Supporting Information

for

Molecular Recognition of Insulin by a Synthetic Receptor

Jordan M. Chinai[†], Alexander B. Taylor[‡,ᵸ], Lisa M. Ryno[†], Nicholas D. Hargreaves[†], Christopher A. Morris[†], P. John Hart[‡,ᵸ,▾], and Adam R. Urbach[*,†]

[†]*Department of Chemistry, Trinity University, One Trinity Place, San Antonio, TX, 78212*

[‡]*Department of Biochemistry,* ᵸ*the X-ray Crystallography Core Laboratory, and* ▾*the Geriatric Research, Education, and Clinical Center, Department of Veteran's Affairs, South Texas Health Care System, The University of Texas Health Science Center at San Antonio, San Antonio, TX, 78229*

[*] To whom correspondence should be addressed: aurbach@trinity.edu

Contents

Materials. The following commercial compounds of analytical purity grade were used without further purification: H–Phe–Gly–Gly–OH (FGG), H–Gly–Phe–Gly–OH (GFG), H–Gly–Tyr–Gly–OH (GYG) (Chem-Impex International); native human insulin (**1**) recombinant from *Saccharomyces cerevisiae* (SAFC Biosciences), immunoglobulin G (IgG) from human serum, bovine serum albumin (BSA), bovine carbonic anhydrase (BCA) mixture of I and II isoforms, H-Gly-Gly-Gly-OH (GGG), acridine orange, sodium phosphate (mono and dibasic), (Sigma-Aldrich); and cucurbit[7]uril (Q7) was synthesized by the group of Dr. Anthony Day (University of New South Wales, Australia) and purchased from Unisearch. The Glu^{B1}, Glu^{B27} insulin mutant (**2**) was donated by Novo Nordisk. Nuclease-free water from US Biochemicals was used to make all buffers.

A stock solution of 1.0 M sodium phosphate buffer was adjusted to pH 7.0 and sterile-filtered. All ITC and fluorescence spectroscopy experiments were carried out in 10 mM sodium phosphate buffer at pH 7.0, which was made as needed by diluting the 1.0 M stock and adjusting the pH to 7.0. The pH was checked periodically. All analytes were massed to ±0.02 mg with an accuracy of at least three significant digits. The concentration of Q7, accounting for acids and waters of crystallization, was determined by titration with methyl viologen. Concentrations were determined by UV-visible spectroscopy in 10 mM sodium phosphate, using molar absorptivities of 6200 $M^{-1}cm^{-1}$ at 276 nm for the insulins, 57,000 $M^{-1}cm^{-1}$ at 280 nm for BCA, 43824 $M^{-1}cm^{-1}$ at 280 nm for BSA, 210,000 $M^{-1}cm^{-1}$ at 280 nm for IgG, 195 $M^{-1}cm^{-1}$ at 257 nm for FGG and GFG, 1405 $M^{-1}cm^{-1}$ at 274 nm for GYG, and 20,400 $M^{-1}cm^{-1}$ at 257 nm for methyl viologen.

Isothermal Titration Calorimetry (ITC). Titration experiments (Figures 2 and S1) were carried out at 300 K on a VP–ITC calorimeter from Microcal, Inc. For the insulins, the protein was in the sample cell at a concentration of ~0.010 mM, and Q7 was in the injection syringe at a concentration of ~0.150 mM. For the tripeptides, Q7 was in the sample cell at a concentration of 0.1-0.5 mM, and the peptide was in the injection syringe at a concentration of 1-5 mM. The titration schedule consisted of consecutive injections of 2-12 μL with an interval of at least 200 s between injections. Heats of dilution, measured by titrating beyond saturation, were subtracted from each data set. All solutions were degassed prior to titration. The data were analyzed using Origin 7.0 software using the 'one-set-of-sites' binding model, which assumes the complexes do not interact with one another. A complete list of binding data is given below in Table S1.

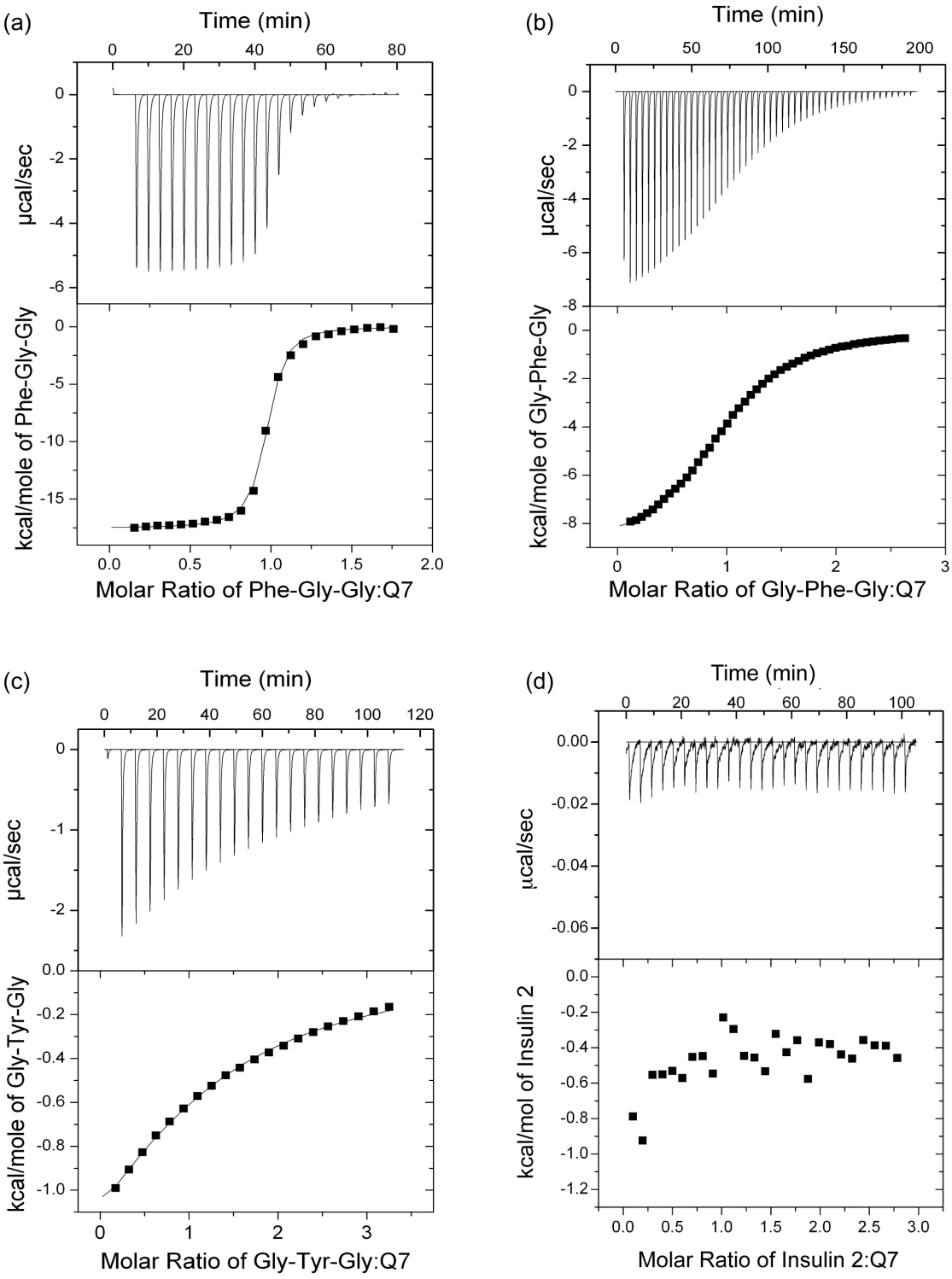

Figure S1. Representative isothermal titration calorimetry experiments at 300 K in 10 mM sodium phosphate, pH 7.0 for the binding of Q7 to (a) Phe-Gly-Gly, (b) Gly-Phe-Gly, (c) Gly-Tyr-Gly, and (d) the $Glu^{B1}Glu^{B27}$ variant insulin **2**.

Table S1. Thermodynamic Binding Data for Q7

Peptide/ Protein	K_a[a] (M^{-1})	ΔG[b] (kcal/mol)	ΔH[c] (kcal/mol)	$-T\Delta S$[d] (kcal/mol)
1	1.5 (±0.4) x 10^6	-8.5 (±0.1)	-10.8 (±0.5)	2.3 (±0.4)
2	< 10^3	nd[e]	nd	nd
3	2.8 (±0.1) x 10^6	-8.9 (±0.1)	-17.5 (±0.1)	8.7 (±0.1)
4	2.2 (±0.1) x 10^4	-6.0 (±0.1)	-9.3 (±0.1)	3.3 (±0.1)
5	2.7 (±0.1) x 10^3	-4.7 (±0.1)	-2.2 (±0.1)	-2.5 (±0.1)

[a] Mean values measured from at least three ITC experiments at 27 °C in 10 mM sodium phosphate, pH 7.0. Standard deviations are given in parentheses. [b] Gibbs free energy values calculated from K_a values. Standard deviations for ΔG values were calculated as the relative error observed in K_a, due to their relationship by a natural logarithm. [c] Enthalpy values measured by ITC. [d] Entropic contributions to ΔG calculated from K_a and ΔH values. [e] not determined.

It is well known that insulin spontaneously forms noncovalent dimers, and in the presence of Zn^{II} these dimers associate into tetramers and hexamers.[S1] Under the conditions of our ITC experiments, we expect some adventitious Zn^{II} in the sample, and thus it is possible that a mixture of monomer, dimer, tetramer, and hexamer exists. However, we observed a single-phase isotherm in the ITC experiment for the binding of Q7 to **1**, with a clear inflection point at a 1:1 molar ratio of Q7:**1** (Figure 2 of the manuscript). The 1:1 ratio shows that there is one molecule of Q7 bound to each molecule of **1** in the sample, and the single-phase isotherm suggests that any potential exchange of **1** between monomer, dimer, tetramer, and hexamer states does not impact the energetics of binding by Q7. In support of this result is the prior observation that the Phe^{B1} position is surface-accessible even when insulin is in the hexameric form.[S2]

Crystallization, Data Collection, Structure Determination and Refinement. Colorless crystals (cubes ~1 mm length) were grown by slow evaporation from a solution containing 40 μM Q7, 40 μM **1**, 10 mM sodium phosphate, 4 mM EDTA, and 1 mM lysine, pH 7.0. The crystals were soaked for 1 min in the same buffer with the addition of 20% ethylene glycol and flash-cooled in liquid nitrogen prior to data collection. Data were collected at the Advanced Photon Source beamline 24-ID-E (Argonne National Laboratory, Argonne, IL) equipped with an ADSC Quantum 315 CCD detector. Data were processed using HKL-2000.[S3] Phases were generated for the complex by the molecular replacement method using PHASER[S4] with the human insulin coordinates in Protein Databank entry 1MSO[S5] as the search model. Coordinates were refined against

the diffraction data using PHENIX,[S6] including simulated annealing, and alternated with manual rebuilding using COOT.[S7] An idealized model for the Q7 molecule was calculated using Macromodel (Schrodinger, Inc.) with the OPLS_2005 force field and continuum aqueous solvent model.

Two conformations were used to refine the structure of the region involving residues Phe^{B1} and Val^{B2} of insulin **1a** (Figure S2). Conformation 1 adopts an orientation in which the aromatic sidechain of Phe^{B1} is modeled inside the cavity of Q7. The electron density for the Phe^{B1} aromatic ring was broadened, consistent with its ability to rotate around the Phe^{B1} C_β-C_γ bond and adopt multiple orientations within the Q7 cavity. In conformation 2, Q7 is rotated approximately 90° with respect to conformation 1, and residues Phe^{B1} and Val^{B2} are not visible in the electron density. Because the adoption of conformation 1 by two adjacent asymmetric units (i.e., 1/1) would lead to steric overlap of the Q7 molecules, it is likely that adjacent asymmetric units in the crystal resemble the configuration of 1/2, as shown in Figure S3. Data collection and refinement statistics are shown in Table S1. The coordinates have been deposited in the Protein Data Bank with accession code 3Q6E.

The structures of the two molecules in the asymmetric unit of the crystal strcuture are analogous to the known T-state conformer[S8,S9] of insulin in that residues B1-B8 adopt an extended, rather than helical, conformation. The two insulin molecules interact along the known monomer•monomer interface, but the resulting dimers do not interact along the known dimer•dimer interface to form tetramers and hexamers, and thus the structure of the complex differs from the known T6 structure.[S2] This result is consistent with the fact that the crystallization buffer contained EDTA, which chelates the Zn^{II} needed to mediate the dimer-dimer interactions.

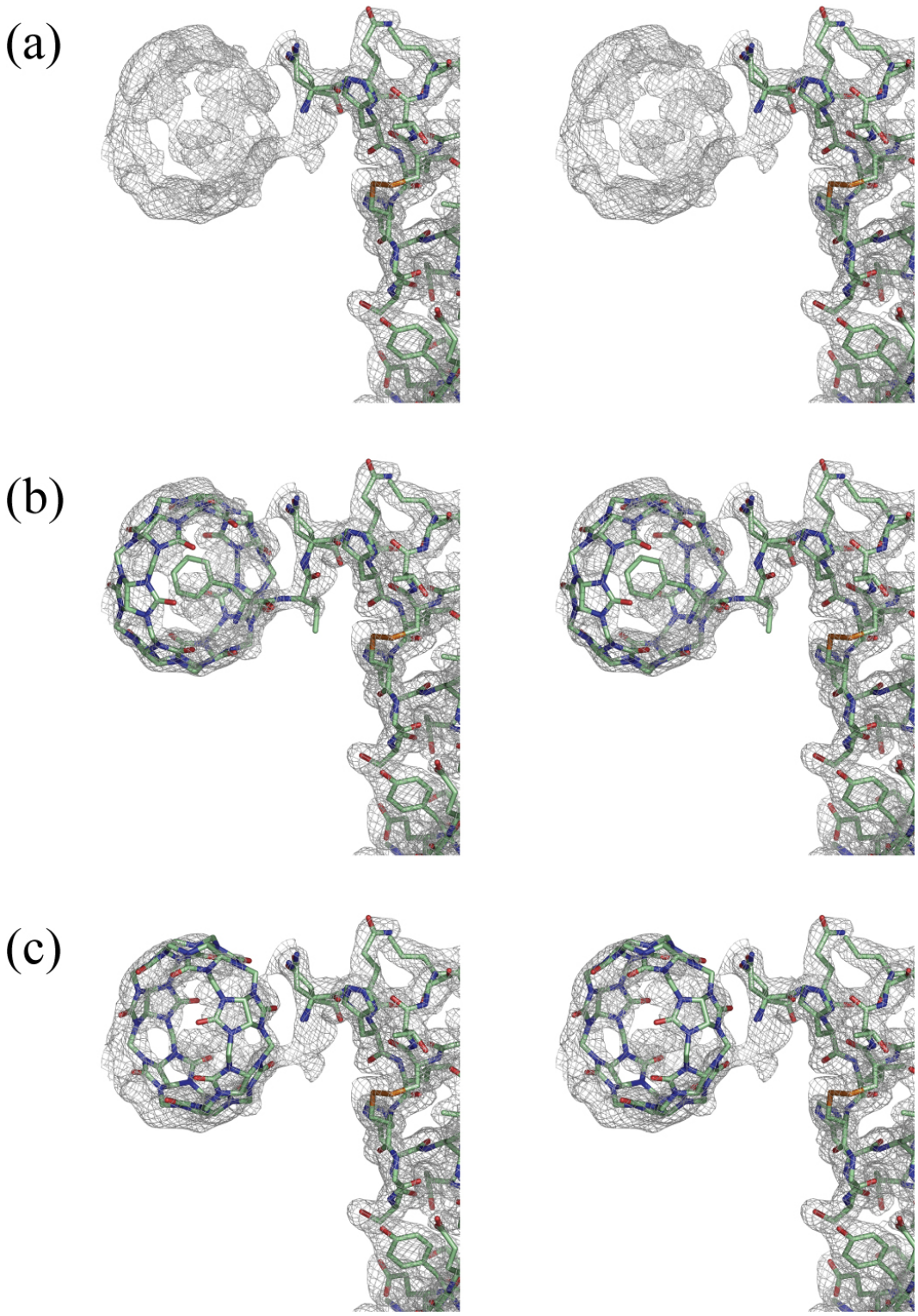

Figure S2. Cross-eyed stereo images of the N-terminal region of the B chain of the insulin•Q7 complex with electron density map shown. The top image (a) shows the density map over the terminal region with no model built in. The middle image (b) shows conformation 1, and the bottom image (c) shows conformation 2, both of which were necessary to accurately model the data.

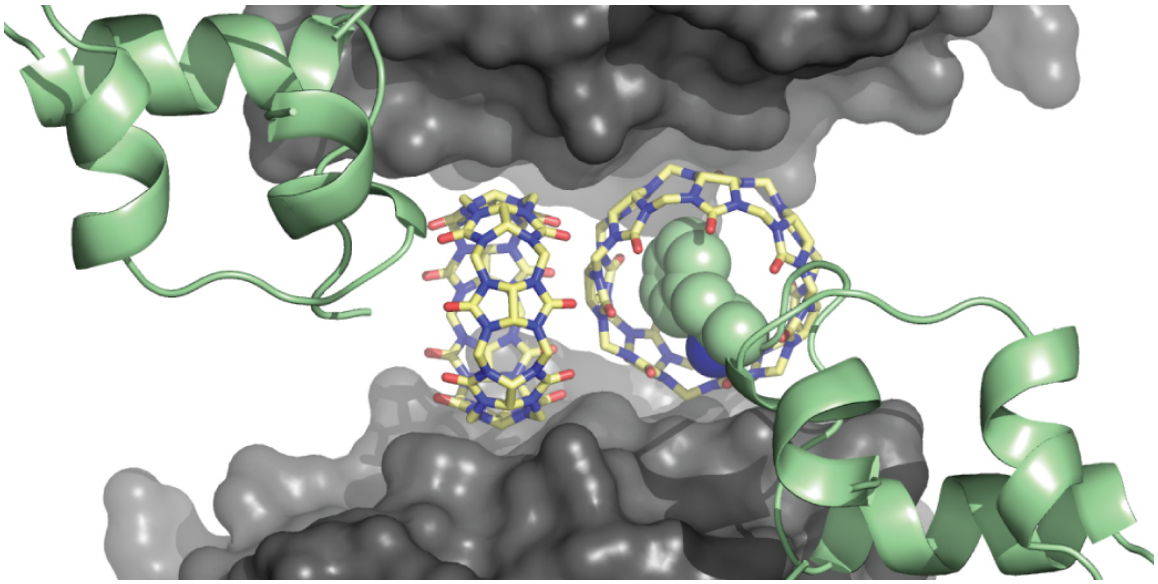

Figure S3. Rendering of the junction between adjacent unit cells at the site of Q7 binding. The unit cell that places Q7 on the right in this figure is rendered in Conformation 1 (Figure S2-b), and the other is in Conformation 2 (Figure S2-c). The insulin that associates with Q7 at its N-terminus (i.e., molecule **1a**) is rendered as green ribbons. The other insulin (i.e., molecule **1b**) in each unit cell is rendered as a black solvent-exposed surface.

Table S2. X-ray data collection and refinement statistics.

Data collection	
Space group	$I4_132$
Cell dimensions	
a, *b*, *c* (Å)	116.5, 116.5, 116.5
α,β,γ (°)	90, 90, 90
Wavelength (Å)	0.97947
Resolution (Å)	30 - 2.05
R_{sym}[a]	0.072 (0.521)
I / sI	35.7 (5.5)
Completeness (%)	99.9 (100)
Redundancy	11.6 (11.9)

Refinement	
Resolution (Å)	27.5 – 2.05
No. reflections	8766
R_{work} / R_{free}	0.206/0.253
No. atoms	
Protein	820
Ligand	168
Solvent	59
R.m.s. deviations[b]	
Bond lengths (Å)	0.007
Bond angles (°)	1.078

[a] Values in parentheses are for the highest resolution shell.
[b] R.m.s. deviations are from idealized values for protein.

Fluorescence Spectroscopy. Fluorescence emission spectra, 492-700 nm, were obtained at 25 °C with a PTI QM-4 spectrofluorometer equipped with a Xe arc lamp and digital photomultiplier tube, exciting at 485 nm, with a step size of 1 nm. Relative fluorescence intensities were determined by the peak height at maximum intensity.

Insulin concentration series (Figure 4a): Samples containing 8 μM acridine orange, 7 μM Q7, and 0, 0.5, 1, 2, 4, 8, and 16 μM insulin **1** were prepared by combining stock solutions containing (i) 8 μM acridine orange and 6 μM Q7 with (ii) 8 μM acridine orange, 6 μM Q7, and 30 μM insulin **1** in the appropriate ratios. Samples were equilibrated overnight in the absence of light. Acridine orange adsorbs nonspecifically to the surfaces of the pipet tips and the fluorescence cuvette. Therefore, the pipet tips and the fluorescence cuvette were soaked overnight in a solution containing 8 μM acridine orange and 6 μM Q7 and rinsed for 30 s in this solution between spectral acquisitions. The slit widths for excitation and emission monochromators were 3 nm.

Selectivity series (Figure 4b): Samples containing 6 μM acridine orange, 4 μM Q7, and 4 μM protein or peptide were prepared. Control samples containing 6 μM acridine orange and 4 μM protein or peptide were prepared. All samples were equilibrated overnight in the absence of light. As discussed in the previous paragraph, the pipet tips and the fluorescence cuvette were soaked overnight in a solution containing 6 μM acridine orange and 4 μM Q7 and rinsed for 30 s in this solution between spectral acquisitions. For measurements in the absence of Q7, the pipet tips and fluorescence cuvette were soaked overnight in a solution containing 6 μM acridine orange and rinsed for 30 s in 6 μM acridine orange between spectral acquisitions. The slit widths for excitation and emission monochromators were 4 nm.

Circular Dichroism Spectroscopy. Spectra for native insulin **1** and mutant insulin **2** were acquired from 200-300 nm on a Jasco J-815 circular dichroism spectropolarimeter at 25 °C with a 0.1 cm path length, 0.5 nm pitch, 3 nm bandwidth, 4 s response time, 100 nm/min scan rate. The samples were at 20 μM protein in a buffer containing 10 mM sodium phosphate, 4 mM EDTA, and 1 mM lysine, pH 7.0.

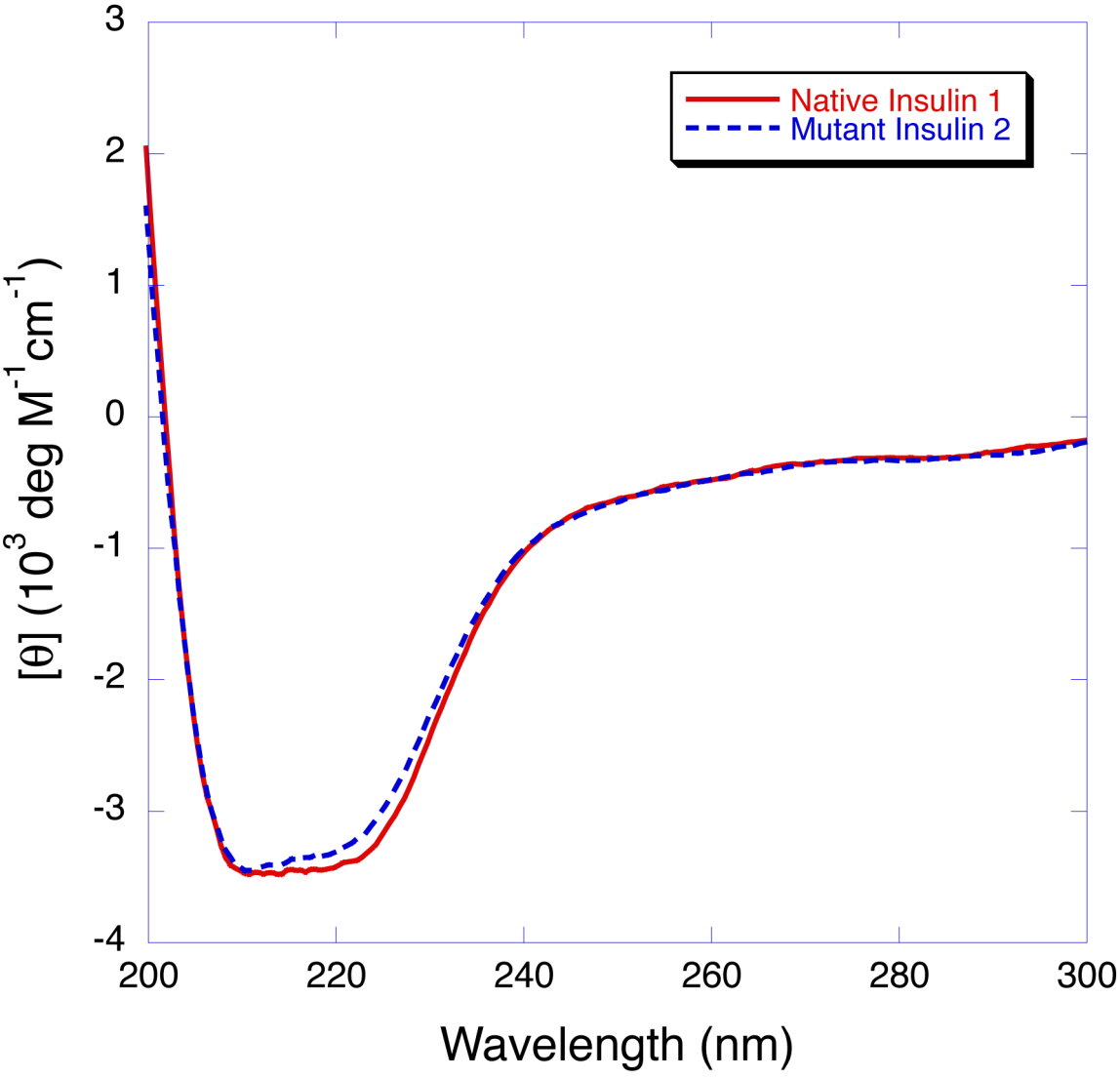

Figure S4. Overlaid circular dichroism spectra of native insulin **1** (red solid line) and mutant insulin **2** (blue dashed line). The raw ellipticity data were divided by the molar concentration (20 μM) and the path length (0.1 cm) to derive data in units of 10^3 deg M^{-1} cm^{-1}. The strong similarity in these spectra suggests that the proteins have very similar secondary and tertiary structures.

References.

S1. *Insulin and Related Proteins – Structure to Function and Pharmacology*, M. Federwisch, M. L. Dieken, P. De Meyts, Eds. (Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002).

S2. G. D. Smith, W. A. Pangborn, R. H. Blessing, *Acta Cryst. Sect. D* **59**, 474 (2003).

S3. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, *Meth. Enzymol.* ***276***, 307-326.

S4. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L.C. and Read, R. J. (2007). *Phaser* crystallographic software, *J. Appl. Cryst.* **40**, 658-674. <http://scripts.iucr.org/cgi-bin/paper?he5368>

S5. Smith, G.D., Pangborn, W.A., and Blessing, R.H. (2003) The structure of T6 human insulin at 1.0 Å resolution. *Acta Crystallogr.* **D59**, 474-482.

S6. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination, *Acta Crystallogr.* **D58**, 1948-1954.

S7. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr.* **D60**, 2126-2132.

S8. N. C. Kaarscholm, H.-C. Ko, M. F. Dunn, *Biochem.* **28**, 4427 (1989).

S9. I. Pittman IV, H. S. Tager, *Biochem.* **34**, 10578 (1995).