Supporting Information

4S-Hydroxylation of insulin at ProB28 accelerates hexamer dissociation and delays fibrillation

Seth A. Lieblich[†,‡], Katharine Y. Fang[†,‡], Jackson K. B. Cahn[†], Jeffrey Rawson[§,‖], Jeanne LeBon[§], H. Teresa Ku[§,‖,¶], David A. Tirrell[†,*]

†Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

[§]Department of Translational Research and Cellular Therapeutics, Diabetes and Metabolism Research Institute, City of Hope, Duarte, CA 91010, USA.

[‖]Beckman Research Institute of City of Hope, Duarte, CA 91010, USA.

[¶]Irell & Manella Graduate School of Biological Sciences, City of Hope, Duarte, CA 91010, USA.

*Corresponding author. Email: tirrell@caltech.edu

Table of Contents

Materials and Methods:

Materials. All canonical amino acids and (4*R*)-hydroxy-L-proline (Hyp) were purchased from Sigma. *(*4*S)*-hydroxy-L-proline (Hzp) was purchased from Bachem Americas. All solutions and buffers were made using double-distilled water (ddH_2O).

Strains and plasmids. The proinsulin (PI) gene with an *N*-terminal hexa-histidine tag (6xHIS), and flanked by *Eco*R1 and *Bam*H1 cut sites was ordered as a gBlock (Integrated DNA Technologies). Both the gBlock and vector pQE80L for IPTG-inducible expression were digested with *Eco*RI and *Bam*HI. Linearized vector pQE80L was dephosphorylated by alkaline phosphatase (NEB). Ligation of the digested PI gene and linearized vector yielded plasmid pQE80PI (to produce ProI). To make plasmid pQE80PI-proS (to produce HzpI and HypI): Genomic DNA was extracted from *E. coli* strain DH10β using DNeasy Blood and Tissue Kit (Qiagen). Primers (Integrated DNA Technologies) were designed to amplify the *E. coli proS* gene, encoding prolyl-tRNA synthetase, under constitutive control of its endogenous promoter, from purified genomic DNA, and to append *Nhe*I and *Nco*I sites. The digested *proS* gene was then inserted into pQE80PI between transcription termination sites by ligation at *Nhe*I and *Nco*I restriction sites. Proline-auxotrophic *E. coli* strain CAG18515 was obtained from the Coli Genetic Stock Center at Yale University. Prototrophic *E. coli* strain BL-21 was used for rich media expression of canonical insulins (ProI, AspI). Site-directed mutagenesis of pQE80PI at B28 was performed to make plasmid pQE80PI-asp, which differs from pQE80PI by three nucleotides that specify a single amino acid mutation to aspartic acid. All genes and plasmids were confirmed by DNA sequencing.

Protein expression. Plasmids pQE80PI and pQE80PI-asp were transformed into BL21 cells and grown on ampicillin-selective agar plates. A single colony was used to inoculate 5 mL of Luria-Bertani (LB) medium and grown overnight; the resulting saturated culture was used to inoculate another 1 L of LB medium. All expression experiments were conducted at 37°C, 200 RPM in shake flasks (Fernbach 2.8 L flasks, Pyrex®). Each culture was induced with 1 mM IPTG at mid-exponential phase (OD_{600} ~0.8). For incorporation of Hyp and Hzp, pQE80PI-proS was transformed into CAG18515 cells, which were grown on ampicillin-selective agar plates. To facilitate growth, a single colony was used to inoculate 25 mL of LB medium and the culture was grown overnight prior to dilution into 1 L of 1X M9, 20 amino acids (8.5 mM NaCl, 18.7 mM NH_4Cl, 22 mM KH_2PO_4, 47.8 mM Na_2HPO_4, 0.1 mM $CaCl_2$, 1 mM $MgSO_4$, 3 mg/L $FeSO_4$, 1 µg/L of trace metals (Cu^{2+}, Mn^{2+}, Zn^{2+}, MoO_4^{2-}), 35 mg/L thiamine hydrochloride, 10 mg/L biotin, 20 mM D-glucose, 200 mg/L ampicillin with 50 mg/L of L-amino acids, each). At an appropriate cell density (OD_{600} ~0.8), the culture was subjected to a medium shift; briefly, cells were centrifuged and washed with saline prior to resuspension into 0.8 L of 1.25X M9, 19 aa (M9, 20 aa medium without L-proline). After cells were further incubated for 30 min to deplete intracellular proline, 200 mL of 5X additives (1.5 M NaCl, 2.5 mM Hyp or Hzp) was added to the culture. After another

15 min of incubation at 37°C to allow amino acid uptake prior to induction, IPTG was added to a final concentration of 1 mM. At the end of 2 h, cells were harvested by centrifugation and stored at -80°C until further use.

Cell lysis and refolding from inclusion bodies. Cells were thawed on the benchtop for 15 min prior to resuspension in lysis buffer (B-PER®, 0.5 mg/mL lysozyme, 50 U/mL benzonase nuclease). Cells were gently agitated at RT for 1 h prior to centrifugation (10 000 g, 10 min, RT); supernatant was discarded and the pellet was washed thrice: once with wash buffer (2 M urea, 20 mM Tris, 1% Triton X-100, pH 8.0) and twice with sterile ddH_2O; centrifugation followed each wash and the supernatant was discarded. The final washed pellet containing inclusion bodies (IBs, ~50% PI) was re-suspended in Ni-NTA binding buffer (8 M urea, 300 mM NaCl, 50 mM NaH_2PO_4, pH 8.0) overnight at 4°C or at RT for 2 h, both with gentle agitation. The suspension was centrifuged to remove insoluble debris; the remaining pellet was discarded and the supernatant was mixed with pre-equilibrated Ni-NTA resin (Qiagen) at RT for 1 h in order to purify PI from the IB fraction. Unbound proteins in the IB fraction were collected in the flow-through (FT), and the resin was washed with Ni-NTA wash buffer (8 M urea, 20 mM Tris base, 5 mM imidazole, pH 8.0) and Ni-NTA rinse buffer (8 M urea, 20 mM Tris base, pH 8.0) prior to stripping PI from the resin with Ni-NTA elution buffer (8 M urea, 20 mM Tris base, pH 3.0). Fractions (IBs, FT, W, elution) were collected and run under reducing conditions on SDS-PAGE (Bis/Tris gels, Novex®); elution fractions containing PI were pooled and solution pH was adjusted to 9.6 with 6 N NaOH in preparation for oxidative sulfitolysis. Oxidative sulfitolysis was performed at RT for 4 h, with the addition of sodium sulfite and sodium tetrathionate (0.2 M Na_2SO_3, 0.02 M $Na_2S_4O_6$); the reaction was quenched by 10-fold dilution with ddH_2O. To isolate PI from the quenched solution, the pH was adjusted to between 3.5 and 4.5 by adding 6 N HCl dropwise; the solution became cloudy. The solution was centrifuged (10 000 g, 10 min, RT) and supernatant discarded. The PI pellet was then re-suspended in refolding buffer (0.3 M urea, 50 mM glycine, pH 10.6) and protein concentration was estimated by the bicinchoninic acid assay (BCA assay, Pierce®). The concentration of PI was adjusted to 0.5 mg/mL. Refolding was initiated by addition of β-mercaptoethanol to a final concentration of 0.5 mM and allowed to proceed at 12°C overnight with gentle agitation (New Brunswick® shaker, 100 RPM). Post-refolding, soluble PI was harvested by adjusting the pH of the solution to 4-5 by dropwise addition of 6 N HCl and by high speed centrifugation to remove insoluble proteins. The supernatant was adjusted to pH 8-8.5 by dropwise addition of 6 N NaOH and dialyzed against fresh PI dialysis buffer (7.5 mM sodium phosphate buffer, pH 8.0) at 4°C with five buffer changes to remove urea. The retentate (PI in dialysis buffer) was then lyophilized and subsequently stored at -80°C until further processing. Typical yields were 25-50 mg PI per L of culture (25-30 mg/L for non-canonical PI, 40-50 mg/L for canonical PI expression in rich media)

Proteolysis and chromatographic (HPLC) purification. The dry PI powder was re-dissolved in water to a final concentration of 5 mg/mL PI (final concentration of sodium phosphate buffer is 100 mM, pH 8.0). Trypsin (Sigma-Aldrich) and carboxypeptidase-B (Worthington Biochemical) were added to final concentrations of 20 U/mL and 10 U/mL, respectively to initiate proteolytic cleavage. The PI/protease solution was incubated at

37°C for 2.5 h; proteolysis was quenched by addition of 0.1% trifluoroacetic acid (TFA) and dilute HCl to adjust the pH to 4. Matured insulin was purified by reversed phase high-performance liquid chromatography (HPLC) on a C_{18} column using a gradient mobile phase of 0.1% TFA in water (solvent, A) and 0.1% TFA in acetonitrile (ACN; solvent, B). Elution was carried from 0% B to 39% B with a gradient of 0.25% B per minute during peak elution. Fractions were collected and lyophilized, and the dry powder was re-suspended into 10 mM sodium phosphate, pH 8.0. Insulin-containing fractions were verified by matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS; Voyager MALDI-TOF, Applied Biosystems) and SDS-PAGE to ensure identify and purity. Typical yields were 5-10 mg insulin per 100 mg PI. Fractions were stored at -80°C in 10 mM phosphate buffer, pH 8.0 until further use.

Verification of Hyp and Hzp incorporation levels and maturation. A 30 µL aliquot of PI solution (8 M urea, 20 mM Tris, pH 8) was subjected to cysteine reduction and alkylation (5 mM DTT, 55°C, 20 min; 15 mM iodoacetamide, RT, 15 min, dark) prior to 10-fold dilution into 100 mM NH_4HCO_3, pH 8.0 (100 µL final volume). Peptide digestion was initiated with 0.6 µL of gluC stock solution (reconstituted at 0.5 µg/µL with ddH_2O, Promega) at 37°C for 2.5 h. The reaction was quenched by adding 10 µL of 5% TFA and immediately subjected to C_{18} ZipTip (Millipore) peptide purification and desalting according to the manufacturer's protocol. Peptides were eluted in 50% ACN, 0.1% TFA; the eluent was then diluted three-fold into matrix solution (saturated α-cyanohydroxycinnamic acid in 50% ACN, 0.1% TFA) and analyzed by mass spectrometry (Voyager MALDI-TOF, Applied Biosystems). Hyp and Hzp incorporation levels were analyzed prior to and after refolding; incorporation percentage was calculated by comparing total AUC (area under the curve, arbitrary units) of the non-canonical peak (m/z = 1573 Da for the proinsulin peptide containing B28Hzp or B28Hyp) with total AUC of its wild-type counterpart (m/z = 1557 Da). Incorporation levels stated in Table S1 were obtained from mass spectra of peptides acquired from at least four different expressions of Hzp-PI and Hyp-PI. Maturation of HypI and HzpI was analyzed after HPLC purification. TFA (1.6 µL, 5%) was added to 15 µL mature insulin solution (10 mM phosphate buffer pH 8.0) and subjected to C_{18} ZipTip (Millipore) peptide purification and desalting according to the manufacturer's protocol. MALDI-MS conditions described above were used to confirm insulin maturation.

Insulin receptor (IR) phosphorylation immunoblot. *In vitro* analysis of insulin receptor (IR) phosphorylation was performed using HEK293 cells according to a previous report[1]. Briefly, HEK293 cells were maintained in a 37°C, 5% CO_2 humidified incubator chamber using Dulbecco's modified Eagle's medium with 4.5 g/L glucose, 2 mM L-glutamine and phenol red (DMEM, Life Technologies) supplemented with 10% fetal bovine serum (FBS, Life Technologies), 5% penicillin/streptomycin (P/S, Life Technologies). Every 3 days, at approximately 80% confluency, cells were subcultured and seeded in a 6-well plate at a cell density of $8x10^3$ cells / cm^2 (or $8x10^4$ cells per well) for 24 h prior to insulin addition. Insulins or vehicle were added directly to the medium at 200 nM (10 µL of a 50 µM solution in vehicle PBS) and incubated for 10 min prior to PBS washes to remove excess medium. HEK293 cells were lysed on-plate using IP Lysis Buffer (ThermoFisher, Pierce) with 50 U/mL benzonase nuclease

(Sigma-Aldrich) for 20 min at 4°C; lysates were precipitated using ice cold acetone and re-suspended in 8 M urea, 20 mM Tris, pH 10.0. The protein concentration in the lysate was quantified by the BCA assay (ThermoFisher, Pierce) according to the manufacturer's protocol and normalized for even protein loading across lanes. Lysates were separated by SDS-PAGE (4-12% Novex Bis/Tris SDS-PAGE gels, Life Technologies) in duplicate and transferred to a nitrocellulose membrane (Hybond ECL, GE Healthcare) using a wet transfer system. The membrane was blocked at RT in 5% nonfat milk in Tris-buffer saline with 0.1% Tween 20 (TBS/Tween) and washed with TBS/Tween prior to blotting with antibodies. Primary antibodies for insulin receptor, phosphorylated insulin receptor (from Cell Signaling Technologies) and β-actin (as loading control, from Invitrogen) were added at 1:1000 dilution in TBS/Tween with gentle agitation either at RT for 4 h or overnight at 4°C. Blots were washed and secondary antibodies (Invitrogen) were added at 1:2000 dilution in TBS/Tween. Blots were washed again prior to fluorescence imaging on a Typhoon Trio (GE Healthcare).

Reduction of blood glucose in diabetic animals. NODscid (NOD.CB17-*Prkdc*^{scid}/J) mice were obtained from Jax Mice (Bar Harbor, Maine). Mice were maintained under specific pathogen-free conditions, and experiments were conducted according to procedures approved by the Institutional Animal Care and Use Committee at the City of Hope. Adult (8-12 week old) male NODscid mice were injected intraperitoneally (50 mg/kg/day for 3 consecutive days) with freshly prepared streptozotocin (STZ) in 0.05 M citrate buffer, pH 4.5 to induce diabetes. Diabetes was confirmed 3 weeks after the last dose of STZ by detection of high glucose levels (defined as >200 mg/dL), measured by using a glucomonitor (FreeStyle; Abbott Diabetes Care, Alameda, CA) in blood (10 µL) sampled from the lateral tail vein. Insulin analogs concentrations were determined from A_{280} measurements using a molar extinction coefficient of 6080 M^{-1} cm^{-1} and diluted to 100 µg/mL into a formulation buffer according to a previous report [2]. Insulin analogs in solution were injected subcutaneously at the scruff and blood glucose was measured at the indicated time points.

Hexamer dissociation assay. Insulins were quantified by both UV absorbance (NanoDrop Lite, ThermoFisher) and BCA assay, and normalized to 125 µM insulin prior to dialysis against 50 mM Tris/perchlorate, 25 µM zinc sulfate, pH 8.0 overnight at 4°C using a D-tube dialyzer (Millipore Corp.) with MWCO of 3.5 kDa. Aliquots of dialyzed insulin solution were mixed with phenol to yield samples of the following composition: 100 µM insulin, 20 µM zinc sulfate, 100 mM phenol. Dissociation was initiated by addition of terpyridine (Sigma-Aldrich) to a final concentration of 0.3 mM from a 0.75 mM stock solution. A Varioskan multimode plate reader (Thermo Scientific) was used to monitor absorbance at 334 nm. Kinetic runs were done at least in triplicate, and the data were fit to a mono-exponential function using Origin software. Post assay insulin samples were pooled and sample quality was determined by SDS-PAGE.

Fibrillation Assay. Insulin samples (60 µM in 10 mM phosphate, pH 8.0) were centrifuged at 22 000 g for 1 h immediately after addition of thioflavin T (ThT) (EMD Millipore) to a final concentration of 1 µM. Samples were continuously shaken at 960 rpm on a Varioskan multimode plate reader at 37°C, and fluorescence readings were

recorded every 15 min for 48 h (excitation 444 nm, emission 485 nm). Assays were run in quadruplicate, in volumes of 200 µL in sealed (Perkin-Elmer), black, clear-bottom 96 well plates (Grenier BioOne).

Circular dichroism. Spectra were collected in a 1 cm quartz cuvette at an insulin concentration of 60 µM in 50 mM sodium phosphate buffer pH 8.0. Data were collected from 185 nm to 250 nm, with step size of 0.25 nm and averaging time of 1 s on a Model 410 Aviv Circular Dichroism Spectrophotometer; spectra were averaged over 3 repeat scans. A reference buffer spectrum was subtracted from the sample spectra for conversion to mean residue ellipticity.

Analytical ultracentrifugation. Sedimentation velocity (SV) and sedimentation equilibrium (SE) experiments were carried out on an XL-1 AUC (BeckmanCoulter). SV experiments were conducted with insulin samples dialyzed against 50 mM Tris, 0.1 mM EDTA, pH 8.0, which also served as the reference buffer. Two sector cells with sapphire windows were filled with sample and reference buffer. These cells were centrifuged at 50,000 rpm with absorbance data collected at 280 nm, or for concentrations above 1 mg/mL, 281 nm or 287 nm. SV data were analyzed in SEDFIT with the c(s) algorithm for a continuous distribution[3]. Buffer density and viscosity were calculated from SEDNTERP; the partial specific insulin volume used was 0.735[4]. SE experiments were conducted with insulin samples dialyzed against 50 mM Tris, 0.1 mM EDTA, pH 8.0, which also served as the reference buffer. Two sector cells with sapphire windows were filled with sample and reference buffer and centrifuged at 15,000, 24,000, 36,000 and 50,000 rpm with absorbance data collected at 280 nm. Equilibrium was ascertained by analysis in SEDFIT and non-equilibrated scan speeds were excluded from data analysis. SE and SV data from multiple concentrations were fitted to a monomer-dimer-hexamer reversible self-association model in SEDPHAT with best model chosen by inspection of residuals as well as critical χ value deviation[5]. Radial dependent baselines were computationally determined using TI noise. Figures were generated using GUSSI[6].

Crystallographic studies. Insulin crystals were obtained from sitting drop trays set using a Mosquito robot (TTP Labtech). Drops were set by mixing 0.4 µL insulin solution with 0.4 µL well solution. Well solution conditions were as follows: 462.5 mM sodium citrate, 100 mM HEPES, pH 8.25 for 5HQI; 300 mM Tris, 0.5 mM zinc acetate, 8.5% acetone, 0.5 M sodium citrate pH 8.0 for 5HPR; 300 mM Tris, 17 mM zinc acetate, 1% phenol, 7.5% acetone, 2.675 M sodium citrate pH 8.0 for 5HRQ; 300 mM Tris, 17 mM zinc acetate, 1% phenol, 7.5% acetone, 1.95 M sodium citrate pH 8.0 for 5HPU. Cells were cryoprotected in a mother liquor containing 30% glycerol prior to looping and flash freezing in liquid nitrogen. Data were collected at SSRL beamline BL12-2 using a DECTRIS PILATUS 6M pixel detector. Initial indexing and scaling was performed with XDS; for some structures, data were re-scaled in alternative space groups using Aimless[7]. Initial phases were generated by molecular replacement in PHASER with 3T2A (5HQI and 5HPR) or 1EV3 (5HRQ and 5HPU)[8]. Structure refinement was carried out in Coot and Refmac5[9-10]. Data were deposited in the PDB with the following codes: 5HQI (T_2-HzpI), 5HPR (T_2-HypI), 5HRQ (R_6-HzpI), 5HPU (R_6-HypI).

Transmission electron microscopy. Insulin samples (60 μM in 10 mM phosphate, pH 8.0) were continuously agitated in microfuge tubes at 42°C, 960 RPM for 48 h in an Eppendorf ThermoMixer to obtain fibrils. Samples were stained on 200-mesh copper grids (formar/carbon coated, plasma cleaned) with 1% uranyl acetate. Imaging was done by Dr. Alasdair McDowell at the Beckman Institute Center for Transmission Electron Microscopy on a Tecnai T12 LaB6 120 eV transmission electron microscope.

Figure S1. Insulin expression and incorporation of hydroxyprolines. (**A, B**) SDS-PAGE of cell lysates with lanes labeled for pre-induction (PRE) and post-induction in minimal media supplemented with either nothing (19aa), Hyp (**A**), Hzp (**B**), or Pro at 0.5 mM. (**C-E**)**:** MALDI-MS traces of isolated proinsulin peptide fragment ^{46}RGFFYT**P**KTRRE^{57} obtained by gluC digestion. Peptide fragment masses correspond to either wild type mass (1558 Da) (**C**) or shifted mass (1574 Da) if Hyp (**D**) or Hzp (**E**) is incorporated. Inset is whole protein MALDI-MS. All MALDI-MS spectra contain ion counts >10^3.

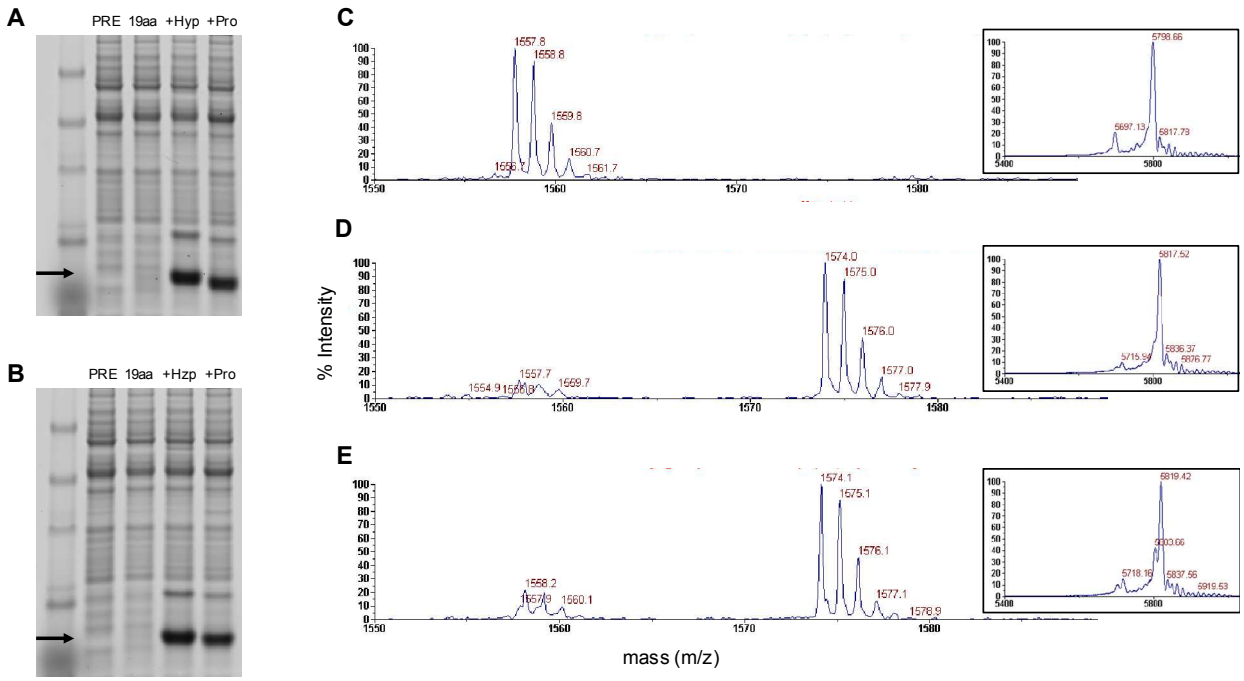
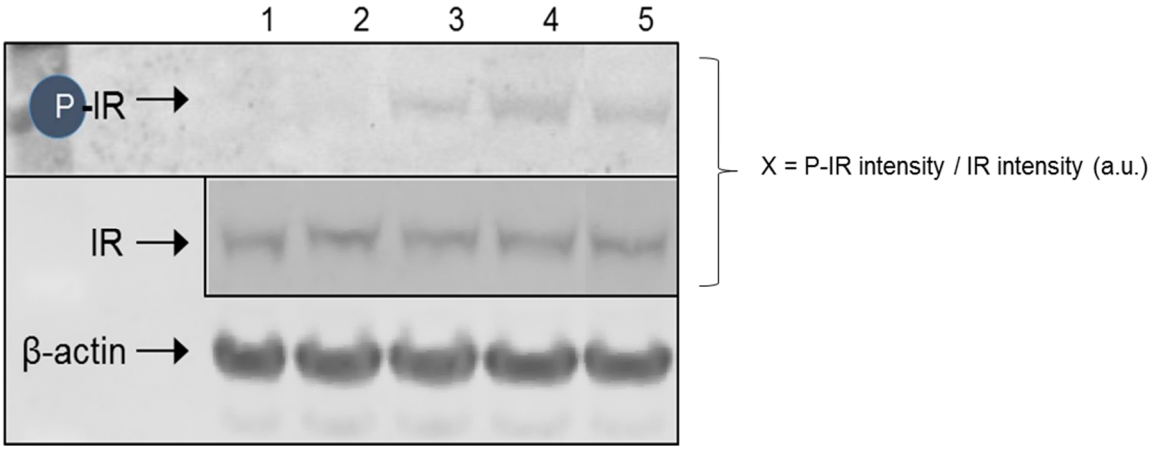


Figure S2. Immunoblot detection of insulin receptor activation. HEK293 cells treated with insulin (200 nM in PBS, pH 7.4) or vehicle. Whole cell lysates were then run on an SDS-PAGE gel and transferred to nitrocellulose membrane to detect insulin receptor (IR) and IR phosphorylation. β-actin immunoblot shown as loading control. Lane 1: Vehicle (PBS); Lane 2: 10% ProI serving as a second negative control due to presence of 10% wt in HzpI and HypI preparations; Lane 3: HzpI; Lane 4: HypI; Lane 5: ProI. Quantification of P-IR and IR bands was done using ImageJ software.



	PBS	10% ProI	HzpI	HypI	ProI
Lane	1	2	3	4	5
X (a.u.)	0.00	0.13	0.77	0.94	0.65
s.d.*	0.0	0.1	0.1	0.1	0.1

*s.d. is calculated from two technical replicates of two biological replicates

Figure S3. Example fits from sedimentation analysis. Insulin samples in 50 mM Tris pH 8.0. (**A**) c(s) curves overlaid for ProI, AspI, HzpI and HypI at 60 µM. (**B, C**) c(s) curves for HzpI (**B**) and HypI (**C**) at indicated concentrations. (**D-F**) Example fits for 60 µM ProI (**D**), 34 µM HzpI (**E**), and 60 µM HypI (**F**), overlaid on top of noise corrected velocity (**D, F**) or equilibrium data (**E**). The SEDPHAT monomer-dimer-hexamer model was utilized over a range of velocity and equilibrium experiments. Global multi-method analysis residuals for the dataset are displayed below each plot.

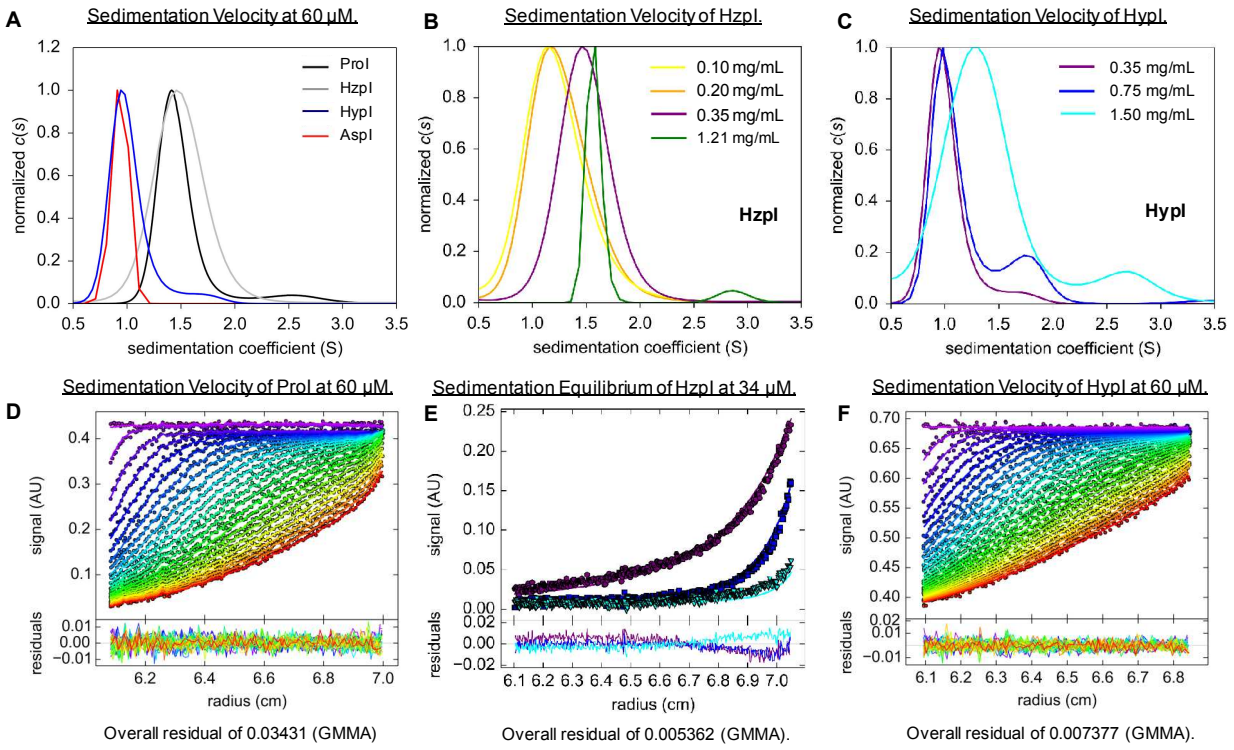


Figure S4. Example fits for analysis of dissociation kinetics. (**A**) Representative dissociation kinetic traces for Zn^{2+} sequestration. Raw data shown and used to fit to a mono-exponential using Origin Software ($y\text{-}y_0 = Ae^{(-t/\tau)}$), where fitted value τ is the characteristic dissociation time constant. (**B, C**) Fitted mono-exponential decay traces for dissociation kinetics, corresponding to (**A**) shown in (**B**). Fitted values for y_0, A used to convert raw data (**A**) to mono-exponential decay representation shown in (**C**). (**D**) Overlay of (**B**) and (**C**) show fitted and raw data to demonstrate quality of fits. *Denotes fitted curves

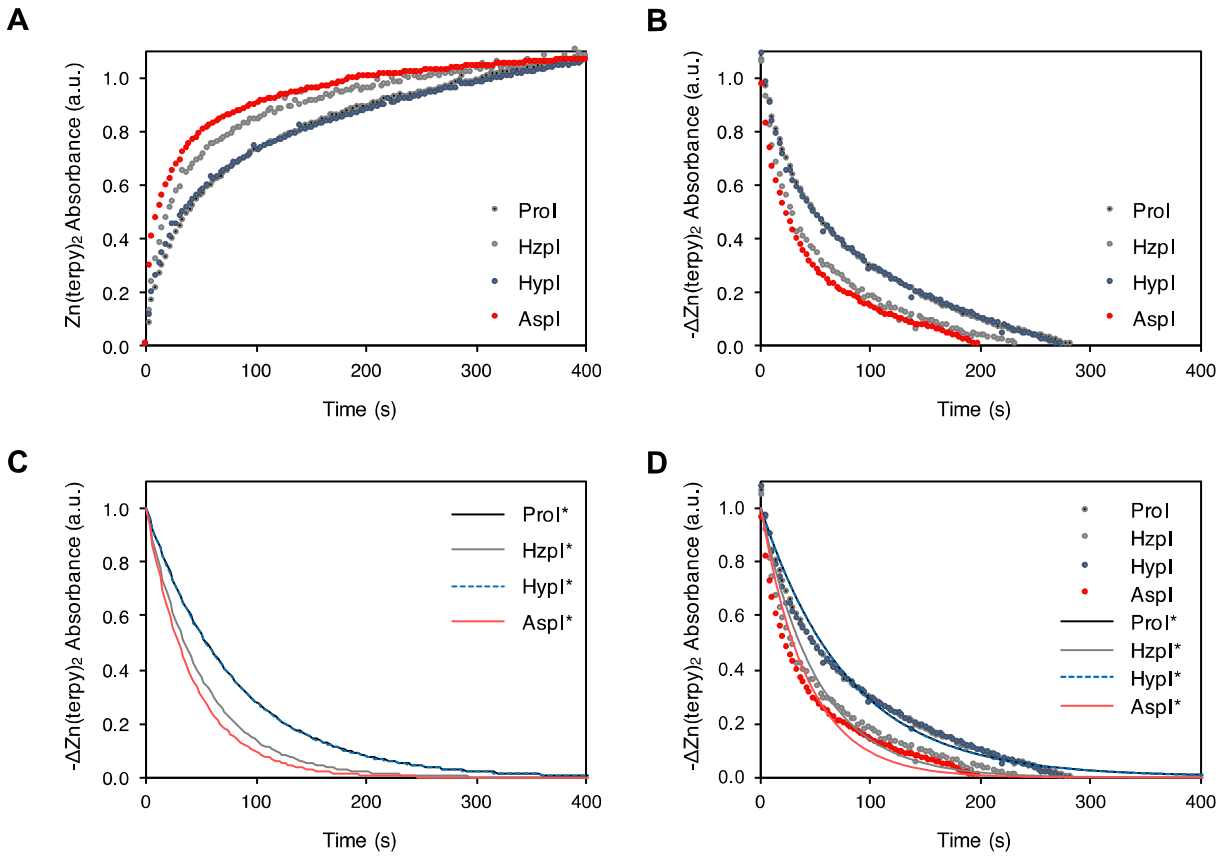


Figure S5. Alignment at position B28. (**A, D**) Alignment of T_2 ProI (tan, PDB:3T2A), and T_2-HzpI (grey) or T_2-HypI (blue) centered on position B28. (**B, E**) Alignment of R_6-ProI (tan) and R_6-HzpI (grey) or R_6-HypI (blue) highlighting the overlap of the backbone at the *C*-terminus. B29 not shown in (**E**) due to lack of electron density. (**C, F**) Alignment of R_6 insulins (ProI, and HzpI or HypI), and AspI (orange, PDB: 1ZEG) centered on position B28 illustrates the similarity of the polypeptide backbones of ProI, HzpI and HypI, and the distinct backbone trajectory of AspI. B29 (**C, F**) and B30 (**A-F**) amino acids not shown for clarify. Arrows denote the N-to-C terminal direction of the backbone originating from carbonyl carbon.

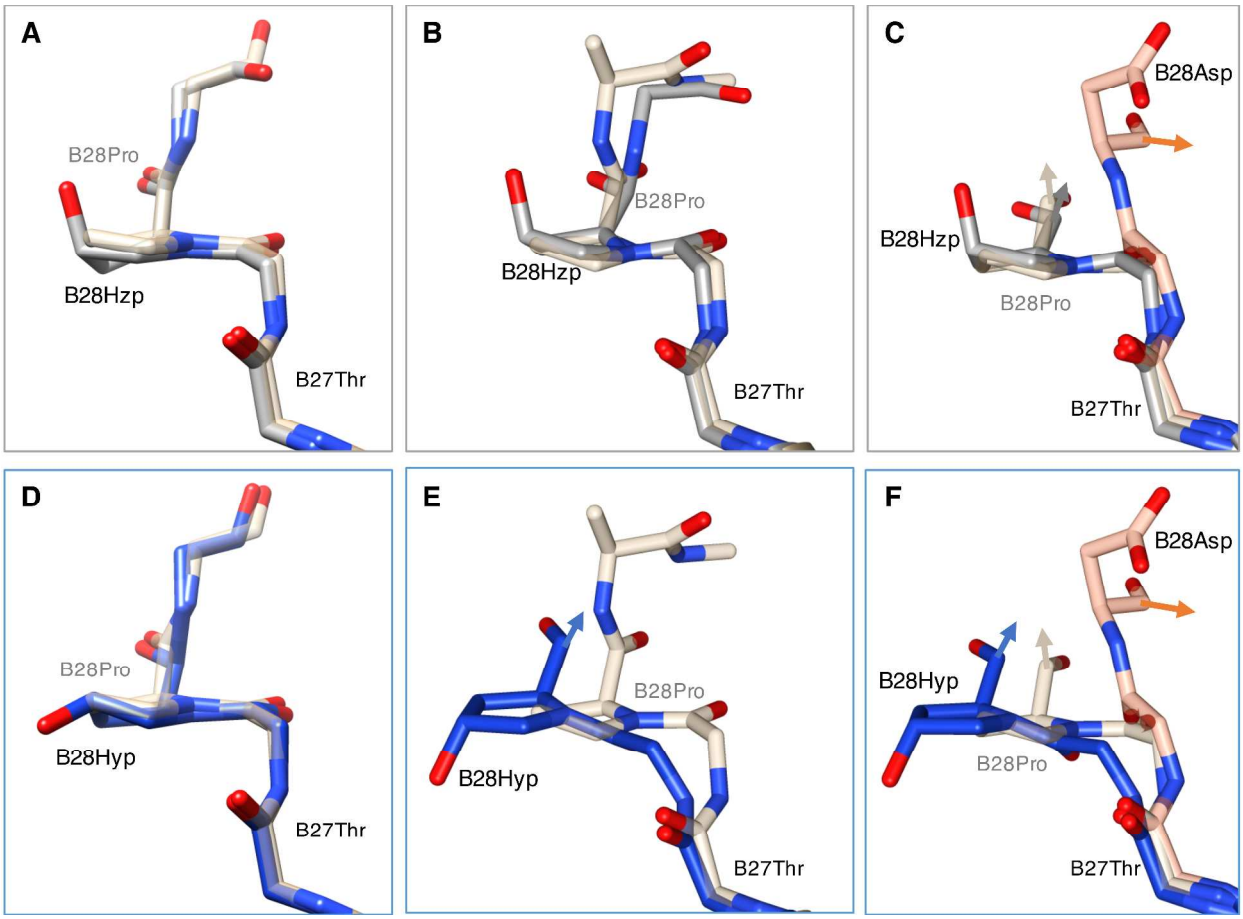


Figure S6. Alignment of R6-AspI and T2-HzpI at position B28. (**A**) R_6-AspI (dark orange; PDB: 1ZEH) does not maintain the backbone trajectory of ProI at position B28. The *C*-terminus of the AspI B-chain is shifted, and a *m*-cresol ligand (light orange) fills the site occupied by B28Pro in ProI. The hydroxyl group of *m*-cresol forms hydrogen bonds with the backbone carbonyl of Glu21′ and a nearby water molecule. (**B**) The same representation of R_6-AspI overlaid with R_6-HzpI (dark grey). Interatomic distances were determined using Chimera. Amino acid B30 is not shown.

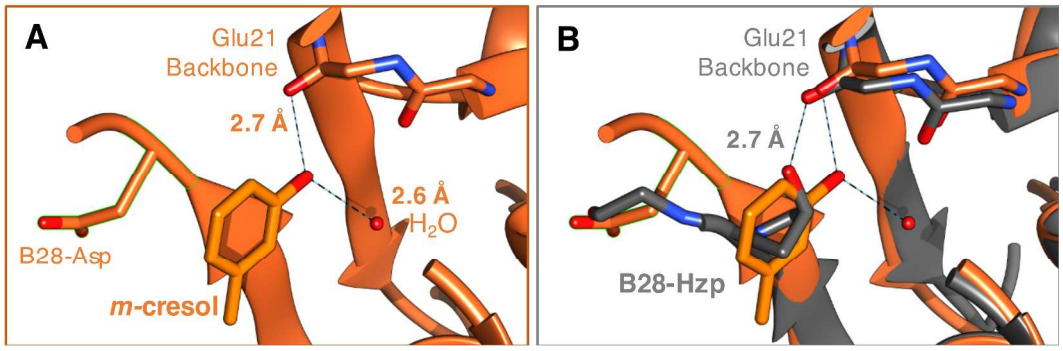


Figure S7. Independent measurement of fibrillation lag time. Representative fibrillation curves for 60 μM insulins (37°C, 960 RPM; n as indicated). Samples were prepared through separate growth, refolding, purification (on a different C18 HPLC column) steps as compared to samples shown in Figure 2C. Insulin fibrils were detected by the rise in Thioflavin T (ThT) fluorescence that accompanies binding to fibrillar aggregates. HzpI did not fibrillate prior to termination of this experiment after 18 h.

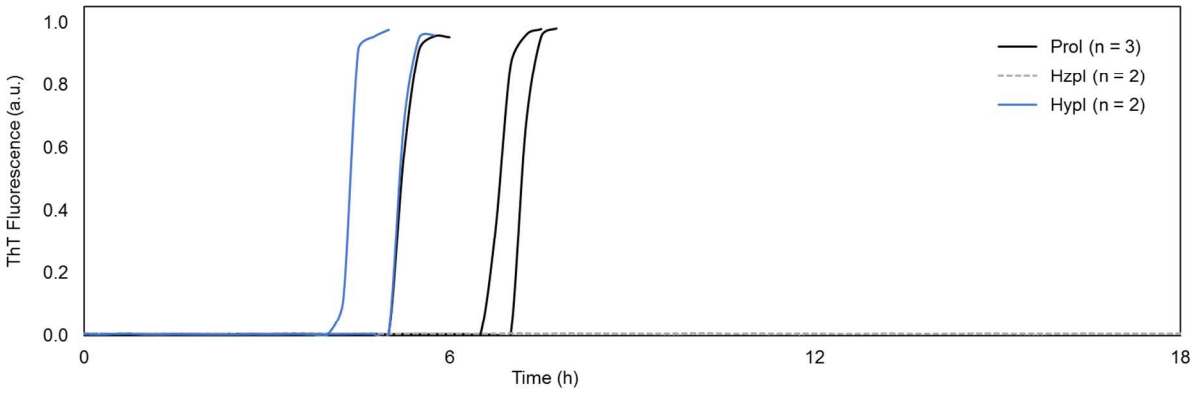


Figure S8. Transmission electron micrographs of insulin fibrils. ProI (**A**), HzpI (**B**), and HypI (**C**). Scale bar 100 nm; staining with 1% uranyl acetate on 200-mesh copper grids.

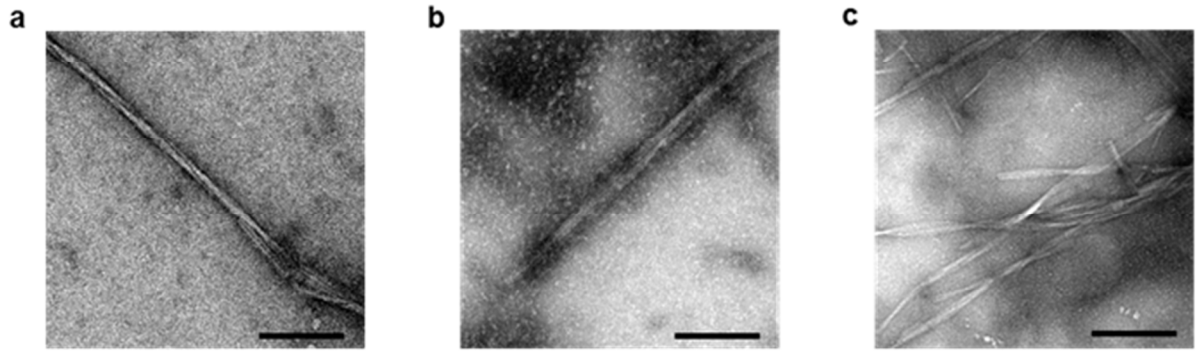


Figure S9. Biophysical characterization of ProI samples prepared in-house and obtained from a commercial source. ProI (Sigma) was purchased from Sigma-Aldrich and purified by HPLC. (**A**) Hexamer dissociation kinetics (**B**) Sedimentation velocity at 60 μM.

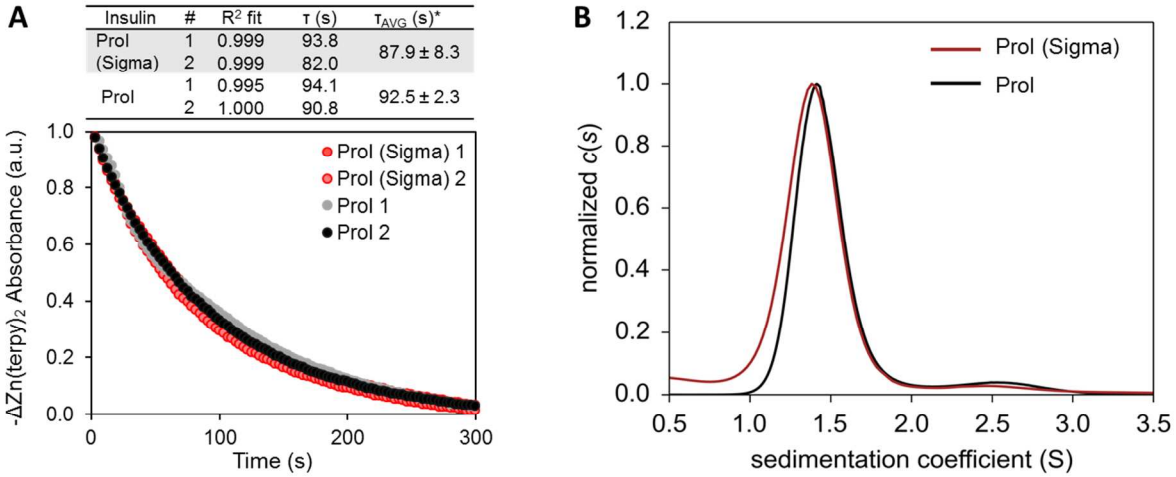


Table S1. Expression Yields and Incorporation Levels of Hydroxyinsulins. [1]Quantified by MALDI-MS from shake flask expressions (n≥4) using proinsulin peptide (^{46}RGFFYTPKTRRE^{57}) obtained by gluC digestion. [2]Yield quantified by BCA assay post-refolding at the proinsulin level

B28 Amino Acid	Incorporation[1]	Yield[2]
L-proline (Pro)	--	50 mg/L
(2S,4S)-4-hydroxy-L-proline (Hzp)	91.3 ± 1.8 %	32 mg/L
(2S,4R)-4-hydroxy-L-proline (Hyp)	88.2 ± 1.2 %	29 mg/L

Table S2. Data Tables and Refinement Values.

	Hypl-T_2 (PDB ID 5HPR)	**Hypl-R_6** (PDB ID 5HPU)	**Hzp-T_2** (PDB ID 5HQI)	**Hzpl-R_6** (PDB ID 5HRQ)
Data collection				
Space group	I 2_1 3	H 3	I 2_1 3	P 1 2_1 1
Cell dimensions				
a, b, c (Å)	78.25, 78.25, 78.25	77.76, 77.76, 40.40	78.23, 78.23, 78.23	47.16, 60.68, 60.66
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 120.00	90.00, 90.00, 90.00	90.00, 110.82, 90.00
Resolution (Å)	55.33-1.33 (1.36-1.33)	38.88-2.20 (2.26-2.20)	55.32-0.97 (0.99-0.97)	56.80-1.28 (1.30-1.28)
R p.i.m. (%)	2.4 (17.1)	4.1 (100.1)	2.6 (173.2)	2.8 (36.5)
CC1/2 (%)	99.9 (90.9)	99.7 (87.8)	99.9 (12.4)	99.8 (66.8)
Mean (I / σ)	17.4 (4.0)	7.4 (0.7)	13.2 (0.4)	13.6 (2.1)
Completeness (%)	99.4 (88.8)	97.0 (87.8)	99.4 (88.2)	96.0 (62.1)
Multiplicity	5.9 (3.3)	3.0 (2.5)	4.7 (2.7)	3.0 (2.2)
Wilson B-factor	13.0	49.0	10.1	14.3
Refinement				
Resolution range (Å)	55.33-1.33 (1.36-1.33)	38.88-2.20 (2.26-2.20)	55.32-0.97 (0.99-0.97)	56.80-1.28 (1.30-1.28)
No. reflections	17298 (1295)	4307 (325)	44555 (3056)	75363 (4122)
Rwork/Rfree (%)	12.5/16.4	16.5/21.6	13.9/15.8	13.5/16.8
No. atoms				

Protein	419	752	414	2451
Ligand/ion	7	18	0	46
Water	60	17	62	160
R.m.s. deviations				
Bond lengths (Å)	0.034	0.006	0.026	0.014
Bond angles (°)	1.386	1.981	1.584	1.259
Ramachandran favored (#)	48	84	49	288
Ramachandran allowed (#)	2	3	1	2
Ramachandran outliers (#)	0	2	0	2

References:

(1) Wharton, J.; Meshulamy, T.; Vallega, G.; Pilch, P. *J. Biol. Chem.* **2005,** *280*, 13483-13486.

(2) Pandyarajan, V.; Smith, B. J.; Phillips, N. B.; Whittaker, L.; Cox, G. P.; Wickramasinghe, N.; Menting, J. G.; Wan, Z. L.; Whittaker, J.; Ismail-Beigi, F.; Lawrence, M. C.; Weiss, M. A. *J. Biol. Chem.* **2014,** *289*, 34709-27.

(3) Schuck, P. *Biophys. J.* **2000,** *78*, 1606-1619.

(4) Laue, T. M.; Shah, B. D.; Ridgeway, T. M.; Pelletier, S. L., *Analytical ultracentrifugation in biochemistry and polymer science*. Royal Society of Chemistry: Cambridge [England], 1992.

(5) Vistica, J.; Dam, J.; Balbo, A.; Yikilmaz, E.; Mariuzza, R. A.; Rouault, T. A.; Schuck, P. *Anal. Biochem* **2004,** *326*, 234-256.

(6) Brautigam, C. A., Chapter five - calculations and publication-quality illustrations for analytical ultracentrifugation data. In *Methods Enzymol.*, James, L. C., Ed. Academic Press: 2015; Vol. Volume 562, pp 109-133.

(7) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2011,** *67*, 235-242.

(8) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J Appl Crystallogr.* **2007,** *40*, 658-674.

(9) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2010,** *66*, 486-501.

(10) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **1997,** *53*, 240-255.