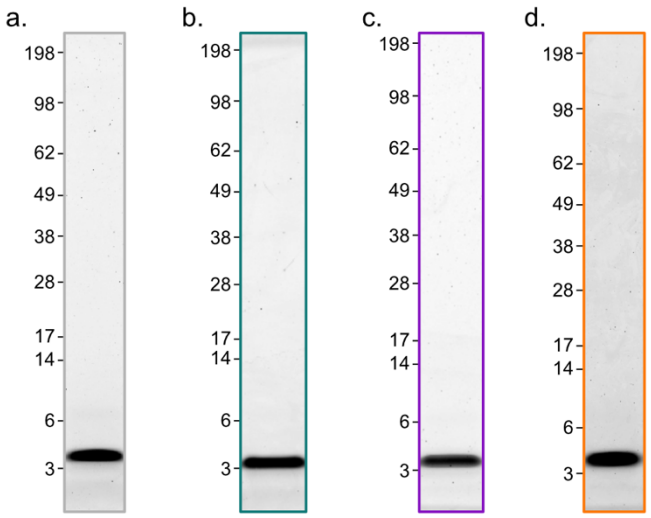

Figure 3.S2. Circular dichroism controls. At 60 µM, human insulin is expected to exist as a dimer at pH 8, monomer in 20% ethanol, and denatured in the presence of 8 M guanidinium chloride. These spectra are overlaid with equilibrium spectra collected before and after lispro dilution for kinetic CD measurements. Spectra below 210-215 nm were omitted for some samples due to high levels of buffer absorbance at these wavelengths.

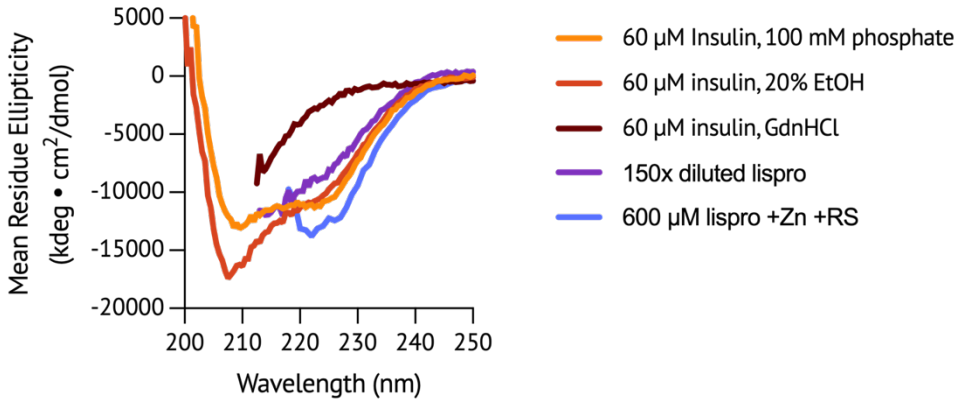

Table 3.S1. Mass spectrometry characterization of lispro variants

	Digested peptide			Mature insulin	
Protein	Expected m/z	Observed m/z	Incorporation efficiency	Expected m/z	Observed m/z
Lispro	1557.78	1558.0	–	5808.6	5807.7 ± 0.4
KP-4*R*-F	1575.82	1575.737 ± 0.008	0.961 ± 0.005	5826.6	5826.31 ± 0.03
KP-4*S*-F	1575.82	1575.53 ± 0.01	0.948 ± 0.033	5826.6	5826.5 ± 0.1
KP-44-diF	1593.81	1593.57 ± 0.01	0.912 ± 0.016	5844.6	5844.1 ± 0.7

Table 3.S2. Lispro yields

Protein	Proinsulin yield (mg L^{-1})[‡]	Approx. insulin yield (mg L^{-1})
Lispro[*]	33	n.d.[#]
KP-4*R*-F	32	3.4
KP-4*S*-F	7	2.2
KP-44-diF	8	2.4

[‡]Yields determined by measuring absorbance (280 nm) after Ni-NTA enrichment following proinsulin refolding.

[*]Expressed in terrific broth (TB)

[#]n.d., not determined

Table 3.S3. Summary of lispro variant characterization

Protein	Ellipticity ratio (208/222 nm)[‡]	Hexamer dissociation $t_{1/2}$ (s)	Fibrillation lag time (h)[‡]	ANS emission maximum (nm)[%]	Sedimentation coefficient (S)[&]
Lispro	1.63 ± 0.07	11.3 ± 3.8	10.9 ± 2.2	466 ± 2	1.0
KP-4*R*-F	1.62 ± 0.06	10.2 ± 1.6	10.3 ± 2.4	464 ± 4	1.0
KP-4*S*-F	1.59 ± 0.06	11.2 ± 3.3	17.9 ± 0.8	463 ± 6	1.1
KP-44-diF	1.61 ± 0.05	9.9 ± 2.9	9.0 ± 1.3	462 ± 5	1.1
Insulin	1.24 ± 0.03[*]	30.2 ± 2.8	16.6 ± 4.1[*]	470 ± 5[*]	n.d.[#]

[‡]60 μM insulin, 100 mM phosphate buffer, pH 8.0

[%]1 μM insulin, 5 μM ANS, 100 mM phosphate buffer, pH 8.0

[&]125 μM insulin, 100 mM phosphate buffer, pH 8.0

[*]Chapter II

[#]n.d., not determined

3.6 References

(1) Haeusler, R. A.; McGraw, T. E.; Accili, D. Biochemical and Cellular Properties of Insulin Receptor Signalling. *Nat Rev Mol Cell Biol* **2018**, *19*, 31–44. https://doi.org/10.1038/nrm.2017.89.

(2) World Heath Organization. *Global Report on Diabetes*; 2016. https://www.who.int/publications/i/item/9789241565257 (accessed 2023-04-02).

(3) 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* **1979**, *76* (1), 106–110. https://doi.org/10.1073/pnas.76.1.106.

(4) Derewenda, U.; Derewenda, Z.; Dodson, E. J.; Reynolds, C. D.; Smith, G. D.; Sparks, C.; Swenson, D. Phenol Stabilizes More Helix in a New Symmetrical Zinc Insulin Hexamer. *Nature* **1989**, *338*, 594–596. https://doi.org/10.1038/338594a0.

(5) Birnbaum, D. T.; Kilcomons, M. A.; DeFelippis, M. R.; Beals, J. M. Assembly and Dissociation of Human Insulin and LysB28-ProB29-Insulin Hexamers: A Comparison Study. *Pharm Res* **1997**, *14* (1), 25–36. https://doi.org/10.1023/a:1012095115151.

(6) Holleman, F.; Hoekstra, J. B. L. Insulin Lispro. *N Engl J Med* **1997**, *337* (3), 176–183. https://doi.org/10.1056/NEJM199707173370307.

(7) 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. https://doi.org/10.1002/pro.5560051215.

(8) 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. https://doi.org/10.1093/protein/5.6.527.

(9) 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. https://doi.org/10.1038/333679a0.

(10) 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. https://doi.org/10.2337/diabetes.52.9.2227.

(11) 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. https://doi.org/10.1038/nrendo.2017.39.

(12) Brems, D. N.; Brown, P. L.; Bryant, C.; Chance, R. E.; Green, L. K.; Long, H. B.; Miller, A. A.; Millican, R.; Shields, J. E.; Frank, B. H. Improved Insulin Stability through

Amino Acid Substitution. *Protein Eng* **1992**, *5* (6), 519–525. https://doi.org/10.1093/protein/5.6.519.

(13) Lepore, M.; Pampanelli, S.; Fanelli, C.; Porcellati, F.; Bartocci, L.; Di Vincenzo, A.; Cordoni, 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. https://doi.org/10.2337/diabetes.49.12.2142.

(14) Ganguly, H. K.; Basu, G. Conformational Landscape of Substituted Prolines. *Biophys Rev* **2020**, *12*, 25–39. https://doi.org/10.1007/s12551-020-00621-8.

(15) Chin, J. W. Expanding and Reprogramming the Genetic Code. *Nature* **2017**, *550*, 53–60. https://doi.org/10.1038/nature24031.

(16) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. Fluoroprolines as Tools for Protein Design and Engineering. *Angew Chem Int Ed* **2001**, *40* (5), 923–925. https://doi.org/10.1002/1521-3773(20010302)40:5<923::AID-ANIE923>3.0.CO;2-#.

(17) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. Code for Collagen's Stability Deciphered. *Nature* **1998**, *392*, 666–667. https://doi.org/10.1038/33573.

(18) 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. https://doi.org/10.1038/nature04130.

(19) Crespo, M. D.; Rubini, M. Rational Design of Protein Stability: Effect of (2*S*,4*R*)-4-Fluoroproline on the Stability and Folding Pathway of Ubiquitin. *PLoS One* **2011**, *6* (5), e19425. https://doi.org/10.1371/journal.pone.0019425.

(20) Naduthambi, D.; Zondlo, N. J. Stereoelectronic Tuning of the Structure and Stability of the Trp Cage Miniprotein. *J Am Chem Soc* **2006**, *128*, 12430–12431. https://doi.org/10.1021/ja0648458.

(21) Rubini, M.; Schärer, M. A.; Capitani, G.; Glockshuber, R. (4*R*)- and (4*S*)-Fluoroproline in the Conserved *Cis*-Prolyl Peptide Bond of the Thioredoxin Fold: Tertiary Structure Context Dictates Ring Puckering. *ChemBioChem* **2013**, *14* (9), 1053–1057. https://doi.org/10.1002/cbic.201300178.

(22) Torbeev, V. Y.; Hilvert, D. Both the *Cis-Trans* Equilibrium and Isomerization Dynamics of a Single Proline Amide Modulate β2-Microglobulin Amyloid Assembly. *Proc Natl Acad Sci* **2013**, *110* (50), 20051–20056. https://doi.org/10.1073/pnas.1310414110.

(23) Torbeev, V.; Ebert, M. O.; Dolenc, J.; Hilvert, D. Substitution of Proline32 by α-Methylproline Preorganizes β2-Microglobulin for Oligomerization but Not for Aggregation into Amyloids. *J Am Chem Soc* **2015**, *137* (7), 2524–2535. https://doi.org/10.1021/ja510109p.

(24) 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. https://doi.org/10.1002/pola.29225.

(25) Lieblich, S. A.; Fang, K. Y.; Cahn, J. K. B.; Rawson, J.; LeBon, J.; Teresa Ku, H.; Tirrell, D. A. 4*S*-Hydroxylation of Insulin at ProB28 Accelerates Hexamer Dissociation and Delays Fibrillation. *J Am Chem Soc* **2017**, *139*, 8384–8387. https://doi.org/10.1021/jacs.7b00794.

(26) Breunig, S. L.; Tirrell, D. A. Incorporation of Proline Analogs into Recombinant Proteins Expressed in *Escherichia coli*. *Methods Enzymol* **2021**, *656*, 545–571. https://doi.org/10.1016/BS.MIE.2021.05.008.

(27) Plum, A.; Agersø, H.; Andersen, L. Pharmacokinetics of the Rapid-Acting Insulin Analog, Insulin Aspart, in Rats, Dogs, Pigs, and Pharcamodynamics of Insulin Aspart in Pigs. *Drug Metab Dispos* **2000**, *28*, 155–160.

(28) 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. https://doi.org/10.1038/nature11781.

(29) Rahuel-Clermont, S.; French, C. A.; Kaarsholm, N. C.; Dunn, M. F. Mechanisms of Stabilization of the Insulin Hexamer through Allosteric Ligand Interactions. *Biochemistry* **1997**, *36* (19), 5837–5845. https://doi.org/10.1021/bi963038q.

(30) Akbarian, M.; Yousefi, R.; Farjadian, F.; Uversky, V. N. Insulin Fibrillation: Toward Strategies for Attenuating the Process. *ChemCommun* **2020**, *56* (77), 11354–11373. https://doi.org/10.1039/d0cc05171c.

(31) Shoulders, M. D.; Raines, R. T. Collagen Structure and Stability. *Annu Rev Biochem* **2009**, *78*, 929–958. https://doi.org/10.1146/annurev.biochem.77.032207.120833.

(32) 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. https://doi.org/10.1002/BIP.360310111.

(33) Brange, J.; Andersen, L.; Laursen, E. D.; Meyn, G.; Rasmussen, E. Toward Understanding Insulin Fibrillation. *J Pharm Sci* **1997**, *86* (5), 517–525. https://doi.org/10.1021/js960297s.

(34) Pandyarajan, V.; Phillips, N. B.; Cox, G. P.; Yang, Y.; Whittaker, J.; Ismail-Beigi, F.; Weiss, M. A. Biophysical Optimization of a Therapeutic Protein by Nonstandard Mutagenesis: Studies of an Iodo-Insulin Derivative. *J Biol Chem* **2014**, *289* (34), 23367–23381. https://doi.org/10.1074/jbc.M114.588277.

(35) Reichel, A.; Rietzsch, H.; Köhler, H.; Pfützner, A.; Gudat, U.; Schulze, J. Cessation of Insulin Infusion at Night-Time during CSII-Therapy: Comparison of Regular Human Insulin and Insulin Lispro. *Exp Clin Endocrinol Diabetes* **1998**, *106* (03), 168–172. https://doi.org/10.1055/s-0029-1211971.

(36) Weiss, M. A. Design of Ultra-Stable Insulin Analogues for the Developing World. *J Health Specialties* **2013**, *1* (2), 59–70. https://doi.org/10.4103/1658-600X.114683.

(37) Min, C. K.; Son, Y. J.; Kim, C. K.; Park, S. J.; Lee, J. W. Increased Expression, Folding and Enzyme Reaction Rate of Recombinant Human Insulin by Selecting Appropriate Leader Peptide. *J Biotechnol* **2011**, *151* (4), 350–356. https://doi.org/10.1016/j.jbiotec.2010.12.023.

(38) 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. https://doi.org/10.1016/J.YMBEN.2014.11.003.

(39) Demeler, B.; Gorbet, G. E. Analytical Ultracentrifugation Data Analysis with Ultrascan-III. *Analytical Ultracentrifugation: Instrumentation, Software, and Applications* **2016**, 119–143. https://doi.org/10.1007/978-4-431-55985-6_8.

(40) Demeler, B.; Van Holde, K. E. Sedimentation Velocity Analysis of Highly Heterogeneous Systems. *Anal Biochem* **2004**, *335* (2), 279–288. https://doi.org/10.1016/J.AB.2004.08.039.