Structure-based stabilization of insulin as a therapeutic protein assembly via enhanced aromatic–aromatic interactions

Received for publication, April 24, 2018, and in revised form, May 30, 2018 Published, Papers in Press, June 7, 2018, DOI 10.1074/jbc.RA118.003650

Nischay K. Rege[‡1], Nalinda P. Wickramasinghe[‡], Alisar N. Tustan[§], Nelson F. B. Phillips[‡2], Vivien C. Yee[‡], Faramarz Ismail-Beigi[§], and Michael A. Weiss[‡¶3]

From the Departments of [‡]Biochemistry and [§]Medicine, Case Western Reserve University, Cleveland, Ohio 44106 and the [¶]Department of Biochemistry, Indiana University School of Medicine, Indianapolis, Indiana 46202

Edited by Jeffrey E. Pessin

Key contributions to protein structure and stability are provided by weakly polar interactions, which arise from asymmetric electronic distributions within amino acids and peptide bonds. Of particular interest are aromatic side chains whose directional π-systems commonly stabilize protein interiors and interfaces. Here, we consider aromatic–aromatic interactions within a model protein assembly: the dimer interface of insulin. Semiclassical simulations of aromatic–aromatic interactions at this interface suggested that substitution of residue Tyr^{B26} by Trp would preserve native structure while enhancing dimerization (and hence hexamer stability). The crystal structure of a [Trp^{B26}]insulin analog (determined as a $T_3R^f_3$ zinc hexamer at a resolution of 2.25 Å) was observed to be essentially identical to that of WT insulin. Remarkably and yet in general accordance with theoretical expectations, spectroscopic studies demonstrated a 150-fold increase in the *in vitro* lifetime of the variant hexamer, a critical pharmacokinetic parameter influencing design of long-acting formulations. Functional studies in diabetic rats indeed revealed prolonged action following subcutaneous injection. The potency of the Trp^{B26}-modified analog was equal to or greater than an unmodified control. Thus, exploiting a general quantum-chemical feature of protein structure and stability, our results exemplify a mechanism-based approach to the optimization of a therapeutic protein assembly.

Weakly polar interactions are ubiquitous among protein structures (1). Among such interactions, the relative packing of aromatic rings is of particular interest in relation to the organization of protein cores and subunit interfaces (2). Aromatic–aromatic interactions are governed by quantum-chemical properties, which underlie dispersion forces and give rise to asymmetric distribution of partial charges. Whereas aromatic stacking is prominent in nucleic acid structures, pairs of aromatic side chains in proteins more often exhibit edge-to-face (ETF)[4] contacts (3). Can such contacts be exploited in therapeutic protein engineering? Here, we have analyzed aromatic–aromatic interactions in the insulin hexamer (4) as a basis for designing improved long-acting (basal) analogs. This class of analogs is central to the treatment of Type 1 and Type 2 diabetes mellitus (5).

Classical crystal structures of insulin hexamers (6, 7) immediately suggested a pathway of assembly (4). Pertinent to the mechanism of storage in the secretory granules of pancreatic β-cells (8), such assembly is also of pharmacologic importance (9). Insulin assembly both protects the hormone from degradation in pharmaceutical formulations and modulates its pharmacokinetic properties (10). Indeed, the first use of insulin analogs in diabetes therapy reflected efforts to *destabilize* the insulin hexamer and thereby accelerate absorption of monomers and dimers from the subcutaneous (SQ) depot (summarized in Fig. S1 (11, 12)). Because it is more straightforward to introduce unfavorable substitutions than favorable ones, such engineering (corresponding to rapid-acting analogs) was more successful than complementary efforts to *enhance* the thermodynamic (and kinetic) stability of the insulin hexamer (13). There are presently three rapid-acting insulin analogs in clinical use (14), but no basal products designed on the basis of enhanced hexamer assembly despite extensive efforts (15). These difficulties were circumvented by alternative mechanisms of protracted action (acylation and pH-dependent SQ precipitation (14); Fig. S2).

The dimer interface of insulin (repeated three times in the hexamer) contains a cluster of eight conserved aromatic rings

This work was supported in part by National Institutes of Health Grant R01 DK040949 (to M. A. W.). M. A. W. has equity in Thermalin, Inc. (Cleveland, OH) where he serves as Chief Innovation Officer; he has also been a consultant to Merck Research Laboratories and DEKA Research & Development Corp. N. F. B. P. is a consultant to Thermalin, Inc. F. I.-B. serves has equity in Thermalin, Inc. and is a consultant to Sanofi and Novo Nordisk. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

We dedicate this article to the memory of the late Prof. Colin Ward (Eliza and Walter Hall Institute, Melbourne, Australia).

This article contains Figs. S1–S12 and Tables S1–S13.

The atomic coordinates and structure factors (code 2CK2) have been deposited in the Protein Data Bank (http://wwpdb.org/).

[1] Predoctoral Fellow of the National Institutes of Health, Medical Scientist Training Program, supported by Grant 5T32GM007250-38 and Fellowship 1F30DK112644.

[2] Supported in part by the American Diabetes Association Grants 7-13-IN-31 and 1-08-RA-149.

[3] To whom correspondence should be addressed. Tel.: 317-274-7151; E-mail: weissma@iu.edu.

[4] The abbreviations used are: ETF, edge-to-face; αCT, carboxyl-terminal domain of IR; μIR, domain-minimized insulin micro-receptor; EEC, entropy-enthalpy compensation; IR, insulin receptor; L1, leucine-rich domain of IR MD, molecular dynamics; MM, molecular mechanics; PD, pharmacodynamics; PDB, Protein Data Bank; r.m.s. deviation, root mean square deviation; SQ, subcutaneous; QM, quantum-mechanical; SEC, size exclusion chromatography; IGF, insulin-like growth factor; TOCSY, total correlated spectroscopy; i.v., intravenous; Orn, ornithine; WT insulin pertains to the human sequence unless otherwise stated.

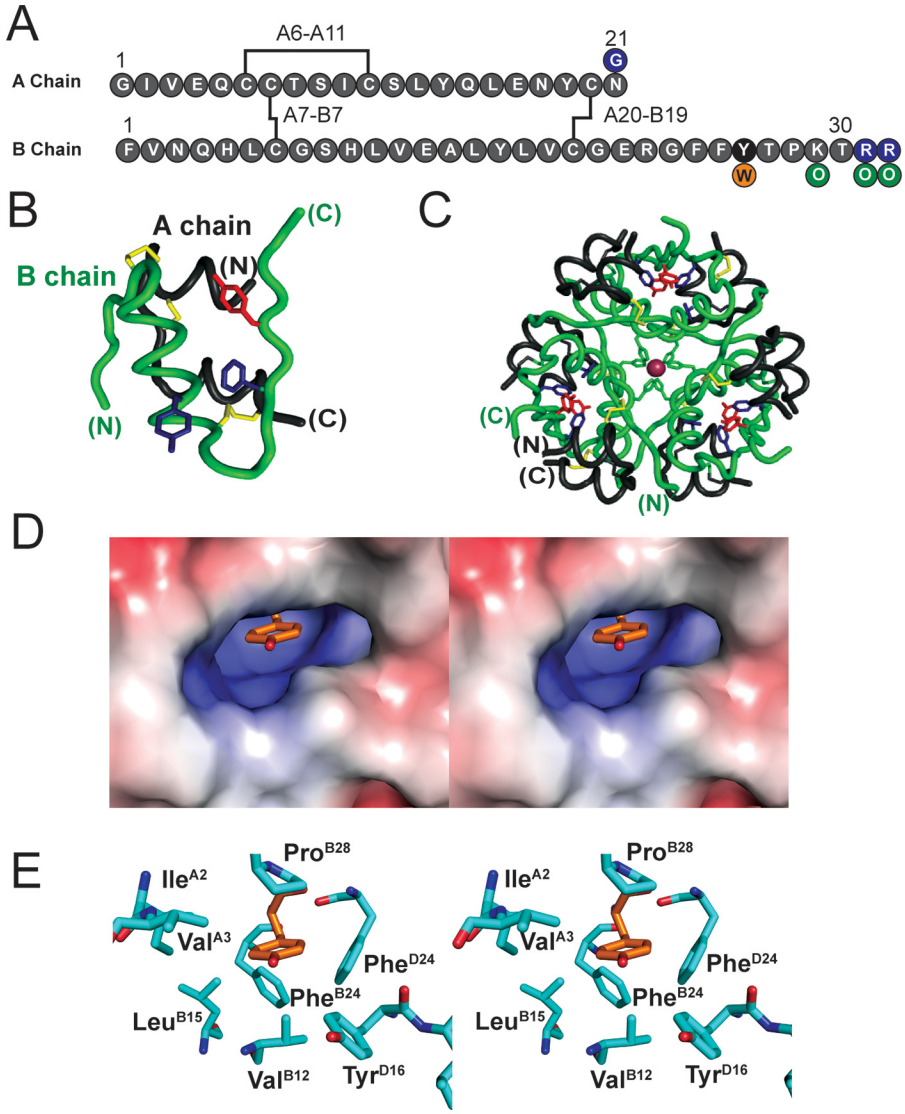

Figure 1. Structural overview of insulin. *A,* sequence of insulin with disulfide bridges in *black*. Tyr^{B26} is highlighted in *black*; the present Tyr^{B26} → Trp substitution is *orange*; and modifications in pI-shifted clinical analog *glargine* in *purple*. Our semisynthetic pI-shifted analog contained Orn (*green*) instead of Lys or Arg. *B,* structure of an insulin monomer; the A chain is shown in *black* and B chain in *green*. Tyr^{B26} is *red*, whereas Phe^{B24} and Tyr^{B16} are *blue* (PDB code 4INS). *C,* structure of zinc-coordinated insulin hexamer (T_6 state), a trimer of dimers; Tyr^{B26}, Phe^{B24}, and Tyr^{B16} are color-coded as in *B*. *D,* stereo view showing Tyr^{B26} (*sticks*) in a cavity within insulin dimer (extracted from $T_3R^f_3$ hexamer, PDB code 1TRZ). *E,* corresponding stick model with residues labeled.

(Tyr^{B16}, Phe^{B24}, Phe^{B25}, Tyr^{B26}, and their dimer-related mates in subunit D; Fig. 1, *A–C*). Of these, successive ETF contacts are formed by B16–D26, B24–B26, B24–D24, and B26–D16; the B25 side chain is peripheral to this network. Whereas variation at these sites is in general constrained by the structure of the hormone–receptor interface (16, 17), our attention focused on the B26 side chain because of its functional tolerance to diverse substitutions (18) and because of its partial exposure in the monomer, dimer, and hexamer (4) (Fig. 1, *C* and *D*). We sought to investigate whether variant B16–D26 and B26–D16 ETF contacts across the dimer interface might in principle modulate, in either direction, the strength of these interactions. We hypothesized that enhanced ETF contacts at this interface might provide the long-sought approach to stabilize the insulin hexamer and so improve the pharmacokinetc properties of basal formulations.

The present study had three parts. The first employed local modeling, using the standard CHARMM empirical energy function, to probe possible effects of a Tyr^{B26} → Trp substitution on aromatic–aromatic interactions within the wildtype (WT) aromatic cluster. These molecular mechanics (MM) calculations suggested that substitution of Tyr^{B26} by Trp could enhance dimer-related ETF contacts and yet otherwise preserve a native-like interface. We next prepared this analog to examine whether this substitution might indeed stabilize the insulin hexamer and retard its disassembly while preserving the biological activity of the monomeric hormone. The crystal structure of a [Trp^{B26}]insulin analog (as a zinc–insulin hexamer) was essentially identical to that of WT insulin. Finally, we undertook studies in diabetic rats to obtain proof of principle that this approach could extend the duration of insulin action on SQ injection.

To our knowledge, our results represent the first exploitation of aromatic–aromatic interactions to enhance the physical and biological properties of a therapeutic protein. Because standard MM calculations employ a simplified model of aromatic systems (*i.e.* approximating their quantum-mechanical (QM) properties via partial atomic charges (19, 20)), *ab initio* QM simulations of the aromatic cluster and their incorporation in QM/MM simulations (21) promise to establish a rigorous foundation for therapeutic protein design, including further optimization through incorporation of modified or nonstandard amino acids (22–24). The present results suggest that insulin's conserved aromatic cluster can provide a natural laboratory for such foundational analysis and its therapeutic translation.

Results

Molecular mechanics calculations suggested that augmented aromatic–aromatic interactions are possible at the Trp^{B26} dimer interface

MM simulations were employed to estimate the strength of aromatic–aromatic interactions of Trp^{B26} at the insulin dimer interface in relationship to those of the native Tyr. These calculations employed the CHARMM empirical energy function in which aromatic rings contain partial atomic charges, parametrized to mimic the electrostatic properties of the π-system (19, 20). Working models of the variant dimer were obtained by local energy minimization (see "Experimental procedures").

Energies of interaction between the eight aromatic residues at the insulin– dimer interface (Tyr^{B16}, Phe^{B24}, Phe^{B25}, Tyr^{B26}, and their symmetry-related mates) were calculated using local models in which Trp^{B26} was substituted within WT T_2, R_2, and TR^f reference dimers (extracted from PDB structures 4INS, 1ZNJ, and 1TRZ) (25). The total interaction energy between B26 and the other aromatic residues at the Trp^{B26} interface, which was calculated using the full CHARMM potential energy function, was augmented by 2.0 and 0.8 kcal/mol relative to the minimized WT interface in the context of R_2 and TR^f structures, respectively, and diminished by 1.5 kcal/mol in the context of the T_2 structure. Results are summarized in Table S1.

A minimal model was utilized to further evaluate the potential impact of a Trp^{B26} substitution on aromatic–aromatic interactions at the dimer interface. To this end, a structural model of the dimer interface (extracted from a T_6 insulin hexamer; PDB code 4INS) was first built containing residue B26 (Tyr^{B26} or Trp^{B26}) and its nearest aromatic neighbors (Phe^{B24}, Phe^{D24}, or Tyr^{D16}). Simulations predicted the orientation of the B26 ring corresponding to free-energy minima of electrostatic interactions between the aromatic residues (the partial-charge parametrization of aromatic residues in CHARMM is shown in Fig. S3) (23). When substituted at position B26, Trp displayed improved electrostatic interactions with its three aromatic neighbors (relative to the WT Tyr) over a broad range of conformations (Fig. 2). This trend extended to conformations that are sterically permitted in the context of the WT insulin hexamer.

Trp^{B26} analog exhibited markedly decreased hexamer dissociation rate

The effect of the Trp^{B26} substitution on the lifetime of insulin hexamers under formulation conditions was assessed in the context of [Orn^{B29}]insulin, a structural equivalent of WT insulin that is amenable to production by trypsin-catalyzed semisynthesis (26). The lifetime of Co^{2+}-substituted, phenol-stabilized (R_6) hexamers of [Trp^{B26},Orn^{B29}]insulin was assessed at equilibrium in relationship to native [Tyr^{B26},Orn^{B29}]insulin. Optical absorbance spectra of these analogs (characteristic of Co^{2+} with tetrahedral coordination) were similar to WT (Fig. 3, *A* and *B*). Assessment of R_6 dissociation rates (summarized in Table 1) revealed a 150-fold increase in hexamer half-life of the Trp^{B26} analog relative to its parent [Orn^{B29}]insulin (Fig. 3, *C* and *D*). This increase is remarkable given that the difference between the half-lives of a rapidly dissociating analog in clinical use (lispro)[5] (27–29) and WT is <2-fold (Table 1).

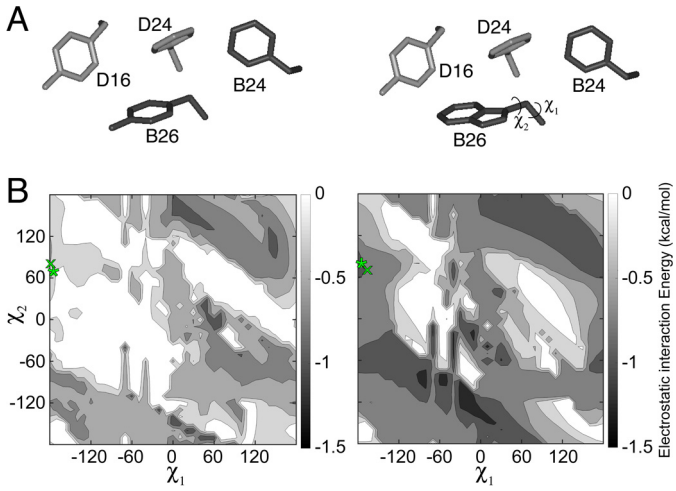

Figure 2. Molecular simulations of aromatic interactions in the insulin dimer. *A,* aromatic–aromatic interactions across insulin's dimer interface involve Phe^{B24}, Tyr^{B16}, Phe^{D24} (*sticks*) and either Tyr^{D26} (*left*) or Trp^{D26} (*right*). Residues were extracted from T_6 structure 4INS. *B,* contour maps depicting empirical interaction energies between B26 (Tyr on *left* and Trp on *right*) at varying χ_1 and χ_2 angles and the other three residues shown in *A*. The orientation of Tyr^{B26} or Trp^{B26} in the WT or variant crystal structure is indicated by a *green "x"*; orientation of Tyr^{B26} or Trp^{B26} in the local model is indicated by a *green asterisk*.

Trp^{B26} imposed kinetic barriers to the dissociation of hexamers into monomers

R_6 dissociation kinetics were further examined by size exclusion chromatography (SEC). The Trp^{B26} analog (formulated in the presence of phenol and zinc ions) was injected onto an SEC column using a zinc- and phenol-free mobile phase. Subsequent dissociation of the R_6 hexamers was monitored in the chromatograms (Fig. 4*A*, Table 1). Absence of a void volume signal (V_0) indicated that none of the proteins formed large nonspecific aggregates. Whereas WT and [Orn^{B29}]insulin eluted as a broad peak representing an association state intermediate between monomer and dimer (9.7 and 8.2 kDa, respectively) and insulin lispro eluted essentially as a monomer (5.1 kDa), [Trp^{B26},Orn^{B29}]insulin eluted in two distinct peaks. The larger peak corresponded to a trimeric or tetrameric association state (molecular mass 28 kDa), and the smaller corresponded to a monomeric state (4 kDa; Fig. 4, *A* and *B*; see Fig. S4

[5] Insulin lispro, containing substitutions Pro^{B28} → Lys and Lys^{B29} → Pro, exhibited an attenuated 574-nm band (69); this is associated with a greater proportion of spectroscopically attenuated $T_3R^f_3$ hexamer and invisible T_6 hexamer (Fig. 3*B*).

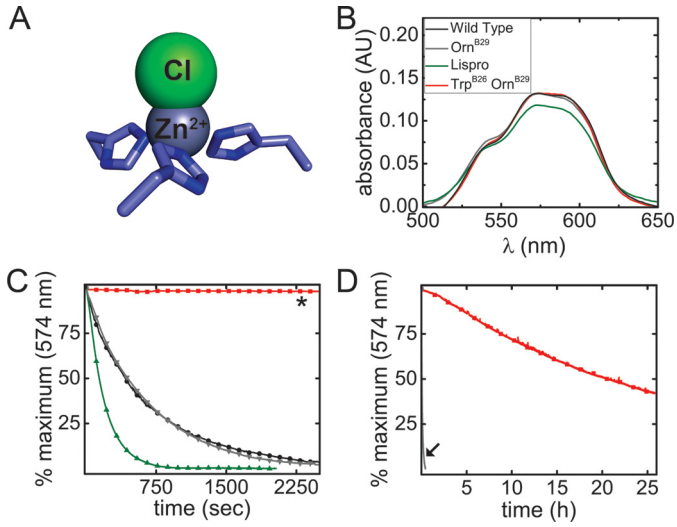

Figure 3. Hexamer dissociation of Trp^{B26} analog. *A,* tetrahedral Zn^{2+}-coordination site in R_6 insulin hexamer (ball-and-stick model): three His^{B10} side chains and one chloride ion. *B,* absorbance spectra of d–d bands in corresponding Co^{2+} complex; the color code is indicated. *C,* hexamer dissociation curves as monitored at 574 nm after addition of excess EDTA; the color code is as in *B*. The lifetime of the [Trp^{B26},Orn^{B29}]insulin hexamer was markedly prolonged (*asterisk*). *D,* dissociation of Trp^{B26},Orn^{B29} hexamer from 0 to 8000 s in relationship to that of parent [Orn^{B29}]insulin (*black arrow*). Half-lives are given in Table 1.

Table 1
Self-association properties of insulin analogs

Analog	$t_{½}$ hexamer dissociation	Calculated molecular mass by SEC[a]
	min ± S.D.	*kDa*
Wildtype	7.7 (± 1.3)	9.7
Lispro[b]	4.6 (± 0.3)	5.1
Orn^{B29}[c]	8.2 (± 0.8)	8.2
Trp^{B26}, Orn^{B29}	1.2 (± 0.3) × 10^3	28.0, 4.0

[a] Proteins were made 0.6 mM in a buffer containing $ZnCl_2$ at a ratio of 2 zinc ions per insulin hexamer and applied to SEC column as described under "Experimental procedures." Masses were calculated from the plot in Fig. 4*B*.
[b] "Lispro" describes insulin analogs containing Pro^{B28} →Lys and Lys^{B29} →Pro substitutions. These substitutions impair dimerization (28, 29).
[c] Use of Orn simplified trypsin-catalyzed semisynthesis (33).

for reference chromatograms of the monomer and hexamer). These findings suggest that oligomers intermediate between hexamer and dimer may delay dissociation of the Trp^{B26} analog; in particular, the absence of broad elution tails implies that Trp^{B26} imposes significant barriers to rapid dissociation.

Trp^{B26} analog retained native biological activity

The *in vitro* affinity of the Trp^{B26} analog for the lectin-purified insulin receptor (IR; isoform B holoreceptor) was determined to be 0.14(±0.02) nM, ~50% that of [Orn^{B29}]insulin (0.07(±0.02) nM) and WT insulin (0.08(±0.03) nM) (Fig. 5*A*). The potencies of these analogs, evaluated by intravenous (i.v.) injection in diabetic rats, were nonetheless, indistinguishable from WT insulin (Fig. 5*B*).

Trp^{B26} analog displayed a zinc-dependent delay in onset of biological activity

Hexamer assembly delays absorption of insulin from its SQ-injection site (11, 12). To assess the onset and duration of [Trp^{B26},Orn^{B29}]insulin relative to [Orn^{B29}]insulin, the pharmacodynamics (PD) profile of these proteins (made 0.15 mg/ml, corresponding to a monomer concentration of 27 μM and a putative hexamer concentration of 4.5 μM) were evaluated as zinc-free solutions or as pre-assembled phenol-stabilized R_6 hexamers in the presence of excess zinc ions (0.30 mM $ZnCl_2$; 70 zinc ions per hexamer). A zinc-dependent delay in onset of activity was observed on SQ injection of [Trp^{B26},Orn^{B29}]insulin but not on injection of [Orn^{B29}]insulin or WT insulin (Fig. 5*C*; see Fig. S5 for WT results). Whereas the latter PD profiles exhibited a nadir at about 120 min irrespective of the zinc ion concentration, the PD profile of [Trp^{B26},Orn^{B29}]insulin occurred at (i) 120 min in the absence of zinc ions and (ii) 150 min in the presence of 0.30 mM zinc ions with corresponding delays in the rate of fall over the first 30 min (Fig. 5*D*). Together with the above *in vitro* results, these findings suggest that the prolonged lifetime of the Trp^{B26} R_6 hexamer (as inferred from the above kinetic studies of the Co^{2+}-substituted hexamer) are responsible for the inferred zinc-dependent delay in SQ absorption.

Trp^{B26} protracted the PD profile of a model pI-shifted analog

Insulin analogs with isoelectric points (pI) shifted to neutral pH generally exhibit prolonged activity due to precipitation in the SQ depot (30). To determine whether Trp^{B26} might further prolong the activity of such analogs, this substitution was introduced into a [Gly^{A21},Orn^{B29},Orn^{B31},Orn^{B32}]insulin. This "glargine-like" framework was designed to recapitulate the pI shift of glargine with greater ease of semisynthesis.[6] The proteins (formulated at 0.6 mM with 0.3 mM $ZnCl_2$, corresponding to 3 zinc ions per hexamer) were each injected SQ in diabetic rats.

The pI-shifted parent analog displayed peak activity at about 120 min with blood glucose levels returning to baseline after about 360 min. By contrast, its Trp^{B26} derivative displayed a prolonged PD profile: peak activity occurred 180 min with slow return to baseline at >800 min (Fig. 5*E*; see Fig. S6 for i.v. potency). Such a marked delay in peak activity was not observed in the parent glargine-like analog or the Trp^{B26} derivative when administered in the absence of zinc (Fig. S7). These results suggest that Trp^{B26} may favorably be incorporated into current basal analogs as a complementary mechanism of prolonged SQ absorption.

Crystal structure of Trp^{B26} analog demonstrated native-like dimer interface

The crystal structure of [Trp^{B26},Orn^{B29}]insulin was determined as a zinc-coordinated hexamer in the presence of phenol to a resolution of 2.25 Å. Diffraction and refinement statistics are provided in Table S2. The asymmetric unit constituted a "TR^f" dimer[7] (31, 32). The overall structures of the T and R^f protomers were essentially identical to those of WT insulin (Fig. S8) with respective r.m.s. deviations of 1.11 ± 0.30 and

[6] This analog simulated pI-shifted insulin glargine, which contains two additional arginine residues at the C-terminal B chain (B31 and B32) and a substitution of Asn^{A21} by Gly (see Fig. 1*A*) (29, 69). Use of Orn permitted semisynthesis.
[7] One protomer in the (dimeric) asymmetric unit adopted the classical "T" conformation (Fig. 1*B*); the other adopted in the frayed R state (31, 32) in which residues B3–B8 extend the B-chain α-helix.

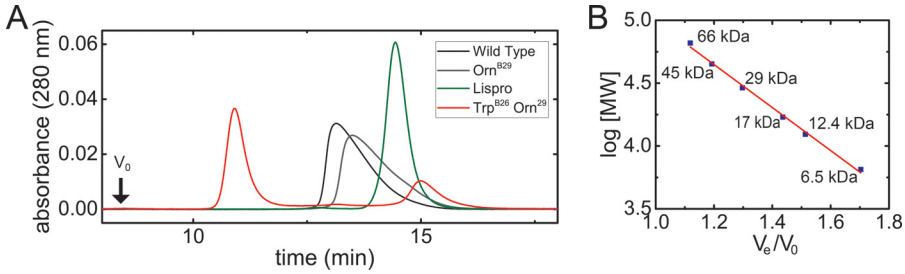

Figure 4. Size exclusion chromatography of Trp^{B26} hexamer. *A,* SEC chromatogram of insulin analogs in the presence of zinc and phenol. The void volume (V_0, *black arrow*) was defined by thyroglobulin (molecular mass 669 kDa). *B,* plot of log(molecular weight) *versus* elution ratio (V_e/V_0) of molecular weight standards. Linear relationship between log[MW] 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 weight] = $-1.71 \times (V_e/V_0) + 6.7012$. Elution times of molecular weight standards are indicated by *blue squares* (labeled by molecular weight). Identity of molecular mass standards is as follows: 66 kDa, BSA; 45 kDa ovalbumin; 20 kDa, carbonic anhydrase, 17 kDa, myosin light chain; 12.4; cytochrome *c*, 6.5 IGF-I. Calculated M_r are given in Table 1.

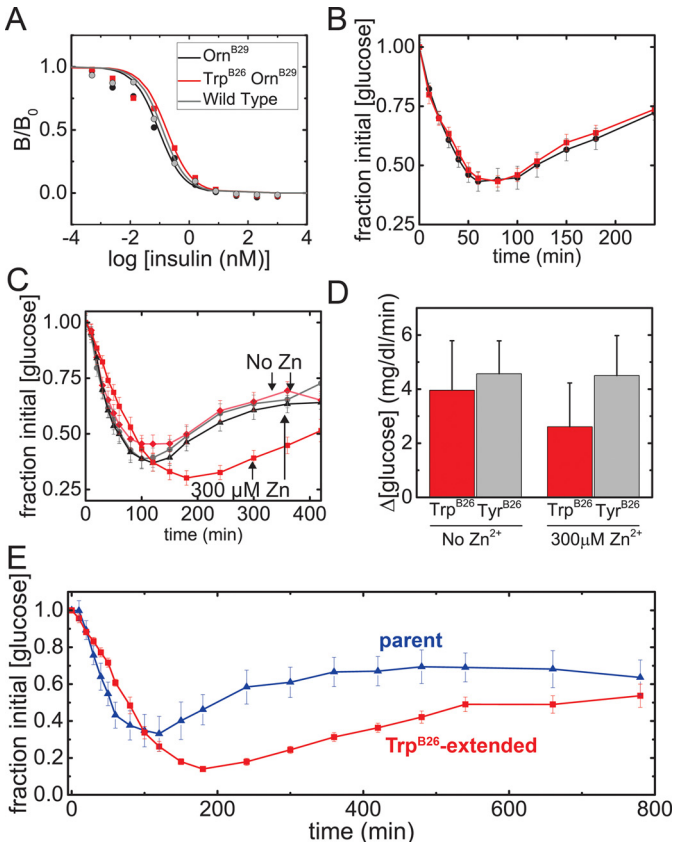

Figure 5. Pharmacology of Trp^{B26} analog. *A,* receptor-binding affinities (isoform B). The affinity of [Trp^{B26},Orn^{B29}]insulin was reduced by 2-fold relative to [Orn^{B29}]insulin (respective K_d estimates 0.14 (±0.03) and 0.07 (±0.02) nM). Color code is indicated in the *inset*. *B,* time course of [blood glucose] following i.v. injection in rats (n = 15); color code as in *A*. *C,* time course of [blood glucose] following SQ injection in the absence or presence of 0.3 mM $ZnCl_2$ (n = 18). *D,* histogram summarizing the rate of fall of [blood glucose] over the first 30 min in *C*, *black bars* indicate S.D. *E*, time course of [blood glucose] following SQ injection of pI-shifted analogs: [Gly^{A21}, Orn^{B29},Orn^{B31},Orn^{32}]insulin, and its Trp^{B26} derivative (n = 6; color code in panel).

1.36 ± 0.30. Additional r.m.s. deviations are given in Tables S3 and S4. Side chain packing near the B26 position was largely unperturbed. In both protomers the Trp^{B26} indole group was oriented with its six-member ring packing against conserved core residues Ile^{A2}, Val^{A3}, and Val^{B12} (Fig. 6); the indole NH group is exposed to solvent in the TR dimer and $T_3R^f_3$ hexamer. The Trp^{B26} side chain in both R and T protomers also displayed dihedral angles within the range of the native Tyr in WT crystal structures (Tables S5 and S6, respectively) with a slight deviation in the positioning of the peptide backbone to accommodate the larger indole side chain.

Spectroscopic probes revealed native-like structure and thermodynamic stability of Trp^{B26} analogs in solution

The native-like crystal structure of [Trp^{B26},Orn^{B29}]insulin is in accordance with its unperturbed circular dichroism (CD) spectrum and thermodynamic stability under monomeric conditions (Fig. 7*A*). Free energies of unfolding (ΔG_u 3.3 ± 0.1) kcal/mol at 25 °C as inferred from two-state modeling of chemical denaturation (33)) were indistinguishable due to small and compensating changes in transition midpoint and slope (m value) (33, 34) (Fig. 7*B*, Table 2). Further evidence that the crystal structure extends to the monomer in solution was provided by 2D ^{1}H NMR studies of Trp^{B26} substituted within an engineered insulin monomer (lispro (35)). Whereas the spectrum of lispro (at pD 7.6 and 37 °C) exhibits sharp resonances for each aromatic spin system (Fig. 8*A*), as expected for a monomeric analog (35), the spectrum of its Trp^{B26} derivative exhibits broadening of resonances at the dimer interface (B16, B24–B26). The latter spin systems can be observed on TOCSY spectrum (Fig. 8*C*) but not in the corresponding DQF-COSY spectrum due to antiphase cancellation. Like the aromatic ring Tyr^{B26} in spectra of insulin lispro (Fig. 8, *A* and *B*), the indole ring exhibited regiospecific nonlocal nuclear Overhauser enhancements (NOEs) from its six-member moiety to the methyl resonances of Val^{B12} and Ile^{A2} (Fig. 8, *B–D*).

The pattern of secondary shifts in the variant is similar to that in the parent monomer. In particular, the aromatic ^{1}H NMR resonances of Trp^{B26} (*red* cross-peaks in Fig. 8*C*) exhibit upfield features (relative to Trp in the isolated B23–B30 octapeptide; *dashed lines*) similar to those of Tyr^{B26} in the parent spectrum (*purple* cross-peaks in Fig. 8*A versus* dotted lines) (36). Dilution of the Trp^{B26} sample partially mitigated resonance broadening but preserved these trends in dispersion. Indole-specific NOEs indicated that the side chain assumes one predominant and asymmetric conformation within a native-like crevice between A- and B-chain α-helices (Table S7). Because Tyr^{B26} undergoes rapid ring rotation about the C_β–C_γ bond axis ("ring flips"), analogous side chain-specific NOEs (inferred in prior studies from molecular modeling) cannot be observed directly (Table S8).

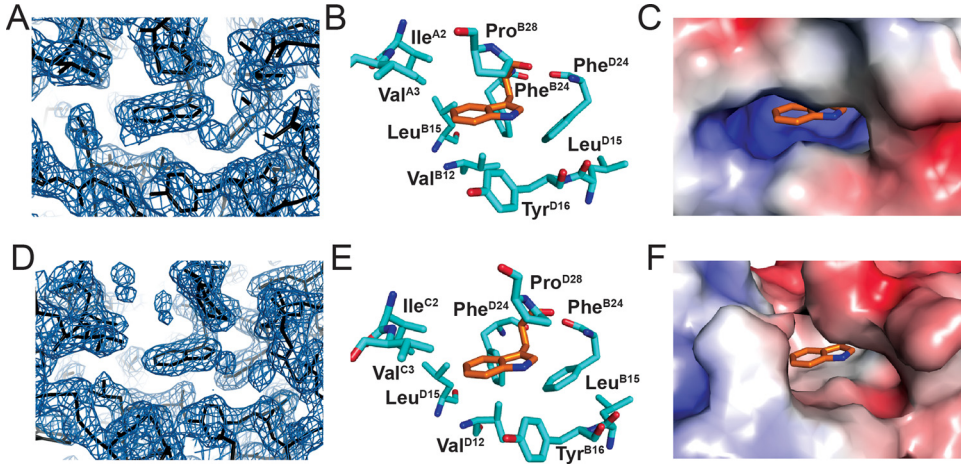

Figure 6. Crystal structure of [Trp^{B26},Orn^{B29}]insulin. *A,* electron density of Trp^{B26} in T-state protomer showing the surrounding density in TR^{f} asymmetric unit (contour level 2.0 Å). *B,* stick model corresponding to map in A; Trp^{B26} is *orange*. *C,* surfaces of residues surrounding Trp^{B26} (*sticks*) as indicated above. *D–F,* corresponding map and models of Trp^{B26} in the R state protomer.

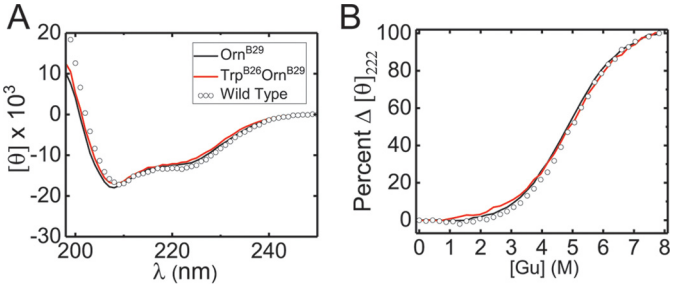

Figure 7. Thermodynamic stability of [Trp^{B26},Orn^{B29}]insulin. *A,* CD spectra of [Trp^{B26},Orn^{B29}]insulin, [Orn^{B29}]insulin, and WT insulin. *B,* guanidine denaturation assays of insulin analogs monitored by ellipticity at 222 nm; color code as in *A*. Stabilities are given in Table 2.

Table 2
Thermodynamic stabilities of insulin analogs

Analog	ΔG^{a}_{u}	C_{mid}	m^{b}
	kcal mol^{-1}	*M*	*kcal mol*$^{-1}$ *M*$^{-1}$
Wildtype	3.4 ± 0.1[c]	5.0 ± 0.1	0.68 ± 0.02
Orn^{B29}	3.3 ± 0.1	4.9 ± 0.1	0.67 ± 0.01
Trp^{B26},Orn^{B29}	3.3 ± 0.1	5.1 ± 0.2	0.64 ± 0.03

[a] Parameters were inferred from CD-detected guanidine denaturation data by application of a two-state model; uncertainties represent fitting errors for a given data set.
[b] The *m*-value (slope Δ(*G*)/Δ(M)) correlates with surface area exposed on denaturation.
[c] Analysis of replicates of [Trp^{B26},Orn^{B29}]insulin, parent [Orn^{B29}]insulin, and WT samples indicated that experimental standard errors were equal to or less than the above fitting errors: ±0.1 kcal mol^{-1} (ΔG_u), ±0.1 M (C_{mid}), and ±0.01 kcal mol^{-1} M^{-1} (*m*).

MM calculations suggested improved aromatic–aromatic interactions within the variant crystal structure

The contribution of aromatic–aromatic interactions involving Trp^{B26} to the stability of the variant dimer interface of the T$_3$R^{f} hexamer was evaluated through calculation of nonbonded interaction energies among aromatic residues B16, B24, B25, and B26 in the TR^{f} dimer. These calculations, which employed the variant crystal structure, were in overall accordance with expectations based on our initial local MM-based modeling (above). In particular, based on aromatic–aromatic interactions alone, the Trp^{B26},Orn^{B29} dimer displayed an increase in interaction energy of 1.4 kcal/mol relative to WT TR^{f} reference structure PDB 1TRZ; the results of these calculations are given in Table S9, *a* and *b*. Although the standard CHARMM empirical energy function, when applied to analyze either crystallographic and MM-minimized models of [Trp^{B26}]insulin, suggested that the electrostatic properties of the Trp side chain were the primary contributors to the increased stability of the dimer, this physical interpretation may reflect the limitations of the partial-charge representation (23, 37). Indeed, preliminary *ab initio* QM simulations of a minimal model (consisting of two aromatic rings *in vacuo*) predicted that enhanced Van der Waals interactions may also make a significant contribution (Fig. S9) (see "Discussion").

Discussion

The physical origins of protein stability and recognition define a foundational problem in biochemistry (38) with central application to molecular pharmacology (39). The zinc–insulin hexamer provides a favorable system for structure-based design due to its long history of crystallographic investigation (40). Indeed, the hexamer's rigidity, as interrogated by NMR spectroscopy (41), renders the overall structure robust to diverse amino acid substitutions (42, 43), even those that destabilize the dimer interface[8] (32, 44). This rigid framework has often enabled analysis of discrete interactions without complications due to the long-range transmission of conformational change (31, 32).

Our studies, stimulated by the seminal recognition of aromatic–aromatic interactions by Burley and Petsko (2) more than 30 years ago, focused on the classical dimer interface, a basic building block of the hexamer (32). Long appreciated as "a thing of beauty" (4, 45), this interface contains eight aromatic residues, six of which engaged in a successive set of aromatic–aromatic interactions. Quantum-chemical simulations of model systems have suggested that nearest-neighbor interactions predominate even in the presence of multiple rings (46). Pairwise dissection of insulin's dimer interface has highlighted

[8] An exception is provided by a substitution at the primary zinc-binding site (His^{B10} → Asp), which converts the hexamer into a novel dodecamer with zinc-coordination mediated by His^{B5} (26). The structure of each protomer is native-like.

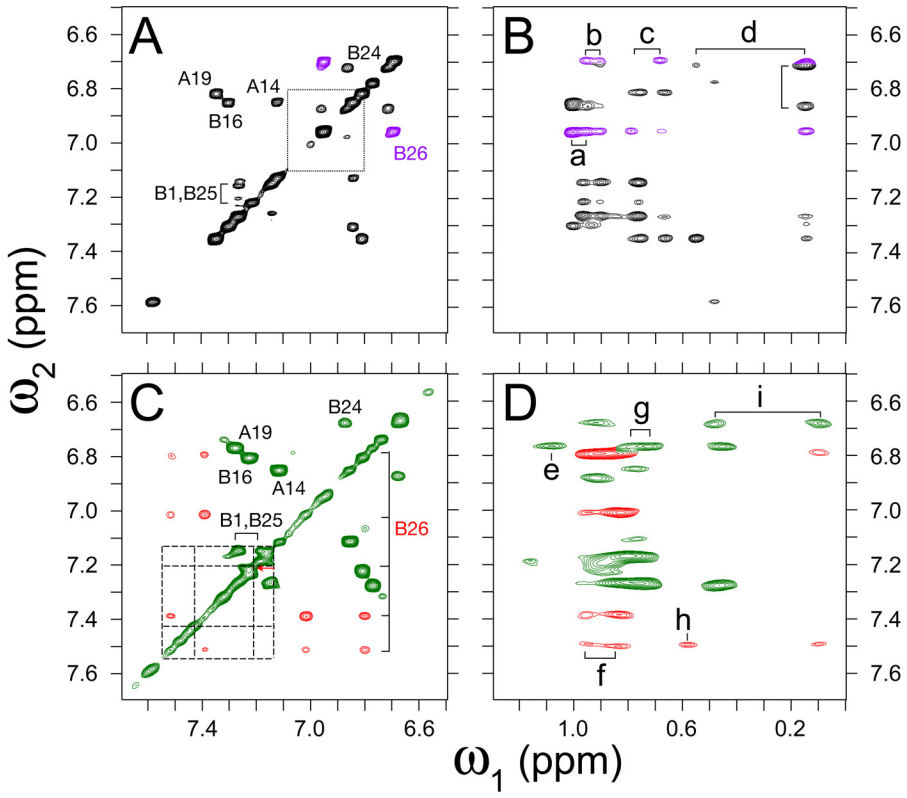

Figure 8. Homonuclear 2D NMR of a Trp^{B26} analog. 2D NMR studies of insulin analogs indicate similar Tyr^{B26}- and Trp^{B26} environments. *A* and *B*, spectra of parent monomer insulin lispro ([Lys^{B28},Pro^{B29}]insulin): *A*, aromatic region of TOCSY spectrum with Tyr^{B26} cross-peaks (*magenta*) shown relative to the Tyr spin system in free octapeptide GFFYTKPT (*dotted lines*); and *B*, region of NOESY spectrum showing contacts between aromatic protons (vertical axis, ω_2) and methyl groups (horizontal axis, ω_1). *C* and *D*, spectra of Trp^{B26} analog of insulin lispro: *C*, aromatic TOCSY spectrum highlighting Trp^{B26} cross-peaks (*red*) relative to the Trp spin system in free octapeptide GFFWTKPT (*dashed lines*) and *D*, region of NOESY spectrum corresponding to *B*. B26-related NOEs are shown in *red*. Cross-peak assignments: (*a*) γ-CH$_3$ Val^{A3}, (*b*) γ-CH$_3$ Val^{B12}, (*c*) γ-CH$_2$, γ-CH$_3$ Ile^{A2}, (*d*) δ-CH$_3$ Leu^{B15}, (*e*) γ-CH$_3$ Val^{A3}, (*f*) γ-CH$_3$ Val^{B12}, (*g*) γ-CH$_2$, γ-CH$_3$ Ile^{A2}, (*h*) δ-CH$_3$ Ile^{A2}, and (*i*) δ-CH$_3$ Leu^{B15}. TOCSY mixing times in spectra *A* and *C* were 55 ms; NOESY mixing times in *B* and *D* were 150 ms.

the potential opportunity to enhance its stability through substitution of Tyr^{B26} by Trp. Whereas our crystallographic analysis verified that this substitution preserves native architecture, a [Trp^{B26}]insulin analog exhibited a dramatic increase in hexamer lifetime *in vitro*. Results of animal testing demonstrated native intrinsic potency (*i.e.* on i.v. bolus injection) but with prolonged activity on SQ injection, presumably due to delayed dissociation of the variant zinc hexamers in the SQ depot.

Protein engineering of insulin analogs is constrained by the complexity of insulin's "conformational lifecycle": from oxidative folding intermediates and self-assembly in the pancreatic β-cell (8) to adoption of an active, "open" conformation on receptor binding (Fig. 9, *A* and *B*) (16). Specific residues may play distinct roles at each stage. In particular, because interfaces within the insulin hexamer overlap the hormone's receptor-binding surface, essentially invariant among vertebrates (47), modifications often impair activity (13). A given WT residue may represent a compromise among competing structural tasks. A recent survey of 18 substitutions at position B26 demonstrated that Tyr is suboptimal with respect to IR-binding affinity but enhances self-assembly relative to more active alternatives (Ala, Ser, or Glu) (18). The latter side chains destabilize the "closed" dimer interface but are favorable at the solvated B26-related edge of the open hormone–receptor interface (Fig. 9*C*).

Protective self-assembly is of key pharmacologic importance

Insulin self-assembly protects the hormone from degradation and toxic misfolding in pancreatic β-cells[9] (8, 48) and in pharmaceutical formulations (49). Because the zinc-insulin hexamer exhibits delayed SQ absorption relative to monomers and dimers, mutational destabilization of the hexamer (10, 12) led to development of rapid-acting insulin analogs (14). Efforts to *stabilize* the insulin hexamer, as a converse strategy to obtain protracted action, were less successful (13). Current long-acting insulin analogs rely instead on higher-order self-association of hexamers within the SQ depot (30) and binding acylated monomer to albumin as a circulating depot (highlighted in Fig. S2) (50, 51).

The problem of how to improve a protein interface is in general more subtle than its opposite, for destabilizing substitutions abound at conserved interfaces, whereas stabilizing substitutions can be rare (13). Structure-based candidate substitutions may encounter entropy–enthalpy compensation (EEC) (52, 53) or cause unintended biological perturbations (54). The challenge posed by insulin is magnified by the struc-

[9] An experiment of nature provided *in vivo* evidence for the relationship between zinc coordination and protection from toxic insulin misfolding: in the rodent species *Octogon degus*, substitution of His^{B10} by Asn is associated with senile amyloidosis of the islets due to insulin fibrillation (70), resulting in DM (48).

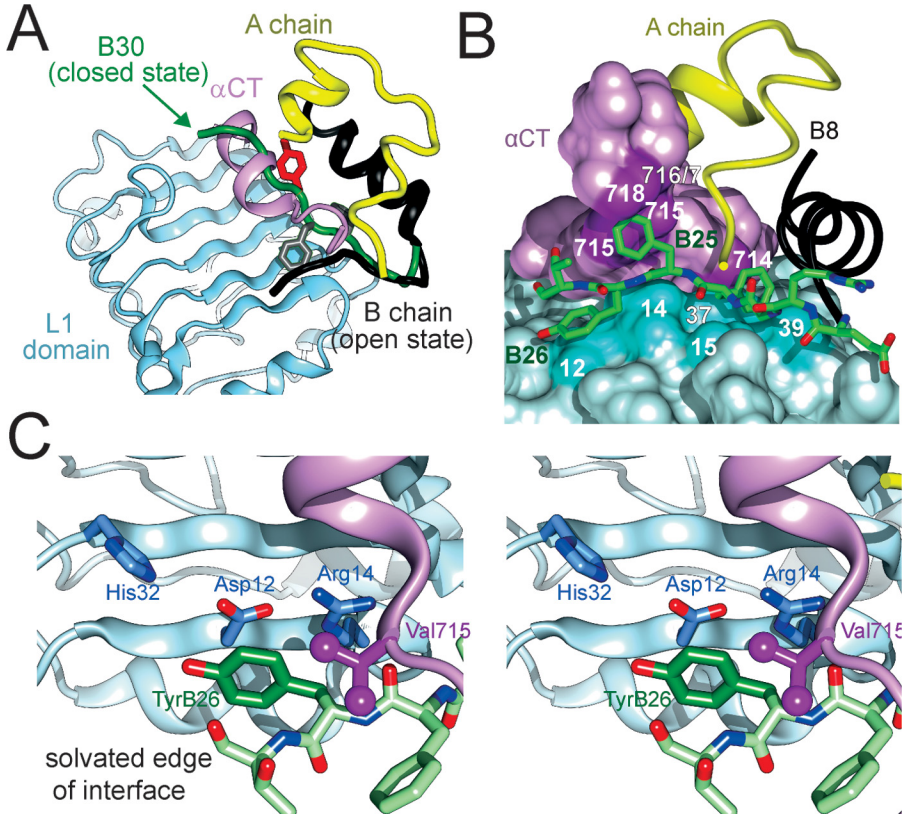

Figure 9. Binding surface of Tyr^{B26} on an IR fragment. *A,* model of WT insulin (in classical T-state) overlaid on the structure of insulin bound to an IR fragment (PDB entry 4OGA). The L1 domain and part of the CR domain are shown in *powder blue,* whereas αCT is shown in *purple*. Phe^{B24} and Tyr^{B26} are, respectively, shown as *gray* and *red sticks*. The B-chain of IR-bound insulin is shown in *dark gray* (B6–B19) or *black* (B20–B27); the *green tube* indicates a classical location within overlay of residues B20–B30 (*green arrow*), thereby highlighting the steric clash of B26–B30 with αCT. Insertion of the insulin B20–B27 segment between L1 and αCT was associated with a small rotation of the B20–B23 β-turn and changes in main chain dihedral angles flanking B24. *B,* stick representation of B-chain residues B20–B27 packed between αCT and the L1 β2 strand. Color code in insulin segment: carbon atoms (*green*), nitrogen (*blue*), and oxygen (*red*). Residues B8–B19 are shown as a *black ribbon*, and the A-chain is shown as a *yellow ribbon*. Key contact surfaces of αCT with B24–B26 are highlighted in *magenta* and L1 with B24–B26 are highlighted in *cyan*. *C,* stereo view of the environment of Tyr^{B26} within its binding site. Neighboring side chains in L1 and αCT are as labeled. This figure was adapted from Ref. 18 with permission of the authors.

tural elegance of its self-assembly (as emphasized by Hodgkin and colleagues (55) in a classic review). Diverse structure-based strategies were previously undertaken with only limited success. Alternate strategies previously used to create basal insulin analogs are depicted in Fig. S2 and summarized in Table S10. One approach focused on the general nonpolar character of the dimer interface: additional hydrophobic substitutions were introduced in an effort to enhance this feature (56). Such designs were not successful and also impaired biological activity, although analogs were identified whose sparing solubility slowed SQ absorption (56). A second approach exploited the classical TR transition among insulin hexamers (57). At the pivot point of this allosteric transition (Fig. S10), an invariant glycine (Gly^{B8}) was substituted by Ser in an effort to stabilize the more stable R state hexamer (58). This analog was unstable as a monomer (54) and exhibited reduced activity (58). Yet another approach sought to stabilize the hexamer by relieving electrostatic repulsion created by the internal clustering of six acidic side chains (Glu^{B13}): their isosteric substitution by Gln indeed promoted assembly of zinc-free hexamers but impaired biological activity (59). Overlap between the self-assembly surfaces of insulin and its receptor-binding surfaces thus compounded the optimization problem.

The present study revisited the architecture of the insulin hexamer in light of recent insights into the hormone's receptor-binding surface (18). Prominent roles are played by a quartet of aromatic residues in the C-terminal B-chain β-strand: Tyr^{B16}, Phe^{B24}, Phe^{B25}, and Tyr^{B26} (33). These four residues, and the clustering of eight dimer-related side chains, have long been the focus of structure-activity relationships (60). Photoactivatable aromatic probes (*para*-azido-Phe) at any of these sites exhibited efficient cross-linking to the IR (61, 62). The co-crystal structure of an insulin monomer bound to a fragment of the IR ectodomain (the μIR model) revealed distinct binding sites at the surface of the L1 domain of the IR β-subunit (B16, B24, and B26) or its αCT element (B24 and B25). Phe^{B24} packs within a nonpolar pocket near aromatic residues in L1 (residues Leu^{37} and Phe^{39}) and αCT (residue Phe^{714}); one wall of this pocket is defined by the aliphatic side chains of Leu^{B15}, Cys^{A20}, and Cys^{B19} as in free insulin (16). This environment differs in detail from those in the insulin dimer but exhibits analogous general features. Although aromaticity is not required to fill the B24-binding pocket (33), its specific size and shape constrain potential substitutions. Tyr^{B16} lies at the periphery of L1 (near Phe^{39} and Lys^{40}). Its substitution by Ala or other aromatic residues preserves activity (63). Phe^{B25} occupies a cleft between αCT

residues (Val^{715}, Pro^{716}, Arg^{717}, and Pro^{718}) that can only accommodate trigonal γ-carbons (22). B25-related aromatic–aromatic interactions are limited due to the peripheral location of this residue (Fig. S11).

We chose to focus on position B26 because of its broad functional tolerance of diverse substitutions (18). As illustrated above (Fig. 8*C*), Tyr^{B26} binds at the solvated periphery of the μIR interface. Indeed, substitution of small, polar, or charged amino acids (such as Ala, Ser, or Glu) enhances receptor affinity, but at the price of impaired self-assembly and decreased thermodynamic stability with heightened susceptibility to physical degradation (16). These findings highlighted the evolutionary importance of the native dimer interface and dual role of Tyr^{B26}. We thus hypothesized that an aromatic substitution at the B26 position might enhance self-assembly without loss of biological activity.

The classical structure of the insulin dimer motivated study of successive aromatic–aromatic interactions as a physical mechanism of stability (6, 64). The increased stability of ETF aromatic–aromatic interactions involving Trp over those involving Tyr contributes to the increased stability of the Trp^{B26} hexamer. The larger size of the delocalized π-orbital of Trp in relationship to that of Tyr causes a stronger negative charge to accumulate on the "face" of the indole ring. For this reason, hydrogen atoms surrounding the aromatic rings of local residues form stronger electrostatic interactions with the face of the indole ring of Trp than with the phenol ring of Tyr (65, 66). Aromatic pairs involving Trp residues are less common than those involving Tyr or Phe (1). However, the strength of aromatic–aromatic interactions involving Trp is evidenced by functional importance of Trp-based interactions; an example is provided by a Trp–Tyr aromatic "lock" that stabilizes the active conformation of the ghrelin receptor (67). Indeed, comparison of electrostatic interactions of Trp^{B26} and Tyr^{B26} within the minimized model of the B16/B24/B26 aromatic network of insulin revealed that interactions involving Trp were favored over those involving Tyr across a broad variety of orientations of the respective aromatic rings. Extension of the aromatic lock metaphor (introduced by Holst *et al.* (67) to describe conformational "trapping" in the GPCR structure) to the insulin hexamer highlights the kinetic effect of the B26 substitution on the rate of hexamer disassembly (as probed by the Co^{2+}-EDTA sequestration assay), which was more dramatic than effects on equilibrium association (as probed by SEC). It would be of future interest to measure activation energies for disassembly. Insight into the structural origins of the prominent Trp^{B26}-associated kinetic lock may be provided by activated molecular dynamics simulations of hexamer disassembly.

Trp^{B26} side chain exhibited an orientation similar to native Tyr

The Trp^{B26} side chain in the crystal structure of [Trp^{B26},Orn^{B29}]insulin displayed an orientation similar to that of the native Tyr. A slight main chain shift in the B24–B28 β-strand (0.2 (±0.03) Å) was sufficient to enable native-like packing of the larger indole ring against the core of a protomer (Fig. S12). In both T and R^f subunits, the six-membered component of the indole side chain was oriented toward Ile^{A2}, Val^{A3}, and Val^{B12}. Based on the classical 3.5–6.5 inter-centroid distance, residue Trp^{B26} (of the T protomer) showed potential interactions with residues Tyr^{D16} and Phe^{D24}. An interplanar angle of 78 degrees between Trp^{B26} and Tyr^{D16} was indicative of classical aromatic–aromatic packing within proteins (2), which generally range from 50 to 90 degrees. Trp^{B26} and Phe^{D24} displayed an interplanar angle of 36 degrees, however, this orientation is less common. Similarly, in the R^f protomer, Trp^{D26} packed near Tyr^{B16} and Phe^{B24}. The interplanar angle between Trp^{D26} and Tyr^{B16} was 62 degrees, whereas between Trp^{D26} and Phe^{B24} was 48 degrees. Previous studies have suggested that Trp residues may form aromatic–aromatic interactions over longer distances than those formed by Tyr–Phe pairs (68). Thus, the two intra-chain aromatic pairs, Trp^{B26}–Phe^{B24} and Trp^{D26}–Phe^{D24} (respectively, separated by 7.8 and 7.1 Å) may interact more efficiently than the corresponding Tyr–Phe pairs in WT insulin (ring geometries are summarized in Table S11, *a* and *b*).

CHARMM calculations of aromatic–aromatic interactions across the dimer interface of [Trp^{B26},Orn^{B29}]insulin revealed 1.4 kcal/mol increased interaction energy. Residue by residue analysis of each component of the aromatic network indicated that the increased strength of interaction across the dimer interface was the result of the interactions involving Trp^{D26} (see Table S9b). This result suggests that Tyr^{B26} → Trp may only display stabilizing properties when in an R state protomer. If so, the T_6 hexamer formed by [Trp^{B26}]insulin would be expected to have dissociation kinetics similar to WT insulin, whereas the corresponding R_6 hexamer may be expected to have markedly increased stability. The R → T transition, which is rapid in WT insulin on release of phenol, may represent the kinetic barrier responsible for the *meta*-stable association state observed in SEC experiments.

The packing of Trp$^{B26/D26}$ within the respective cores of T- and R^f crystallographic protomers is similar in each case to the WT Tyr$^{B26/D26}$ and oriented such that the indole's nonpolar six-membered portion projects more deeply into a crevice between A- and B-chains than does its proximal heterocycle. Our ^{1}H NMR studies of Trp^{B26} within an engineered monomer (35) provided evidence that this overall conformation does not require self-assembly. Analogous partial burial of Tyr^{B26} and Trp^{B26} in the respective protein structures would in itself be expected to augment the variant's stability due to enhanced solvation free energy[10] (69) (*i.e.* as predicted by water–octanol transfer studies of free Tyr and Trp (69); Table S12). Guanidine denaturation nonetheless, indicated that their stabilities are indistinguishable.[11]

We speculate that the predicted residue-specific differences in solvation free energy are attenuated by differences in protein dynamics leading to EEC (70). Although the two B26 side chains each exhibit upfield secondary ^{1}H NMR shifts and analogous inter-residue NOEs, it is possible that the substitution is associated with local or nonlocal differences in protein dynam-

[10] The identification of a destabilizing Trp → Tyr mutation in *Escherichia coli* thioesterase-I demonstrated, in breach, the improved cavity-filling properties of Trp relative to Tyr (72).

[11] That [Trp^{B26},Orn^{B29}]insulin exhibited a greater exposure of nonpolar surfaces on denaturation in guanidine HCl than did [Orn^{B29}]insulin was suggested by a change in *m* values in two-state modeling (Table 2).

ics. It would be of future interest to investigate dynamic features by amide proton $^{1}H-^{2}H$ exchange and heteronuclear NMR relaxation methods (70). Because Trp^{B26} promotes partial dimerization of insulin lispro under NMR conditions (as indicated by concentration-dependent ^{1}H NMR resonance broadening), such studies may require use of an alternative monomeric template.

Given the ubiquity of EEC as a confounding general aspect of protein design (71), the profound effects of Trp^{B26} on the properties of the insulin hexamer seem all the more remarkable. We envision that EEC is circumvented in this case by the rigidity of the insulin hexamer (including the interlocked aromatic residues at its dimer interfaces) with efficient burial of the WT and variant B26 side chains (72). With similar internal structures and external solvation properties, the variant hexamer would gain (relative to WT) two advantages from the Tyr → Trp substitution: (i) greater B26-related solvation transfer free energy (73) (Table S12) augmented by (ii) an uncompensated enthalpic advantage arising from more favorable ETF aromatic interactions as next discussed.

Molecular mechanics calculations rationalized physical and pharmacologic properties

The structural rigidity of the R_6 insulin hexamer (42, 74) motivated local MM-based modeling to assess the interactions contributing to the stability of the [Trp^{B26}]insulin hexamers. Analysis of insulin oligomers by Raman spectroscopy revealed dampened conformational fluctuations of the R_6 hexamer in relationship to lower-order oligomers and T_6 hexamers (72). Moreover, the thermodynamic stability of the assembly was evidenced by the lack of conformational changes visualized by NMR spectroscopy over a temperature range of 10 – 80 °C (42). The resistance of the R_6 hexamer to structural perturbation has also been shown in the context of mutant insulin analogs: native-like X-ray crystal structures have been reported of R_6 containing a broad range of substitutions (Table S13) (17, 75, 76). Even substitutions that were shown to destabilize the dimer interface of insulin, such as the substitution of Phe^{B24} by the nonaromatic cyclohexylalanine, were shown to have little impact on the global structure of the R_6 hexamer. For this reason, the Tyr^{B26} → Trp substitution was expected to affect only the local structure of the insulin hexamer. Thus, the effects of the mutation were amenable to initial analysis in simplified (eight-ring) models of the dimer interface of insulin.

Energy minimization of a local model of [Trp^{B26}]insulin (*i.e.* as substituted into a WT insulin dimer extracted from a representative crystal structure of a $T_3R^f_3$ zinc hexamer) yielded a native-like framework with enhanced nearest-neighbor B26-related aromatic–aromatic interactions. The partial-charge (monopole) model of the aromatic rings in the CHARMM empirical energy function, parametrized in accordance with *ab initio* simulations (77), predicted an increase of 0.8 kcal/mol. Although this calculation could in principle have been confounded by transmitted conformational perturbations and did not consider potential changes in conformational entropy or solvation, its conservative features were verified by X-ray crystallography. That the structure of [Trp^{B26},Orn^{B29}]insulin is essentially identical to WT suggests that local properties of Trp^{B26} directly underlie the observed increase in hexamer stability and lifetime.

The general asymmetry of the variant TR^f dimer in the crystallographic hexamer was associated with differences in the details of corresponding aromatic–aromatic interactions across the dimer interface. Although CHARMM calculations predicted an increase of a 1.4 kcal/mol in interaction energies (0.7 kcal/mol per protomer) in accordance with our initial modeling, residue by residue decomposition ascribed this increase primarily to Trp^{D26} (in the R^f protomer) and not to Trp^{B26} (in the T protomer). It is formally possible that Trp^{B26} is only stabilizing in an R state, but additional studies would be required to resolve this issue. Our cobalt-EDTA sequestration studies focused on the R_6 state as the preferred storage vehicle in a pharmaceutical formulation. On SQ injection, rapid diffusion of phenolic ligands from the depot leads to a TR transition. That Trp^{B26} was found to delay subsequent absorption into the bloodstream suggests that this substitution also enhances the kinetic stability and lifetime of the T_6 hexamer.

A seeming paradox is posed by the evolutionary exclusion of Trp at position B28 of vertebrate insulins despite the evident compatibility of this aromatic side chain with native structure and function. We speculate that this exclusion reflects the biological importance of the rapid disassembly of zinc insulin hexamers on their secretion by pancreatic β-cells. Whereas the enhanced thermodynamic and kinetic stabilities of [Trp^{B26}]insulin hexamers would seem favorable for storage within secretory granules (as within pharmaceutical formulations) (8), delayed disassembly of the variant hexamers in the portal circulation would be predicted to reduce the hormone's bioavailability on first pass through the liver, as insulin dimers and hexamers cannot bind to the IR (78). Such delayed disassembly might also decrease the delivery of free zinc ions to the liver, recently predicted to constitute a regulatory signal in its own right (for review, see Ref. 79).

The resulting impairment in hormonal regulation of hepatic metabolism (and accompanying systemic hyperinsulinemia (80)) could in principle have imposed a selective disadvantage in the course of vertebrate evolution. Although to our knowledge, such a kinetics-based mechanism has not been observed *in vivo*, the converse, abnormally rapid clearance of insulin, has been found on release from zinc-deficient secretory granules; presumably zinc-free insulin oligomers more rapidly dissociate on dilution in the portal circulation and so are more efficiently cleared than WT zinc insulin hexamers (81). Similarly, the exclusion of Trp at position B26 of vertebrate IGFs may reflect a disadvantageous competition between self-assembly and binding to IGF-binding proteins, which are critical to the integrated physiology of IGF function (for review, see Ref. 82). Such functional complexity may impose hidden constraints on the evolution (and so divergence) of protein sequences. The conserved "aromatic triplet" of vertebrate insulins and IGFs provides a natural laboratory to uncover such evolutionary constraints (45). These considerations further suggest that structural features of a protein pertinent to its endogenous function may in general be distinct from those biophysical properties that may be re-engineered to optimize molecular pharmacology.

Comparison of [Trp^{B26}]insulin and a corresponding iodo-Tyr analog

Design of [Trp^{B26}]insulin was in part motivated by prior studies of an analog containing 3-iodo-Tyr at the B26 (17, 23, 83). The latter analog ("3I-Y-insulin") exhibited a variety of favorable properties: increased affinity for the IR (83), increased thermodynamic stability, and augmented resistance to fibrillation (17). Moreover, when formulated as an R_6 hexamer, this analog exhibited a decreased dissociation rate (17). Because modified amino acids raise the cost and complexity of protein manufacture, we wondered whether a natural amino acid might mimic, at least in part, the structure of 3I-Y-insulin and so confer favorable pharmacologic properties with conventional manufacturing.

This translational goal motivated our initial MM-based local modeling of [Trp^{B26}]insulin. Analysis of the crystal structure of 3I-Y-insulin, determined as an R_6 zinc hexamer (23), revealed that the large, nonpolar iodo-substituent packed within the core of the insulin protomer with preservation of a native-like dimer interface. Molecular dynamics simulations, undertaken with a multipolar electrostatic model of the modified aromatic ring (23), rationalized this conformation: the iodine atom efficiently filled a cryptic packing defect in WT insulin, lined by the conserved side chains of Ile^{A2}, Val^{A3}, and Tyr^{A19}. The enhanced packing efficiency of the modified insulin and novel network of halogen-specific electrostatic interactions ("weakly polar" interactions (1)) appear to underlie the analog's increased thermodynamic stability and resistance to fibrillation. Whereas the standard partial-charge model of aromatic systems failed to account for the conformation of iodo-Tyr observed in the crystal structure, a multipolar electrostatic model rationalized thermodynamic stabilization of the dimer interface by this halogen "anchor" (23). Subtle changes in the geometry of aromatic–aromatic interactions were observed both in the simulations and in the crystal structure. Although activated MD simulations were not undertaken to probe the process of dimer dissociation, we presume that the above mechanisms of ground-state stabilization, enhanced core packing efficiency, and a halogen-specific weakly polar network, also underlie the increased barrier to dissociation as indicated by the prolonged lifetime of the variant R_6 hexamer (23).

Although the profound QM effects of halo-aromatic substitutions, including weakly polar interactions and "halogen bonding" (84), cannot be recapitulated by natural amino acids, it seemed possible that enhanced core packing efficiency might be achieved by analogy to 3I-Y-insulin. Indeed, the steric profile of the asymmetric indole side chain of Trp is similar in size to 3-iodo-Tyr. Accordingly, we imagined that the offset six-membered portion of the bicyclic indole ring might pack within the core of insulin in a manner similar to the iodine atom. In this intuitive picture the extended π-system of Trp^{B26} was envisioned to interact with neighboring residues to recapitulate, at least in part, the favorable electrostatic properties of iodo-Tyr^{B26} (23, 37).

The above line of reasoning led to the present set of studies. In accordance with our original intuition and local MM-based modeling, the crystal structures of 3I-Y-insulin and [Trp^{B26},Orn^{B29}]insulin exhibited similar features. Nevertheless, salient differences between the quantum-chemical properties of iodo-Tyr and Trp might make the similar structures of these B26 analogs a fortuitous outcome of distinct mechanisms. Whereas effects of 3-iodo-Tyr on aromatic–aromatic interactions at the insulin dimer interface are subtle, presumably reflecting indirect inductive effects of iodo-substitutent (23), substitution of Tyr by Trp introduces a larger aromatic surface at this interface. It is possible that this feature underlies the more marked impact of Trp^{B26} on insulin oligomerization relative to 3-iodo-Tyr.

Aromatic–aromatic interactions exemplify limitations of classical models

The quantum origins of aromatic–aromatic interactions are complex, and so classical electrostatic models (such as partial-charge model of Tyr in the standard CHARMM empirical energy function) can be incomplete (85). Although QM calculations more fully capture this complexity, a tradeoff is encountered between rigor and computational feasibility, especially in complex systems such as proteins (86), but even in *ab initio* simulations of benzene– benzene interactions (46). Standard MM and MD methods thus employ parametrized force fields to approximate quantum-chemical interactions (21). Although parameters (*e.g.* partial charges assigned to an aromatic ring) have been chosen to provide reasonable protein models, such use of "monopolar" electrostatics neglects the polarizability of aromatic systems and so omits dispersion forces (87). Parametrized classical models may thus mischaracterize the strength or directionality of intermolecular interactions, particularly those involving more than one aromatic group or an aromatic group and a charged moiety. Although the pioneering studies of Burley and Petsko (88) suggested that the partial-charge model of aromatic rings in proteins was sufficient for characterizing aromatic–aromatic interactions, more recent work has highlighted its limitations with respect to delineating underlying physical mechanisms (85, 89). The limits of parametrized classical force fields are of particular importance when applications are sought in nonstandard systems (such as unnatural protein mutagenesis (90)) for which the parameters were not intended. Examples are provided by halogen-modified aromatic systems (widely employed in medicinal chemistry) (91), which are associated with marked changes in quantum-chemical properties (92). Formal QM/MM simulations of proteins may nonetheless be circumvented through force fields incorporating multipolar electrostatic models of aromatic–aromatic networks (85).

The correspondence between our initial local MM-based modeling and our experimental findings, however striking, may be coincidental (93): rigorous elucidation of the physics of the variant aromatic cluster at insulin's dimer interface may require application of free-energy MD-based simulations with explicit inclusion of water molecules ("molecular alchemy" (94)). This approach may also provide insights into whether or how EEC may be circumvented. Nevertheless, standard MM calculations can guide initial biochemical analysis of protein structure as a guide for protein engineering. In the present study such initial modeling reinforced our structural intuition by highlighting the plausibility that substitution of Tyr^{B26} by Trp might pre-

serve a native-like interface and in fact enhance its weakly polar properties. Because the rigid hexameric framework provided a favorable context for local modeling of amino acid substitutions at subunit interfaces and yet it is the monomer that functions as a hormone in the bloodstream, it would be of future interest to apply more sophisticated computational techniques to simulate the structure and dynamics of [Trp^{B26}]insulin as a monomer in solution. Such predictions may in principle be tested through biophysical studies of an engineered insulin monomer containing Trp^{B26}. Although this substitution is favorable in the context of the hexamer (and so of potential pharmacologic benefit), design of a monomeric NMR model will need to overcome the *confounding* effects of Trp^{B26}, as defined in this experimental context, to promote self-association.

Concluding remarks

To our knowledge, this study represents the first exploitation of aromatic–aromatic interactions to enhance the biophysical properties of a therapeutic protein (95). Our approach may be broadly applicable in protein engineering (as ETF interactions are ubiquitous) and generalizable to nonstandard aromatic moieties. The latter would be likely to require QM/MM methods rather than classical force fields parametrized with partial charges (90). Overall effects of such substitutions on protein stability and self-assembly will require an integrated analysis of solvation free energies (86), changes in protein dynamics (96), and potential EEC (97). Three- and four-dimensional heteronuclear NMR experiments would be expected to provide higher-resolution information regarding the local and nonlocal interactions of the optimized aromatic system. In the present application, analyses of ^{15}N relaxation and $^{1}H–^{2}H$ amide–proton exchange are expected to improve understanding of the molecular dynamics of the Trp^{B26}-modified aromatic cluster and so provide a more rigorous biophysical context for its enhanced self-association properties (70, 96).

The present application to insulin demonstrates a direct relationship between stabilization of the insulin hexamer and prolonged activity of a basal analog. Continuous and flat 24-h insulin activity ("peakless" basal formulations) is of clinical interest in the treatment of Type 1 and Type 2 diabetes mellitus to reduce the risk of hypoglycemia (especially at night) at a given level of glycemic control (98). Because this mechanism is unrelated to present strategies to achieve protracted action, we envision that Trp^{B26}-related enhancement of dimer-specific aromatic–aromatic interactions could favorably be introduced into current basal insulin formulations. In the future a combination of orthogonal molecular strategies might enable development of a once-a-week basal insulin therapy analogous to that of GLP-1 agonists (99). Optimization of weakly polar interactions may thus assume a central place in the toolkit of molecular pharmacology.

Experimental procedures

Materials

Insulin was purchased from BioDel® (Danbury, CT). Insulin glargine was obtained from Lantus® (Sanofi-Aventis, Paris, FR). Reagents for peptide synthesis were as described (100).

Preparation of insulin analogs

Variant insulins were prepared by semisynthesis (101). In brief, synthetic peptides were coupled to a tryptic fragment of insulin (*des*-octapeptide [B23-B30]insulin) in aqueous/organic solvent using trypsin as a catalyst. Following RP-HPLC purification, predicted molecular masses were confirmed by MS (33).

Hexamer disassembly assays

Disassembly of phenol-stabilized (R_6) Co^{2+}-substituted insulin hexamers was monitored as described (102). In brief, WT insulin or variants were made 0.6 mM in buffer containing 50 mM Tris-HCl (pH 7.4), 50 mM phenol, and 0.2 mM $CoCl_2$ (18) and incubated overnight at room temperature to attain conformational equilibrium. Spectra (400–750 nm) were obtained to monitor tetrahedral Co^{2+} coordination (27) through its signature absorption band at 574 nm (27). Co^{2+} sequestration was initiated by addition of EDTA to a concentration of 2 mM. Dissociation was probed via attenuation of a 574-nm band (27); data were fit to a monoexponential decay equation (29).

Protein crystallography

Crystals were obtained by hanging drop vapor diffusion at room temperature in the presence of a 1:1.7 ratio of Zn^{2+} to protein monomer and a 3.5:1 ratio of phenol to protein monomer in Tris-HCl. Diffraction was observed using synchrotron radiation at a wavelength of 0.9795 Å at the Stanford Synchrotron Radiation Light Source (beamline BL7-1); crystals were flash frozen to 100 K. Structure determination was carried out using molecular replacement using CCP4 (103) and Phenix structure-determination suites (104). The resulting structure was validated using a PDB Redo server (105). The lattice contained one TR^f dimer per asymmetric unit. The main chain conformations of the 97 residues in the refined model of the TR^f dimer in the asymmetric unit (excluding 2 Thr, 1 Orn, and 1 Phe residues) each resided in a most favored Ramachandran region.

Receptor-binding assays

Analog affinities for detergent-solubilized IR-B holoreceptor were measured by a competitive-displacement assay (18). Successive dilutions of WT insulin or analogs were incubated overnight with wheat germ agglutinin-SPA beads (PerkinElmer Life Sciences), receptor, and radiolabeled tracer before counting (18). To obtain dissociation constants, competitive binding data were analyzed by nonlinear regression by the method of Wang (106).

Rat studies

Male Lewis rats (mean mass ~300 g) were rendered diabetic by streptozotocin. Effects of insulin analogs formulated in Lilly® buffer (18) on blood glucose concentration following SQ injection were assessed in relationship to WT or [Orn^{B29}]insulin (18). [Orn^{B29}]Insulin and [Trp^{B26},Orn^{B29}] insulin were made in the above buffer. Gly^{A21}, Orn^{B29}, Orn^{B31}, Orn^{B32}, Trp^{B26} and Gly^{A21}, Orn^{B29}, Orn^{B31}, and Orn^{B32} were dissolved in diluted HCl (pH 4) containing *meta*-cresol, glycerol, and a 1:2 ratio of $ZnCl_2$:insulin monomer. Rats

were injected SQ with 3.44 nmol of insulin or insulin analogs (~12–13.7 nmol).

Animals used in this study were housed in the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC)-accredited facilities of Case Western Reserve University (CWRU) School of Medicine. All procedures were approved by the Institutional Animal Care and Use Committee (IACUC) Office at CWRU, which provided Standard Operating Procedures and reference materials for animal use (in accordance with the NIH Guide for the Care and Use of Laboratory Animals). The animal health program for all laboratory animals was directed by the CWRU Animal Resource Center. Animal care and use was further monitored for Training and Compliance issues by Veterinary Services.

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 (33). Insulin samples (20 μl) were loaded on an Enrich® SEC70 column (10 × 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 (18).

Spectroscopy

CD spectra were acquired in 10 mM potassium phosphate (pH 7.4) and 50 mM KCl (33). Free energies of unfolding (ΔG_u) were inferred at 25 °C from two-state modeling of protein denaturation by guanidine HCl (33, 34). ^{1}H NMR spectra were acquired at 700 MHz at pH 8.0 or pD 7.6 (direct meter reading) at 37 °C (36). Chemical shifts of aromatic protons in residues B24–B24 (Phe-Phe-Tyr or Phe-Phe-Trp) were evaluated in relationship to corresponding chemical shifts in respective octapeptides B23–B30, presumed to represent random-coil shifts.

Molecular mechanics calculations

Calculations were performed using CHARMM (kindly provided by Prof. M. Karplus). Its standard empirical energy function was employed (in whose development aromatic rings were parametrized by partial atomic charges in accordance with *ab initio* QM simulations (21)). Representative WT insulin dimers were obtained from PDB entries 4INS, 1ZNJ, and 1TRZ. These structures and corresponding Trp^{B26} homology models were subjected to local energy minimization (100 steps of steepest descent followed by Adopted Basis Newton-Raphson with gradient tolerance *tolg* 0.0008/10,000 steps). Minimizations were halted either at 1,000 steps or when the above tolerance was reached. Changes in conformation were allowed only to eight side chains (Tyr^{B16}, Phe^{B24}, Phe^{B25}, Tyr^{B26} (or Trp^{B26}), and their dimer-related mates; the remaining atoms in the respective dimers were fixed. Total interaction energies and respective electrostatic components were obtained between the side chain of residue B26 and the neighboring three aromatic side chains (Phe^{B24} and dimer-related Tyr^{D16} and Phe^{D24}). Following this survey of crystallographic dimers, such energies were further evaluated in a simplified molecular model that contained only the side chains of residues Tyr^{B26} (or Trp^{B26}) and the same neighboring three aromatic residues as extracted from PDB entry 4INS; this yielded the electrostatic interaction energy map shown in Fig. 2, in which B26 χ_1 and χ_2 dihedral angles were systematically varied without energy minimization.

Quantum mechanics calculations

Electron density and molecular electrostatic potential of Tyr and Trp side chains were calculated with B3LYP and 6-31G(d) basis sets using Gaussian utility Cubegen in Gaussian09 (77). The isosurface map was generated using Jmol (107). *Ab initio* energies of interaction between pairs of isolated aromatic rings were determined by calculating interaction energies using the MP2 method with the aug-cc-pVDZ basis set in Gaussian 09 (77).

Author contributions—N. K. R., N. F. P., and M. A. W. conceptualization; N. K. R., N. P. W., A. N. T., and V. C. Y. data curation; N. K. R. validation; N. K. R., N. P. W., A. N. T., V. C. Y., and F. I.-B. investigation; N. K. R. and M. A. W. writing-original draft; N. P. W. and V. C. Y. formal analysis; N. F. P., V. C. Y., and M. A. W. supervision; N. F. P. and M. A. W. funding acquisition; N. F. P. methodology; N. F. P., F. I.-B., and M. A. W. writing-review and editing; F. I.-B. and M. A. W. resources.

Acknowledgments—We thank K. El-Hage, M. Lawrence, M. Meuwly, and V. Pandyarajan for helpful discussion; K. Carr, R. Grabowski, P. Macklis, and M. Swain for assistance with rat studies; V. Kumar and F. van den Akker for advice regarding X-ray crystallography; J. Whittaker and L. Whittaker for advice regarding IR-binding assays; P. De Meyts (Novo Nordisk) for the gift of radiolabeled insulin; and M. C. Lawrence for assistance preparing Fig. 9. Use of the Stanford Synchrotron Radiation Lightsource at the SLAC National Accelerator Laboratory is supported by United States Department of Energy (DOE), Office of Science and Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and National Institutes of Health NIGMS Grant P41GM103393.

References

1. Burley, S. K., and Petsko, G. A. (1988) Weakly polar interaction in proteins. *Adv. Protein Chem.* **39,** 125–189 Medline
2. Burley, S. K., and Petsko, G. A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **229,** 23–28 CrossRef Medline
3. Waters, M. L. (2002) Aromatic interactions in model systems. *Curr. Opin. Chem. Biol.* **6,** 736–741 CrossRef Medline
4. Blundell, T. L., Dodson, G. G., Hodgkin, D. C., and Mercola, D. A. (1972) Insulin: the structure in the crystal and its reflection in chemistry and biology. *Adv. Protein Chem.* **26,** 279–402 CrossRef
5. Salsali, A., and Nathan, M. (2006) A review of types 1 and 2 diabetes mellitus and their treatment with insulin. *Am. J. Ther.* **13,** 349–361 CrossRef Medline
6. Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., Hardine, M. M., Hodgkin, D. C., Rimer, B., and Sheet, S. (1969) Structure of rhombohedral 2 zinc insulin crystals. *Nature* **224,** 491–495 CrossRef
7. Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., and Mercola, D. (1976) Structure of insulin in 4-zinc insulin. *Nature* **261,** 166–168 CrossRef Medline
8. Dodson, G., and Steiner, D. (1998) The role of assembly in insulin's biosynthesis. *Curr. Opin. Struct. Biol.* **8,** 189–194 CrossRef Medline

9. Brange, J., and Langkjaer, L. (1997) Insulin formulation and delivery. *Pharm. Biotechnol.* **10,** 343–409 Medline
10. Brange, J., Owens, D. R., Kang, S., and Volund, A. (1990) Monomeric insulins and their experimental and clinical implications. *Diab. Care* **13,** 923–954 CrossRef
11. DeFelippis, M. R., Chance, R. E., and Frank, B. H. (2001) Insulin self-association and the relationship to pharmacokinetics and pharmacodynamics. *Crit. Rev. Ther. Drug Carrier Syst.* **18,** 201–264 Medline
12. Berger, M., Halban, P. A., Girardier, L., Seydoux, J., Offord, R., and Renold, A. E. (1979) Absorption kinetics of subcutaneously injected insulin. *Diabetologia* **17,** 97–99 CrossRef Medline
13. Brange, J., and Vølund, A. (1999) Insulin analogs with improved pharmacokinetic profiles. *Adv. Drug Deliv. Rev.* **35,** 307–335 CrossRef Medline
14. Mayer, J. P., Zhang, F., and DiMarchi, R. D. (2007) Insulin structure and function. *Biopolymers* **88,** 687–713 CrossRef Medline
15. Markussen, J., Jonassen, I., Havelund, S., Brandt, J., Kurtzhals, P., Hansen, P. H., and Kaarsholm, N. C. (September 16, 2003) U. S. patent US 10620651, Novo Nordisk AS
16. Menting, J. G., Yang, Y., Chan, S. J., Phillips, N. B., Smith, B. J., Whittaker, J., Wickramasinghe, N. P., Whittaker, L., Pandyarajan, V., Wan, Z., Yadav, S. P., Carroll, J. M., Srokes, N., Roberts, C. T., Ismail-Beigi, F., *et al.* (2014) Protective hinge in insulin opens to enable its receptor engagement. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 1–48 CrossRef
17. 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., and Weiss, M. A. (2014) Aromatic anchor at an invariant hormone-receptor interface function of insulin residue B24 with application to protein design. *J. Biol. Chem.* **289,** 34709–34727 CrossRef Medline
18. Pandyarajan, V., Phillips, N. B., Rege, N., Lawrence, M. C., Whittaker, J., and Weiss, M. A. (2016) 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 CrossRef Medline
19. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. A., and Karplus, M. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4,** 187–217 CrossRef
20. Kramer, C., Gedeck, P., and Meuwly, M. (2012) Atomic multipoles: electrostatic potential fit, local reference axis systems, and conformational dependence. *J. Comput. Chem.* **33,** 1673–1688 CrossRef Medline
21. Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., *et al.* (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30,** 1545–1614 CrossRef Medline
22. Nakagawa, S. H., and Tager, H. S. (1986) Role of the phenylalanine B25 side chain in directing insulin interaction with its receptor: steric and conformational effects. *J. Biol. Chem.* **261,** 7332–7341 Medline
23. El Hage, K., Pandyarajan, V., Phillips, N. B., Smith, B. J., Menting, J. G., Whittaker, J., Lawrence, M. C., Meuwly, M., and Weiss, M. A. (2016) Extending halogen-based medicinal chemistry to proteins: iodo-insulin as a case study. *J. Biol. Chem.* **291,** 27023–27041 CrossRef Medline
24. Lieblich, S. A., Fang, K. Y., Cahn, J. K., Rawson, J., LeBon, J., Ku, H. T., and Tirrell, D. A. (2017) 4S-Hydroxylation of insulin at Pro^{B28} accelerates hexamer dissociation and delays fibrillation. *J. Am. Chem. Soc.* **139,** 8384–8387 CrossRef Medline
25. Ciszak, E., and Smith, G. D. (1994) Crystallographic evidence for dual coordination around zinc in the T_3R_3 human insulin hexamer. *Biochemistry* **33,** 1512–1517 CrossRef Medline
26. Riemen, M. W., Pon, L. A., and Carpenter, F. H. (1983) Preparation of semisynthetic insulin analogues from bis(tert-butyloxycarbonyl)-des-octapeptide-insulin phenylhydrazide: importance of the aromatic region B24–B26. *Biochemistry* **22,** 1507–1515 CrossRef Medline
27. Roy, M., Brader, M. L., Lee, R. W., Kaarsholm, N. C., Hansen, J. F., and Dunn, M. F. (1989) Spectroscopic signatures of the T to R conformational transition in the insulin hexamer. *J. Biol. Chem.* **264,** 19081–19085 Medline
28. Brems, D. N., Alter, L. A., Beckage, M. J., Chance, R. E., DiMarchi, R. D., Green, L. K., Long, H. B., Pekar, A. H., Shields, J. E., and Frank, B. H. (1992) Altering the association properties of insulin by amino acid replacement. *Protein Eng.* **5,** 527–533 CrossRef Medline
29. Birnbaum, D. T., Kilcomons, M. A., DeFelippis, M. R., and Beals, J. M. (1997) Assembly and dissociation of human insulin and $Lys^{B28}Pro^{B29}$-insulin hexamers: a comparison study. *Pharm. Res.* **14,** 25–36 CrossRef Medline
30. Gillies, P. S., Figgitt, D. P., and Lamb, H. M. (2000) Insulin glargine. *Drugs* **59,** 253–260; discussin 261–262 CrossRef Medline
31. Chothia, C., Lesk, A. M., Dodson, G. G., and Hodgkin, D. C. (1983) Transmission of conformational change in insulin. *Nature* **302,** 500–505 CrossRef Medline
32. Dodson, G. G., Dodson, E. J., Turkenburg, J. P., and Bing, X. (1993) Molecular recognition in insulin assembly. *Biochem. Soc. Trans.* **21,** 609–614 CrossRef Medline
33. Pandyarajan, V., Phillips, N. B., Cox, G. P., Yang, Y., Whittaker, J., Ismail-Beigi, F., and Weiss, M. A. (2014) Biophysical optimization of a therapeutic protein by non-standard mutagenesis: studies of an iodo-insulin derivative. *J. Biol. Chem.* **289,** 23367–23381 CrossRef Medline
34. Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) Application of circular dichroism to study RNA folding transitions. *Methods Enzymol.* **317,** 393–409 CrossRef Medline
35. Holleman, F., and Hoekstra, J. B. (1997) Insulin lispro. *N. Engl. J. Med.* **337,** 176–183 CrossRef Medline
36. Hua, Q. X., Jia, W., and Weiss, M. A. (2011) Conformational dynamics of insulin. *Front. Endocrinol.* **2,** 48 Medline
37. El Hage, K., Bereau, T., Jakobsen, S., and Meuwly, M. (2016) Impact of quadrupolar electrostatics on atoms adjacent to the σ-hole in condensed-phase simulations. *J. Chem. Theory Comp.* **12,** 3008–3019 CrossRef
38. Sikosek, T., and Chan, H. S. (2014) Biophysics of protein evolution and evolutionary protein biophysics. *J. R. Soc. Interface* **11,** 20140419 Medline
39. Baron, R., and McCammon, J. A. (2013) Molecular recognition and ligand association. *Annu. Rev. Phys. Chem.* **64,** 151–175 CrossRef Medline
40. Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A., and Vijayan, M. (1971) Atomic positions in rhombohedral 2-zinc insulin crystals. *Nature* **231,** 506–511 CrossRef Medline
41. Chang, X., Jorgensen, A. M., Bardrum, P., and Led, J. J. (1997) Solution structures of the R6 human insulin hexamer. *Biochemistry* **36,** 9409–9422 CrossRef Medline
42. Jacoby, E., Hua, Q. X., Stern, A. S., Frank, B. H., and Weiss, M. A. (1996) Structure and dynamics of a protein assembly: ^{1}H-NMR studies of the 36 kDa R_6 insulin hexamer. *J. Mol. Biol.* **258,** 136–157 CrossRef Medline
43. Keller, D., Clausen, R., Josefsen, K., and Led, J. J. (2001) Flexibility and bioactivity of insulin: an NMR investigation of the solution structure and folding of an unusually flexible human insulin mutant with increased biological activity. *Biochemistry* **40,** 10732–10740 CrossRef Medline
44. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., and Snel, L. (1988) Monomeric insulins obtained by protein engineering and their medical implications. *Nature* **333,** 679–682 CrossRef Medline
45. Weiss, M. A., and Lawrence, M. C. (2018) A thing of beauty: structure and function of insulin's "aromatic triplet." *Diabetes Obes. Metab.* **20,** 51–63
46. Sinnokrot, M. O., and Sherrill, C. D. (2006) High-accuracy quantum mechanical studies of π-π interactions in benzene dimers. *J. Phys. Chem. A* **110,** 10656–10668 CrossRef Medline
47. Conlon, J. M. (2001) Evolution of the insulin molecule: insights into structure-activity and phylogenetic relationships. *Peptides* **22,** 1183–1193 CrossRef Medline
48. Jekl, V., Hauptman, K., and Knotek, Z. (2011) Diseases in pet degus: a retrospective study in 300 animals. *J. Small Anim. Prac.* **52,** 107–112 CrossRef
49. Brange, J., and Havelund, S. (1984) Stabilized insulin preparations and method for their production. in U. S. Patent and Trademark Office patent full-text database, A61K 3726 Ed., Novo Industries A/S

50. Jonassen, I., Havelund, S., Hoeg-Jensen, T., Steensgaard, D. B., Wahlund, P. O., and Ribel, U. (2012) Design of the novel protraction mechanism of insulin degludec, an ultra-long-acting basal insulin. *Pharm. Res.* **29,** 2104–2114 CrossRef Medline
51. Kurtzhals, P., Havelund, S., Jonassen, I., Kiehr, B., Larsen, U. D., Ribel, U., and Markussen, J. (1995) Albumin binding of insulins acylated with fatty acids: characterization of the ligand-protein interaction and correlation between binding affinity and timing of the insulin effect *in vivo*. *Biochem. J.* **312,** 725–731 CrossRef Medline
52. Chodera, J. D., and Mobley, D. L. (2013) Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.* **42,** 121–142 CrossRef Medline
53. Birnbaum, D. T., Dodd, S. W., Saxberg, B. E., Varshavsky, A. D., and Beals, J. M. (1996) Hierarchical modeling of phenolic ligand binding to 2Zn-insulin hexamers. *Biochemistry* **35,** 5366–5378 CrossRef Medline
54. Hua, Q. X., Nakagawa, S., Hu, S. Q., Jia, W., Wang, S., and Weiss, M. A. (2006) Toward the active conformation of insulin: stereospecific modulation of a structural switch in the B chain. *J. Biol. Chem.* **281,** 24900–24909 CrossRef Medline
55. Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., and Mercola, D. A. (1972) Three-dimensional atomic structure of insulin and its relationship to activity. *Diabetes* **21,** 492–505 CrossRef Medline
56. Markussen, J., Hougaard, P., Ribel, U., Sørensen, A. R., and Sørensen, E. (1987) Soluble, prolonged-acting insulin derivatives: I. degree of protraction and crystallizability of insulins substituted in the termini of the B-chain. *Protein Eng.* **1,** 205–213 CrossRef Medline
57. Weiss, M. A. (2009) The structure and function of insulin: decoding the TR transition. *Vitam. Horm.* **80,** 33–49 CrossRef Medline
58. Markussen, J., Diers, I., Engesgaard, A., Hansen, M. T., Hougaard, P., Langkjaer, L., Norris, K., Ribel, U., Søorensen, A. R., and Sørensen, E. (1987) Soluble, prolonged-acting insulin derivatives: II. degree of protraction and crystallizability of insulins substituted in positions A17, B8, B13, B27, and B30. *Protein Eng.* **1,** 215–223 CrossRef Medline
59. Bentley, G. A., Brange, J., Derewenda, Z., Dodson, E. J., Dodson, G. G., Markussen, J., Wilkinson, A. J., Wollmer, A., and Xiao, B. (1992) Role of B13 Glu in insulin assembly: the hexamer structure of recombinant mutant (B13 Glu → Gln) insulin. *J. Mol. Biol.* **228,** 1163–1176 CrossRef Medline
60. Mirmira, R. G., Nakagawa, S. H., and Tager, H. S. (1991) Importance of the character and configuration of residues B24, B25, and B26 in insulin-receptor interactions. *J. Biol. Chem.* **266,** 1428–1436 Medline
61. Xu, B., Hu, S. Q., Chu, Y. C., Huang, K., Nakagawa, S. H., Whittaker, J., Katsoyannis, P. G., and Weiss, M. A. (2004) Diabetes-associated mutations in insulin: consecutive residues in the B chain contact distinct domains of the insulin receptor. *Biochemistry* **43,** 8356–8372 CrossRef Medline
62. Huang, K., Xu, B., Hu, S. Q., Chu, Y. C., Hua, Q. X., Qu, Y., Li, B., Wang, S., Wang, R. Y., Nakagawa, S. H., Theede, A. M., Whittaker, J., De Meyts, P., Katsoyannis, P. G., and Weiss, M. A. (2004) How insulin binds: the B-chain α-helix contacts the L1 β-helix of the insulin receptor. *J. Mol. Biol.* **341,** 529–550 CrossRef Medline
63. Hu, S. Q., Burke, G. T., and Katsoyannis, P. G. (1993) Contribution of the B16 and B26 tyrosine residues to the biological activity of insulin. *J. Protein Chem.* **12,** 741–747 CrossRef Medline
64. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., and Reynolds, C. D. (1988) The structure of 2Zn pig insulin crystals at 1.5-Å resolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **319,** 369–456 CrossRef Medline
65. Guvench, O., and Brooks, C. L., 3rd. (2005) Tryptophan side chain electrostatic interactions determine edge-to-face *vs* parallel-displaced tryptophan side chain geometries in the designed β-hairpin "trpzip2." *J. Am. Chem. Soc.* **127,** 4668–4674 CrossRef Medline
66. Samanta, U., Pal, D., and Chakrabarti, P. (1999) Packing of aromatic rings against tryptophan residues in proteins. *Acta Crystallogr. D* **55,** 1421–1427 CrossRef Medline
67. Holst, B., Nygaard, R., Valentin-Hansen, L., Bach, A., Engelstoft, M. S., Petersen, P. S., Frimurer, T. M., and Schwartz, T. W. (2010) A conserved aromatic lock for the tryptophan rotameric switch in TM-VI of seven-transmembrane receptors. *J. Biol. Chem.* **285,** 3973–3985 CrossRef Medline
68. Bhattacharyya, R., Samanta, U., and Chakrabarti, P. (2002) Aromatic-aromatic interactions in and around α-helices. *Protein Eng.* **15,** 91–100 CrossRef Medline
69. Lee, L. C., Chou, Y. L., Chen, H. H., Lee, Y. L., and Shaw, J. F. (2009) Functional role of a non-active site residue Trp(23) on the enzyme activity of *Escherichia coli* thioesterase I/protease I/lysophospholipase L(1). *Biochim. Biophys. Acta* **1794,** 1467–1473 CrossRef Medline
70. Glidden, M. D., Yang, Y., Phillips, N. B., Carr, K., Smith, N. A., Wickramasinghe, N. P., Smith, B. J., Ismail-Beigi, F., Lawrence, M. C., and Weiss, M. A. (2017) Solution structure and dynamics of a single-chain insulin analog. *J. Biol. Chem.* **293,** 69–88 Medline
71. Frederick, K. K., Marlow, M. S., Valentine, K. G., and Wand, A. J. (2007) Conformational entropy in molecular recognition by proteins. *Nature* **448,** 325–329 CrossRef Medline
72. Dong, J., Wan, Z., Popov, M., Carey, P. R., and Weiss, M. A. (2003) Insulin assembly damps conformational fluctuations: Raman analysis of amide I line widths in native states and fibrils. *J. Mol. Biol.* **330,** 431–442 CrossRef Medline
73. Wimley, W. C., and White, S. H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Mol. Biol.* **3,** 842 CrossRef
74. Hassiepen, U., Federwisch, M., Mülders, T., and Wollmer, A. (1999) The lifetime of insulin hexamers. *Biophys. J.* **77,** 1638–1654 CrossRef Medline
75. Derewenda, U., Derewenda, Z., Dodson, G., and Brange, J. (1987) The crystal-structure of the B-12 Ile human insulin prepared by site-directed mutatgenesis. *Protein Eng.* **222,** 425–433
76. Liu, M., Wan, Z. L., Chu, Y. C., Aladdin, H., Klaproth, B., Choquette, M., Hua, Q. X., Mackin, R. B., Rao, J. S., De Meyts, P., Katsoyannis, P. G., Arvan, P., and Weiss, M. A. (2009) Crystal structure of a "non-foldable" insulin: impaired folding efficiency and ER stress despite native activity. *J. Biol. Chem.* **284,** 35259–35272 CrossRef Medline
77. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., *et al.* (2009) Gaussian 09, revision A.02., Gaussian, Inc., Wallingford, CT
78. Shoelson, S. E., Lu, Z. X., Parlautan, L., Lynch, C. S., and Weiss, M. A. (1992) Mutations at the dimer, hexamer, and receptor-binding, surfaces of insulin independently affect insulin-insulin and insulin-receptor interactions. *Biochemistry* **31,** 1757–1767 CrossRef Medline
79. O'Halloran, T. V., Kebede, M., Philips, S. J., and Attie, A. D. (2013) Zinc, insulin, and the liver: a ménage à trois. *J. Clin. Investig.* **123,** 4136–4139 CrossRef Medline
80. Michael, M. D., Kulkarni, R. N., Postic, C., Previs, S. F., Shulman, G. I., Magnuson, M. A., and Kahn, C. R. (2000) Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol. Cell* **6,** 87–97 CrossRef Medline
81. Tamaki, M., Fujitani, Y., Hara, A., Uchida, T., Tamura, Y., Takeno, K., Kawaguchi, M., Watanabe, T., Ogihara, T., Fukunaka, A. *et al.* (2013) The diabetes-susceptible gene SLC30A8/ZnT8 regulates hepatic insulin clearance. *J. Clin. Investig.* **123,** 4513–4524 CrossRef Medline
82. Forbes, B., McCarthy, P., and Norton, R. (2012) Insulin-like growth factor binding proteins: a structural perspective. *Front. Endocrinol.* **3,** 38 Medline
83. Heidenreich, K. A., Yip, C. C., Frank, B. H., and Olefsky, J. M. (1985) The preparation and characterization of mono-iodinated photoreactive analogs of insulin. *Biochem. Biophys. Res. Commun.* **126,** 1138–1145 CrossRef Medline
84. Scholfield, M. R., Zanden, C. M., Carter, M., and Ho, P. S. (2013) Halogen bonding (X-bonding): a biological perspective. *Protein Sci.* **22,** 139–152 CrossRef Medline
85. Marshall, G. R. (2013) Limiting assumptions in molecular modeling: electrostatics. *J. Comput. Aided Mol. Des.* **27,** 107–114 CrossRef Medline
86. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D.,

Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., *et al.* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616 CrossRef Medline

87. MacKerell, A. D., Brooks, B., Brooks, C. L., Nilsson, L., Roux, B., Won, Y., and Karplus, M. (2002) CHARMM: the energy function and its parameterization. in *Encyclopedia of Computational Chemistry* (Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., eds) pp. 271–277, John Wiley & Sons, Ltd., Chichester, United Kingdom
88. Burley, S., and Petsko, G. (1986) Dimerization energetics of benzene and aromatic amino acid side chains. *J. Am. Chem. Soc.* **108,** 7995–8001 CrossRef
89. Kamerlin, S. C., Haranczyk, M., and Warshel, A. (2009) Progress in ab initio QM/MM free-energy simulations of electrostatic energies in proteins: accelerated QM/MM studies of pK_a, redox reactions and solvation free energies. *J. Phys. Chem. B* **113,** 1253–1272 Medline
90. Reha, D., Kabelác, M., Ryjácek, F., Sponer, Sponer, J. E., Suhai, M., and Hobza, P. (2002) Intercalators: 1. nature of stacking interactions between intercalators (ethidium, daunomycin, ellipticine, and 4′,6-diaminide-2-phenylindole) and DNA base pairs: *ab initio* quantum chemical, density functional theory, and empirical potential study. *J. Am. Chem. Soc.* **124,** 3366–3376 CrossRef Medline
91. Steward, L. E., and Chamberlin, A. R. (1998) Protein engineering with non-standard amino acids. *Methods Mol. Biol.* **77,** 325–354 Medline
92. Dougherty, D. A. (2000) Unnatural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* **4,** 645–652 CrossRef Medline
93. van Gunsteren, W. F., Daura, X., Hansen, N., Mark, A. E., Oostenbrink, C., Riniker, S., and Smith, L. J. (2018) Validation of molecular simulation: an overview of issues. *Angew. Chem. Int. Ed.* **57,** 884–902 CrossRef
94. Aleksandrov, A., Thompson, D., and Simonson, T. (2010) Alchemical free energy simulations for biological complexes: powerful but temperamental. *J. Mol. Recognit.* **23,** 117–127 Medline
95. Kannan, N., and Vishveshwara, S. (2000) Aromatic clusters: a determinant of thermal stability of thermophilic proteins. *Protein Eng.* **13,** 753–761 CrossRef Medline
96. Szyperski, T., Luginbühl, P., Otting, G., Güntert, P., and Wüthrich, K. (1993) Protein dynamics studied by rotating frame nitrogen-15 spin relaxation times. *J. Biomol. NMR* **3,** 151–164 Medline
97. Freire, E. (2009) A thermodynamic approach to the affinity optimization of drug candidates. *Chem. Biol. Drug Des.* **74,** 468–472 CrossRef Medline
98. Cryer, P. E. (2002) Hypoglycaemia: the limiting factor in the glycaemic management of type I and type II diabetes. *Diabetologia* **45,** 937–948 CrossRef Medline
99. Christensen, M., and Knop, F. K. (2010) Once-weekly GLP-1 agonists: how do they differ from exenatide and liraglutide? *Curr. Diab. Rep.* **10,** 124–132 CrossRef Medline
100. Wan, Z. L., Huang, K., Hu, S. Q., Whittaker, J., and Weiss, M. A. (2008) The structure of a mutant insulin uncouples receptor binding from protein allostery: an electrostatic block to the TR transition. *J. Biol. Chem.* **283,** 21198–21210 CrossRef Medline
101. Inouye, K., Watanabe, K., Morihara, K., Tochino, Y., Kanaya, T., Emura, J., and Sakakibara, S. (1979) Enzyme-assisted semisynthesis of human insulin. *J. Am. Chem. Soc.* **101,** 751–752 CrossRef
102. Rahuel-Clermont, S., French, C. A., Kaarsholm, N. C., Dunn, M. F., and Chou, C. I. (1997) Mechanisms of stabilization of the insulin hexamer through allosteric ligand interactions. *Biochemistry* **36,** 5837–5845 CrossRef Medline
103. 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., McCoy, A., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67,** 235–242 CrossRef Medline
104. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 CrossRef Medline
105. Joosten, R. P., Womack, T., Vriend, G., and Bricogne, G. (2009) Re-refinement from deposited X-ray data can deliver improved models for most PDB entries. *Acta Crystallogr. D* **65,** 176–185 CrossRef Medline
106. Wang, Z. X. (1995) An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.* **360,** 111–114 CrossRef Medline
107. Willighagen, E., and Howard, M. (2007) Fast and scriptable molecular graphics in web browsers without Java3D. *Nat. Proceed.* CrossRef

Structure-based stabilization of insulin as a therapeutic protein assembly via enhanced aromatic–aromatic interactions

Nischay K. Rege, Nalinda P. Wickramasinghe, Alisar N. Tustan, Nelson F. B. Phillips, Vivien C. Yee, Faramarz Ismail-Beigi and Michael A. Weiss

J. Biol. Chem. 2018, 293:10895-10910.
doi: 10.1074/jbc.RA118.003650 originally published online June 7, 2018

Stabilization of Insulin as a Therapeutic Protein Assembly via Enhanced Aromatic-Aromatic Interactions

Nischay K. Rege, Nalinda P. Wickramasinghe, Alisar N. Tustan, Nelson B. Phillips, Vivien C. Yee, Faramarz Ismail-Beigi, & Michael A. Weiss

Purpose of Supplement

We provide 12 Supplemental Figures and 12 Supplemental Tables as cited in the main text. The information pertains to five aspects: (i) tabulation of *ab initio* and molecular mechanics calculations; (ii) details of the crystal structure of Trp^{B26}, Orn^{B29}-insulin, corresponding NMR studies, and comparison with WT insulin; (iii) additional control SEC studies; (iv) additional rat studies; and (iv) details regarding the structural properties of insulin as exploited in protein engineering. The figures and tables are as follows

Table of Contents

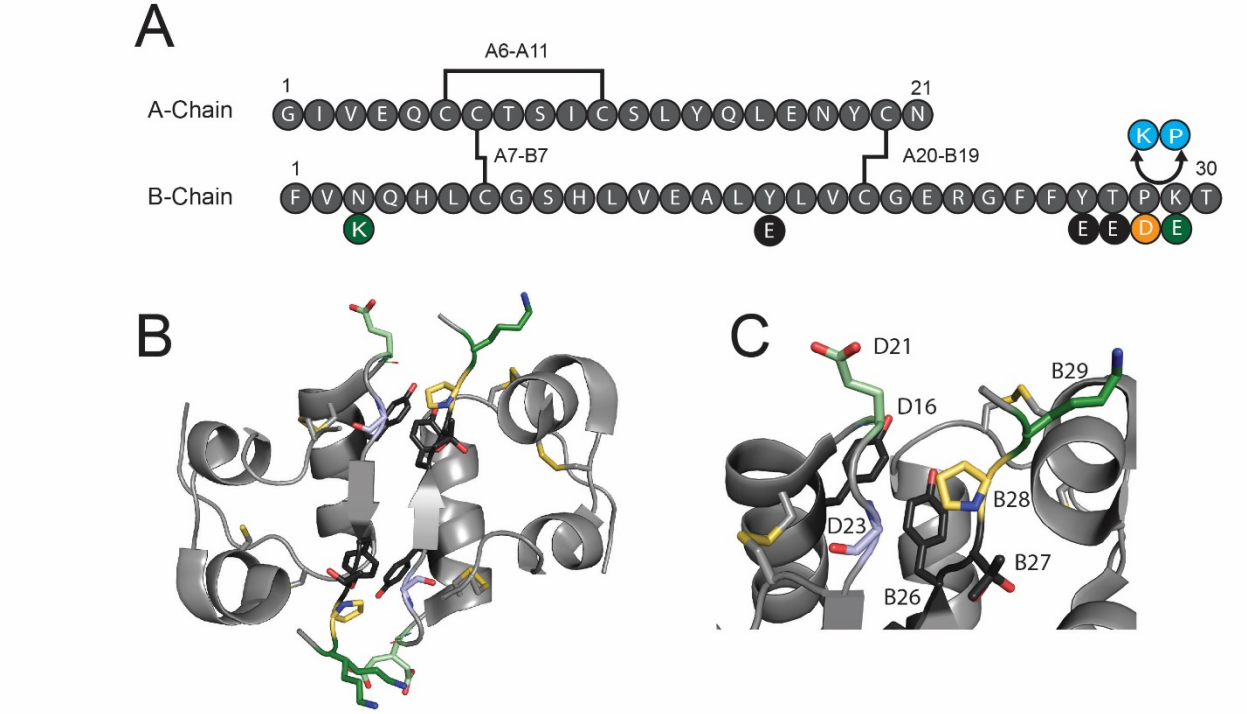

Figure S1. Structural summary of rapid-acting insulin analogs. (*A*) Representation of modifications to the sequence of insulin to create rapid-acting insulin analogs. Insulin lispro (developed by analogy to insulin-like growth factor-I [IGF-I]), highlighted the Pro^{B28}/Gly^{D23} dimer contact (highlighted in *sky blue* in panels *B* and *C*) (1). Other designs introduced polar or charged substitutions at the dimer interface. Among these were substitution of Thr^{B27}, Tyr^{B26}, or Tyr^{B16} by Glu (2,3) (shown as *black circles* in *A*). Clinical analogs insulin glulisine (Apidra®) and insulin aspart (Novolog®) contain acidic residues at positions B29 (*green*) and B28 (*yellow*), respectively (4). These residues were hypothesized to repel residue Glu^{D21}, impairing formation of dimers and higher-order oligomers (3). (*B*) The native residues of the positions listed above are highlighted in a 3D structure of an insulin dimer (PDB entry 5INS with color code as above). (*C*) Expanded view of panel *B* with residue positions as labeled.

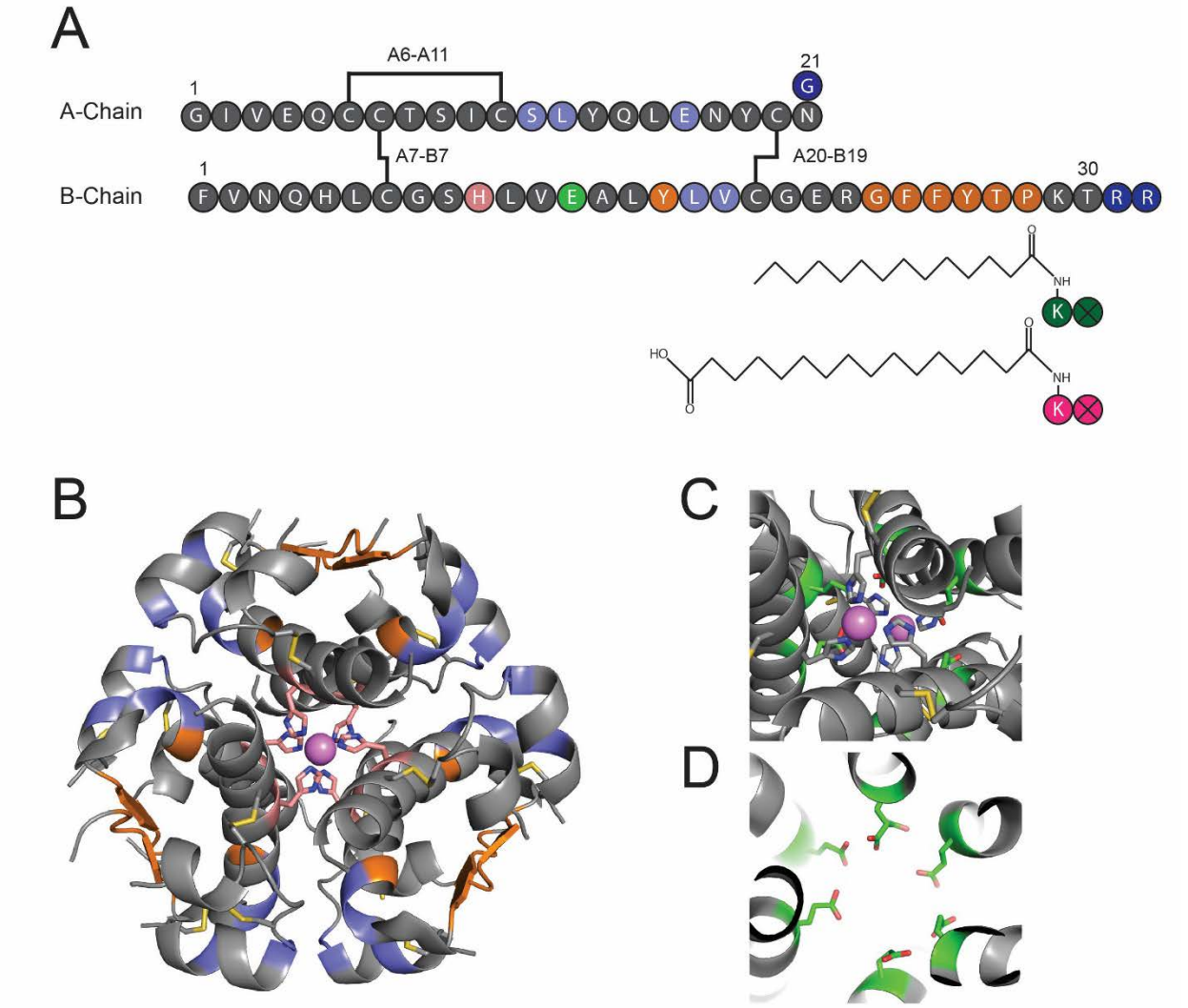

Figure S2. Past strategies to create basal insulin analogs (see also Table S10). (*A*) Sites in insulin modified in long-acting insulins. Insulin glargine (Lantus®; purple) represents the most successful product (5) with pI shifted towards neutrality due to a basic extension of the B chain. Such pI-shifted analogs are soluble as formulated at acidic but precipitate in the neutral environment of the SQ space to create a long-lived depot (6). A second class of basal insulins are derivatized with long acyl chains at Lys^{B29} (7). These include insulin detemir (Levemir®; *green circles*) and degludec (Tresiba®; *magenta circles*) (8), which bind to albumin as a "circulating depot (7)." Insulin degludec forms an acyl-linked dimer of zinc hexamers in its formulation but undergoes a switch to form an acyl-bridged, extended multihexameric linear complex in the SQ depot (8,9). These mechanisms are orthogonal to efforts to enhance the intrinsic stability of the insulin hexamer itself. Amino acid residues forming the dimer and hexamer interfaces of insulin in *orange* and *light blue*, respectively. His^{B10}, which coordinates with divalent zinc ions in hexamers, is highlighted in *pink*. (*B*) Dimer- and hexamer interfaces in the zinc insulin hexamer (T_6; color-coded as above). Analogs, designated "hydrophobic insulins," were engineered to contain hydrophobic substitutions at dimer or hexamer interfaces in an effort to stabilize the hexamer. Yet another class of analogs introduced hydrophobic substitutions to the surface of insulin hexamers to create non-specific multi-hexameric precipitates. However, the poor solubility of such analogs

limited their clinical utility (5). (*C*) Electrostatic stabilization of the insulin hexamer: view of the negatively charged ring formed by the six Glu^{B13} side chains (highlighted in *green*). An unobstructed view is shown in panel *D*. The most successful effort to stabilize the insulin hexamer involved substitution of Glu^{B13} by Gln, which eliminated repulsion within the Glu residues; however, the Gln^{B13} analog had reduced biological activity (3,10).

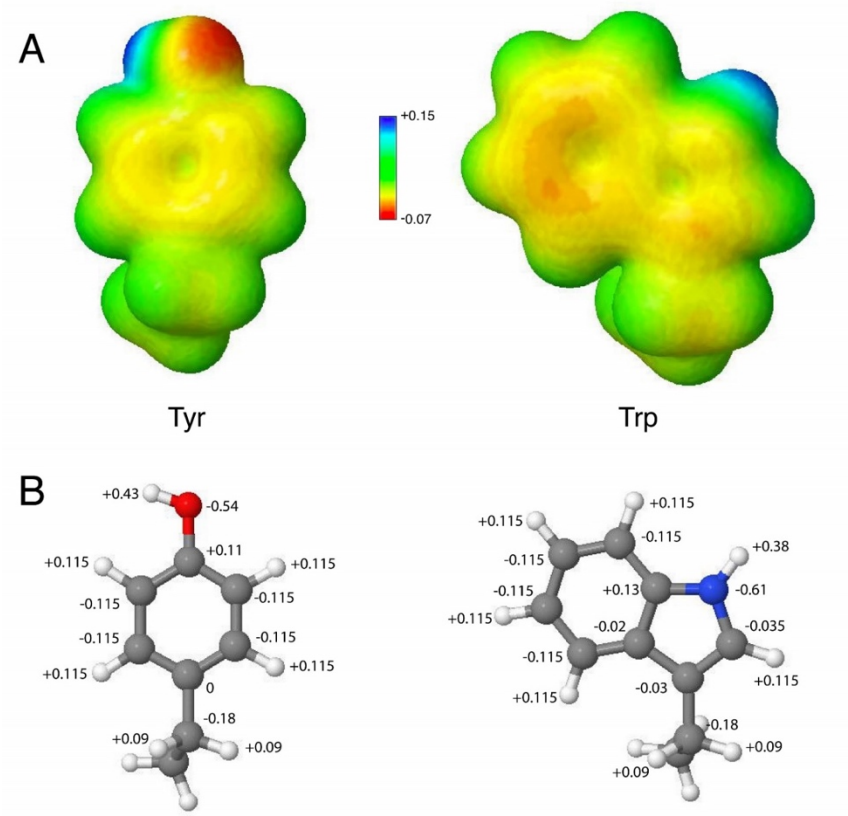

Figure S3. Comparison of *Ab initio* electrostatics and CHARMM parameters of Tyr and Trp side chains. (*A*) Isosurface representation of electron density and molecular electrostatic potential (MEP) map of Tyr (*left*) and Trp (*right*). Electron density and MEP were calculated using B3LYP method and 6-31G(d) basis set using Gaussian utility Cubegen on Gaussian 09 (11). The isosurface map was then generated using Jmol (12). (*B*) Ball-and-stick models of Tyr (*left*) and Trp (*right*) side chains. Point charges of each atom as implemented in CHARMM22 are indicated (13,14).

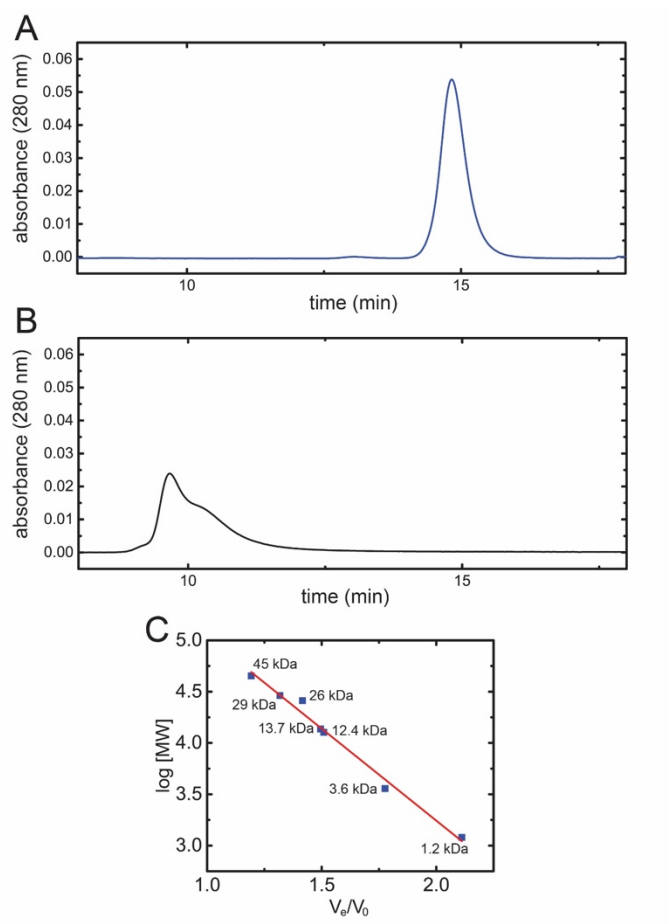

Figure S4. Solvent exclusion chromatography (SEC). Elution times of insulin in monomeric and hexameric states in SEC provide context for species identified in Figure 4. (*A*) SEC profile of monomeric lispro (Zn^{2+}-and phenol-free). The molecular weight (MW; or molecular mass) of the species was 3.1 kDa (calculated as in Fig. 4). (*B*) WT insulin formulated with 0.3 mM $ZnCl_2$ and phenol was run in a mobile-phase containing 50 mM cyclohexanol and 0.3 mM $ZnCl_2$. The sample eluted as a hexamer (Calculated molecular mass 48 kDa) with dissociation intermediates constituting the "tail" of the peak. (*C*) Calibration plot of SEC column with mobile phase used in panel *B*. Linear fit (*red line*) of log(MW) to V_e/V_0 of MW standards (*blue squares*). The equation of the line is log(MW) = -1.79 (V_e/V_0) +6.83 (R^2 = 0.986). The following standards were used for calibration: thyroglobulin (669 kDa, V_0), ovalbumin (45 kDa), carbonic anhydrase (29 kDa), elastase (26 kDa), ribonuclease A (13.7 kDa), cytochrome C (12.3 kDa) and synthetic peptides (3.6 and 1.2 kDa).

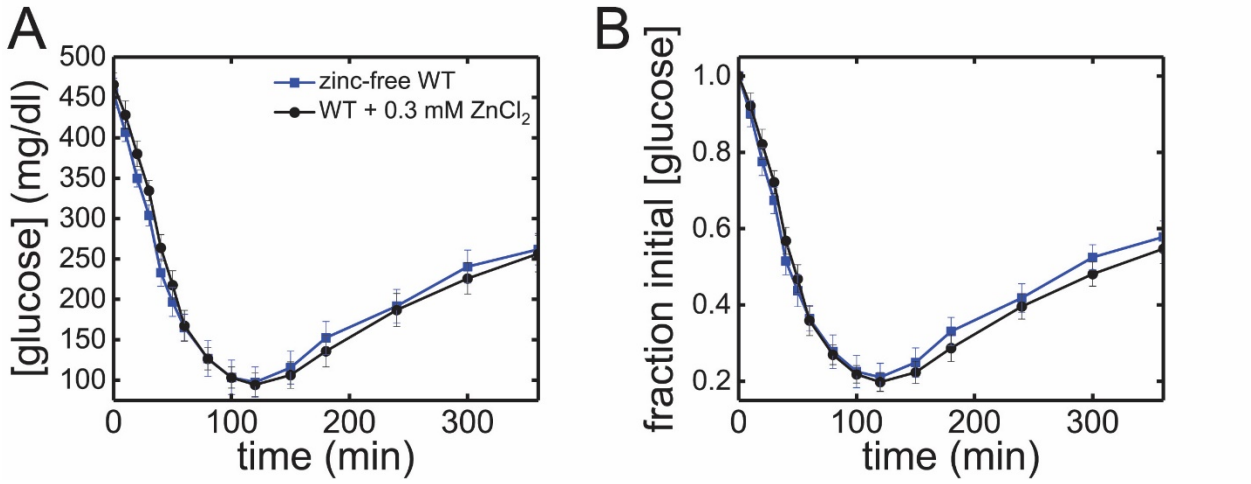

Figure S5. Effects of Zn^{2+} on PD profile of WT insulin. (*A*) Time course of [blood glucose] after SQ injection of WT insulin formulated in absence of Zn^{2+} (*blue*; N=8) or in presence of 0.3 mM $ZnCl_2$ (*black*; N=9). (*B*) Normalized curves from panel *A*.

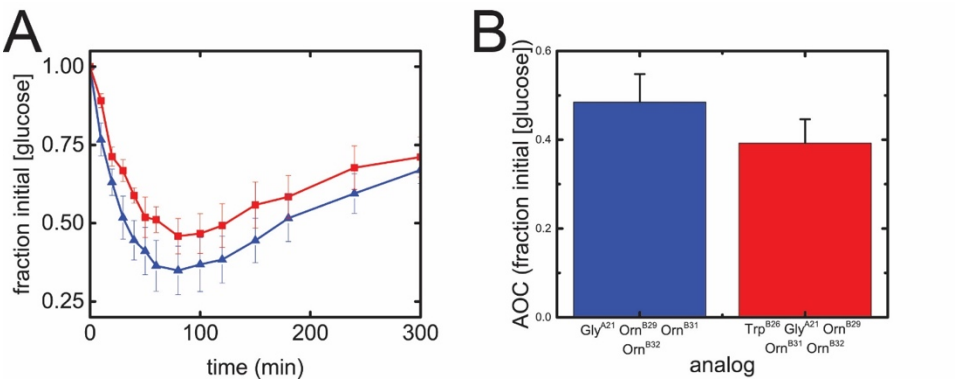

Figure S6. (*A*) Time course of [glucose] after IV injection of parent pI-shifted insulin analog (blue; N=6) and its Trp^{B26} derivative (red; N=6). (*B*) Bar graph showing area over curve (AOC) of curves from panel *A*. *Black bars* indicate S.E.M. The Trp^{B26} derivative displayed 82 ± 6% potency relative to the parent analog but is a complete agonist on injection of higher doses

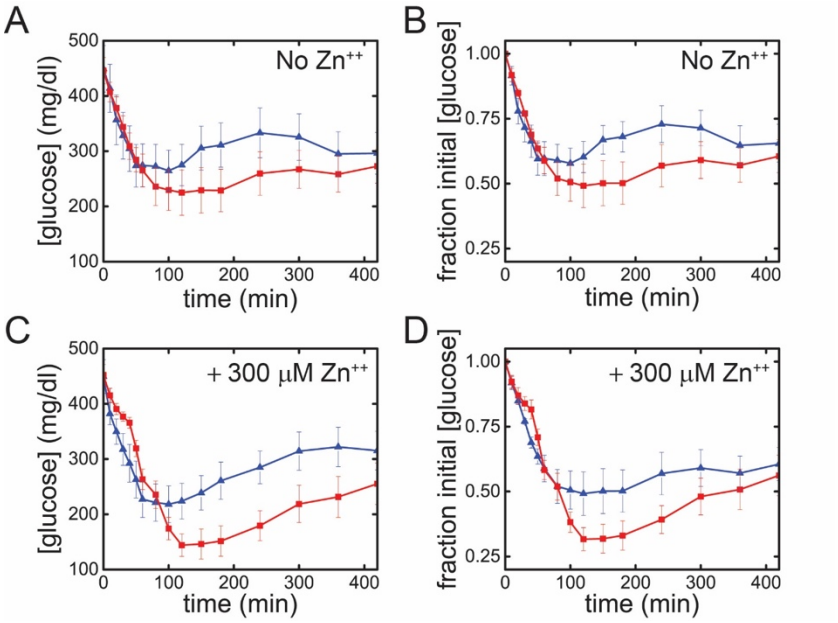

Figure S7. (*A*) Time course of [blood glucose] after SQ injection of zinc-free parent pI-shifted insulin analog (*blue*; N=6) and Trp^{B26} derivative (*red*; N=6). Normalized data are shown in *B*. (*C-D*) Corresponding plots of analogs administered after formulation in presence of 0.3 mM $ZnCl_2$, color-coded as above (N=6).

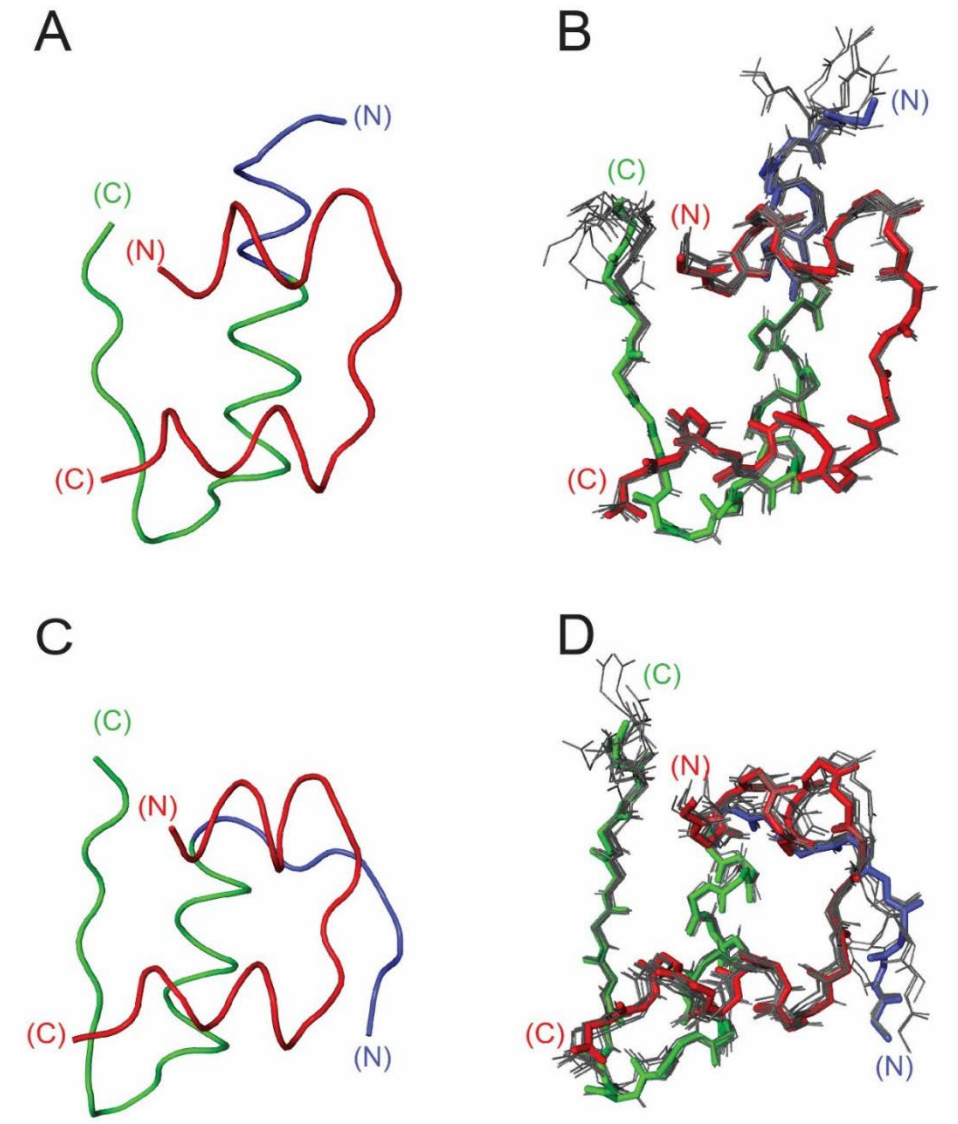

Figure S8. Comparison of the structure of Trp^{B26}, Orn^{B29}-insulin to a collection of WT insulin structures. (*A*) R^f-state protomer of Trp^{B26}, Orn^{B29}-insulin. The A chain is shown in *red*, and the B chain in *blue* (B1-B8) or *green* (B9-B30). (*B*) R^f-state protomer of Trp^{B26}, Orn^{B29}-insulin (*sticks*, color coded as above) in relation to extensive set of crystal structures of insulin and insulin analogs (PDB entries: 1BEN, 1G7A, 1RWE, 1EV3, 1EV6, 1MPJ, 1TRZ, 1TYL, 1MPJ, 1ZEG, 1ZNJ and 1ZNI; *gray sticks*). Structures are aligned with respect to the main-chain atoms of residues A1-A21 and B3- B28. RMSD values are given in Table S3. (*C-D*) Corresponding representations of the T-state protomer of Trp^{B26}, Orn^{B29}-insulin, color code as above. PDB entries used for alignment are as follows: 1APH, 1DPH, 1BEN, 1MPJ, 1TRZ, 1TYL, 1TYM, 1RWE, 1G7A, 1ZNI, 2INS and 4INS. RMSD values are given in Table S4.

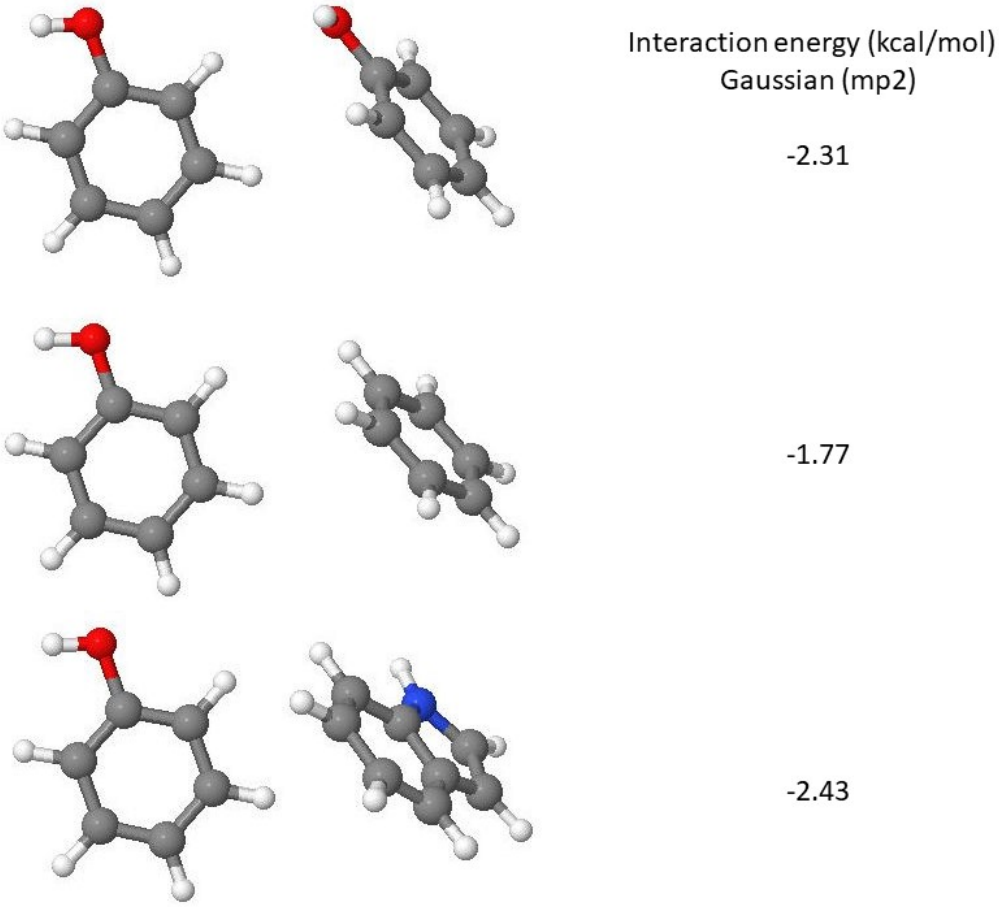

Figure S9. *Ab initio* calculations of energy of interaction between pairs of isolated aromatic molecules: phenol-phenol (*top*), phenol-benzene (*middle*), and phenol-indole (*bottom*). The phenol-indole pair was determined to form the most stable ETF interaction as a result of Van der Waals forces. Interaction energies were calculated using MP2 method and aug-cc-pVDZ basis set using Gaussian 09 (15).

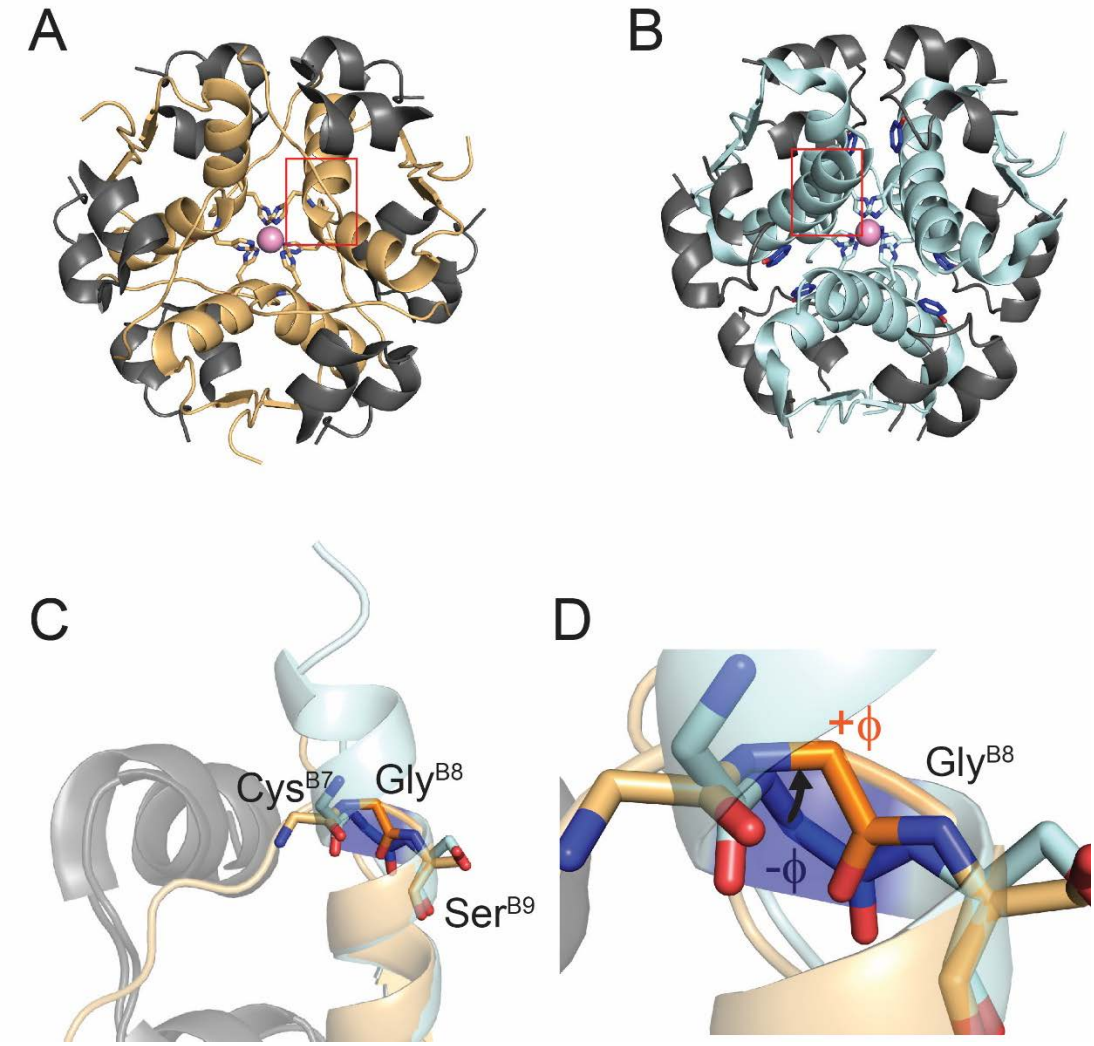

Figure S10. Structural representation of the classical T→R transition of the zinc insulin hexamer. (*A*) T_6 insulin hexamer. The eight N-terminal residues of the B chain (*orange*) are in an extended conformation (*red box*). (*B*) R_6 hexamer is stabilized by bound phenolic ligands (blue). Residues B1-B8 form an α-helix (*red box*). (*C*) Gly^{B8} serves as the pivot point of the transition between the R- (*blue*) and T-states (*orange*). (*D*) Gly^{B8} adopts a right-handed conformation (*i.e.*, with positive φ angle) in the T-state (*orange sticks*) and a left-handed conformation (negative φ angle) in the R-state (*blue sticks*).

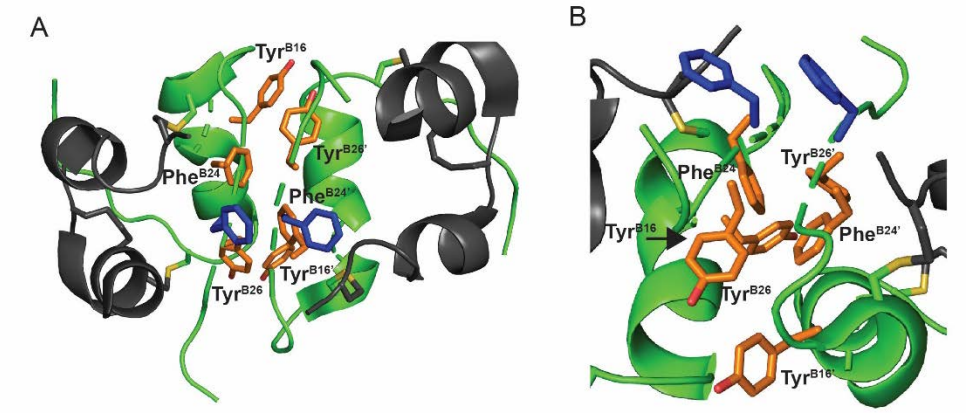

Figure S11. Ribbon model displaying the orientation of Phe^{B25} in the insulin dimer (PDB 4INS) in two views (*A, B*). The A chain is shown as dark gray ribbon, and the B chain is shown as *green ribbon*. Phe^{B25} (*blue sticks*) is peripheral to the aromatic network formed by Tyr^{B16}, Phe^{B24}, Tyr^{B26} and their symmetry-related mates (*orange sticks*).

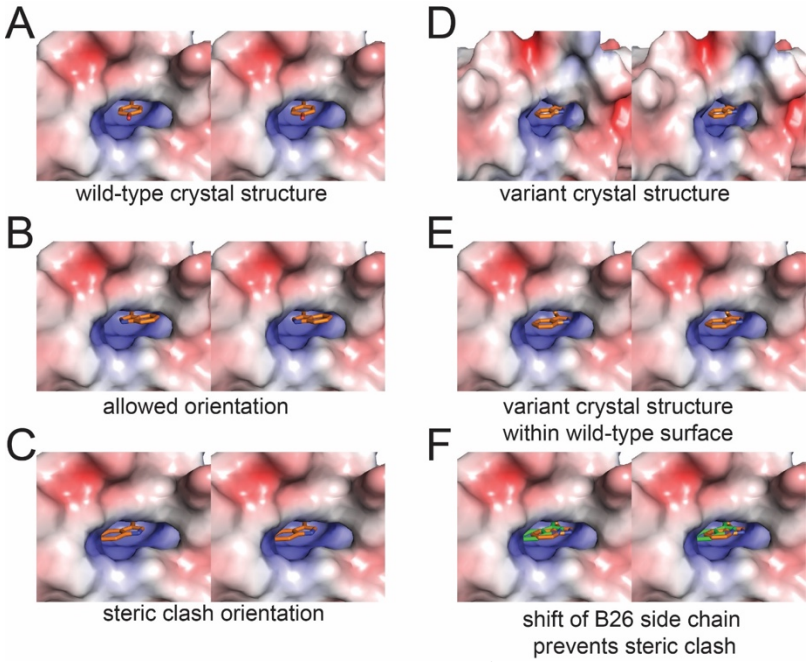

Figure S12. Comparison of B26 crevice within TR^{f} dimers of Trp^{B26}, Orn^{B29}-insulin to WT (1TRZ). (*A*) Stereo view of Tyr^{B26} (*sticks*) of 1TRZ T-state protomer within an electrostatic potential surface (generated using APBS plug-in to Pymol® (16)) formed by the surrounding residues (Fig. 1*D, E*). Positively-charged surfaces are represented in *blue*, negatively charged surfaces in *red*, and neutral surfaces in *white*. (*B, C*) Two possible orientations of the Trp^{B26} that retain χ_1 and χ_2 angles of the WT structure shown in panel *A*. Due to the asymmetric structure of the Trp side chain, two possible χ_2 angles of correspond in principle to the native Tyr. However, Trp^{B26} encounters a steric clash with residues in the core of insulin in the orientation shown in *C*. (*D*) Trp^{B26} of the R^{f}-state protomer from the Trp^{B26}, Orn^{B29}-insulin crystal structure depicted within an electrostatic potential surface formed by surrounding residues. (*E*) Depiction of Trp^{B26} from panel *D* within the B26 crevice from panel *A* (WT). The Trp^{B26} side chain does not encounter a steric clash. (*F*) Alignment of the naïve model of Trp^{B26} from panel *C* (*green sticks*) to the Trp^{B26},

Orn^{B29}-insulin structure (*orange sticks*). Residues are depicted within the WT crevice (*panels A, B, C,* and *E*). The steric clash predicted in the naïve model is mitigated in the Trp^{B26}, Orn^{B29}-insulin structure by (a) a local shift (0.2 Å) in the backbone of the C-terminal B-chain and (b) a slight difference in the χ_1 torsion angle of Trp^{B26}.

Table S1a. Interaction energy between the B26 residue of an energy minimized model of Trp^{B26} insulin relative to native Tyr^{B26} in the context of a TR^{f} dimer.

Residue	Interacting Residue	Total Energy of Interaction (kcal/mol)	
		1TRZ Tyr^{B26}	1TRZ Trp^{B26}
B26	Phe^{B24}	-0.344	-0.502
	Phe^{B25}	-0.100	-0.113
	Tyr^{B16}	-0.006	0.007
	Phe^{D24}	-0.887	-1.465
	Phe^{D25}	-0.029	-0.120
	Tyr/Trp^{D26}	0.000	-0.032
	Tyr^{D16}	-1.254	-1.672
D26	Phe^{B24}	-0.828	-1.436
	Phe^{B25}	0.005	-0.025
	Tyr/Trp^{B26}	0.000	-0.032
	Tyr^{B16}	-1.594	-0.503
	Phe^{D24}	-0.524	-0.699
	Phe^{D25}	-0.036	-0.084
	Tyr^{D16}	0.006	-0.005
Total		-6.986	-7.811

Table S1b. Local energy-minimized model of Trp^{B26} insulin relative to native Tyr^{B26} where polar atoms of Tyrosine and Tryptophan rings are replaced with nonpolar synthetic residues.

Residue	Interacting Residue	Total Energy of Interaction (kcal/mol)	
		1TRZ Tyo^{B26}	1TRZ Tro^{B26}
B26	Phe^{B24}	-0.358	-0.503
	Phe^{B25}	-0.090	-0.138
	Tyr^{B16}	0.006	0.009
	Phe^{D24}	-0.882	-1.501
	Phe^{D25}	-0.054	-0.095
	Tyr^{D26}	0.000	0.001
	Tyr^{D16}	-1.416	-1.527
D26	Phe^{B24}	-0.821	-1.478
	Phe^{B25}	0.005	-0.039
	Tyr^{B26}	0.000	-0.001
	Tyr^{B16}	-0.938	-0.209
	Phe^{D24}	-0.566	-0.708
	Phe^{D25}	-0.049	-0.061
	Tyr^{D16}	0.006	0.003
		-6.531	-7.438

[a] Tabulation of non-bonded interaction energy at the insulin dimer interface in the energy-minimized naïve model of Tyr^{B26}→Trp displays interactions most improved by the substitution.

[b] This table demonstrates the impact of π-π interactions on non-bonded interaction energy across the insulin dimer interface. Polar oxygen and nitrogen atoms are replaced, *in silico*, with non-polar synthetic atoms eliminating the confounding effects of polar interactions involving those atoms on free-energy calculations.

Table S1c. Interaction energy between the B26 residue of a local energy-minimized model of Trp^{B26} insulin relative to native Tyr^{B26} in the context of an R_2 dimer.

Residue	Interacting Residue	Total Energy of Interaction (kcal/mol)	
		1ZNJ Tyr^{B26}	1ZNJ Trp^{B26}
B26	Phe^{B24}	-0.501	-0.571
	Phe^{B25}	-0.142	-0.146
	Tyr^{B16}	0.008	0.008
	Phe^{D24}	-1.032	-1.308
	Phe^{D25}	-0.002	-0.027
	Tyr/Trp^{D26}	-0.011	-0.003
	Tyr^{D16}	-1.569	-1.996
D26	Phe^{B24}	-0.738	-1.283
	Phe^{B25}	0.004	-0.034
	Tyr/Trp^{B26}	-0.011	-0.003
	Tyr^{B16}	-0.613	-1.326
	Phe^{D24}	-0.683	-0.635
	Phe^{D25}	-0.120	-0.145
	Tyr^{D16}	0.004	0.008
Total		-5.406	-7.461

Table S1d. Interaction energy between the B26 residue of a local energy-minimized model of Trp^{B26} insulin relative to native Tyr^{B26} in the context of an T_2 dimer.

Residue	Interacting Residue	Total Energy of Interaction (kcal/mol)	
		4INS Tyr^{B26}	4INS Trp^{B26}
B26	Phe^{B24}	-0.397	-0.549
	Phe^{B25}	-0.027	-0.058
	Tyr^{B16}	0.003	0.006
	Phe^{D24}	-1.405	-1.181
	Phe^{D25}	-0.008	-0.017
	Tyr/Trp^{D26}	-0.008	-0.054
	Tyr^{D16}	-1.643	-1.868
D26	Phe^{B24}	-1.242	-0.679
	Phe^{B25}	-0.016	-0.022
	Tyr/Trp^{B26}	-0.016	-0.054
	Tyr^{B16}	-2.710	-1.261
	Phe^{D24}	-0.365	-0.466
	Phe^{D25}	-0.049	-0.140
	Tyr^{D16}	0.004	0.006
Total		-7.879	-6.337

Table S2. Data collection and refinement statistics pertaining to the crystal structure of Trp^{B26}, Orn^{B29}-insulin.

Data Collection	
Source	SSRL BL7-1
wavelength (Å)	0.9795
space group	R3
Unit-cell dimensions	a = 79.7552 Å, b = 79.7552 Å, c = 37.61 Å, α = 90°, β = 90°, γ = 120°
resolution (Å)	33.04-2.25
R_{merge} (%)	3.698 (10.28)
$I/\sigma I$	45.78 (8.99)
completeness (%)	99.5 (99.8)
redundancy	10.1 (9.6)
no. of reflections/used in refinement	4273/4226
Refinement	
resolution (Å)	33.04-2.25
R-factor/R_{Free} (%)	16.8/24.9
r.m.s.d. bond length (Å)	0.036
r.m.s.d. bond angle (°)	.971
median B-factors ($Å^2$)	
chain A	28
chain B	22
chain C	27
chain D	18
Ramachandran plot (%)	
most favored region	100
additionally allowed regions	0
generously allowed regions	0
disallowed regions	0

Table S3. RMSD values of main-chain and side-chains of the R-state protomer in the crystal structures of Trp^{B26}, Orn^{B29}-insulin and selected WT R_6 and T_3R_3 hexamers.

PDB ID	main-chain RMSD (Å)	Side-chain RMSD (Å)
1RWE R^f $(T_3R_3)^a$	0.70	1.25
1BEN R^f (T_3R_3)	0.39	0.97
1TRZ R^f (T_3R_3)	0.43	0.95
1TYL R^f (T_3R_3)	0.46	1.04
1TYM R^f (T_3R_3)	0.44	1.10
1ZNI R^f (T_3R_3)	0.45	0.89
1MPJ R^f (T_3R_3)	0.45	0.91
3MTH R^f (T_3R_3)	0.52	1.18
1LPH R^f (T_3R_3)	0.53	1.07
1G7A $R1^f$ (rhom T_3R_3)	0.61	1.06
$R2^f$	0.79	1.53
1EV3 R1 (rhom R_6)^b	0.71	1.59
R2	0.56	1.32
1ZEG R1 (rhom R_6)	0.52	1.35
R2	0.61	1.82
1ZEH R1 (rhom R_6)	0.55	1.25
R2	0.66	1.85
1ZNJ R1 (monoclinic R_6)^c	0.62	1.37
R2	0.56	1.66
R3	0.76	1.79
R4	0.61	1.18
R5	0.63	1.47
R6	0.66	1.23

Continued on next page

1EV6 R1 (monoclinic R_6)[b]	0.68	1.53
R2	0.58	1.56
R3	0.82	1.94
R4	0.59	1.27
R5	0.65	1.83
R6	0.58	1.54
Average	0.59 ± 0.11	1.36 ± 0.30

[a] R^f or "R-frayed" indicates that the first three residues of the insulin B chain (B1-B3) are in a coil conformation rather than part of the B1-B16 helix, as in standard R-state structures.

[b] Rhombohedral crystals contain two unique R-state protomers in the asymmetrical unit that are related by 3-fold symmetry to the two other R-state dimers comprising an R_6 hexamer.

[c] Monoclinic R_6 crystals contain 6 individual R-state protomers in the asymmetrical unit.

Table S4. RMSD values of main-chain and side-chains of the T-state protomer of the crystal structures of Trp^{B26}, Orn^{B29}-insulin and selected WT T_6 and T_3R_3 hexamers

PDB ID	Main-chain RMSD (Å)	side-chain RMSD (Å)
1RWE T (T_3R_3)	0.302	1.18
1BEN T (T_3R_3)	0.289	0.67
1TRZ T (T_3R_3)	0.316	1.09
1TYL T (T_3R_3)	0.319	1.38
1TYM T (T_3R_3)	0.295	1.07
1ZNI T (T_3R_3)	0.283	0.68
1MPJ T (T_3R_3)	0.382	0.84
3MTH T (T_3R_3)	0.318	0.94
1LPH T (KP T_3R_3)	0.486	0.95
1G7A T1 (rhom T_3R_3)[a]	0.307	0.75
T2	0.334	1.16
4INS T1 (2ZN T_6)	1.127	1.94
T2	0.686	1.49
1APH (cubic T_2)[b]	0.620	1.00
1BPH (cubic T_2)	0.653	1.12
1CPH (cubic T_2)	0.667	1.32
1DPH (cubic T_2)	0.659	1.21
Average	0.47 ± 0.11	1.11 ± 0.30

[a] Rhombohedral crystals contain two unique T-state protomers in the asymmetrical unit that are related by 3-fold symmetry to the two other TR dimers comprising an T_3R_3 hexamer.

[b] Cubic T_2 crystals contain two T-state protomers forming an insulin dimer. They are zinc-free crystals.

Table S5. Comparison of χ_1 and χ_2 dihedral-angle values of the T protomer of the Trp^{B26}, Orn^{B29}-insulin crystal structure to the B26 side chains of WT T_6 or T_3R_3 insulin crystal structures[a]

PDB ID	chain	χ_1	χ_2
Trp^{B26}, Orn^{B29}-insulin	B	167.131	-108.457
1TRZ	B	-179.531	78.679[b]
4INS	B	170.768	66.198
	D	175.602	-113.722
2INS	B	170.867	63.894
	D	170.708	-111.824
1ZNI	D	178.670	67.486
1BEN	B	178.543	67.578
	D	172.629	60.183
3MTH	D	178.321	70.378

[a] Tabulation of χ_1 and χ_2 angles demonstrates the native-like orientation of the Trp^{B26} ring of the T protomomer of the Trp^{B26}, Orn^{B29}-insulin crystal structure.
[b] Due to the symmetric structure of the Tyr side chain, χ_2 and χ_2-180° are equivalent. For example, 78.679° is equivalent to -101.321°.

Table S6. Comparison of χ_1 and χ_2 dihedral angle values of the R protomomer of Trp^{B26}, Orn^{B29}-insulin crystal structure to the B26 side chains of WT R_6 or R^f insulin crystal structures[a]

PDB ID	chain	χ_1	χ_2
WB26	D	-179.108	-102.751
1TRZ	D	179.520	81.505[b]
1ZNJ	B	-173.898	82.727
	D	168.865	-90.454
	F	-176.674	69.837
	H	176.540	74.677
	J	-173.848	80.139
	L	175.389	72.440
	B	-178.950	77.319
1BEN	D	-162.232	84.877
1MPJ	B	-177.306	-89.456
3MTH	B	-173.822	-87.442

[a] Tabulation of χ_1 and χ_2 angles demonstrates the native-like orientation of the Trp^{B26} ring of the R protomomer of the Trp^{B26}, Orn^{B29}-insulin crystal structure.

[b] Due to the symmetric structure of the Tyr side chain, χ_2 and χ_2-180° are equivalent.

Table S7. Comparison of predicted proton-proton distances involving Trp^{B26} from Trp^{B26}, Orn^{B29}-insulin crystal structure and Trp^{B26}-associated NOEs in Trp^{B26} lispro

Trp Proton	Residue	Proton	Predicted Distance[a] (Å)	NOE Strength[b]
H-ε3	Ile^{A2}	H-δ1	3.82	w
H-ζ3	Val^{B12}	H-γ1[c]	6.93	s
H-ζ3	Val^{B12}	H-γ2	4.61	s
H-η2	Val^{B12}	H-γ1	7.61	m
H-η2	Val^{B12}	H-γ2	4.68	m
H-ζ2	Val^{B12}	H-γ1	7.38	w
H-ζ2	Val^{B12}	H-γ2	4.40	w
H-ε3	Val^{B12}	H-γ1	8.85	w
H-ε3	Val^{B12}	H-γ2	4.24	w
H-ε3	Leu^{B15}	H-δ1	3.24	w
H-ε3	Leu^{B15}	H-δ2	6.04	w
H-ζ3	Leu^{B15}	H-δ1	4.74	w
H-ζ3	Leu^{B15}	H-δ2	6.85	w

[a] Predicted distances were obtained from the AB (T-state) protomer of the Trp^{B26}, Orn^{B29}-insulin crystal structure.
[b] NOEs were categorized as strong (s, <4.0 Å), medium (m, 4.4-5.5 Å), and weak (w, >5.5 Å).
[c] Predicted proton-proton distances are provided for protons at equivalent carbon positions although associated NOEs are difficult to differentiate.

Table S8. Comparison of predicted proton-proton distances involving Tyr^{B26} from WT crystal structure and Tyr^{B26}-associated NOEs from lispro NMR spectrum.

Tyr Proton	Residue	Proton	Predicted Distance[a] (Å)	NOE Strength[b]
H-ε	Ile^{A2}	H-δ1	5.17	w
H-ε	Val^{A3}	H-γ1[c]	6.82	w
H-ε	Val^{A3}	H-γ2	4.05	w
H-δ	Val^{A3}	H-γ1	4.12	w
H-δ	Val^{A3}	H-γ2	6.25	w
H-ε	Leu^{B11}	H-β	4.61	w
H-ε	Leu^{B11}	H-γ	3.13	w
H-ε	Leu^{B11}	H-δ1	5.42	w
H-ε	Leu^{B11}	H-δ2	3.67	w
H-δ	Leu^{B11}	H-δ1	5.75	w
H-δ	Leu^{B11}	H-δ2	5.24	w
H-ε	Val^{B12}	H-γ1	7.21	w
H-ε	Val^{B12}	H-γ2	4.72	w
H-ε	Leu^{B15}	H-δ1	5.12	w
H-ε	Leu^{B15}	H-δ2	7.80	w
H-δ	Leu^{B15}	H-δ1	3.14	m
H-δ	Leu^{B15}	H-δ2	6.14	w
H-δ	Phe^{B24}	H-ε	4.55	w
H-δ	Thr^{B27}	H-γ2	6.88	w

[a] Predicted distances were obtained from the AB (T-state) protomer of WT structure 1TRZ.
[b] NOEs were categorized as strong (s, <4.0 Å), medium (m, 4.4-5.5 Å) or weak (w, >5.5 Å).
[c] Predicted proton-proton distances are provided for protons at equivalent carbon positions although associated NOEs are difficult to differentiate.

Table S9a. Energy of non-bonded interaction of Tyr^{B26} with local aromatic residues[a]

Protomer	Residue	Interaction Energy (kcal/mol)
Tyr^{B26}	B24F	-0.323
	B25F	-0.111
	B16Y	0.007
	D24F	-1.010
	D25F	-0.014
	D26Y	0.000
	D16Y	-1.310
Tyr^{D26}	B24F	-1.063
	B25F	0.001
	B26Y	0.000
	B16Y	-1.391
	D24F	-0.443
	D25F	-0.033
	D16Y	0.006
Total		-7.252

[a] As calculated from WT structure 1TRZ.

Table S9b. Energy of non-bonded interaction of Trp^{B26} with neighboring aromatic residues[a]

Protomer	Residue	Interaction Energy (kcal/mol)
Trp^{B26}	B24F	-0.375
	B25F	-0.162
	B16Y	0.009
	D24F	-1.266
	D25F	-0.023
	D26Y	0.015
	D16Y	-1.411
Trp^{D26}	B24F	-1.517
	B25F	-0.013
	B26Y	0.015
	B16Y	-2.073
	D24F	-0.733
	D25F	-0.073
	D16Y	0.007
Total		-8.670

[a] Table (calculated from Trp^{B26} Orn^{B29} structure) show specific aromatic-aromatic aromatic interactions improved in the TR dimer of Trp^{B26} Orn^{B29}-insulin in relation to WT insulin.

Table S10. Summary of previous designs of basal insulin analogs[a]

Modification	Mechanism	Drawbacks
Human Proinsulin	C-peptide slows absorption, prolongs activity (17)	Safety concerns regarding hyperproinsulinemia and cardiovascular side effects Not a truly basal insulin (17)
Neutral Protamine Hagedorn / Neutral Protamine Lispro	Forms crystals with partner molecule protamine (18)	Consistent dosing difficult (19) Intermediate acting (20) Can cause allergic reactions (21)
Gln^{B13}	Forms zinc-free hexamers (10)	Dramatically reduced biological activity (5)
Lys^{B27}	Positive charge of stabilizes packing of multiple hexamers (22) Shifts pI of insulin	Effective only in complement with other modifications (22)
Gln^{B13} Gln^{A17} Lys^{B30}	Iso-electric precipitation Stabilization of Hexamer	Dramatically reduced biological activity (23)
Gly^{A21},Arg^{B27},Thr^{B30}-NH2	Iso-electric precipitation in subcutaneous depot, Gly^{A21} prevents chemical degradation in acidic formulations (23)	Macrophage invasion and inflammation at injection site (7,24)
Gly^{A21} $Arg^{B30}Arg^{B31}$	Iso-electric precipitation in subcutaneous depot, Gly^{A21} prevents chemical degradation in acidic formulations	Reports of association with neoplasms (25) PK/PD profile less predictable than other basal analogs (26) Unknown effect of enzymatic degradation products (25,27)
Myristic acid-coupled Lys^{B29}	Slows absorption from SubQ depot, binds to albumin (24)	Reduced potency (24)
Hexadodecandioic acid-coupled Lys^{B29}	Slows absorption from SubQ depot, forms multihexameric complexes, binds to albumin (8,28)	Reduced potency, has not yet been used long-term clinically(8)

[a] This table underscores the challenges associated with the development of insulin analogs with protracted PD profiles. Trp^{B26} insulin analogs are unique in their ability to protract the PD profile of insulin by stabilizing the hexamer without significantly compromising the potency of the analog.

Table S11a. Intercentroid distances and angles between Tyr^{B26} and local aromatic residues[a]

	B26		D26	
Residue	r	Dih.A	r	Dih.A
B16	13.6	76	6.0	52
B24	7.3	85	5.9	87
B25	10.9	34	10.6	70
B26	-	-	12.0	77
D16	5.8	58	13.1	89
D24	5.8	38	7.5	78
D25	12.6	38	9.5	45
D26	12.0	77	-	-

Table S11b. Intercentroid distances and angles between Tyr^{B26} and local aromatic residues[a]

	B26		D26	
Residue	r	Dih.A	r	Dih.A
B16	12.8	90	5.5	20
B24	7.8	79	5.9	25
B25	9.7	37	13.0	44
B26	-	-	12.2	73
D16	5.6	51	13.7	66
D24	6.6	46	7.0	89
D25	10.9	73	11.1	10
D26	12.2	73	-	-

[a] Tables S11a and S11b (respectively calculated from WT structure 1TRZ and the present crystal structure of the TrpB26 analog) indicate the character of aromatic-aromatic interactions involving Tyr or Trp^{B26} at the insulin dimer interface. Whereas some interactions are classical ETF interactions, others show some deviation in ring-to-ring dihedral angle (Dih.A).

Table S12a. Thermodynamic stabilization associated with the B26 side chain during folding of insulin analogs[a]

Side Chain	ΔSA	ΔG_u [b] (kcal/mol)	Side Chain	ΔSA	ΔG_u (kcal/mol)
Tyr^{B26}	68.9%	0.49 [a]	Trp^{B26}	69.3%	1.4
Tyr^{D26}	78.6%	0.56	Trp^{D26}	70.1%	1.5
Mean $\Delta\Delta G_u$: 0.9 kcal/mol (not observed)					

Table S12b. Thermodynamic stabilization associated with the B26 side chain during dimerization of insulin analogs[a]

Side Chain	ΔSA	ΔG_u (kcal/mol)	Side Chain	ΔSA	ΔG_u (kcal/mol)
Tyr^{B26}	25.9%	0.18	Trp^{B26}	26.6%	0.55
Tyr^{D26}	78.6%	0.56	Trp^{D26}	26.5%	0.55
Mean $\Delta\Delta G_u$: 0.37 + 0.42 = 0.8 kcal/mol/dimer interface					

[a] The relative thermodynamic stabilities of Tyr^{B26} and Trp^{B26} insulin monomers calculated using hydrophobic transfer free energies contrasted with those determined by guanidine-denaturation assays. However, NMR data suggests dimerization may be favored in Trp^{B26} lispro to a greater extent than in native lispro.

[b] Tabulated values of respective H_2O/octanol transfer free energies are -0.71 kcal/mol (Tyr) and -2.09 kcal/mol (Trp) as described (29). This calculation pertains only to changes in exposure of residue B26 and does not consider secondary changes in exposure of neighboring side chains.

Table S13. Native-like crystal structures of R_6 hexamers formed by mutant insulin.

PDB ID or Reference	Modification
1EV6	WT Human Insulin
1ZNJ	Porcine insulin: Ala^{B30} at chain terminus
4E7V	Bovine Insulin: Ala^{A8}, Val^{A10}, Ala^{B30} at surfaces
3GKY	His^{A8}, Val^{A16} at surface (A8) or in core (A16)
4P65	Cha^{B24} at dimer interface
3ROV	D-Ala^{B20}, D-Ala^{B23}, lispro within β-turn
5HRQ	cis-hydroxyproline^{B28} at edge of dimer interface
5HPU	trans-hydroxyproline^{B28} at edge of dimer interface
5URU	Dihydroxyproline^{B28} at edge of dimer interface
5UQA	4-fluoro-proline^{B28} at edge of dimer interface
2WS6	N-methyl-Tyr^{B26} at dimer interface and inter-chain
2WS7	Pro^{B26} at dimer interface
3ZS2	Tyr^{B25}, N-methyl-Phe^{B26}, lispro at dimer interface
3ZQR	N-methyl-Phe^{B25} at dimer interface
5EMS	3-iodotyrosine^{B26}, Norleucine^{B29} at dimer interface
1QIY	Tyr^{B5} at inter-chain crevice
1ZEG	Asp^{B28} at edge of dimer interface
3ZU1	*Des*-B30, Nε-ω-carboxyheptadecanoyl-Lys^{B29}
Derewenda, 1987 (30)	$Val^{B12} \rightarrow$ Ile at dimer interface

Supplemental References:

1. Holleman, F., and Hoekstra, J. B. (1997) Insulin lispro. *N. Engl. J. Med.* **337**, 176-183
2. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., and Snel, L. (1988) Monomeric insulins obtained by protein engineering and their medical implications. *Nature* **333**, 679-682
3. Brange, J., and Vølund, A. (1999) Insulin analogs with improved pharmacokinetic profiles. *Adv. Drug Deliv. Rev.* **35**, 307-335
4. van Bon, A. C., Bode, B. W., Sert-Langeron, C., DeVries, J. H., and Charpentier, G. (2011) Insulin glulisine compared to insulin aspart and to insulin lispro administered by continuous subcutaneous insulin infusion in patients with type 1 diabetes: a randomized controlled trial. *Diabetes Technol. Ther.* **13**, 607-614
5. Markussen, J. M., Hougaard, P., Ribel, U., Sørensen, A. R., and Sørensen, E. (1987) Soluble, prolonged-acting insulin derivatives. I. Degree of protraction and crystallizability of insulins substituted in the termini of the B-chain. *Protein Eng.* **1**, 205-213
6. Gillies, P. S., Figgitt, D. P., and Lamb, H. M. (2000) Insulin glargine. *Drugs* **59**, 253-260
7. Markussen, J., Havelund, S., Kurtzhals, P., Andersen, A., Halstrøm, J., Hasselager, E., Larsen, U., Ribel, U., Schäffer, L., and Vad, K. (1996) Soluble, fatty acid acylated insulins bind to albumin and show protracted action in pigs. *Diabetologia* **39**, 281-288
8. Gough, S., Harris, S., Woo, V., and Davies, M. (2013) Insulin degludec: overview of a novel ultra long-acting basal insulin. *Diab. Obes. Metab.* **15**, 301-309
9. Steensgaard, D. B., Schluckebier, G., Strauss, H. M., Norrman, M., Thomsen, J. K., Friderichsen, A. V., Havelund, S., and Jonassen, I. (2013) Ligand-controlled assembly of hexamers, dihexamers, and linear multihexamer structures by the engineered acylated insulin degludec. *Biochemistry* **52**, 295-309
10. Bentley, G. A., Brange, J., Derewenda, Z., Dodson, E. J., Dodson, G. G., Markussen, J., Wilkinson, A. J., Wollmer, A., and Xiao, B. (1992) Role of B13 Glu in insulin assembly. The hexamer structure of recombinant mutant (B13 Glu→Gln) insulin. *J. Mol. Biol.* **228**, 1163-1176
11. Frisch, M. J., Trucks, G. W., Schlegal, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vrenven, T., Montogomery, J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. (2009) Gaussian09 Revision D.01. Gaussian Inc., Wallingford, CT
12. Willighagen, E., and Howard, M. (2007) Fast and Scriptable Molecular Graphics in Web Browsers without Java3D. in *Nature Precedings*
13. Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545-1614

14. MacKerell, A. D., Brooks, B., Brooks, C. L., Nilsson, L., Roux, B., Won, Y., and Karplus, M. (2002) CHARMM: The Energy Function and Its Parameterization. in *Encyclopedia of Computational Chemistry* (Schleyer, P. v. R. C., N. L. A. T. , Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R. S. ed.), John Wiley & Sons, Ltd, Chichester. pp 271-277

15. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J., J. A. , Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. (2009) Gaussian 09, Revision A.02. in *Gaussian, Inc., Wallingford CT*, Revision A.02 Ed.
16. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Nat. Acad. Sci.* **98**, 10037-10041
17. Galloway, J. A., Hooper, S. A., Spradlin, C. T., Howey, D. C., Frank, B. H., Bowsher, R. R., and Anderson, J. H. (1992) Biosynthetic human proinsulin. Review of chemistry, *in vitro* and *in vivo* receptor binding, animal and human pharmacology studies, and clinical trial experience. *Diab. Care* **15**, 666-692
18. Hagedorn, H., Jensen, B. N., Krarup, N., and Wodstrup, I. (1936) Protamine insulinate. *JAMA* **106**, 177-180
19. Lauritzen, T., Pramming, S., Gale, E., Deckert, T., and Binder, C. (1982) Absorption of isophane (NPH) insulin and its clinical implications. *Brit. Med. J.* **285**, 159-162
20. Owens, D. R., and Bolli, G. B. (2008) Beyond the era of NPH insulin—long-acting insulin analogs: chemistry, comparative pharmacology, and clinical application. *Diabetes Technol. Ther.* **10**, 333-349
21. Stewart, W. J., Mcsweeney, S. M., Kellett, M. A., Faxon, D. P., and Ryan, T. J. (1984) Increased risk of severe protamine reactions in NPH insulin-dependent diabetics undergoing cardiac catheterization. *Circulation* **70**, 788-792
22. Markussen, J., Diers, I., Engesgaard, A., Hansen, M. T., Hougaard, P., Langkjaer, L., Norris, K., Ribel, U., Sorensen, A. R., Sorensen, E., and et al. (1987) Soluble, prolonged-acting insulin derivatives. II. Degree of protraction and crystallizability of insulins substituted in positions A17, B8, B13, B27 and B30. *Protein Eng.* **1**, 215-223
23. Markussen, J., Diers, I., Hougaard, P., Langkjaer, L., Norris, K., Snel, L., Sørensen, A. R., Sørensen, E., and Voight, H. O. (1988) Soluble, prolonged-acting insulin derivatives. III. Degree of protraction, crystallizability and chemical stability of insulins substituted in positions A21, B13, B23, B27 and B30. *Protein Eng.* **2**, 157-166
24. Kurtzhals, P., Havelund, S., Jonassen, I., Kiehr, B., Larsen, U., Ribel, U., and Markussen, J. (1995) Albumin binding of insulins acylated with fatty acids: characterization of the ligand-protein interaction and correlation between binding affinity and timing of the insulin effect in vivo. *Biochem. J.* **312**, 725-731
25. Suissa, S., Azoulay, L., Dell'Aniello, S., Evans, M., Vora, J., and Pollak, M. (2011) Long-term effects of insulin glargine on the risk of breast cancer. *Diabetologia* **54**, 2254-2262
26. Yamamoto, C., Miyoshi, H., Fujiwara, Y., Kameda, R., Ichiyama, M., Nomoto, H., Kameda, H., Nakamura, A., and Atsumi, T. (2016) Degludec is superior to glargine in terms of daily glycemic variability in people with type 1 diabetes mellitus. *Endocr. J.* **63**, 53-60

27. Sommerfeld, M. R., Muller, G., Tschank, G., Seipke, G., Habermann, P., Kurrle, R., and Tennagels, N. (2010) *In vitro* metabolic and mitogenic signaling of insulin glargine and its metabolites. *PLoS One* **5**, e9540
28. Jonassen, I., Havelund, S., Hoeg-Jensen, T., Steensgaard, D. B., Wahlund, P. O., and Ribel, U. (2012) Design of the novel protraction mechanism of insulin degludec, an ultra-long-acting basal insulin. *Pharm. Res.* **29**, 2104-2114
29. Wimley, W. C., and White, S. H. (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Molec. Biol.* **3**, 842
30. Derewenda, U., Derewenda, Z., Dodson, G., and Brange, J. (1987) The Crystal-Structure of the B-12 Ile Human Insulin Prepared by Site-Directed Mutatgenesis. *Protein Eng.* ***222****, 425-433*