PNAS
www.pnas.org

Supplementary Information for

Evolution of Insulin at the Edge of Foldability and its Medical Implications

Nischay K. Rege, Ming Liu, Yanwu Yang, Balamurugan Dhayalan, Nalinda P. Wickramasinghe, Yen-Shan Chen, Leili Rahimi, Huan Guo, Leena Haataja, Jinhong Sun, Faramarz Ismail-Beigi, Nelson B. Phillips, Peter Arvan, & Michael A. Weiss*

Corresponding Author: Michael A. Weiss
Email: weissma@iu.edu

This PDF file includes:

Supplementary Information Text

1. Supplemental Results

Expanded description of the NMR analysis. Engineered insulin monomers contained substitutions Asp^{B10} (D), Lys^{B28} (K) and Pro^{B29} (P) and so are designated "DKP." These analogs are more amenable to biosynthetic labeling with ^{13}C and ^{15}N in yeast *Pichia pastoris* due to the favorable effects of Asp^{B10} on recombinant expression and yield. The present studies also employed insulin *lispro,* which is the active component of the rapid-acting insulin formulation Humalog® (Lilly, Indianapolis, IN), as a template to investigate protein thermostability through amide-proton ^{1}H-^{2}H exchange (at pH 3 and 25 °C in the presence of 10 mM deuterio-acetic acid). This therapeutic analog contains Lys^{B28} (K) and Pro^{B29} (P) and so is designated "KP." All NMR data were acquired using a BRUKER 700 MHz spectrometer.

2D homonuclear ^{1}H-NMR spectra indicated similar, but not identical, chemical-shift dispersion and native-like packing of Tyr^{B24} against the hydrophobic core of insulin (Fig. S14). Patterns of ^{1}H-^{15}N main-chain chemical shifts and ^{1}H-^{13}C chemical shifts in respective HSQC spectra were also similar (Fig. S15). The secondary ^{1}H-NMR chemical shifts of the upfield-shifted side chain of Leu^{B15} (a characteristic signature of the B24 ring current in the native B-chain super-secondary structure) were attenuated in the Tyr^{B24} variant (Fig. S14A and Table S4a); corresponding signature nuclear Overhauser effects (NOEs) between Leu^{B15} methyl protons and those of the B24 aromatic side chain were observed in spectra of each analog (Figure S14B,C and Table S4a). These features were associated with a subtle pattern of attenuation of neighboring ^{1}H-^{15}N secondary shifts (Fig. S15*B* and Table S4b). These qualitative ^{1}H, ^{13}C and ^{15}N spectra features provide evidence of a native-like structure with a local change in structure and/or dynamics near the site of substitution transmitted to the central B-chain α-helix and to proximal portions of the A chain (expected at Val^{A3} via native-like Tyr^{B26}-Val^{A3} packing and expected at Asn^{A21} via its inter-chain hydrogen bond to the carbonyl oxygen of Gly^{B23} and via cystine B19-A20).

The overall pattern of inter-residue NOEs in spectra of Tyr^{B24}-DKP-insulin is similar to that of its parent DKP-insulin. Subtle differences were observed in the precise set of B24-related NOEs within a shared framework of contacts (Table S5a). Subtle differences in inter-residue distances were observed in the respective solution structures (Table S5b,c).

Statistics pertaining to NMR-derived restraints and distance-geometry calculations are provided in Table S6a (Tyr^{B24}-DKP-insulin) and Table S6b (Phe^{B24}-DKP-insulin). These structures were determined via heteronuclear multidimensional NMR spectroscopy (the latter represents a refinement of a previously reported structure; (S1)). 3D ^{13}C-edited NOESY and 4D time-shared NOESY experiments were used to extract distance constraints; dihedral angle constraints were derived from chemical shift data using TALOS+ program. The solution structure of the Tyr^{B24} analog was calculated by simulated annealing (SA) based on 925 NMR distance restraints and 83 dihedral-angular restraints (Table S6). 20 NMR structures were chosen from 100 calculated structures. The best-fit superposition of the polypeptide backbone (C_a, N and C') is shown in Figure 9A (main text). The average root-mean-square difference (RMSD, residues A2-A20 and B5-B26) between this set of structures is 0.21 Å for main-chain atoms and 0.77 Å for all heavy atoms. The solution structure of the Phe^{B24} parent analog was also determined by parallel analysis of NOESY spectra. The corresponding ensemble of 20 SA structures (Fig. 8B) exhibits a precision of 0.23 Å (main-chain atoms) and 0.77 Å (all heavy atoms). The parent and Tyr^{B24} ensembles are essentially identical in their ordered regions with subtle differences in side-chain dihedral angles at B24 and a few neighboring sites (generally < 10°; Table S6a,b).

Ramachandran plots of the Phe^{B24} and Tyr^{B24} ensembles indicated that all residues were located in allowed regions with no residues in disallowed regions. Statistical assessment of the ensembles indicated that in each case there are no distance violations greater than 0.3 Å and no dihedral-angle violations greater than 5º. The structural features of residues surrounding B24 position were further analyzed (Fig. 8C,D with attention to Ile^{A2}, Tyr^{A19}, Leu^{B11}, Val^{B12}, Leu^{B15}, Tyr^{B16}, Tyr^{B24} [or Phe^{B24}] and Tyr^{B26}) in relation to (χ_1, χ_2) angles (Table S7). This result also reveals high structural

similarity between the Tyr^{B24} analog and its parent Phe^{B24} analog. This similarity in core packing is in accordance with the similarity between respective 4D ^{13}C-^{13}C edited NOESY spectra (Fig. S16).

Amide-proton exchange kinetics. NMR-monitored ^{1}H-^{2}H amide proton-exchange was used to probe the thermodynamic stability of Tyr^{B24}-KP independently of the two-state model standardly employed in analysis of CD-unfolding data (S2). Our protocol was described in Hua, Q.-X., et al. (S1). Protection factors (PF) were calculated for individual protons based on rates of exchange of intrinsic ^{1}H protons for ^{2}H in D_2O with 10 mM deuterio-acetic acid. Whereas the exchange of protons with rapid exchange rates was measured using successive 1D ^{1}H-NMR spectra (Fig. S17) in D_2O solution, that of more slowly exchanging protons was monitored using serial 2D TOCSY spectra (Fig. S18) (see Methods) (S3) (Figure 10A in main text). Amide protons were categorized as sites of global, sub-global, or local exchange, which respectively require complete, partial, or local unfolding of the insulin molecule to enable ^{1}H-^{2}H exchange to (Table S7, Fig. 9B-D). The overall pattern of global and sub-global sites was preserved between Tyr^{B24}- and parent-KP analogs: protons within the hydrophobic core in the C-terminal α-helix of the insulin A chain and the internal surface of the B-chain α-helix constituted a majority of global- and sub-global sites of exchange (Fig. 9A, Table S7). However, in relation to the parent analog, PFs were generally decreased in the Tyr^{B24} analog corresponding to a calculated ΔG_u of 2.55(±0.02) kcal/mol compared to 2.89(±0.02) kcal/mol for the parent Phe^{B24} analog.

We note in passing a seeming paradox: it is possible that there are *more* sites of global exchange kinetics in a destabilized variant than in the WT protein. This paradox is illusory. Because global exchange provides an upper bound on observable PFs, decreasing this upper bound via a destabilizing mutation (such as Tyr^{B24}) can convert a site of subglobal exchange in the parent into a site of global exchange in the less-stable analog. This can happen when the rate of global unfolding equals the former rate of subglobal conformational fluctuation. Conversely, stabilizing mutations can in favorable cases expand the scale for detection of subglobal motions

2. Supplemental Biophysical Discussion

The solution structure of a Tyr^{B24} insulin analog is essentially identical to WT. Nonetheless, detailed comparison of (a) Tyr^{B24}-insulin *lispro* to parent monomer insulin *lispro* and (b) of dimerization-competent Tyr^{B24}, Orn^{B29}-insulin to its parent Orn^{B29}-insulin highlighted a range of subtle perturbations. A similar range of properties is observed in the natural history of vertebrate insulins and IGFs.

Thermodynamic stability. The $Phe^{B24}\rightarrow$Tyr substitution modestly impairs stability as inferred from guanidine titration at pH 7.4 ($\Delta\Delta G_u$ 0.9(±0.2) kcal/mole for the Orn^{B29} analogs and $\Delta\Delta G_u$ 0.6(±0.2) kcal/mole for the *lispro* monomeric analogs; Table 2) and ^{1}H-^{2}H exchange at pD 3.0 (*lispro* analogs; $\Delta\Delta G_u$ 0.35 kcal/mole; Table S8). However, changes in stability in this range may not impose an evolutionary constraint, as suggested by control studies of species variants. We speculate that a threshold stability is required for physiological function--above the redox potential in the blood stream such that disulfide bridges remain intact--but that a broad range of stabilities above this threshold may be allowed given the brief life time of the hormone in the circulation (half-life < 10 min).

Dimer affinity. The $Phe^{B24}\rightarrow$Tyr substitution impairs dimerization as inferred from CD studies and shown by SEC. The basis for this is not unclear. Although in molecular models the dimer-related *para*-OH groups would be nearby (S4, S5), Tyr^{B24} analogs form stable R_6 hexamers whose lifetime is similar to that of the Orn^{B29} parent analog (Table 3). Changes in dimer affinity are unlikely to impose an evolutionary constraint excluding Tyr^{B24} as species variants exist with similarly altered self-association properties (S6).

Toxic misfolding. The fibrillation lag time of Tyr^{B24}, Orn^{B29}-insulin was reduced relative to WT, which may be a side consequence of impaired dimerization (Table 4). However, the analog is less susceptible to fibrillation than a species variant (bovine insulin), which does not form amyloid *in vivo* (S7). Change in lag time is therefore unlikely to have imposed an evolutionary constraint excluding Tyr^{B24}.

NMR Features and Protein Dynamics. Subtle differences were observed between Phe^{B24} and Tyr^{B24} insulin analogs in NMR features. Despite the native-like solution structure of a monomeric Tyr^{B24} insulin analog, the upfield-shifted ^{1}H-NMR chemical shifts of the Leu^{B15} side chain are attenuated in the variant, even as respective native-like long-range inter-residue NOEs are maintained (including to the aromatic ring of Tyr^{B24}; *SI Appendix*, Table S5a). Such attenuation of secondary ^{1}H-NMR chemical shifts (*SI Appendix*, Table S3) may reflect either subtle changes in B24 ring orientation and distance relative to Leu^{B15} (*SI Appendix*, Table S5bc) or increased conformational averaging of its ring current. Also attenuated are the secondary amide chemical shifts of residues in the central B-chain α-helix (*SI Appendix*, Fig. S12). The lower subglobal ^{1}H-^{2}H PFs in this helix (*SI Appendix*, Table S6), although reflecting a distinct physical process at a longer time scale (hours), also provide evidence for enhanced conformational fluctuations in the Tyr^{B24} analog. Although such dynamic perturbations may in themselves be of limited significance, we imagine that they provide a faint echo, in the native state, of critically frustrated side-chain interactions in an on-pathway protein-folding intermediate.

Despite the native-like solution structure of a monomeric Tyr^{B24} insulin analog, the upfield-shifted ^{1}H-NMR chemical shifts of the Leu^{B15} side chain are attenuated in the variant, even as respective native-like long-range inter-residue NOEs are maintained (including to the aromatic ring of Tyr^{B24}; *SI Appendix*, Table S5a). Such attenuation of secondary ^{1}H-NMR chemical shifts (*SI Appendix*, Table S3) may reflect either subtle changes in B24 ring orientation and distance relative to Leu^{B15} (*SI Appendix*, Table S5bc) or increased conformational averaging of its ring current. Also attenuated are the secondary amide chemical shifts of residues in the central B-chain α-helix (*SI Appendix*, Fig. S12). The lower subglobal ^{1}H-^{2}H PFs in this helix (*SI Appendix*, Table S6), although reflecting a distinct physical process at a longer time scale (hours), also provide evidence for enhanced conformational fluctuations in the Tyr^{B24} analog.

3. Supplemental Methods

Materials. Human insulin was purchased from BioDel®. Bovine insulin was purchased from Sigma Aldrich® (St. Louis, MO). Reagents for peptide synthesis were as described (S8).

Preparation of Insulin Analogs. WT human insulin was obtained from commercial vials of Humulin® (Eli Lilly) and repurified by reverse-phase (rp) HPLC. Variant human insulins were prepared by semisynthesis. In brief, synthetic peptides were coupled to a tryptic fragment of insulin (*des*-octapeptide [B23-B30] insulin; DOI) in aqueous/organic solvent using trypsin as a peptide ligase. Following rp-HPLC purification, predicted molecular masses were confirmed by mass spectrometry (S8). To prepare *des*-B30 porcupine insulin, a corresponding 50-residue single-chain synthetic precursor was prepared by solid-phase peptide synthesis (SPPS) as described by DiMarchi and colleagues (S9, S10). The precursor has Lys^{B29}-Gly^{A1} linkage, which can be cleaved after folding/disulfide bond formation to the two-chain molecule (Scheme S1).

Preparation of *des*B30 Porcupine Insulin in Detail.

To enable the synthetic access, porcupine insulin was initially made as a *des*-B30 analog (a 50-residue single-chain precursor) that has Lys^{B29}-Gly^{A1} linkage which can be cleaved after folding/disulfide bond formation to the two-chain molecule.

Single-chain *des*-B30 porcupine insulin analog was synthesized using a Tribute Peptide Synthesizer with preprogrammed SPPS Fmoc protocol. The synthesis of 50-residue polypeptide was carried out on a 0.1-mmol scale using H-Asn(Trt)-HMPB-ChemMatrix resin (0.46 mmol/g loading). Side-chain protections were Asn(Trt), Arg(Pbf), Asp(OtBu), Cys(Trt), Gln(Trt), Glu(OtBu), His(Trt), Lys(Boc), Ser(tBu), Thr(tBu) and Tyr(tBu). N.N'-Diisoproylcarbodiimide(DIC)/6-chloro-1-hydroxybenzotriazole (HOBt) in 1:2 dichloromethane/DMF was used for the coupling and 20% piperidine in DMF for N$^{\alpha}$-Fmoc-deprotection. The product peptide was then cleaved from the resin and simultaneously deprotected by subjecting it to trifluoroacetic acid (TFA)/triisopropyl

silane/water/3,6-dioxa-1,8-octane-dithiol (DODT)/anisole (90:2.5:2.5:2.5:2.5 v/v) for 2.5 h at ambient temperature. The crude product was precipitated with ice-cold ether, washed and dried.

For folding, 100-mg *des*-B30 porcupine 50-mer crude polypeptide (corresponding to 18 μmol), cysteine hydrochloride (43 mg, 2 mmol/L), and glycine (270 mg, 20 mmol/L) were placed in a 600-mL glass beaker, and distilled, deionized water was added to a 0.1 mmol/L final concentration. The pH of the solution was adjusted to 10.5, and the reaction was stirred open to the air at 4 °C for 16 h. The desired folded product was purified by preparative high-performance liquid chromatography (HPLC) on a C4 column and the purity was tested by analytical rp-HPLC on a C8 Proto (4.6x250 mm) 300 Å, 5 mm, Higgins Analytical Inc. column, using 25-50% Solvent B (0.1% TFA in acetonitrile) in Solvent A (0.1% TFA in water) over 35 min at a flow rate of 1.0 ml/min (1525; Waters). Protein elution was monitored at 215 and 280 nm (2489 absorbance detector; Waters), and masses were confirmed by ESI mass spectrometry (Fig. S22).

The single-chain *des*-B30 porcupine insulin precursor was then treated with Endo Lys-C in a buffer containing 25 mM Tris and 100 mM Urea (pH 8.5) for 3 h at 12 °C. After HPLC showed complete consumption of single-chain molecule, the reaction was acidified and purified by semi-preparative rp-HPLC (Fig. S22).

Hexamer Disassembly Assays. Disassembly of phenol-stabilized Co^{2+}-substituted R_6 insulin hexamers was monitored as described (S11). Spectra (400-750 nm) were obtained to monitor tetrahedral Co^{2+} coordination through its d-d absorption band at 574 nm (S12). Co^{2+} sequestration was initiated by addition of EDTA to a concentration of 2 mM. Data were fit to a monoexponential decay equation (S13).

Receptor-Binding Assays. Analog affinities for detergent-solubilized IR-B, IR-A, or IGF-1R holoreceptor were measured by a competitive-displacement assay (S11). To obtain dissociation constants, competitive binding data were analyzed by non-linear regression (S14).

Rat Studies. Male Lewis rats (mean mass ~300 g) were rendered diabetic by streptozotocin. Effects of insulin analogs formulated in Lilly® buffer (S11) on blood-glucose concentration following intravenous injection were assessed in relation to Orn^{B29}-insulin (S11). Rats were anesthetized with isoflurane, and insulin or analog was administered via tail-vein injection.

Cell Signaling Assays. MCF-7 cell lysates were dissolved in Laemmli buffer (Bio-Rad) with 10% β-mercaptoethanol, heated at 100 °C for 8 min and then centrifuged at 9300 *g* for 1 min (Biofuge A; Baxter Scientific, Soddy Daisy, TN). Samples were loaded into 10% MiniPROTEAN TGX gels (Bio-Rad). Proteins were transferred to PVDF membrane and then blocked in 5% bovine serum albumin for 1 h. Membranes were incubated overnight at 4 °C with antibodies reactive to phosphorylated IR, total IR, phosphorylated AKT, or total AKT as described (S3). After the primary antibody incubation, membranes were washed with Tris-buffered saline with 0.1% Tween 20 and incubated in goat anti-rabbit HRP-conjugated secondary antibody diluted 1:10000 in 5% bovine serum albumin for 1–2 h at room temperature. All membranes were washed three times with 1X Tris-buffered saline (TBS; pH 7.4) containing 0.1% Tween 20 and incubated in HRP Substrate (Merck Millipore, Billerica, MA) for 45 s and developed.

Size-Exclusion Chromatography. Analogs were made 0.6 mM in 10 mM Tris-HCl (pH 7.4), 1.6% glycerol (v/v), 0.3 mM $ZnCl_2$, and 7 mM phenol (S8). Samples (20 μl) were loaded on an Enrich® SEC70 column (10 mm x 300 mm with fractionation range 3-70 kDa); the mobile phase consisted of 10 mM Tris-HCl (pH 7.4), 140 mM NaCl, and 0.02% sodium azide. Elution times were monitored by absorbance at 280 nm. Molecular masses and void volume (V_0) were inferred in reference to standard proteins (S11).

Circular Dichroism Spectroscopy. CD spectra were acquired in 10 mM potassium phosphate (pH 7.4) and 50 mM KCl (S15). Free energies of unfolding (ΔG_u) were inferred at 25 °C from two-state modeling (S15).

Fibrillation Assays. Insulin or analogs were made 60 μM in PBS with 0.02% sodium azide and 16 μM thioflavin T (ThT). Samples were plated in a Costar® plate (250μl/well) and incubated at 37 °C with continuous shaking at 400 RPM in a Biotech plate reader. ThT fluorescence at 480 nm (excitation 450 nm) was assessed at 5-min intervals. The time of initial THT fluorescence defined lag time.

NMR Spectroscopy ^{1}H-NMR and triple-resonance NMR spectra were acquired at 700 MHz at pH or pD 7.4 (direct meter reading) at 25°C as described (S3); see also Supplemental Results above. Chemical-shift assignments were accomplished by triple resonance through-bond correlation experiments (S16); distance restraints were derived from through-space correlation experiments, such as 3D ^{13}C-NOESY (S17, S18) and 4D time-shared NOESY (S18). Main chain dihedral angle restraints were generated using *TALOS+* based on ^{1}H, ^{13}C and ^{15}N chemical shifts. Solution structures were calculated using *XPLOR-NIH* (S19) as described (S3). Initial structure was calculated by employing only unambiguous NOEs; calculation was iteratively performed on further NOE analysis. 20 best-fit structures were chosen from 100 calculated structures. Structure ensemble was visualized using *insightII* and *molmol* software (S20). ^{1}H-^{2}H exchange was performed in 100% D_2O with 10 mM deuterio-acetic acid (Cambridge Isotopes, Tewksbery, MA). Lyophilized samples of Tyr^{B24}-KP or parent KP analogs (lyophilized from H_2O with 10 mM deuterio-acetic acid, pH 3.0) were placed on ice followed by addition of chilled D_2O with 10 mM deuterio-acetic acid. A series of 1D proton and 2D TOCSY spectra were collected to measure exchange of amide protons as a function of peak intensity; 1D ^{1}H spectra were collected at 2-minute intervals to assess very fast exchange of amide protons, while 2D TOCSY spectra were acquired at 2-hr intervals to assess amide protons with slower exchange rates. Protection factors were calculated from the ratio of measured exchange rate and intrinsic exchange rate (S3).

Cell-Based Assays of Proinsulin Folding and Secretion. HEK-293T cells were plated into 12-well plates 1 d before transfection. 1-2 μg plasmid DNA were transfected using Lipofectamine (Invitrogen®, Carlsbad, CA). At 48-h post-transfection, the cells were labeled with ^{35}S-Met/Cys for 30 min followed by 0 or 90 min chase as indicated. At the end of the chase, the media were collected and the cells were lysed with cell lysis buffer containing a proteinase inhibitor mixture. Samples were pre-cleared with Zysorbin and immunoprecipitated with anti-insulin antibodies (S21) and protein-A agarose overnight at 4 °C. Newly synthesized proinsulin immunoprecipitated with anti-insulin was analyzed using tris-tricine-urea-SDS-PAGE under non-reducing conditions *versus* reducing conditions (S21). The secretion of proinsulin was analyzed using 4-12% NuPAGE (Invitrogen) under reducing condition.

Cell-Based Assessment of *trans*-Dominance HEK293T cells were plated as above and co-transfected with plasmid encoding untagged human proinsulin-WT and plasmids encoding Myc-tagged human proinsulin-WT or variants as previously described (S22). Cells were pulse-labeled as above for 30 min followed by 0- or 90 min chase. Cell lysates and chase media were immunoprecipitated (IP) with anti-insulin followed by analysis using 4-12% NuPAGE (Invitrogen) under reducing condition (S23).

BiP-Driven Luciferase Assay. Min6 cells were plated into 24-well plates 1 d before transfection. Using Lipofectamine 2000 (Invitrogen), cells were co-transfected with BiP-firefly-luciferase reporter plasmid, CMV-renilla-luciferase plasmid (Promega), and human WT or mutant proinsulin at a DNA ratio of 1:2:5, respectively. At 48-h post-transfection, cell extracts were prepared for the dual-luciferase reporter assay (Promega) with BiP-luciferase normalized to Renilla luciferase activity (S21).

Immunofluorescence Studies of Proinsulin Biosynthesis. Rat pancreatic β-cell line INS1 were transiently transfected with human proinsulin WT or variants using Lipofectamine 2000 (Invitrogen) and at 48 h fixed with 3.7% formalin in PBS (pH 7.4) for 20 min, permeabilized with TBS containing 0.4% Triton X-100, blocked with TBS containing 3% BSA and 0.2% Triton-X100, and then stained overnight (4 °C) with primary mAb anti-human proinsulin (not cross-reacting with rodent proinsulin and not cross-reacting with human insulin) and rabbit anti-calnexin diluted in TBS containing 3% BSA and 0.2% Tween. Thereafter, sections were rinsed and incubated with secondary antibodies

conjugated to Alexa Fluor 488 or 568 (Invitrogen). Slides were mounted with Prolong Gold with DAPI (Invitrogen) and imaged by epifluorescence in an Olympus FV500 confocal microscope with a X60 objective.

Western blot-based insulin signaling assays. Signaling activities of insulin analogs were tested in human breast-adenocarcinoma cell line MCF-7 (which expresses IR-A, IR-B and IGF-1R (S24). The cells (American Type Culture Collection, Manassas, VA) were cultured in DMEM medium supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin and sodium 1 mM pyruvate. A 24-hr serum-starvation protocol using an FBS-free culture medium was applied at 70–75% cell confluence (~6 x 10^6 per 10-cm dish). A serum-free medium containing a given insulin analog at one of three ligand concentrations (5, 10 or 20 nM) was added to each well; control wells received medium with no added insulin. Media were removed after 15 min, followed by cell lysis using RIPA buffer with protease- and phosphatase inhibitors (Cell Signaling Technology, Danvers, MA). Overall protein concentrations in the cell lysates were determined with a Pierce BCA assay kit (Thermo Fisher Scientific).

WB protocols were adapted from Glidden et al. (2018) (S3). Cell lysates were dissolved in Laemmli buffer (Bio-Rad) with 10% β-mercaptoethanol, heated at 100 °C for either 5 min (for p-Akt/Akt blotting) or 8 min (for p-IR/IR). Samples were loaded into 10% Mini-PROTEAN TGX precast electrophoresis gels (Bio-Rad). Proteins were transferred and then blocked in 5% bovine serum albumin (BSA) or 10% milk for p-Akt assays for 1 hr. All antibodies were purchased from Cell Signaling Technology unless otherwise stated. Membranes were incubated overnight at 4 °C with IR subunit-β (4B8) rabbit monoclonal antibody (mAb) or an equal mixture of anti-pIR-β ($Tyr^{1150/1151}$) rabbit mAb; anti-pIR (Tyr^{1158}) polyclonal antibody (Thermo Fisher Scientific, Waltham, MA); anti-pIR (Tyr^{1334}) polyclonal antibody (Thermo Fisher Scientific); anti-pIR-β (Tyr^{1345}) rabbit mAb; and anti-pIR (Tyr^{972}) antibody (Abcam, Cambridge, UK). Dilutions for these antibodies were made 1:5000 in 5% BSA. Antibodies for AKT blotting were p-AKT1/2/3 antibody (Ser^{473}) (1:1000) and AKT1/2/3 antibody (H-136) (1:2000). After primary antibody incubations, membranes were washed and incubated in goat anti-rabbit HRP-conjugated secondary antibody, which was diluted 1:10,000 in 5% BSA (or 10% milk for p-AKT data) for 2 hrs at room temperature. All membranes were washed and incubated in Western HRP Substrate (Cell Signaling Technology) for 45 sec and developed.

Plate-based fluorescence in-cell immunoblotting assay. The assay was designed to assess hormone-induced IR autophosphorylation via in-cell illumination readouts. Human liver-cancer derived HepG2 cells were seeded (~8000 cells per well) into a 96-well assay black plate with clear bottom, and grown in tissue culture (Fisher). After starving in 100 µl of plain Hank's Balanced Salt Solution (HBSS) for 2 hrs at 37 °C, 100 µl of a given insulin analog at a specific dose (in the range 0.5-100 nM) were applied to each well and incubated for 20 min at 37 °C. After removing the insulin solution, 150 µl of 3.7% formaldehyde (Fisher) was added to each well, and plates were incubated for 20 min at 37 °C; 200 µl of 0.1 % Triton-X-100 detergent solution (Sigma) was then added to mediate cell permeabilization under same incubation conditions. 100-µl Odyssey Blocking Buffer (LI-COR) was added to fixed cells after permeabilization. The blocking procedure was implemented for 1 hr at room temperature on gentle agitation in an orbital shaker.

After blocking, the fixed cells were incubated with the primary antibody (10 µl anti-pTyr 4G10 into 20 mL Blocking Buffer) overnight at 4°C. After a wash, the secondary antibody (anti-mouse-IgG-800-CW antibody (Sigma) into 25 mL Blocking Buffer) was added. This enabled detection at 800 nm (red emission frequency) to probe the extent of pTyr modification. DRAQ5 (Fisher) was also applied to enable 700 nm emission (green) as a control to estimate cell number. These respective fluorescence signals were determined at 700 and 800 nm using an LI-COR Infrared Imaging system (Odyssey) with offset set to 4 mm and intensity set to "auto."

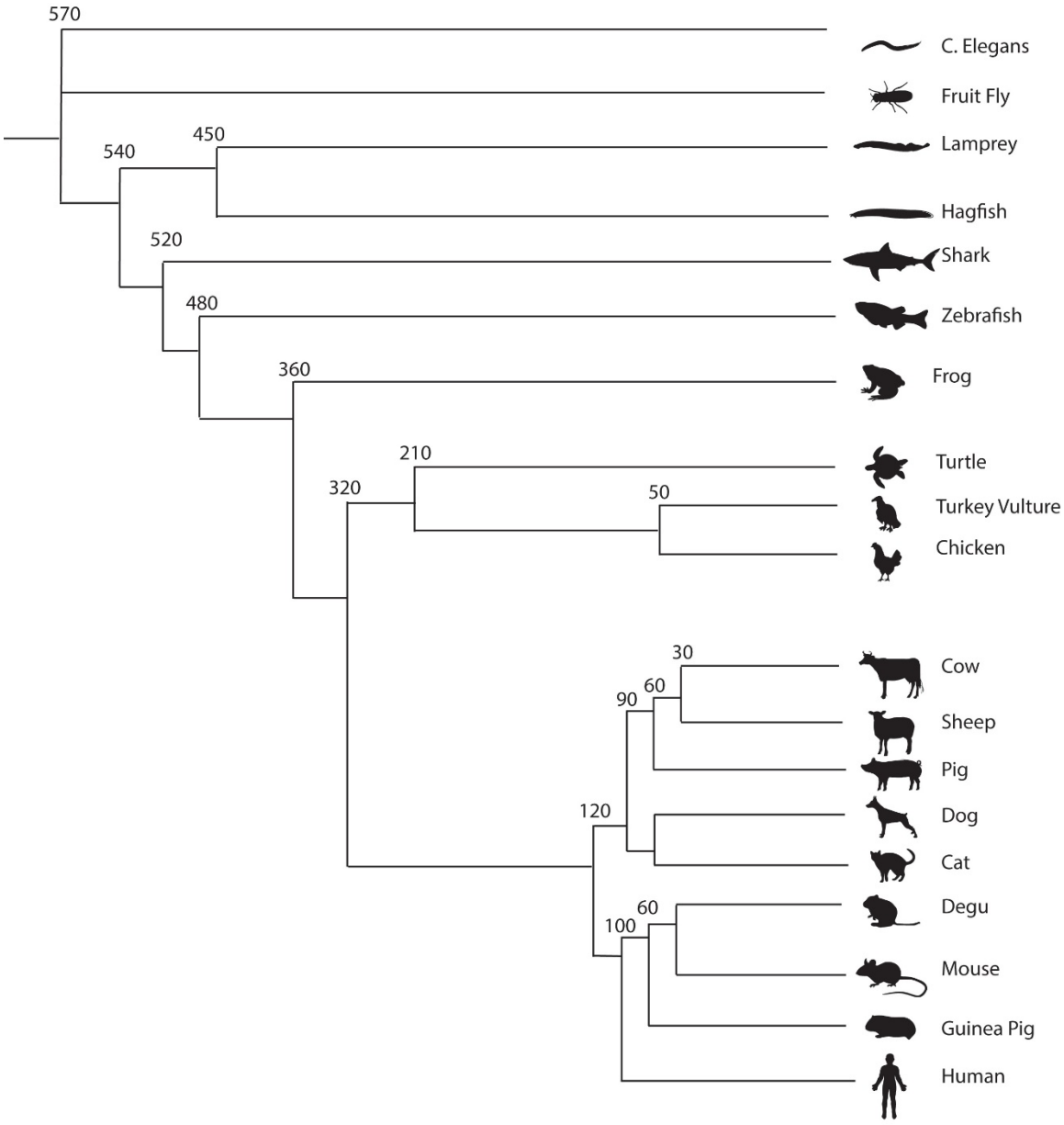


Fig. S1. Phylogenetic tree of vertebrate evolution. Various taxa within the vertebrate subphylum are displayed; *Caenorhabditis elegans* and *Drosophila melanogaster* are shown as outgroups. The insulin gene arose during the metazoan radiation ("Cambrian explosion") 540-575 million years ago. The B24-B26 aromatic triplet is a feature only of vertebrate insulins and insulin-like growth factors.

A alignment of vertebrate insulin sequences

	20	.	.	.	.	25	.	.	.	.	30
Human	G	E	R	G	F	F	Y	T	P	K	T
Turkey Vulture	G	E	R	G	F	F	Y	S	P	K	A
Zebrafish	G	P	T	G	F	F	Y	N	P	K	R
Guinea Pig	Q	D	D	G	F	F	Y	I	P	K	D
Sea Turtle	G	E	R	G	F	F	Y	S	P	K	A
Chicken	G	E	R	G	F	F	Y	S	P	K	A
Cattle	G	E	R	G	F	F	Y	T	P	K	A
Bullfrog	G	D	R	G	F	F	Y	S	P	R	S
Hagfish	G	V	R	G	F	F	Y	D	P	T	K
Sheep	G	E	R	G	F	F	Y	T	P	K	A
Dog	G	E	R	G	F	F	Y	T	P	K	A
Rat	G	E	R	G	F	F	Y	T	P	M	S
Pig	G	E	R	G	F	F	Y	T	P	K	A
Cat	G	E	R	G	F	F	Y	T	P	K	A
Shark	G	E	R	G	F	F	Y	S	P	K	Q
Rattlesnake	G	E	R	G	F	Y	Y	S	-	-	-
Lamprey	G	D	R	G	F	F	Y	T	P	-	-
Octodon degus	G	R	S	G	F	Y	R	-	P	H	D
Microcavia australis	K	D	K	G	F	F	S	R	P	K	-

B alignment of vertebrate IGF sequences

	19	.	.	.	.	24	.	.	.	.	29
Human IGF-1	G	D	R	G	F	Y	F	N	P	T	G
Rat IGF-1	G	P	R	G	F	Y	F	N	P	T	G
Pig IGF-1	G	D	R	G	F	Y	F	N	P	T	G
Chicken IGF-1	G	D	R	G	F	Y	F	S	P	T	G
Dog IGF-1	G	D	R	G	F	Y	F	N	P	T	G
Cattle IGF-1	G	D	R	G	F	Y	F	N	P	T	G
Sheep IGF-1	G	D	R	G	F	Y	F	N	P	T	G
Clawed Frog IGF-1	G	D	R	G	F	Y	F	S	P	T	G
Zebrafish IGF-1	G	D	R	G	F	Y	F	S	P	T	G
Guinea Pig IGF-1	G	D	R	G	F	Y	F	N	P	T	G
Cat IGF-1	G	D	R	G	F	Y	F	N	P	T	G
Snapping Turtle IGF-1	G	D	R	G	F	Y	F	S	P	T	G
Turkey Vulture	G	D	R	G	F	Y	F	S	P	T	G
Shark IGF-1	G	N	R	G	F	Y	F	N	P	T	G
Hagfish IGF-1	G	D	R	G	F	F	F	V	P	P	R
	22	.	.	.	.	27	.	.	.	.	32
Human IGF-2	G	D	R	G	F	Y	F	S	P	A	S
Lamprey IGF-2	G	D	R	G	F	Y	F	V	H	P	S
Pig IGF-2	G	D	R	G	F	Y	F	S	P	A	S
Sheep IGF-2	G	D	R	G	F	Y	F	S	P	S	S
Cat IGF-2	G	D	R	G	F	Y	F	S	P	A	S
Cattle IGF-2	G	D	R	G	F	Y	F	S	P	S	S
Rat IGF-2	S	D	R	G	F	Y	F	S	P	S	S
Zebrafish IGF-2	G	E	D	G	F	Y	I	S	P	N	R
Clawed Frog IGF-2	G	D	R	G	F	Y	F	S	N	N	G
Shark IGF-2	G	E	R	G	F	Y	F	S	A	T	G
Dog IGF-2	G	D	R	G	F	Y	F	S	P	A	S
Snapping Turtle IGF-2	G	D	R	G	F	Y	F	S	P	V	G
King Cobra IGF-2	G	E	R	G	F	Y	F	S	P	V	G
Chicken IGF-2	G	D	R	G	F	Y	F	S	P	V	G
Guinea Pig IGF-2	G	D	R	G	F	Y	F	S	P	A	S

Fig. S2. Alignment of vertebrate insulin and IGF sequences. (A) Residues B20-B30 in the B chain of vertebrate insulins. (B) Homologous residues 19-29 in IGF sequences. Residue B24 in insulin and its homologue among IGFs are highlighted in *red*.

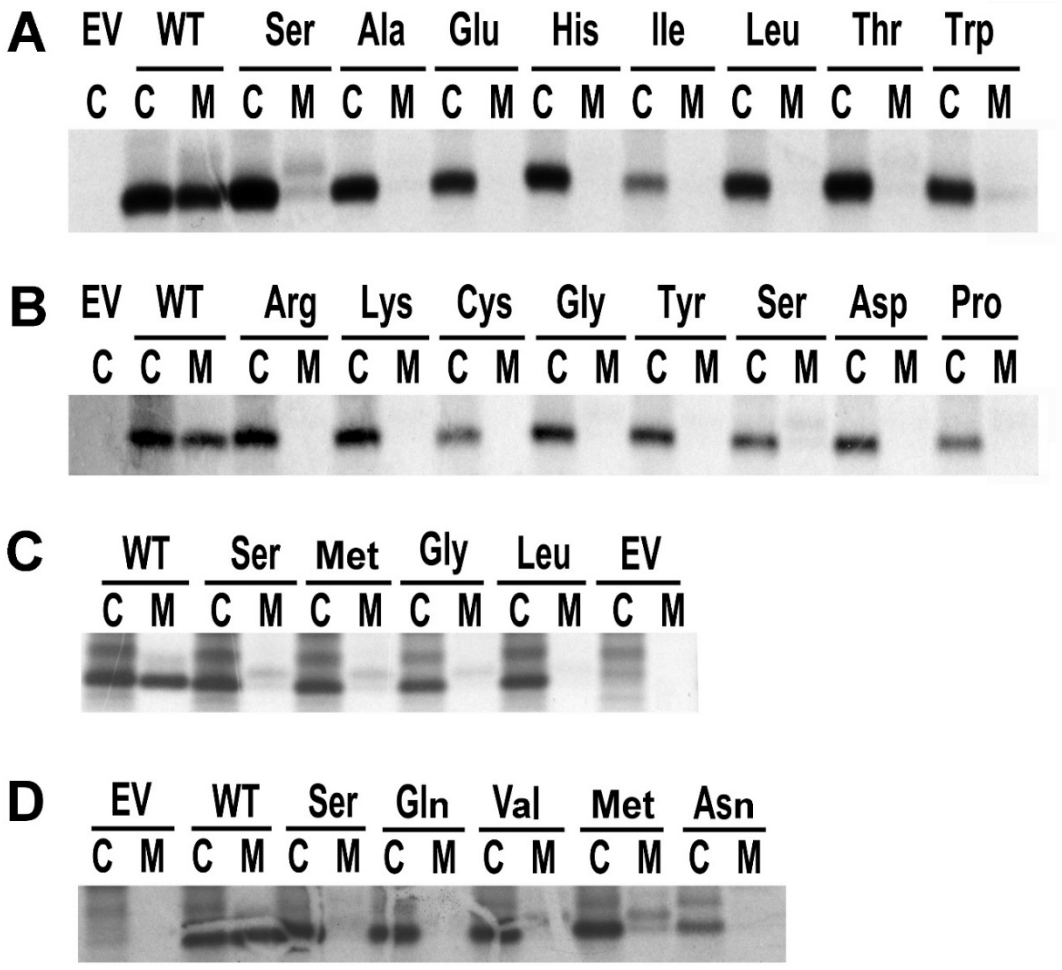


Fig. S3. Complete SDS-PAGE gels from Main Text Fig. 3. At position B24, Phe uniquely enables secretability of proinsulin: systematic screen of B24-proinsulin variants. The empty vector, WT human proinsulin and clinical variant Ser^{B24}-proinsulin were included in all four gels (A, B, C, and D). Panels C and D highlight effects of non-polar substitutions of varying size at the B24 position.

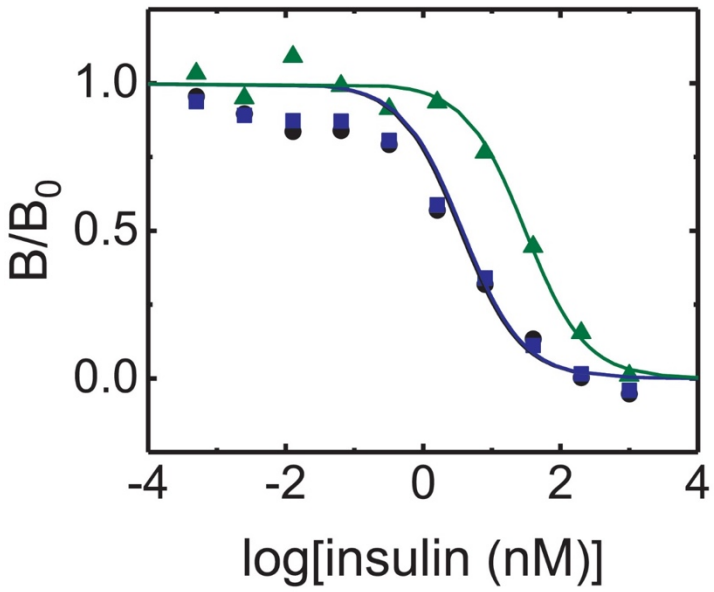


Fig. S4. IGF-1R receptor-binding assays. The affinity of Tyr^{B24}, Orn^{B29}-insulin (*green triangles*) is reduced relative to Orn^{B29}-insulin (*blue squares*) or WT human insulin (*black circles*). Fitted curves are shown as lines color-coded to match symbols. See Table 1 for inferred dissociation constants.

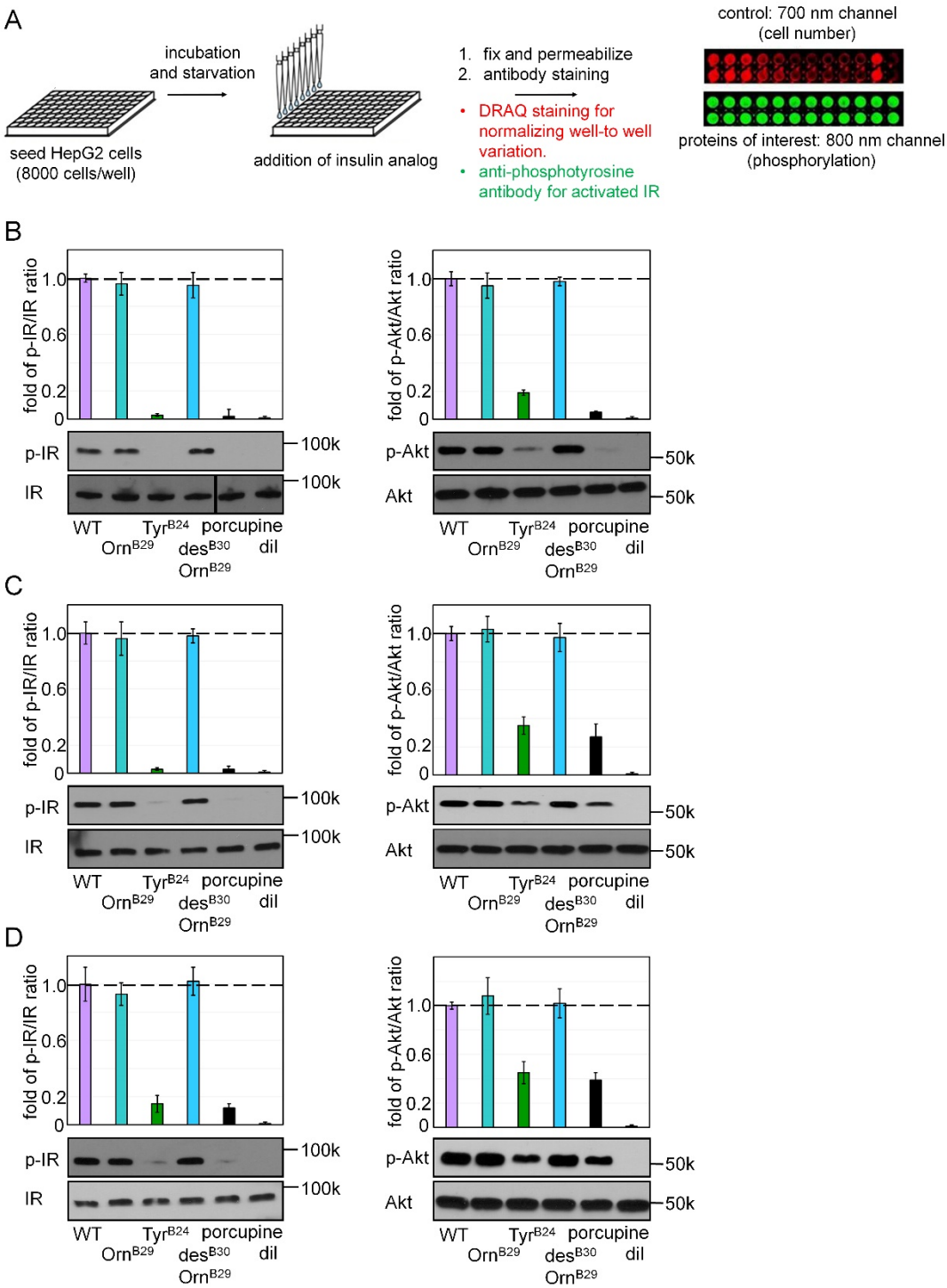


Fig. S5. (A) IR activation probed by pTyr status in HepG2 cells via in-cell immunoblotting. Flow chart illustrates protocol to assess hormone-induced IR phosphorylation at a range of insulin doses (0.05-900 nM): (immunoconjugate-based 800 nm emission; red) pIR level; and (700 nm emission; green) DRAQ staining as cell-number control. (B-D) Western blot-based assays of insulin signaling at three hormone doses: 5 nM (B), 10 nM (C), and 20 nM ligand (D). Insulin-stimulated IR phosphorylation (p-IR/IR ratio) and downstream activation of AKT kinase (p-AKT/AKT ratio) were

determined by WB in breast-cancer-derived MCF-7 cells by indicated ligand relative to diluent control ("dil"). Color code from right to left: (*purple*) WT human insulin; (*turquoise*) Orn^{B29}-insulin; (*green*) Tyr^{B24}, Orn^{B29}-insulin; (*aquamarine*) *des*-B30 insulin; (*black*) des-B30 porcupine insulin; and (*negligible*) diluent control. Errors bars in histograms indicate standard error of mean (SEM). Representative gel images corresponding to p-IR and IR (left) and p-AKT and AKT (right) are displayed underneath the corresponding columns in each histogram. Assays were performed in triplicate. MCF-7 cells express IR-A, IR-B, IGF-1R and presumed hybrid receptors.

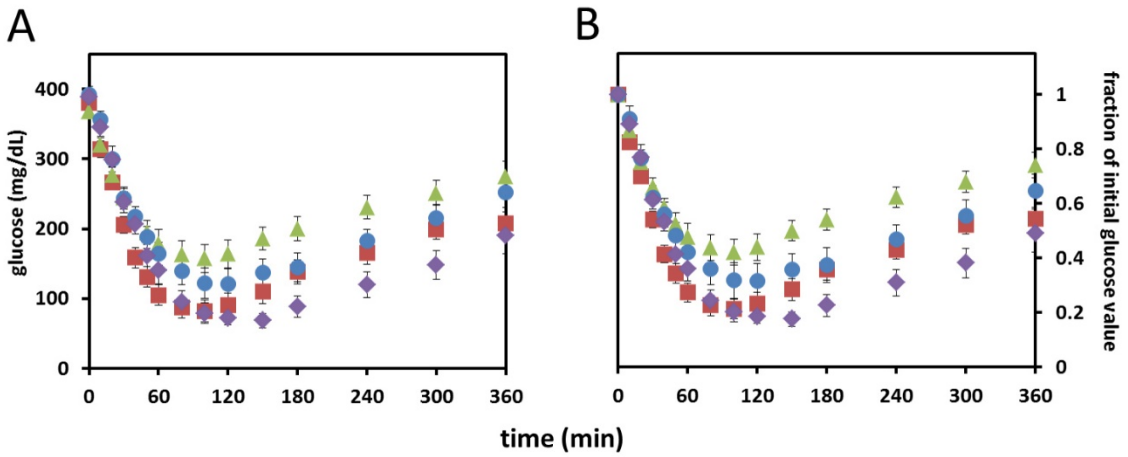


Fig. S6. Testing of biological potency in rats. Biological activity of Tyr^{B24}-insulin *lispro* was evaluated on intravenous injections in male Lewis rats made diabetic by streptozotocin (STZ; mean glycemia 400 mg/dl; mean body mass 300 g). (A) Dose-response studies of Tyr^{B24}-insulin *lispro* relative to parent analog at a dose 10 µg/300 g (*red squares*; N=11). Doses of the Tyr^{B24} analog: 10 µg (*green triangles*; N=11), 20 µg (*blue circles*; N=5) and 30 µg (*violet diamonds*; N=5). (B) Data are normalized relative to initial level of glycemia in each animal (defined as 1.0). Standard errors of mean are given by *black vertical error bars.*

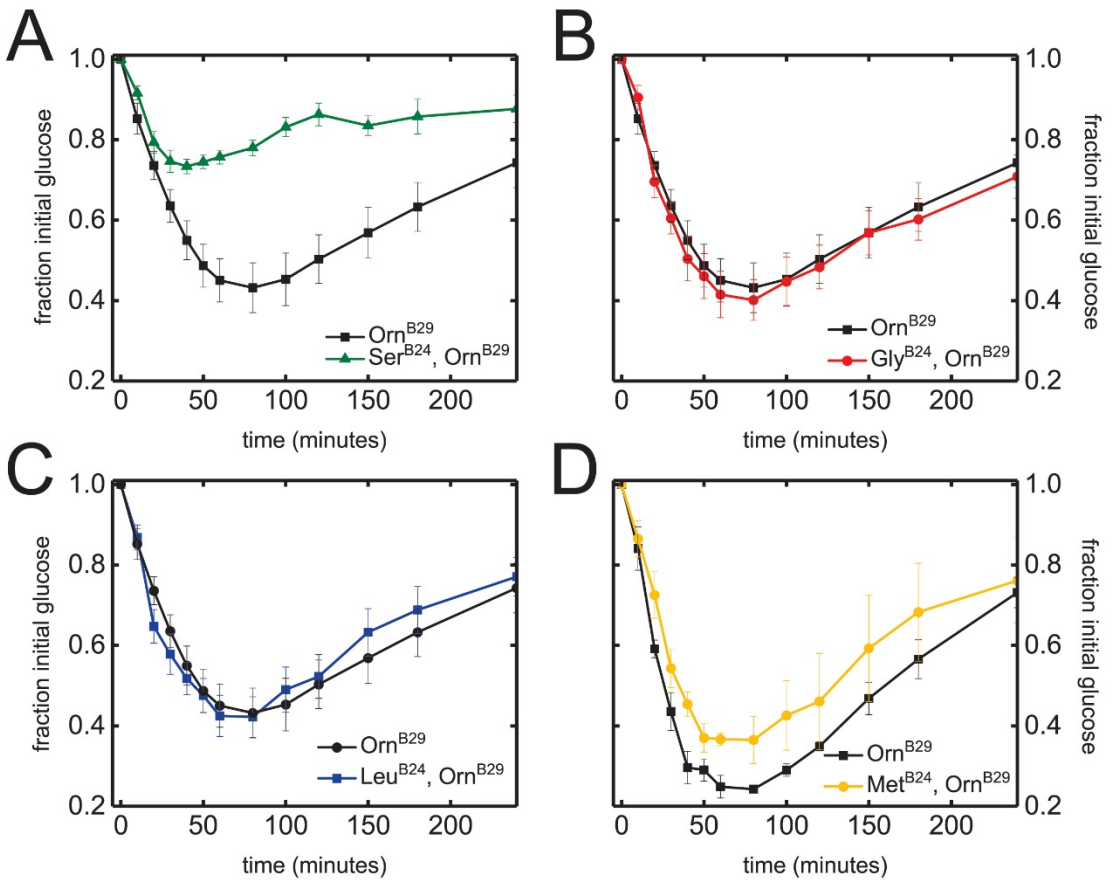


Fig. S7. Biological potencies of four B24 variants in context of Orn^{B29}-insulin. Activities were evaluated in male STZ rats (mean initial glycemia *ca*. 400 mg/dl and mean body mass *ca*. 300 grams) on intravenous injections of 10-μg doses. Data were normalized relative to initial level of glycemia in each animal (defined as 1.0). Panels A-C (respectively testing Ser^{B24}, Gly^{B24}, and Leu^{B24} analogs; color codes as inset) were taken from Rege, N., *et al*. *J. Biol. Chem.* (2020) (S21) reference 35). Panel D, representing Met^{B24}, Orn^{B29}-insulin *versus* parent Orn^{B29}, is reproduced from Pandyarajan, V., *et al. J. Biol. Chem.* (S15).

Foldability and Function as Separable Evolutionary Constraints. "Non-foldable" insulin variants containing B24 substitutions can nonetheless exhibit substantial *in vivo* activity. This study provides evidence that the strict exclusion of Met, Leu and Gly at this position in the vertebrate insulin-IGF family reflects impaired biosynthesis of the variant proinsulin, leading to β-cell dysfunction. Interference in *trans* of WT proinsulin biosynthesis would cause impaired glucose tolerance and thereby incur purifying selection. Ser^{B24}, the least active in this set of analogs, is associated with onset of diabetes mellitus in the third decade (proband) but with incomplete penetrance. This mild phenotype (relative to permanent neonatal-onset DM as triggered by Cys^{B24}) may be ascribed to Ser^{B24} as a "ER stress allele," presumably representative of a polygenic set of ER stress alleles that contribute in aggregate to enhanced risk for non-syndromic Type 2 DM.

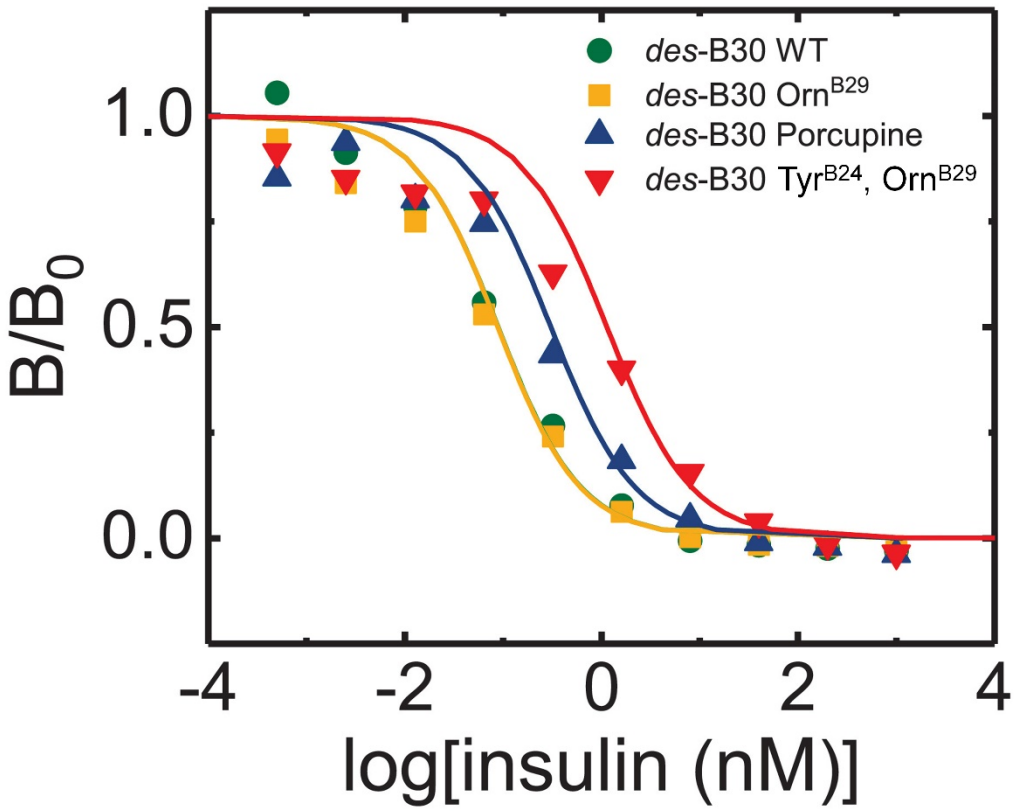


Fig. S8. *In vitro* studies of a species variant: porcupine insulin. Affinities of *des*-B30 insulin analogs for IR (isoform B). Data points are represented by colored symbols (color and symbol code are given in the inset) whereas corresponding lines indicate fitted curves; affinities are given in Table S2. Porcupine insulin was chosen based on its partial sequence divergence relative to other placental mammals. Analogs were made *des*-B30 for synthetic convenience (the C-terminal residue of B chain does not contribute to the hormone's activity).

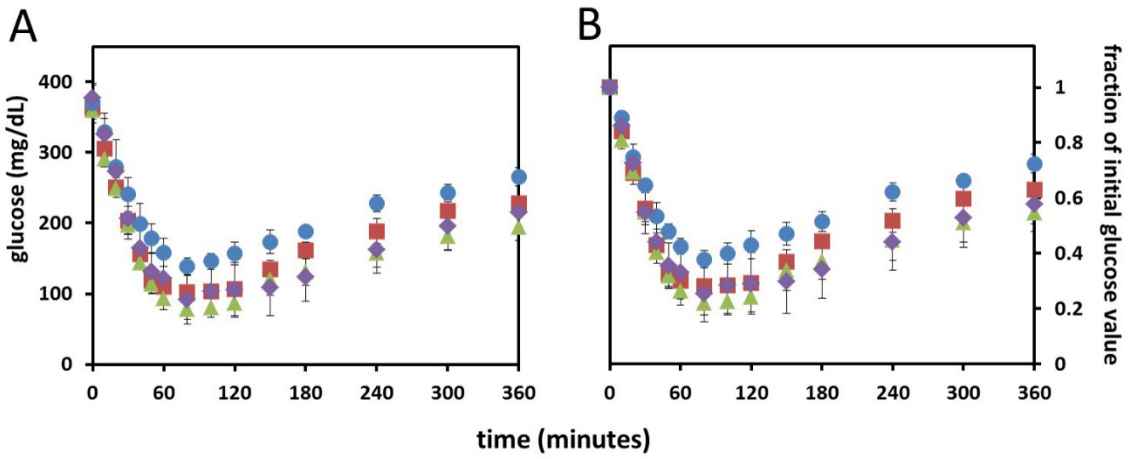


Fig. S9. *In vivo* studies of a species variant: porcupine insulin. (A) Biological activity of *des*-B30 porcupine insulin was tested following intravenous injection in male STZ rats in relation to WT human insulin (*red squares*; N=5) and *des*-B30 human insulin (*green triangles*; N=5) at control doses of 10 µg per 300 g rat. Doses of *des*-B30 porcupine insulin: 10 µg (*blue circles*; N=4) and 25 µg (*violet diamonds*; N=5). (B) Data are normalized relative to initial level of glycemia in each animal (defined as 1.0). Standard errors of mean are given by *black vertical error bars.* The analogs were made *des*-B30 for synthetic convenience (the C-terminal residue of B chain does not contribute to the hormone's activity).

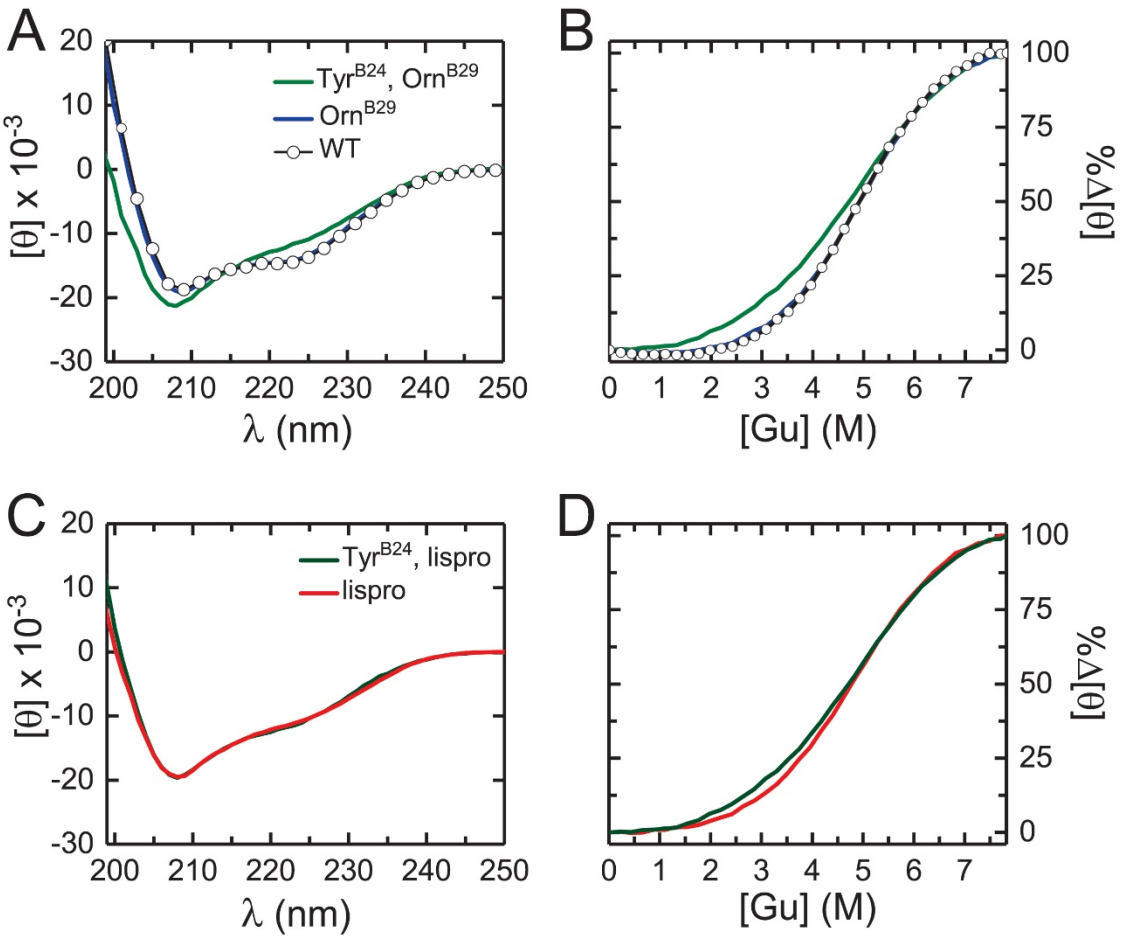


Fig. S10. CD and denaturation studies of Tyr^{B24} insulin analogs. (A, B) Far-UV spectra of WT-, Orn^{B29}-, and Tyr^{B24}, Orn^{B29}-insulin analogs (A) and corresponding CD-detected unfolding transitions (B). (C, D) Spectra of monomeric analogs insulin *lispro* and Tyr^{24}-insulin *lispro* (*C*) and corresponding unfolding transitions (D). Symbols and color codes are defined in panels *A* and *C*. Studies were performed at 25°C. Guanidine-induced protein unfolding was monitored by ellipticity at 222 nm; inferred thermodynamic parameters are given in Table 2.

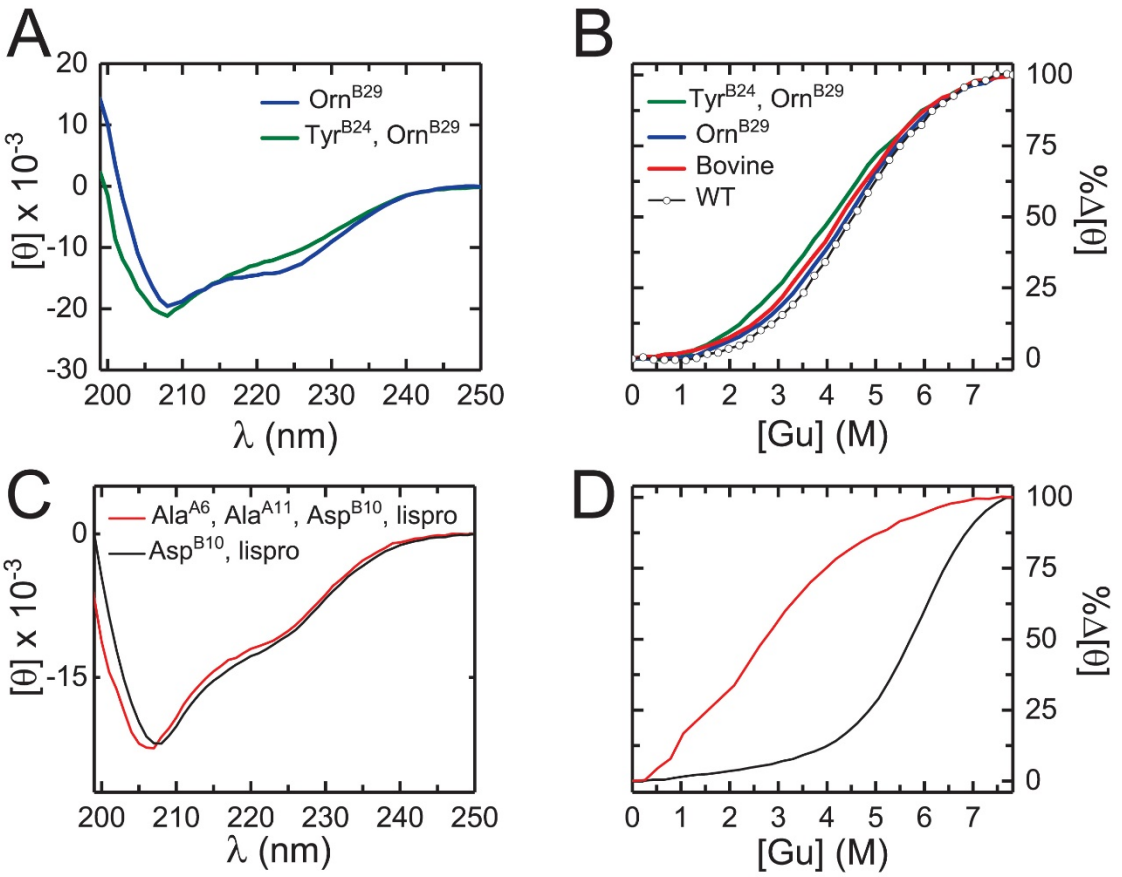


Fig. S11. CD and denaturation studies of insulin analogs. (A, C) Far-UV CD spectra of insulin analogs in relation to controls: Tyr^{B24}, Orn^{B29}-insulin (*green*) and native Orn^{B29}-insulin (*blue*) at 37°C; (C) Asp^{B10}-insulin *lispro* ("DKP") and Ala^{A6}, Ala^{A11}, Asp^{B10}-insulin *lispro* at 4°C. (B, D) Guanidine denaturation assays monitored by CD ellipticity at 222 nm: samples correspond to analogs in panels *A* and *C*, respectively. Thermodynamic parameters inferred from a two-state unfolding model are given in Table S3a-d. *Background*. Asp^{B10} stabilizes the insulin monomer for NMR studies; substitution of His^{B10} blocks trimer association and Zn^{2+} coordination as described by Shoelson, S.E., et al. (S25). Pairwise substitution of cystine A6-A11 by Ala destabilizes insulin as a molten partial fold as described by Hua, Q.-X., et al. (S1). In mammalian cells this molten globule is normally expressed, folded, trafficked and secreted—indicating that the ER and Golgi apparatus do not "measure" ΔG_u as a check-point.

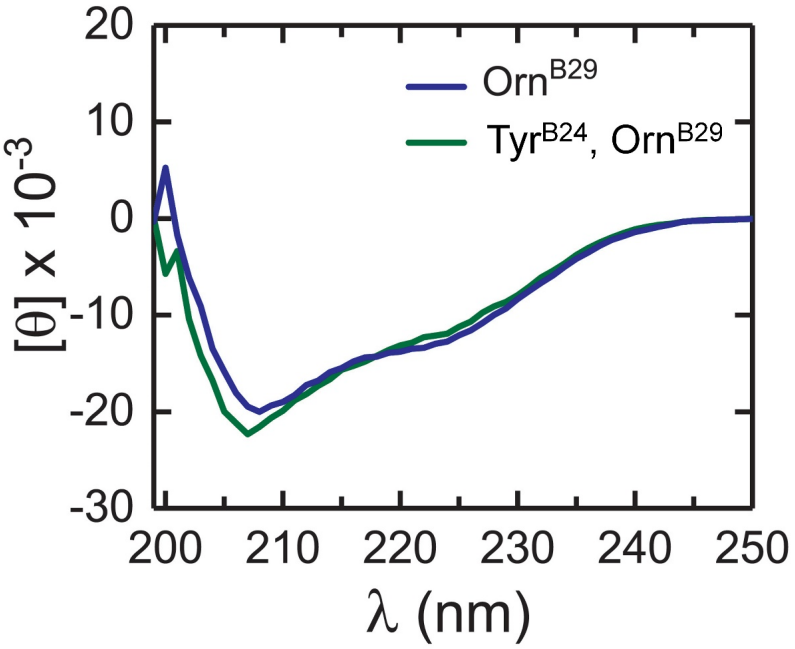


Fig. S12. CD spectra of dilute insulin analogs. On tenfold dilution relative to the spectra shown in Figures 8 (main text) and S10 (Supplement), Orn^{B29}-insulin (*blue*) and Tyr^{B24}, Orn^{B29}-insulin (*green*) exhibit essentially identical far-UV spectra. The proteins were made 5 μM in 50 mM KCl and 10 mM potassium phosphate buffer at 25 °C.

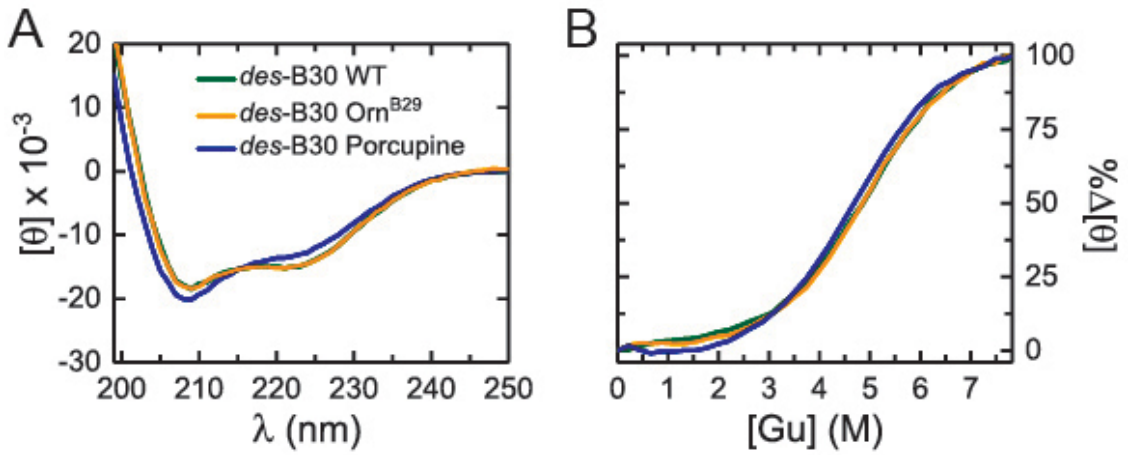


Fig. S13. CD studies of an insulin species variant: *des*-B30 porcupine insulin *versus* controls. (A) Spectra of *des*-B30 Orn^{B29}-insulin (*yellow*), *des*-B30 porcupine insulin (*blue*), and *des*-B30 wild-type (*green*). (B) Guanidine denaturation studies of analogs in panel A as monitored by ellipticity at 222 nm; color code as in panel A. Thermodynamic stabilities and parameters are given in Table S3b. Porcupine sequence: A chain (GIV**D**QCCT**GV**CSLYQL**Q**NYCN) and B chain (FVNQHLCGSHLVEALYLVCG**ND**GFFY**R**PK**A**; B30 underlined). Invariant Phe^{B24} is highlighted in red. Differences relative to human insulin are in bold.

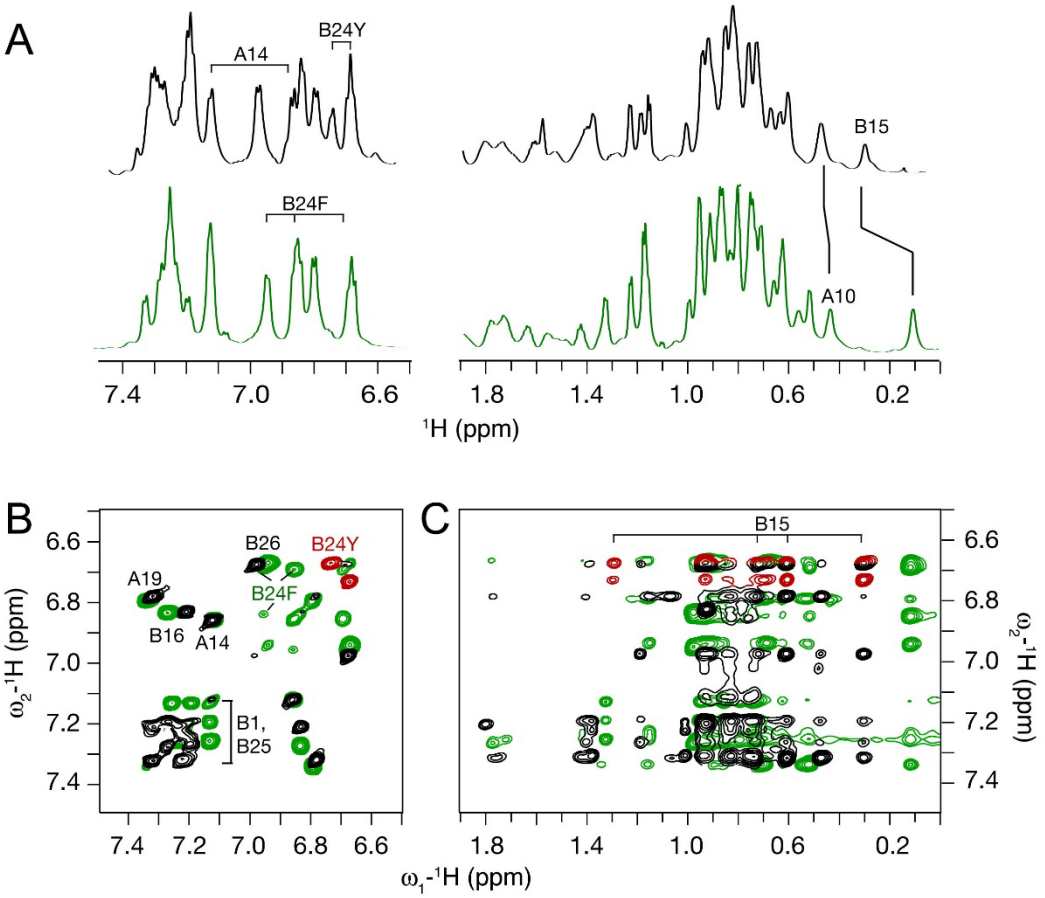


Fig. S14. NMR studies of an Tyr^{B24}-insulin as an engineered monomer. (A) 1D ^{1}H-NMR spectra of Tyr^{B24}-DKP insulin (*top, black*) and parent analog DKP-insulin (*bottom, green*). The Tyr^{B24} analog exhibits a reduction in chemical-shift dispersion as observed in the upfield-shifted methyl region (*far right*). (B) Overlay of aromatic resonances of 2D TOCSY spectra of Tyr^{B24}-DKP-insulin in *black* (Tyr^{B24} is highlighted in *red*) and parent DKP-insulin in *green*. (C) Overlay of long-range NOEs from aromatic to aliphatic protons in 2D NOESY spectra of Tyr^{B24}-DKP-insulin (*black*) and DKP-insulin (*green*). Spectra in A-C were acquired at pD 7.6 (direct meter reading) at 32 °C. TOCSY and NOESY spectra in B and C were acquired with respective mixing times of 55 and 200 ms. The monomeric "DKP" template contains substitutions Asp^{B10} (D), Lys^{B28} (K) and Pro^{B29} (P).

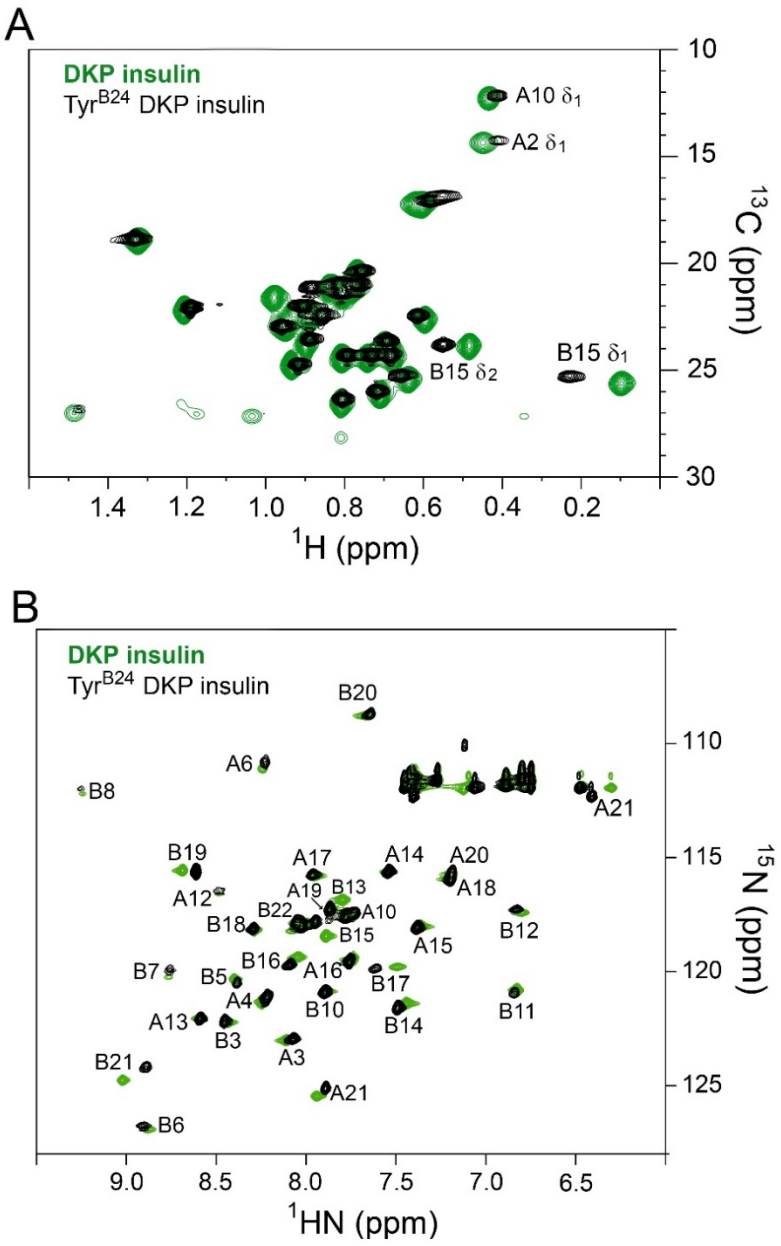


Fig. S15. ^{1}H-^{13}C HSQC and ^{1}H-^{15}N spectra of insulin analogs. (A) Overlay of ^{1}H-^{13}C-HSQC spectra of B24Y-DKP-insulin (*black*) and DKP-insulin (*green*). (B) Overlay of ^{1}H-^{15}N-HSQC spectra of Tyr^{B24}-DKP insulin (*black*) and DKP-insulin (*green*). Spectra were collected using samples uniformly enriched in ^{13}C and ^{15}N in the α-helical domain (residues A1-A21 and B1-B22) via recombinant expression of a single-chain precursor in *Pichia pastoris*. ^{15}N- and ^{13}C-HSQC spectra were acquired at 25 °C and pH 7.4 in 90% H_2O and 10% D_2O. Selected resonance assignments are as indicated. The monomeric "DKP" template contains substitutions Asp^{B10} (D), Lys^{B28} (K) and Pro^{B29} (P).

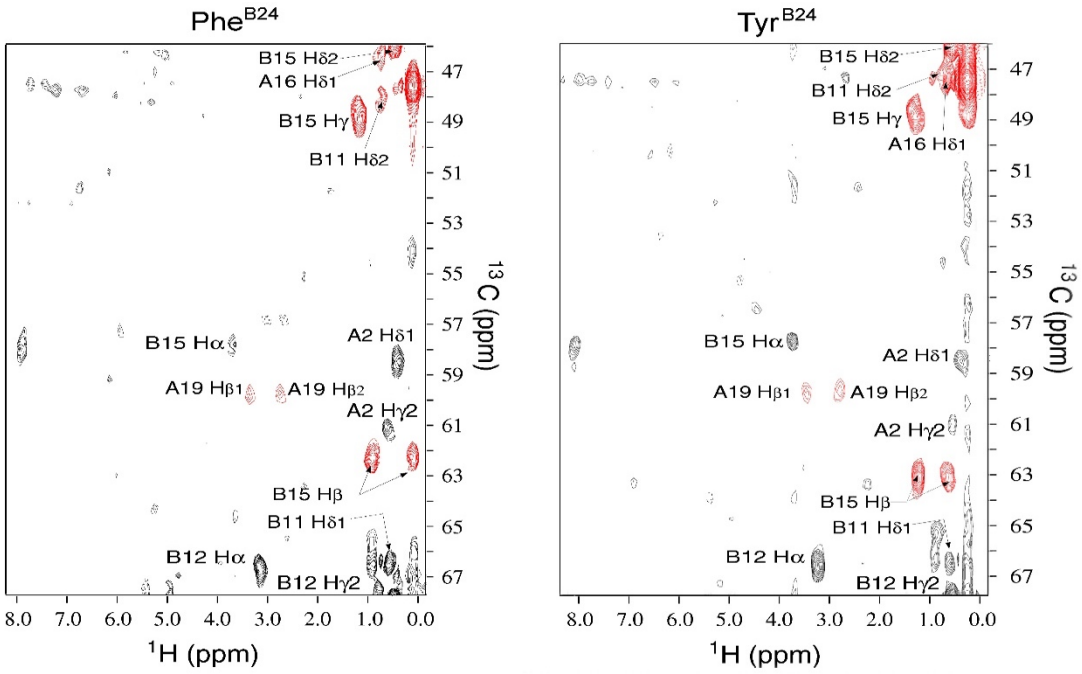


Fig. S16. Heteronuclear multidimensional NMR studies. 4D ^{13}C-^{13}C edited NOESY spectra of Tyr^{B24}-DKP-insulin (*left*) and parent DKP-insulin (*right*) showing respective contacts involving the Leu^{B15} δ_1 methyl group. Corresponding long-range NOEs similar core packing in the two proteins, including analogous Tyr^{B24}/Leu^{B15} and Phe^{B24}/Leu^{B15} contacts. The monomeric "DKP" template contains substitutions Asp^{B10} (D), Lys^{B28} (K) and Pro^{B29} (P).

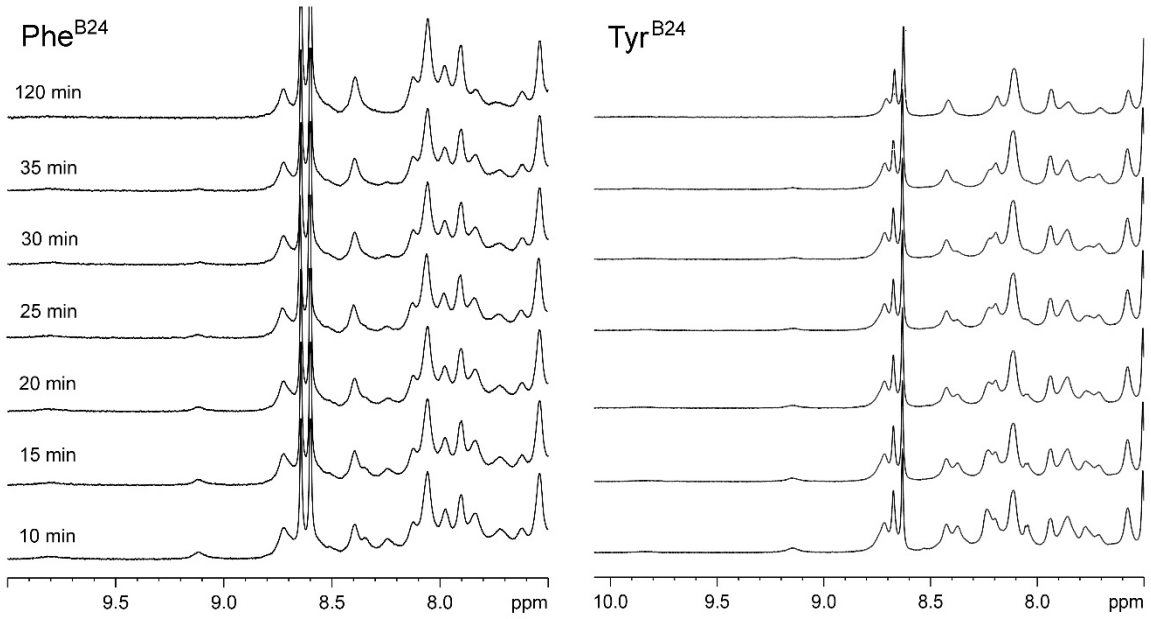


Fig. S17. 1D Amide-proton ^{1}H-^{2}H exchange kinetics in D_2O. Successive 1D ^{1}H-NMR spectra of Tyr^{B24}-KP-insulin and parent KP-insulin in 10 mM deuterio-acetic acid (pD 3.0) at 25 °C at the time points indicated. KP-insulin (substitutions Lys^{B28} (K) and Pro^{B29} (P)) is also designated insulin *lispro*, the active component of Humalog® (Lilly), a rapid-acting insulin formulation in current clinical use. Under acidic conditions the protonated His^{B10} side chain is associated with sharper ^{1}H-NMR spectra than the protonated Asp^{B10} as would be present in DKP-insulin at this pD or pH due to partial self-associated of DKP-insulin analogs under these conditions.

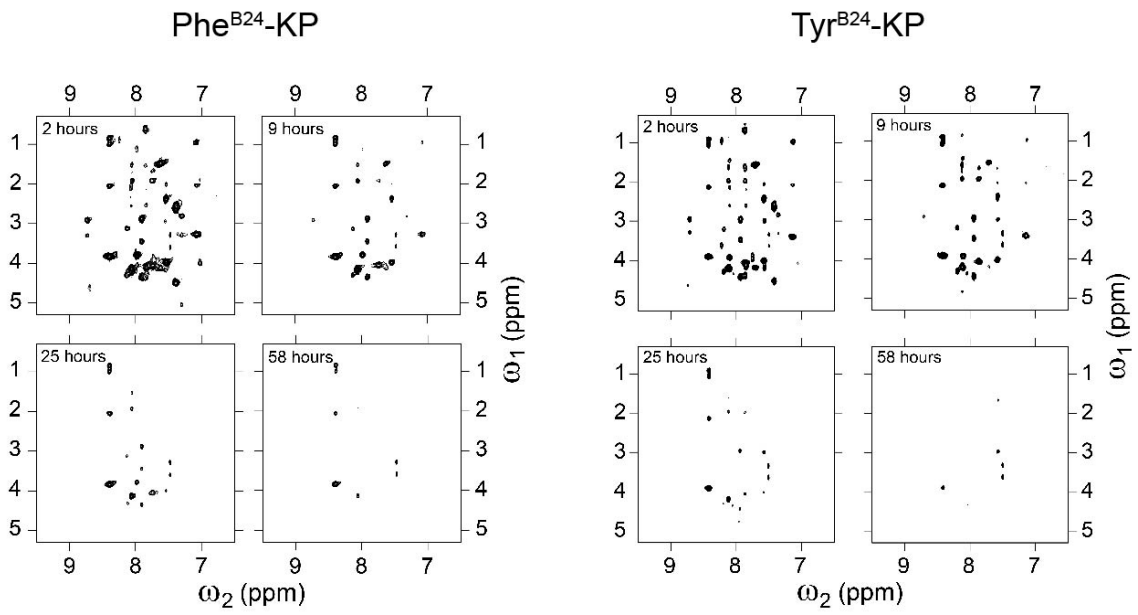


Fig. S18. 2D Amide-proton ^{1}H-^{2}H exchange kinetics in D_2O. Successive ^{1}H-^{1}H-TOCSY spectra of Tyr^{B24}-KP-insulin are shown relative to parent KP-insulin at the stated time points. Spectra were obtained after dissolving the proteins in 100% D_2O with 10 mM deuterio-acetic acid at pD 3.0 and 25 °C. KP-insulin (substitutions Lys^{B28} (K) and Pro^{B29} (P)) is also designated insulin *lispro*, the active component of Humalog® (Lilly), a rapid-acting insulin formulation in current clinical use.

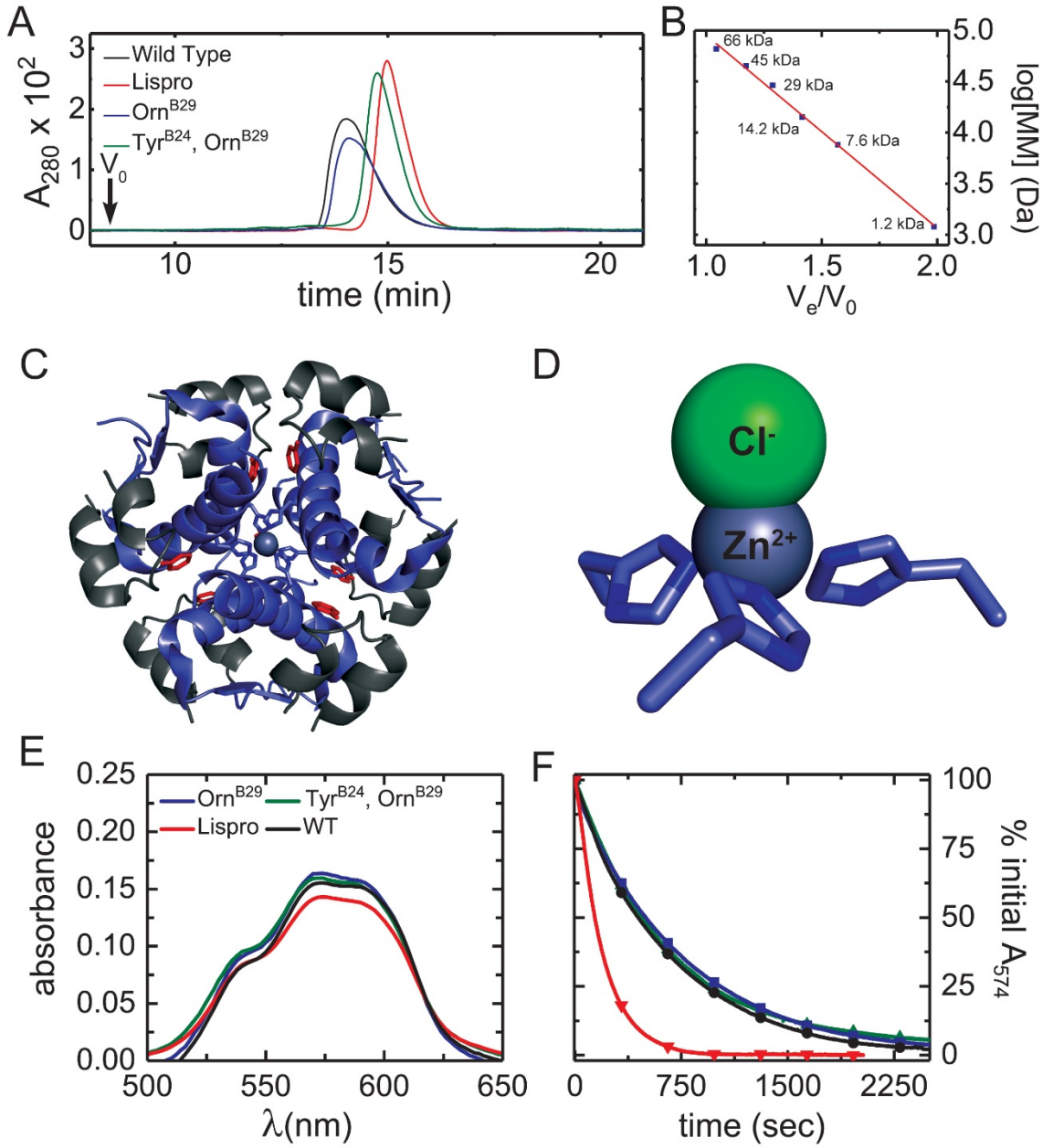


Fig. S19. Self-assembly and dissociation kinetics. (A) SEC chromatogram of insulin analogs under zinc-free conditions. Calculated MM are given in Table 3. The void volume (V_0, *black arrow*) was defined by thyroglobulin (669 kDa). (B) Control plot of log(molecular mass) vs elution ratio (V_e/V_0) of molecular-mass standards. Linear relationship between log[MM] to elution ratio (V_e/V_0) is indicated by the red line with coefficient of determination (R^2) 0.996 and parameters log[molecular mass] = -1.71*(V_e/V_0)+6.7012. Elution times of standards are indicated by *blue squares* (labeled by molecular mass): (BSA) 66 kDa; (ovalbumin) 45 kDa; (carbonic anhydrase) 29 kDa, (myosin light chain) 17 kDa; (cytochrome C) 12.4 kDa; and (IGF-I) 6.5 kDa. (C) Ribbon model of R_6 insulin hexamer (PDB ID: 1ZNJ) with A chain in *dark gray*, B chain in *blue* and phenol in *red*. Tetrahedral coordination of a Zn^{2+} ion (*gray sphere*) by His^{B10} (*blue sticks*) is shown. (D) An expanded view of Zn^{2+}-coordination site (ball-and-stick model): The Zn^{2+} ion is coordinated by three His^{B10} side chains and one chloride ion. (E) Visible absorbance spectra in corresponding Co^{2+}complex; the color code is indicated. (F) Hexamer dissociation curves as monitored at 574 nm after addition of excess EDTA to sequester released divalent metal ion (color code as in E).

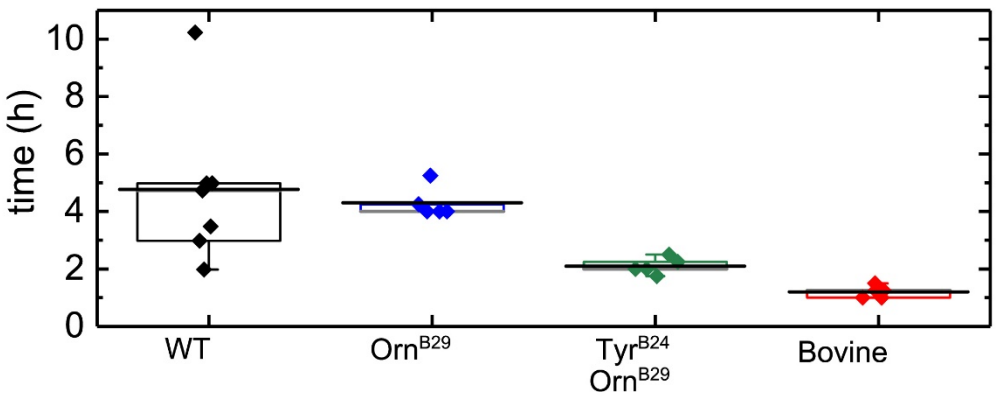


Fig. S20. Fibrillation lag-time analysis. Dot plot of lag times in studies of Tyr^{B24}, Orn^{B29}-insulin *vs*. control analogs. Mean lag times are indicated by the *black horizontal lines*. Upper and lower borders of boxes respectively denote third and second quartiles; whiskers indicate 1.5-fold interquartile range from median. Onset of fibrillation was defined by initial increase in ThT fluorescence.

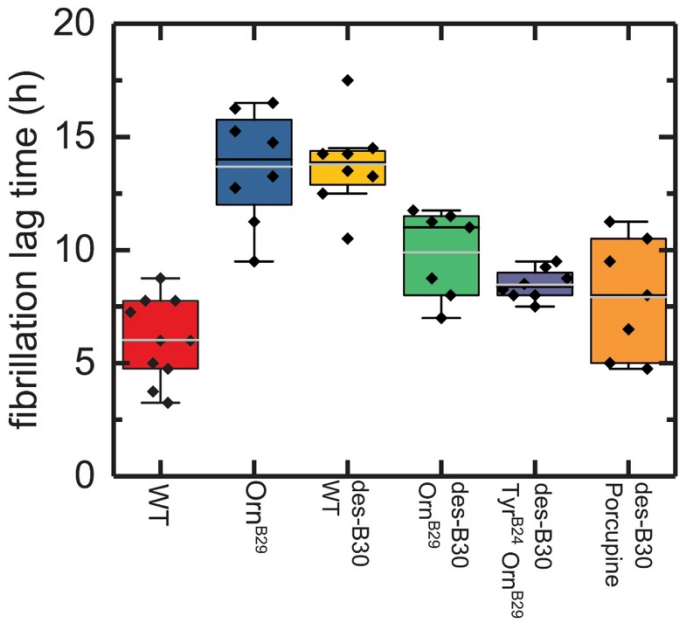


Fig. S21. Summary of fibrillation assays. Box plot of lag times prior to fibril formation of *des*-B30 insulin analogs in relation to controls. Upper- and lower boundaries of boxes respectively represent first and third quartiles. Upper and lower whiskers demarcate the highest data point within 1.5 times the interquartile range of the upper quartile and the lowest data point within 1.5 times the interquartile range of the lower quartile, respectively. Mean lag times are represented by *gray horizontal lines*, and median lag times are represented by *black horizontal lines*. Lag times of individual replicates are represented by *black diamonds*.

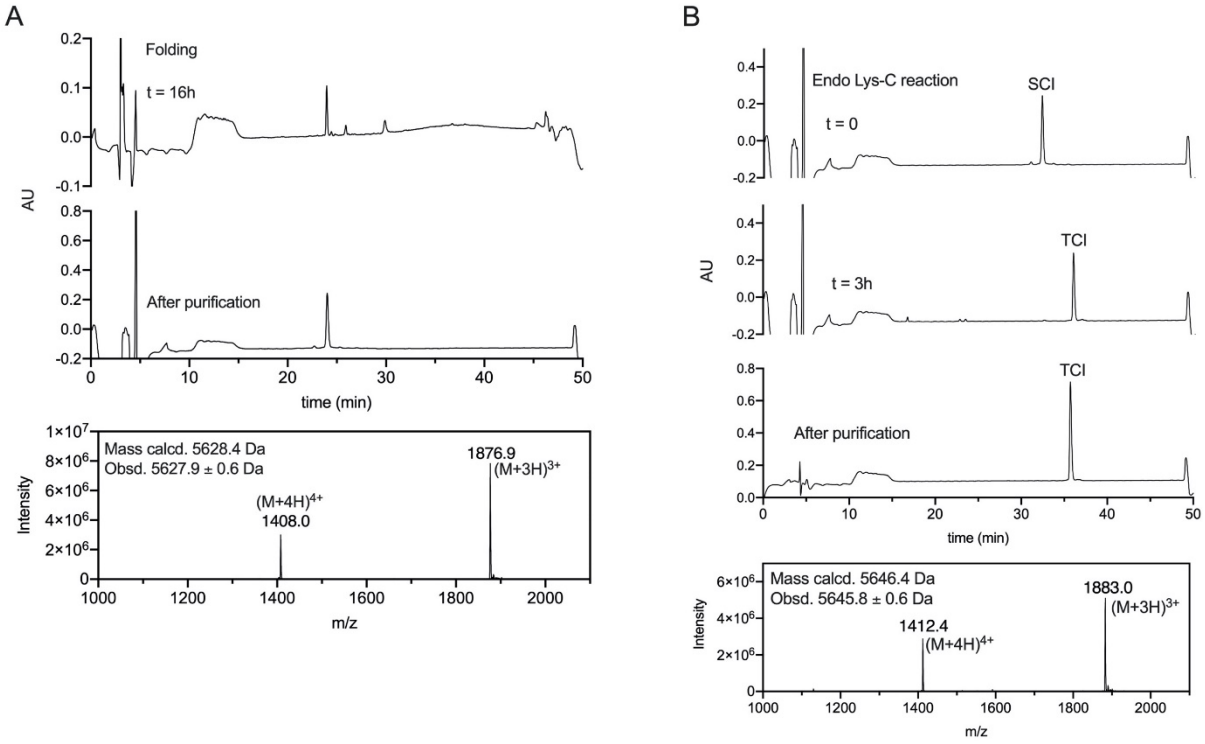
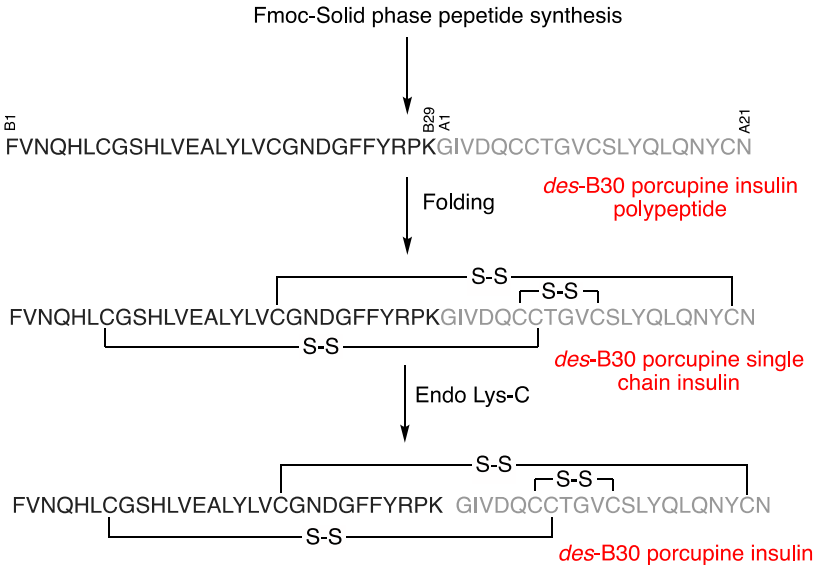


Fig. S22. Chemical synthesis of *des*-B30 porcupine insulin. (A) Folding and purification of single-chain precursor. ESI-MS spectra of the purified *des*-B30 porcupine single-chain insulin (SCI) molecule is shown in bottom panel. (B) Conversion of single-chain precursor to the two-chain insulin (TCI) molecule. ESI-MS spectra of the final two-chain molecule is shown in bottom panel.



Scheme S1. Synthetic strategy to prepare *des*-B30 porcupine insulin. A-chain (*grey*), B-chain (*black*).

Table S1. Synthetic and semi-synthetic insulin analogs

analog	purpose in study	mol. mass (Da)	
		calculated	observed
A. Control Analogs			
human insulin	wild type (WT)[a]	5807.6	5807.0 ± 0.6
insulin *lispro*	engineered monomer[b]	5807.6	5807.4 ± 0.6
Orn^{B29}-insulin	WT model	5793.6	5793.6 ± 0.6
des-B30 insulin	Control for des^{B30}-porcupine insulin	5706.5	5706.3 ± 0.6
des-B30 Orn^{B29}-insulin	Control for des^{B30}-porcupine insulin	5692.5	5691.8 ± 0.6
B. Tyr^{B24} Analogs			
Tyr^{B24}, Orn^{B29}-insulin	probe B24 "evolutionary paradox"	5809.6	5809.1 ± 0.6
des-B30 Tyr^{B24}-insulin	Comparison with des^{B30}-porcupine insulin	5708.5	5708.3 ± 0.6
Tyr^{B24}-insulin *lispro*	Tyr^{B24} in an engineered monomer	5823.6	5823.4 ± 0.6
C. Species Variants			
des-B30porcupine insulin	natural variation among mammals [c]	5646.4	5645.8 ± 0.6
bovine insulin	natural variation among mammals	5733.5	5733.4 ± 0.6

[a]Use of human insulin as a canonical "WT" is arbitrary from an evolutionary perspective but offers the advantages of (i) extensive prior mutational, biochemical, biophysical and biological characterization, (ii) activity similar to the majority of mammalian insulins and (iii) prior application of protein engineering to obtained active, monomeric variants (such as insulin *lispro*). Recombinant human insulin was purified from Humulin® (Eli Lilly) and not synthesized here.

[b]Insulin lispro contains paired substitutions $Pro^{B28} \rightarrow Lys$ and $Lys^{B29} \rightarrow Pro$; it is the active component of the Lilly product Humalog®.

[c]Residue B30 in mammalian insulins is typically not well ordered and does not contribute to receptor binding, self-assembly or stability; this residue is deleted in two clinical analogs (acylated basal products insulin *detemir* [Levemir®] and insulin *degludec* [Tresiba®]; Novo-Nodisk).

Table S2. Affinities of *des*-B30 insulin analogs[a,b]

analog	K_d (nM)
des-B30 wild-type	0.07 ± 0.01
des-B30 Orn^{B29}	0.07 ± 0.01
des-B30 porcupine	0.25 ± 0.04
des-B30 Tyr^{B24}, Orn^{B29}	0.96 ± 0.15

[a]Assays employed IR isoform B.

[b]Residue B30 was not included in the preparation of porcupine insulin for synthetic convenience; corresponding *des*-B30 controls were then prepared.

Table S3a. Thermodynamic stabilities of insulin analogs at 37 °C

Analog	ΔG_u (kcal mol^{-1})	C_{mid} (M)	m (kcal mol^{-1} M^{-1})
WT human insulin (ins)	2.5 ± 0.1	4.5 ± 0.2	0.57 ± 0.02
Orn^{B29} -ins	2.6 ± 0.1	4.6 ± 0.2	0.57 ± 0.02
Tyr^{B24}, Orn^{B29}-ins	1.7 ± 0.1	3.8 ± 0.3	0.45 ± 0.03
bovine ins	2.4 ± 0.1	4.4 ± 0.1	0.54 ± 0.02

Table S3b. Thermodynamic stabilities of *des*-B30 analogs at 37 °C

Analog	ΔG_u (kcal mol^{-1})	C_{mid} (M)	m (kcal mol^{-1} M^{-1})
des-B30 WT insulin	3.6 ± 0.1	5.0 ± 0.2	0.72 ± 0.03
des-B30 Orn^{B29}-ins	3.5 ± 0.1	4.9 ± 0.1	0.71 ± 0.02
des-B30 porcupine-ins	2.9 ± 0.1	5.5 ± 0.7	0.53 ± 0.06

Table S3c. Thermodynamic stability of IGF-1 at 25 and 37 °C

Analog (temperature in °C)[a]	ΔG_u (kcal mol^{-1})	C_{mid} (M)	m (kcal mol^{-1} M^{-1})
WT insulin (25°C)	3.6 ± 0.1	4.9 ± 0.1	0.73 ± 0.01
IGF-1 (25°C)	2.9 ± 0.1	5.2 ± 0.3	0.56 ± 0.03
WT insulin (37°C)	2.9 ± 0.1	4.4 ± 0.1	0.65 ± 0.01
IGF-1 (37°C)	2.2 ± 0.1	4.8 ± 0.1	0.46 ± 0.01

[a]Experiments used phosphate-buffered saline (pH 7.4) as sample buffer.

Table S3d. Thermodynamic stabilities of insulin analogs at 4 °C

analog	ΔG_u (kcal mol^{-1})	C_{mid} (M)	m (kcal mol^{-1} M^{-1})
Asp^{B10}-KP-insulin	4.7 ± 0.1	6.2 ± 0.1	0.76 ± 0.01
Ala^{A6}, Ala^{A11}, Asp^{B10}-KP-ins	1.8 ± 0.5	3.2 ± 0.6	0.57 ± 0.11

Table S4a. Selected ^{1}H-NMR chemical shifts (ppm) and chemical-shift differences ($|\delta| > 0.05$ ppm)[a, b]

Residue	Resonance	Phe^{B24} shift[c] (ppm)	Tyr^{B24} shift (ppm)
Val^{B2}	$\gamma_{1,2}$ -CH_3/Hγ	0.77/0.78/1.92	0.76/0.82/1.87
Leu^{B6}	$\delta_{1,2}$ -CH_3/HγH/$\beta_{1,2}$/Hα	0.70/0.81/1.59/0.69,1.69/4.56	0.69/0.80/1.59/0.70,1.70/4.54
Leu^{B11}	$\delta_{1,2}$ -CH_3/Hγ/H$\beta_{1,2}$/Hα	0.59/0.64/1.17/1.77,1.12/3.87	0.62/0.66/1.23/1.76,1.13/3.89
Val^{B12}	$\gamma_{1,2}$ -CH_3/H$_\beta$/H$_\alpha$	0.98/0.91/2.08/3.14	0.88/0.90/2.06/3.24 ppm
Ala^{B14}	β -CH_3/Hα	1.32/4.04	1.33/4.04
Leu^{B15}	$\delta_{1,2}$ -CH_3/Hγ/H$_{\beta1,2}$ /H$_\alpha$	0.93/0.90/1.70/1.90,1.73/4.06	0.91/0.88/1.76/1.88,1.68/3.96
Tyr^{B16}	H$\beta_{1,2}$/Hα	3.12/3.12/4.38	3.10/3.22/4.29
Leu^{B17}	$\delta_{1,2}$ -CH_3/Hγ/H$_{\beta1,2}$ /H	0.93/0.90/1.70/1.90,1.73/4.06	0.91/0.88/1.76/1.88,1.68/3.96
Val^{B18}	$\gamma_{1,2}$ -CH_3/H$_\beta$//H$_\alpha$	0.95/0.81/1.87/3.70	0.96/0.81/1.95/3.70
[d]Thr^{B27}	γ -CH_3/H$_\beta$	1.17/4.08	1.18/4.08
[d]Thr^{B30}	γ -CH_3/H$_\beta$	1.17/4.22	1.18/4.23
Ile^{A2}	δ, γ' -CH_3/H$\gamma_{1,2}$	0.45/0.62/0.80,1.06	0.42/0.56/0.78,1.01
Val^{A3}	$\gamma_{1,2}$ -CH_3/Hγ	0.84/0.88/1.90	0.77/0.86/1.88
Thr^{A8}	γ -CH_3/H$_\beta$	1.20/4.36	1.19/4.36
Ile^{A10}	δ, γ' -CH_3/H$\gamma_{1,2}$	0.43/0.61/0.34,1.03	0.42/0.59/0.32,1.01
Leu^{A13}	$\delta_{1,2}$ -CH_3/Hγ/H$\beta_{1,2}$/Hα	0.81/0.74/1.48 /1.37,1.44/3.82	0.79/0.73/1.48 /1.37,1.41/2.82
Leu^{A16}	$\delta_{1,2}$ -CH_3/Hγ/H$\beta_{1,2}$/Hα	0.67/0.71/1.72/1.94,1.34/4.07	0.69/0.72/1.70/1.92,1.38/4.04

[a]Residues in gray did not exhibit resonances with differences in chemical shift greater than 0.05 ppm in magnitude.

[b]Chemical shifts are provided for Asp^{B10}-insulin *lispro* (also designated DKP-insulin) and Tyr^{B24} analog of Asp^{B10}-insulin *lispro* (Tyr^{B24}-DKP-insulin).

[c]Physical origin of an attenuated ring-current shift (rcs) may in principle either be (a) a static change in the distance or orientation of the aromatic ring relative to the aliphatic side chain or (b) an increase in conformational averaging in this inter-residue distance or orientation; the latter would reflect a dynamic process at a submillisecond time scale.

[d]Chemical shifts at pH 7.0 and 32 °C obtained by homonuclear TOCSY spectra.

Table S4b. Tyr^{B24}-associated differences in $^{1}H_{N}$- and ^{15}N-NMR chemical shifts[a]

residue	location	Phe^{B24} shift[b] (ppm)	Tyr^{B24} shift (ppm)
Asn^{A21}	inter-chain H-bond[c]	8.02, 125.56	7.96, 125.26
Leu^{B6}	β-turn 1	7.87, 116.95	7.85, 117.72
Glu^{B13}	central α-helix	7.87, 116.95	7.85, 117.72[d]
Leu^{B15}	central α-helix	8.02, 125.56	7.96, 125.26
Tyr^{B16}	central α-helix	8.12, 119.47	8.16 , 119.84
Leu^{B17}	central α-helix	7.57, 119.89	7.68, 120.00
Cys^{B19}	central α-helix	8.76, 115.66	8.68, 115.85
Glu^{B21}	β-turn 2	9.10, 124.86	8.96, 124.32[d]

[a]Chemical shifts were obtained from ^{1}H-^{15}N HSQC spectra in which the respective α-helical domains of the insulin analogs (residues A1-A21 and B1-B22) were uniformly enriched with ^{13}C and ^{15}N via biosynthetic expression of a single-chain precursor in *Pichia pastoris* (see Fig. S15). Sites listed above exhibit changes in ^{1}H- and/or ^{15}N chemical shifts greater than or equal to 0.05 ppm in magnitude (^{1}H) or 0.3 ppm in magnitude (^{15}N). Tyr^{B24}-associated perturbations in chemical shifts are generally in the direction toward random-coil values.

[b]The parent monomer contains three substitutions: Asp^{B10} (D), Lys^{B28} (K) and Pro^{B29} (P); hence designation DKP-insulin.

[c]The main-chain amide proton of Asn^{A21} participates in an inter-chain hydrogen bond to the carbonyl oxygen of Gly^{B23}, adjacent to the site of Tyr^{B24} substitution.

[d]To illustrate analysis of NMR chemical shifts, we consider two glutamic-acid residues in different structural environments: Glu^{B13} and Glu^{B21}. (i) The former's amide ^{15}N resonance lies at 116.95 ppm in DKP=insulin and at 117.72 ppm in Tyr^{B24}-DKP-insulin. Since the tabulated random-coil frequency of Glu is 120.8 ppm (S26), respective secondary ^{15}N chemical shifts are -3.8 ppm and -3.1 ppm—hence attenuated by 18% in the variant. By contrast, its H_N resonance exhibits a chemical shift of 7.87 ppm in DKP-insulin and 7.85 ppm in the variant spectrum; this represents a negligible chemical-shift difference ($\Delta\delta$ 0.02 ppm). (ii) Similarly, Glu^{B21} exhibits a change in ^{15}N secondary shift from +4.1 ppm to +3.5 ppm (attenuated by 15%) and a chan in ^{1}H secondary shift from +0.68 ppm to 0.54 ppm (attenuated by 21%). Residue-type ^{1}H-NMR chemical shifts are given by Wishart, D.S., et al. (S27); cystine-specific random-coil shifts (as distinct from cysteine) are given by Wang, C.C., et al. (28); and pH-dependent corrections to the random-coil chemical shifts of C-terminal residues are given by Platzer, G. et al. (S29).

Table S5a. B24-Related NOEs involving Phe^{B24}-DKP-insulin and Tyr^{B24}-DKP-insulin[a]

Phe^{B24}/Tyr^{B24} proton	shared	Phe^{B24} analog only[b,d]	Tyr^{B24} analog only[c]
H_α	Cys^{A20}-H_α; Asn^{A21}-H_N	Asn^{A21}-H_α[#]	Cys^{A20}-H_β*
H_β	Cys^{A20}-H_α; Asn^{A21}-H_N; Cys^{B19}-H_β; Leu^{B15}- H_β & H_δ	Tyr^{A19}-H_β	Cys^{A20}-H_β*
H_δ	Leu^{B15}-H_δ; Tyr^{B16}- H_α; Cys^{B19}-Hβ; Gly^{B23}- $H_{\alpha1,2}$	Tyr^{A19}-H_β; Cys^{A20}-Hα[#]; Leu^{B15}-H_β; Gly^{B23}-H_N	-
H_ε	Val^{B12}-H_α & $H_{\gamma1,2}$; Leu^{B15}-Hδ; Tyr^{B16}-H_α & H_β	Leu^{B15}-H_β; Tyr^{B16}-Hδ & H_ε; Cys^{B19}-H_β; Gly^{B23}-$H_{\alpha2}$	-
H_N	Gly^{B23}-Hα & H_N	-	-

[a]The parent monomer contains three substitutions: Asp^{B10} (D), Lys^{B28} (K) and Pro^{B29} (P); hence its designation as DKP-insulin.

[b]NOESY spectra were acquired using two kind of samples: uniformly labeled Phe^{B24}-DKP-insulin (as obtained from labeled DKP-insulin; Yang, Y., *et al*. (S30)) and a domain-labeled Phe^{B24}-DKP-insulin in which the α-helix portion (residues A1-A21 and B1-B23) was uniformly enriched with ^{13}C and ^{15}N.

[c]NOESY spectra of Tyr^{B24}-DKP-insulin were obtained as homonuclear 2D spectra of the unlabeled protein and as 3D/4D spectra of a sample uniformly labeled in the Asp^{B10}-DOI portion.

[d]These NOEs were observed in the parent Phe^{B24} analog, but were obscured in the spectrum of the Try^{B24} analog due to either chemical-shift degeneracy, line broadening or absence of ^{13}C labeling.

*In spectrum of parent DKP-insulin chemical-shift degeneracy was observed between Cys^{A20} H_β and Phe^{B24} H_β protons, precluding rigorous identification of these two NOEs observed in Tyr^{B24} analog.

[#]Corresponding distances in respective solution structures are provided in Table S5b,c: mean distances ± standard deviaitons (SD) among 20 NMR structures in each ensemble.

Table S5b. Distances from B24 protons to B-chain protons in DG ensembles

proton pair	distance (Å)	
	Phe^{B24}	Tyr^{B24}
$H_{\varepsilon2}$-$B_{15}H_{\beta1}$	3.92±0.03 (Å)	4.55±0.04 (Å)
$H_{\varepsilon2}$-$B_{15}H_{\beta2}$	4.12±0.12	4.66±0.03
$H_{\delta2}$-$B_{15}H_{\beta1}$	3.20±0.03	3.63±0.05
$H_{\delta2}$-$B_{15}H_{\beta2}$	4.38±0.03	4.65±0.06
$H_{\varepsilon2}$-$B_{16}H_{\delta1}$	2.31±0.01	2.48±0.03
$H_{\varepsilon2}$-$B_{16}H\varepsilon_1$	4.41±0.03	4.16±0.03
$H_{\varepsilon2}$-$B_{19}H_{\beta1}$	5.51±0.04	5.29±0.04
$H_{\varepsilon2}$-$B_{19}H_{\beta2}$	4.05±0.05	4.17±0.04
$H_{\varepsilon2}$-$B_{23}H_{\alpha2}$	3.85±0.04	4.30±0.03
$H_{\delta2}$-$B_{23}H_N$	3.84±0.06	3.79±0.09

Table S5c. Distances from B24 protons to A-chain protons in DG ensembles

Proton pair	Distance (Å)	
	Phe^{B24}	Tyr^{B24}
$H_{\beta1}$-$A_{19}H_{\beta2}$	4.08±0.06 (Å)	5.15±0.05 (Å)
$H_{\delta1}$-$A_{19}H_{\beta1}$	4.07±0.02	4.55±0.19
H_α-A20$H_{\beta1}$	4.99±0.06	5.14±0.03
$H_{\beta1}$-A20$H_{\beta1}$	4.34±0.01	4.52±0.05
$H_{\delta2}$-A20H_α	4.19±0.02	5.12±0.09
H_α-A21H_α	4.16±0.05	5.33±0.03

Table S6a. Statistics of experimental data and structure calculations of Tyr^{B24} analog

parameter	SA ensemble[a]
RMSD from experimental distance restraints (Å)	
all (925)	0.041 ± 0.009
intraresidue, i=j (325)	0.036 ± 0.009
sequential, |i-j|=1 (229)	0.040 ± 0.009
medium range, 1<|i-j|<5 (178)	0.039 ± 0.007
long range, |i-j|≥5 (180)	0.055 ± 0.013
RMSD from idealized covalent geometry	
bonds (Å)	0.0066± 0.0015
angles (°)	0.92 ± 0.20
impropers (°)	2.32 ± 0.52
E_{L-J} (kcal/mol)[b]	-180.6 ± 40.6
Ramachandran plot[c]	
most favored regions (%)	89.6
additionally and generously allowed regions (%)	10.4
disallowed regions (%)	0.0
Coordinate precision of native insulin-like region[d]	
rmsd of backbone atoms to the mean (Å)	0.21 ± 0.05
rmsd of all heavy atoms to the mean (Å)	0.77 ± 0.17

[a]Mean ± standard error where applicable.
[b]Lennard-Jones potential energy function, calculated with CHARMM19 (S27) empirical energy parameters.
[c]Residues 2-20 (A2-A20), 26-47 (B5-B26) for native insulin region. Glycine residues were not included.
[d]Calculated using molmol program (reference) for residues 2-20 (A2-A20), 26-47 (B5-B26).

Table S6b. Statistics of experimental data and structure calculations of Phe^{B24}-DKP-insulin

parameter	SA ensemble [a]
RMSD from experimental distance restraints (Å)	
all (1013)	0.032 ± 0.007
intraresidue, i=j (390)	0.034 ± 0.008
sequential, |i-j|=1 (232)	0.040 ± 0.009
medium range, 1<|i-j|<5 (193)	0.024 ± 0.006
long range, |i-j|≥5 (198)	0.023 ± 0.005
RMSD from idealized covalent geometry	
bonds (Å)	0.0064± 0.0014
angles (°)	0.92 ± 0.21
impropers (°)	2.33 ± 0.52
E_{L-J} (kcal/mol)[b]	-210.0 ± 45.3
Ramachandran plot[c]	
most favored regions (%)	85.7
additionally and generously allowed regions (%)	14.3
disallowed regions (%)	0.0
Coordinate precision of native insulin-like region[d]	
rmsd of backbone atoms to the mean (Å)	0.23 ± 0.07
rmsd of all heavy atoms to the mean (Å)	0.77 ± 0.12

[a]Mean ± standard error where applicable.
[b]Lennard-Jones potential energy function, calculated with CHARMM19 (S27) empirical energy parameters.
[c]Residues 2-20 (A2-A20), 26-47 (B5-B26) for native insulin region. Glycine residues were not included.
[d]Calculated using molmol program (reference) for residues 2-20 (A2-A20), 26-47 (B5-B26).

Table S7. Side-chain dihedral angles of residue B24 and its surrounding residues[a]

	χ_1		χ_2	
	Phe^{B24}	Tyr^{B24}	Phe^{B24}	Tyr^{B24}
Ile^{A2}	-170.7±2.8	-178.2±3.2	177.2±2.1	168.9±1.7
Tyr^{A19}	-54.6±1.2	-65.3±2.2	-89.4±0.5	-80.6±1.6
Leu^{B11}	171.8±1.1	167.2±1.6	50.4±1.2	54.6±1.9
Val^{B12}	178.3±1.2	164.1±1.8		
Leu^{B15}	-75.6±0.5	-68.3±1.7	163.0±0.4	157.3±0.5
Tyr^{B16}	-172.9±4.9	-164.8±2.6	86.8±10.9	94.9±4.6
*Phe^{B24}	47.4±2.0	57.9±1.9	100.5±1.1	103.1±1.2
Tyr^{B26}	176.5±2.4	177.1±2.0	79.0±6.9	83.3±2.1

[a]χ_1 and χ_2 angles were determined for Phe^{B24} in the parent analog or for Tyr^{B24} in the variant.

Table S8. ^{1}H-^{2}H exchange rates of KP-insulin and Tyr^{B24} KP-insulin (pD 3.0 and 25 °C)

KP control				Tyr^{B24} KP		
Residue	Kobs (10^{-5})	PF	ΔGu	Kobs (10^{-5})	PF	ΔGu
Global effected						
A16 HA	1.02±0.01	109.2±1.4	2.78±0.01	1.83±0.03	60.5±1.2	2.43±0.01
A16HB1	0.93±0.02	119.7±3.2	2.83±0.02	1.68±0.14	65.9±5.5	2.48±0.05
A16HB2	0.95±0.04	116.3±5.1	2.81±0.03	1.68±0.04	66.1±1.6	2.48±0.01
A16HG	1.07±0.02	103.3±1.8	2.74±0.01	1.60±0.23	69.2±8.6	2.51±0.07
A19 HA	1.48±0.04	111.8±2.9	2.79±0.02	2.68±0.06	61.9±1.4	2.44±0.01
A19HB1	1.44±0.01	114.8±0.9	2.81±0.01	2.52±0.03	65.8±0.8	2.48±0.01
A19HB2	1.33±0.03	125.2±2.8	2.86±0.01	2.43±0.01	68.3±0.3	2.50±0.01
B18HA	0.51±0.02	176.8±7.0	3.06±0.02	0.97±0.01	93.5±0.9	2.69±0.01
B18HB	0.52±0.01	173.5±4.0	3.05±0.01	0.94±0.19	96.8±1.9	2.71±0.01
B18HG1	0.51±0.03	178.3±9.9	3.07±0.03	0.96±0.06	94.0±6.4	2.69±0.04
B18HG2	0.53±0.01	170.7±4.1	3.04±0.01	1.07±0.01	84.4±1.1	2.63±0.01
Sub-global effected						
A15HA	3.24±0.07	52.8±1.2		4.05±0.09	42.2±0.9	
A15HB1	4.35±0.27	39.3±2.4		3.94±0.07	43.4±0.8	
A15HB2	3.30±0.09	51.8±1.5		5.24±0.27	32.6±1.6	
A17HA	2.73±0.16	53.8±3.2		3.08±0.16	47.6±2.4	
A17HB1	3.51±0.14	41.8±1.7		8.72±0.93	16.8±1.7	
A17HB2	4.43±0.46	34.3±3.7		6.69±0.22	21.9±0.7	
A17HG	4.30±0.57	34.2±4.5		6.91±1.06	21.2±3.2	
A18HA	10.1±0.45	55.3±2.5		13.0±0.02	42.7±0.6	
A18HB1	12.1±0.10	46.2±3.9		12.8±0.26	43.4±0.9	
A18HB2	11.3±0.05	49.1±2.2		12.1±0.37	45.9±1.4	
A20HA	5.19±0.34	67.8±4.5		5.35±0.24	65.6±0.8	
A20HB1	4.23±0.44	83.1±8.7		6.83±1.97	51.5±14.8	
B12HA	2.48±0.11	36.7±1.6		4.38±0.10	20.7±0.5	
B12HB	2.73±0.34	33.2±4.2		4.80±0.36	18.9±1.4	
B12HG	3.48±0.09	26.1±0.7		6.56±0.91	13.8±1.9	
B13HA	7.79±0.74	19.9±1.9		12.46±0.93	12.4±0.9	
B13HB1	9.59±0.93	16.2±1.6		12.87±0.44	12.0±0.4	
B14HA	4.24±0.17	60.1±2.4		7.72±0.29	33.0±1.2	
B14HB	4.75±0.04	53.7±0.5		7.27±0.18	35.1±0.9	
B15HA	1.58±0.06	67.1±2.5		2.99±0.05	35.5±0.6	
B15HB1	1.57±0.05	67.3±2.0		3.13±0.18	33.9±2.0	
B15HB2	N/A			5. 69±0.87	18.6±2.84	
B16HA	1.41±0.03	79.3±1.5		2.07±0.05	53.9±1.3	
B16HB	1.22±0.07	91.9±4.9		2.04±0.09	54.7±2.5	
B17HA	1.31±0.06	77.6±3.5		2.05±0.09	49.7±2.2	
B17HB1	1.30±0.08	78.1±4.7		1. 56±0.23	65.2±9.5	
B17HB2	1.25±0.07	81.8±4.9		2.12±0.29	48.1±6.6	
B19HB1	3.69±0.43	69.4±8.1		6.82±0.38	37.6±2.1	
B19HB2	3.63±0.17	70.4±3.3		7. 07±0.98	36.2±5.0	
Local effected						
A3HA	N/A			46.70±4.32	1.8±0.2	
A3HB	N/A			53.05±2.22	1.6±0.1	
A3HG1	N/A			47.28±8.45	1.8±0.3	
A3HG2	57.50±2.34	1.5±0.6		47.86±7.95	1.8±0.3	
A3HN[c]	186.22±2.02	1.0±0.0		87.72±8.31	1.0±0.1	
A10HA	32.26±1.73	3.2±0.2		22.83±4.62	4.6±0.9	
A10HB	31.87±1.79	3.3±0.2		N/A		
A10HG1	N/A			35.84±3.76	2.9±0.3	
A10HG2	33.23±7.55	3.1±0.7		31.70±1.46	3.3±0.2	
A10HD1	N/A			39.30±2.88	2.6±0.2	
A11HN[c]	282.48±56.65	1.0±0.0		223.00±48.10	1.0±0.7	
A12HA	23.31±1.23	23.3±1.2		35.31±1.46	15.4±3.7	
A14HA	65.15±2.67	1.7±0.1		25.67±3.06	4.4±0.5	
B6HN	111.35±15.22	1.8±0.3		161.03±14.52	1.3±0.1	
B11HA	94.61±13.07	2.1±0.3		123.45±32.01	1.6±0.1	
B11HB1	83.26±12.00	2.4±0.3		N/A		
B20HA1	N/A			35.50±7.99	12.7±2.8	
B20HA2	N/A			40.53±8.41	11.1±2.3	

[a]In previous studies (Hua, Q.-X., et al. *(S31)*) the baseline thermodynamic stability of insulin *lispro* (KP-insulin), as inferred from CD-detected guanidine denaturing studies at pH 3.0 or pD 3.0 at 25 °C, was ΔG_u = 3.3(±0.1) kcal/mol in accordance with an ^{1}H-^{2}H estimate based of global exchange of ΔG_u = 3.26 kcal/mole. Global PFs were lower in the present study, presumably due to a systematic difference in conditions.

[b]In TOCSY spectra, amide protons exhibit H_N-H_α, H_N-H_β correlations and also cross peaks with side-chain protons, which enable calculation of multiple separated values for one residue. Values

of H_N were obtained from serial 1D experiments.
[c]Measured exchange rates were smaller than intrinsic rates.
[d]Definition of exchange categories. (i) *Global sites*: PFs at sites of global exchange in insulin *lispro* (KP-insulin) is similar to estimates based on guanidine denaturation and two-state modeling (see Hua, Q.-X., et al. *(S31)*). Global sites were thus defined as residues whose PFs were in within 10% of the CD-derived estimate of ΔG_u. (ii) *Subglobal sites*: PFs in the range 20-90% relative to the CD-derived estimate of ΔG_u. (iii) *Local sites*: PFs < 20% relative to the CD-derived estimate of ΔG_u.

Table S9. Fibrillation lag times of *des*-B30 insulin analogs

analog (n)[a]	fibrillation lag time (h ± SD)
human insulin (ins) (10)	6.0 ± 1.8
Orn^{B29}-insulin (8)	13.7 ± 2.5
des-B30 human insulin (8)	13.8 ± 2.0
des-B30 Orn^{B29}-insulin (7)	9.9 ± 1.9
des-B30 Tyr^{B24}, Orn^{B29}-ins (8)	8.5 ± 0.7
des-B30 porcupine insulin[b] (7)	7.9 ± 2.6

[a]The number of replicates is given by (n).

[b]The sequence of porcupine insulin differs from human insulin by the following eight substitutions: $Glu^{B21}\rightarrow$Asn, $Arg^{B22}\rightarrow$Asp, $Thr^{B27}\rightarrow$Arg, $Thr^{B30}\rightarrow$Ala, $Glu^{A3}\rightarrow$Asp, $Ser^{A9}\rightarrow$Gly, $Ile^{A10}\rightarrow$Val and $Glu^{A17}\rightarrow$Gln. Such natural variation (which includes interchange of similar amino acids) highlights in the breach the anomalous exclusion of Tyr^{B24} among vertebrate insulin and IGF sequences.

Table S10. Biological activities of low-potency species variants of insulin

species	*In vitro* activity[a]	*cellular* activity[b]	reference
porcupine	25%	4%	Horuk, 1980 (6)
guinea pig	3.1%	-	Wood, 1975 (32)
	4.5%	4-9%	Zimmerman, 1973 (33)
chinchilla	62%	-	Wood 1975 (32)
coypu	5-8%	3%	Bajaj, 1986 (34)
Atlantic hagfish	25%	-	Cutfield, 1979 (35)
	4.5%	9.4%	Sajid, 2009 (36)
lamprey	8%	8.4%	Sajid, 2009 (36)

[a]Potencies were determined in relation to human insulin, porcine insulin, or bovine insulin, which have equivalent IR affinities, using competition assays that utilized either purified IR, whole cells expressing IR, or IR-containing cell membranes (S6, S32-S36).

[b]Relative *in vivo* activities were determined in relation to controls given in table footnote *a* using rat adipocyte lipogenesis assays, glucose uptake assays in whole cells, or glucose-lowering assays after injection in rodents (S6, S32-S36).

Table S11a. Thermodynamic stabilities of B24 variant insulin analogs[a]

analog	ΔG_u (kcal mol^{-1})	C_{mid} (M)	*m* (kcal mol^{-1} M^{-1})
human insulin	3.4 ± 0.1	5.0 ± 0.2	0.7 ± 0.03
Orn^{B29}-insulin	3.6 ± 0.2	4.9 ± 0.2	0.7 ± 0.03
Gly^{B24}, Orn^{B29}-insulin	2.3 ± 0.1	4.6 ± 0.28	0.5 ± 0.03
Ser^{B24}, Orn^{B29}-insulin	2.3 ± 0.1	4.5 ± 0.10	0.5 ± 0.01
Leu^{B24}, Orn^{B29}-insulin	2.3 ± 0.1	4.6 ± 0.09	0.5 ± 0.01

[a]Data were originally published in Rege, N., *et al. J. Biol. Chem* (2020) (S21).

Table S11b. Thermodynamic stabilities of B24 variant insulin analogs[b]

analog	ΔG_u (kcal mol^{-1})	C_{mid} (M)	*m* (kcal mol^{-1} M^{-1})
human insulin	3.4 ± 0.1	4.8 ± 0.1	0.69 ± 0.02
Orn^{B29}-insulin	3.5 ± 0.2	4.8 ± 0.2	0.73 ± 0.03
Met^{B24},Orn^{B29}-insulin	2.2 ± 0.1	4.1 ± 0.48	0.53 ± 0.02
Ile^{B24},Orn^{B29}-insulin	2.2 ± 0.1	4.2 ± 0.10	0.53 ± 0.02

[b]Data were originally published in Pandyarajan, V., *et al. J. Biol. Chem* (2014) (S15).

SI References

Additional supplemental references are provided in footnotes to Supplemental Tables.

S1. Q. X. Hua, W. Jia, B. H. Frank, N. F. Phillips, M. A. Weiss, A protein caught in a kinetic trap: structures and stabilities of insulin disulfide isomers. *Biochemistry* **41**, 14700-14715 (2002).

S2. T. R. Sosnick, X. Fang, V. M. Shelton, Application of circular dichroism to study RNA folding transitions. *Methods Enzymol.* **317**, 393-409 (2000).

S3. M. D. Glidden *et al.*, Solution structure of an ultra-stable single-chain insulin analog connects protein dynamics to a novel mechanism of receptor binding. *J. Biol. Chem.* **293**, 69-88 (2018).

S4. E. N. Baker *et al.*, The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **319**, 369-456 (1988).

S5. J. Badger *et al.*, Structure of the pig insulin dimer in the cubic crystal. *Acta Crystallogr. B* **47 (Pt 1)**, 127-136 (1991).

S6. R. Horuk *et al.*, A monomeric insulin from the porcupine (*Hystrix cristata*), an Old World hystricomorph. *Nature* **286**, 822-824 (1980).

S7. G. L. Devlin *et al.*, The component polypeptide chains of bovine insulin nucleate or inhibit aggregation of the parent protein in a conformation-dependent manner. *J. Mol. Biol.* **360**, 497-509 (2006).

S8. V. Pandyarajan *et al.*, Biophysical optimization of a therapeutic protein by non-standard mutagenesis. Studies of an iodo-insulin derivative. *J. Biol. Chem.* **289**, 23367-23381 (2014).

S9. A. N. Zaykov, J. P. Mayer, V. M. Gelfanov, R. D. Dimarchi, Chemical synthesis of insulin analogs through a novel precursor. *ACS Chem. Biol.* 10.1021/cb400792s (2014).

S10. F. Liu, A. N. Zaykov, J. J. Levy, R. D. DiMarchi, J. P. Mayer, Chemical synthesis of peptides within the insulin superfamily. *J. Pept. Sci.* **22**, 260-270 (2016).

S11. V. Pandyarajan *et al.*, Contribution of Tyr^{B26} to the function and stability of insulin. Structure-activity relationships at a conserved hormone-receptor interface. *J. Biol. Chem.* **291**, 12978-12990 (2016).

S12. M. Roy *et al.*, Spectroscopic signatures of the T to R conformational transition in the insulin hexamer. *J. Biol. Chem.* **264**, 19081-19085 (1989).

S13. D. T. Birnbaum, M. A. Kilcomons, M. R. DeFelippis, J. M. Beals, Assembly and dissociation of human insulin and Lys^{B28}, Pro^{B29}-insulin hexamers: a comparison study. *Pharm. Res.* **14**, 25-36 (1997).

S14. Z. X. Wang, An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule *FEBS Lett.* **360**, 111-114 (1995).

S15. V. Pandyarajan *et al.*, Aromatic anchor at an invariant hormone-receptor interface function of insulin residue B24 with application to protein design. *J. Biol. Chem.* **289**, 34709-34727 (2014).

S16. A. Bax, Multidimensional nuclear-magnetic-resonance methods for protein studies. *Curr. Opin. Struct. Biol.* **4**, 738-744 (1994).

S17. L. E. Kay, NMR methods for the study of protein structure and dynamics. *Biochem. Cell Biol.* **75**, 1-15 (1997).

S18. Y. Xu, D. Long, D. Yang, Rapid data collection for protein structure determination by NMR spectroscopy. *J. Am. Chem. Soc.* **129**, 7722-7723 (2007).

S19. C. D. Schwieters, J. J. Kuszewski, G. M. Clore, Using Xplor-NIH for NMR molecular structure determination. *Prog. NMR Spec.* **48**, 47-62 (2006).
S20. R. Koradi, M. Billeter, K. Wuthrich, MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55, 29-32 (1996).
S21. N. K. Rege *et al.*, "Register-shift" insulin analogs uncover constraints of proteotoxicity in protein evolution. *J. Biol. Chem.* **295**, 3080-3096 (2020).
S22. J. Sun *et al.*, Role of proinsulin self-association in mutant *INS* gene–induced diabetes of youth. *Diabetes* **69**, 954-964 (2020).
S23. M. Liu, Y. Li, D. Cavener, P. Arvan, Proinsulin disulfide maturation and misfolding in the endoplasmic reticulum. *J. Biol. Chem.* **280**, 13209-13212 (2005).
S24. N. Listov-Saabye *et al.*, MCF-7 human mammary adenocarcinoma cells exhibit augmented responses to human insulin on a collagen IV surface. *J. Appl. Toxicol.* **29**, 470-477 (2009).
S25. S. E. Shoelson, Z. X. Lu, L. Parlautan, C. S. Lynch, M. A. Weiss, Mutations at the dimer, hexamer, and receptor-binding, surfaces of insulin independently affect insulin-insulin and insulin-receptor interactions. *Biochemistry* **31**, 1757-1767 (1992).
S26. D. Braun, G. Wider, K. Wuethrich, Sequence-corrected 15N "random coil" chemical shifts. *J. Am. Chem. Soc.* **116**, 8466-8469 (1994).
S27. D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, ^{1}H, ^{13}C, and ^{15}N random coil NMR chemical-shifts of the common amino-acids. 1. Investigation of nearest-neighbor effects. *J. Biomol. NMR* **5**, 67-81 (1995).
S28. C. C. Wang, J. H. Chen, S. H. Yin, W. J. Chuang, Predicting the redox state and secondary structure of cysteine residues in proteins using NMR chemical shifts. *Proteins* **63**, 219-226 (2006).
S29. G. Platzer, M. Okon, L. P. McIntosh, pH-dependent random coil ^{1}H, ^{13}C, and ^{15}N chemical shifts of the ionizable amino acids: a guide for protein pKa measurements. *J. Biomolec. NMR* **60**, 109-129 (2014).
S30. Y. Yang *et al.*, Solution structure of proinsulin: connecting domain flexibility and prohormone processing. *J. Biol. Chem.* **285**, 7847-7851 (2010).
S31. Q. X. Hua, W. Jia, M. A. Weiss, Conformational dynamics of insulin. *Front Endocrinol. (Lausanne)* **2**, 48 (2011).
S32. S. P. Wood, T. L. Blundell, A. Wollmer, N. R. Lazarus, R. W. Neville, The relation of conformation and association of insulin to receptor binding; X-ray and circular-dichroism studies on bovine and hystricomorph insulins. *Eur. J. Biochem.* **55**, 531-542 (1975).
S33. A. E. Zimmerman, M. L. Moule, C. C. Yip, Guinea pig insulin II. Biological activity. *J. Biol. Chem.* **249**, 4026-4029 (1974).
S34. M. Bajaj *et al.*, Coypu insulin. Primary structure, conformation and biological properties of a hystricomorph rodent insulin. *Biochem. J.* **238**, 345-351 (1986).
S35. J. F. Cutfield *et al.*, Structure and biological activity of hagfish insulin. *J. Mol. Biol.* **132**, 85-100 (1979).
S36. W. Sajid *et al.*, Structural basis of the aberrant receptor binding properties of hagfish and lamprey insulins. *Biochemistry* **48**, 11283-11295 (2009).