Protective hinge in insulin opens to enable its receptor engagement

John G. Menting[a,1], Yanwu Yang[b,1], Shu Jin Chan[c], Nelson B. Phillips[b], Brian J. Smith[d], Jonathan Whittaker[b], Nalinda P. Wickramasinghe[b], Linda J. Whittaker[b], Vijay Pandyarajan[b], Zhu-li Wan[b], Satya P. Yadav[e], Julie M. Carroll[f], Natalie Strokes[g], Charles T. Roberts, Jr.[f,h], Faramarz Ismail-Beigi[i], Wieslawa Milewski[c], Donald F. Steiner[c,2], Virander S. Chauhan[j], Colin W. Ward[a], Michael A. Weiss[b,i,k,2,3], and Michael C. Lawrence[a,l,2,3]

[a]Structural Biology Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia; Departments of [b]Biochemistry, [g]Physiology, [i]Medicine, and [k]Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44106; [c]Department of Medicine, University of Chicago, Chicago, IL 60637; [d]La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC 3086, Australia; [e]Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, OH 44195; [f]Oregon National Primate Research Center, Beaverton, OR 97006; [h]Department of Medicine, Oregon Health and Science University, Portland, OR 97239; [j]International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi 110067, India; and [l]Department of Medical Biology, University of Melbourne, Parkville, VIC 3010, Australia

Contributed by Donald F. Steiner, July 9, 2014 (sent for review May 25, 2014)

Insulin provides a classical model of a globular protein, yet how the hormone changes conformation to engage its receptor has long been enigmatic. Interest has focused on the C-terminal B-chain segment, critical for protective self-assembly in β cells and receptor binding at target tissues. Insight may be obtained from truncated "microreceptors" that reconstitute the primary hormone-binding site (α-subunit domains L1 and αCT). We demonstrate that, on microreceptor binding, this segment undergoes concerted hinge-like rotation at its B20-B23 β-turn, coupling reorientation of Phe^{B24} to a 60° rotation of the B25-B28 β-strand away from the hormone core to lie antiparallel to the receptor's L1–β_2 sheet. Opening of this hinge enables conserved nonpolar side chains (Ile^{A2}, Val^{A3}, Val^{B12}, Phe^{B24}, and Phe^{B25}) to engage the receptor. Restraining the hinge by nonstandard mutagenesis preserves native folding but blocks receptor binding, whereas its engineered opening maintains activity at the price of protein instability and nonnative aggregation. Our findings rationalize properties of clinical mutations in the insulin family and provide a previously unidentified foundation for designing therapeutic analogs. We envisage that a switch between free and receptor-bound conformations of insulin evolved as a solution to conflicting structural determinants of biosynthesis and function.

diabetes mellitus | signal transduction | receptor tyrosine kinase | metabolism | protein structure

How insulin engages the insulin receptor has inspired speculation ever since the structure of the free hormone was determined by Hodgkin and colleagues in 1969 (1, 2). Over the ensuing decades, anomalies encountered in studies of analogs have suggested that the hormone undergoes a conformational change on receptor binding: in particular, that the C-terminal β-strand of the B chain (residues B24–B30) releases from the helical core to expose otherwise-buried nonpolar surfaces (the detachment model) (3–6). Interest in the B-chain β-strand was further motivated by the discovery of clinical mutations within it associated with diabetes mellitus (DM) (7). Analysis of residue-specific photo–cross-linking provided evidence that both the detached strand and underlying nonpolar surfaces engage the receptor (8).

The relevant structural biology is as follows. The insulin receptor is a disulfide-linked $(\alpha\beta)_2$ receptor tyrosine kinase (Fig. 1*A*), the extracellular α-subunits together binding a single insulin molecule with high affinity (9). Involvement of the two α-subunits is asymmetric: the primary insulin-binding site (site 1*) comprises the central β-sheet (L1–β_2) of the first leucine-rich repeat domain (L1) of one α-subunit and the partially helical C-terminal segment (αCT) of the other α-subunit (Fig. 1*A*) (10). Such binding initiates conformational changes leading to trans-phosphorylation of the β-subunits' intracellular tyrosine kinase (TK) domains. Structures of wild-type (WT) insulin (or analogs) bound to extracellular receptor fragments were recently described at maximum resolution of 3.9 Å (11), revealing that hormone binding is primarily mediated by αCT (receptor residues 704–719); direct interactions between insulin and L1 were sparse and restricted to certain B-chain residues. On insulin binding, αCT was repositioned on the L1–β_2 surface, and its helix was C-terminally extended to include residues 711–714. None of these structures defined the positions of C-terminal B-chain residues beyond B21. Support for the detachment model was nonetheless provided by entry of αCT into a volume that would otherwise be occupied by B-chain residues B25–B30 (i.e., in classical insulin structures; Fig. 1*B*) (11).

Significance

Insulin provides a model for analysis of protein structure and evolution. Here we describe in detail a conformational switch that enables otherwise hidden nonpolar surfaces in the hormone to engage its receptor. Whereas the classical closed conformation of insulin enables its stable storage in pancreatic β cells, its active conformation is open and susceptible to nonnative aggregation. Our findings illuminate biophysical constraints underlying the evolution of an essential signaling system and provide a structural foundation for design of therapeutic insulin analogs.

Author contributions: D.F.S., C.W.W., M.A.W. and M.C.L. designed research; J.G.M., Y.Y., S.J.C., N.B.P., J.W., N.P.W., L.J.W., V.P., S.P.Y., J.M.C., N.S., C.T.R., F.I.-B., W.M., and V.S.C. performed research; J.G.M. performed crystallography experiments and analyzed diffraction data; B.J.S. and Z.-l.W. analyzed data; M.A.W. and M.C.L. wrote the paper; Y.Y., N.P.W., and M.A.W. performed NMR experiments; N.B.P. prepared synthetic peptides and semisynthetic insulin analogs and performed fibrillation assays; Y.Y., B.J.S., N.P.W., and Z.-l.W. performed calculations; S.J.C., J.W., L.J.W., and W.M. performed receptor-binding studies; Y.Y., N.P.W., and M.A.W. analyzed solution structures; V.P., S.P.Y., and V.S.C. prepared synthetic peptides and semisynthetic insulin analogs; J.M.C., N.S., C.T.R., and F.I.-B. performed biological studies; and M.C.L. analyzed diffraction data.

Conflict of interest statement: F.I.-B. and M.A.W. hold stock in Thermalin Diabetes, LLC (Cleveland, OH), for which N.B.P. and J.W. are consultants; M.A.W. is Chief Scientific Officer and a Director. Part of M.C.L.'s research is funded by Sanofi (Germany).

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (insulin/μIR/Fab 83-7 complex, PDB ID code 4OGA; insulin analog 1, PDB ID code 4NIB; analog 2, PDB ID code 2MLI; and analog 4, PDB ID code 2MPI).

[1]J.G.M. and Y.Y. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: dfsteine@midway.uchicago.edu, michael.weiss@case.edu, or lawrence@wehi.edu.au.

[3]M.A.W. and M.C.L. contributed equally to this work.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1412897111/-/DCSupplemental.

*Whereas site 1 contacts insulin's classical receptor-binding surface, site 2 (proposed to reside at the junction of the α-subunit's fibronectin homology domains 1 and 2) is defined in relation to cognate substitutions in the hormone's trimer-forming surface that affect the lifetime of the holoreceptor complex (9).

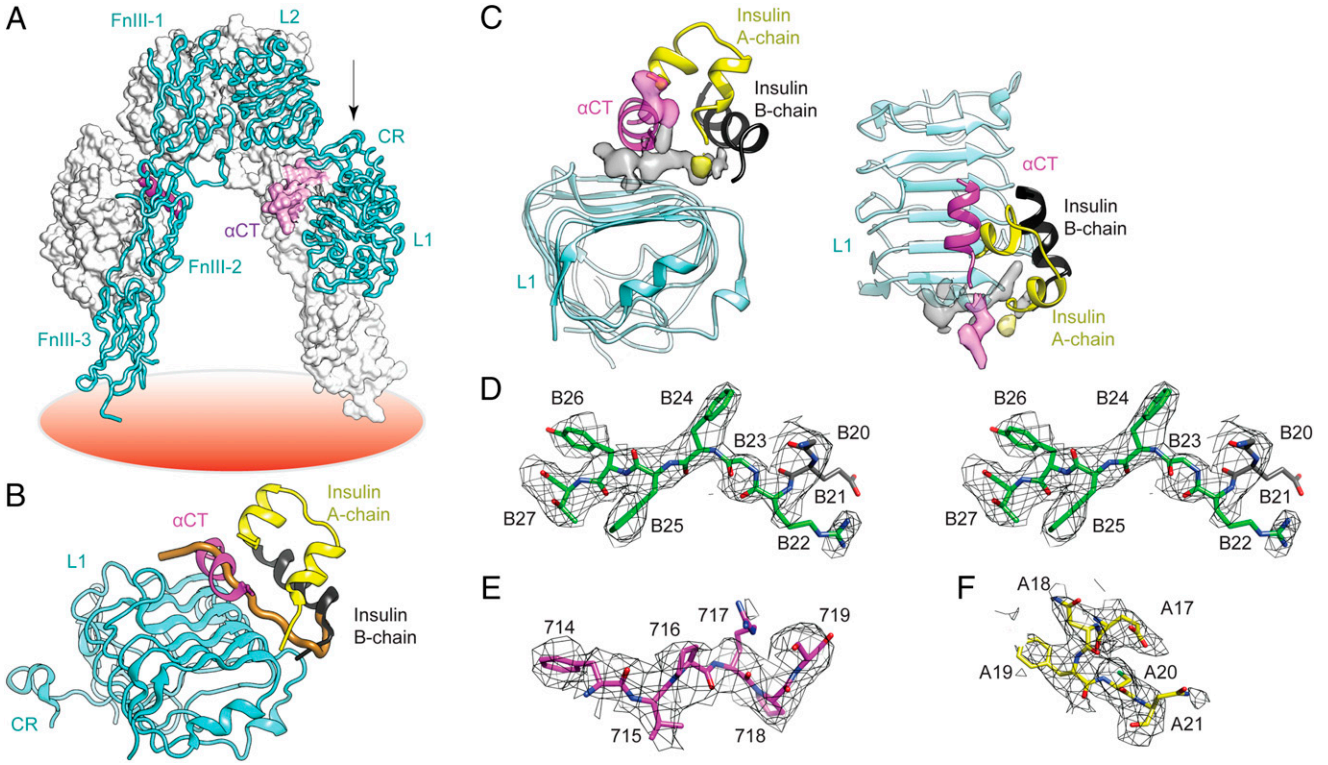

Fig. 1. Insulin B-chain C-terminal β-strand in the μIR complex. (*A*) Structure of *apo*-receptor ectodomain. One monomer is in tube representation (labeled), the second is in surface representation. L1, first leucine-rich repeat domain; CR, cysteine-rich domain; L2, second leucine-rich repeat domain; FnIII-1, -2 and -3; first, second and third fibronectin type III domains, respectively; αCT, α-subunit C-terminal segment; coral disk, plasma membrane. (*B*) Insulin bound to μIR; the view direction with respect to L1 in the *apo*-ectodomain is indicated by the arrow in *A*. Only B-chain residues indicated in black were originally resolved (11). The brown tube indicates classical location of residues B20-B30 in free insulin, occluded in the complex by αCT. (*C*) Orthogonal views of unmodeled $2F_{obs}$-F_{calc} difference electron density (*SI Appendix*), indicating association of map segments with the αCT C-terminal extension (transparent magenta), insulin B-chain C-terminal segment (transparent gray), and Asn^{A21} (transparent yellow). Difference density is sharpened (B_{sharp} = −160 Å^2). (*D*–*F*) Refined models of respective segments insulin B20–B27, αCT 714–719, and insulin A17-A21 within postrefinement $2F_{obs}$-F_{calc} difference electron density (B_{sharp} = −160 Å^2). *D* is in stereo.

We describe here the structure and interactions of the detached B-chain C-terminal segment of insulin on its binding to a "microreceptor" (μIR), an L1–CR domain-minimized version of the α-subunit (designated IR310.T) plus exogenous αCT peptide 704–719 (11). Our analysis defines a hinge in the B chain whose opening is coupled to repositioning of αCT between nonpolar surfaces of L1 and the insulin A chain. To understand the role of this hinge in holoreceptor binding and signaling, we designed three insulin analogs containing structural constraints (Table 1): [D-Ala^{B20}, D-Ala^{B23}]-insulin, ΔPhe^{B25}-insulin, and ΔPhe^{B24}-insulin, where ΔPhe is (α,β)-dehydrophenylalanine (Fig. 2) (12). The latter represents, to our knowledge, the first use of ΔPhe—a rigid "β-breaker" with extended electronic conjugation between its side chain and main chain (*SI Appendix*, Fig. S1)—as a probe of induced fit in macromolecular recognition. In addition, a fourth analog, active but with anomalous flexibility in the B chain (5, 6) (Table 1), was used to investigate the relationship between the hinge and insulin's susceptibility to misfolding.

Despite the limitations of domain minimization, our structure of the μIR complex illuminates the properties of DM-associated mutations in insulin and rationalizes a wealth of prior biochemical data. Of broader importance, our findings demonstrate that hidden within insulin sequences lie multiple layers of structural information, encoding a complex conformational life cycle from biosynthesis to function. As such, they provide a structural foundation for design of therapeutic analogs.

Results

Insulin Binding to μIR vs. Holoreceptor. To distinguish whether the apparent disorder of insulin segment B22–B30 in the extant μIR crystal structure (11) was a consequence of receptor truncation or a technical limitation of prior crystals, we tested whether Phe^{B24}, Phe^{B25}, and Tyr^{B26} contribute to μIR binding in a fashion equivalent to holoreceptor binding (2, 4). Each residue was singly substituted by alanine (Ala), and assembly of the μIR

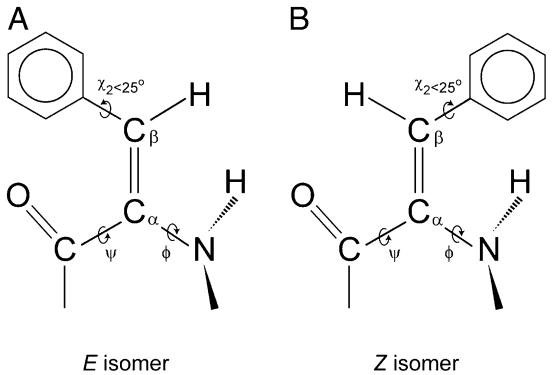

Fig. 2. Structure of ΔPhe. (*A* and *B*) Respective line drawings of *E* and Z configurational isomers of (α,β)-dehydro-Phe. The present studies use the more stable *Z* isomer (23).

complex was monitored as a function of the added concentration of insulin or analog. Relative stabilities of the variant μIR complexes correlated with analog affinities for the holoreceptor (*SI Appendix*, Fig. S2 *A* and *B*). Consistency was also obtained in assays of insulin variants containing single Ala substitutions at classical receptor contact sites resolved in the original μIR structure (*SI Appendix*, Fig. S2 *C* and *D* and Table S1) (2, 13–15). Together, these data imply that residues B24–B26 make similar contributions to the binding of insulin to the μIR and holoreceptor.

Transverse Relaxation-Optimized NMR Spectroscopy. Given the critical role of the B24–B26 segment in insulin binding to the holoreceptor, the above findings suggested that this segment is ordered in the μIR complex. We investigated this via transverse relaxation-optimized NMR spectroscopy (TROSY) (16). Complexes were prepared containing ^{13}C-labeled insulin and complementary labels in αCT; TROSY-defined ^{1}H- and ^{13}C chemical shifts provided probes of the labeled sites in the free molecules and in the μIR complex. A monomeric insulin analog [Lys^{B28}, Pro^{B29}]-porcine insulin (KP-p-insulin) (17) was prepared containing uniformly [^{13}C, ^{15}N]-enriched amino acids at positions B23, B24, B26, and B30 (the ^{15}N labels were not exploited). An αCT peptide derived from receptor α-subunit sequences 703–717 (with an N-terminal Arg and C-terminal Lys added to enhance solubility) was likewise labeled at positions Tyr708, Leu709, Val713, and Phe714. [^{1}H-^{13}C]-TROSY spectra in D_2O were obtained from a complex comprising unlabeled IR310.T, labeled KP-p-insulin, and labeled αCT peptide (complex 1; Fig. 3 and *SI Appendix*, Fig. S3) and from a complex comprising unlabeled IR310.T, unlabeled KP-insulin, and labeled αCT (complex 2; used to distinguish αCT cross-peaks in complex 1). Chemical shifts were assessed in relation to spectra of the isolated labeled components.

Each of the labeled sites in insulin (except flexible chain terminus Ala^{B30}) and in αCT underwent significant changes in ^{1}H and/or ^{13}C chemical shifts on μIR assembly (*SI Appendix*, Table S2). In general, labeled αCT sites exhibited complexation shifts away from random-coil values (i.e., enhanced secondary shifts) in accordance with stabilization of the α-helical secondary structure on assembly. Although aliphatic ^{1}H-^{13}C heteronuclear single-quantum coherence (HSQC) signals from the labeled KP-p-insulin were attenuated in complex 1 (due to resonance broadening; *SI Appendix*, Fig. S3), the aromatic cross-peaks of Phe^{B24} and Tyr^{B26} remained sharp and exhibited large changes in chemical shift (Fig. 3). Depending on the site, secondary shifts were enhanced or attenuated, suggesting local differences in magnetic environments between free and bound hormone. These findings imply that, on μIR binding, Phe^{B24} and Tyr^{B26} adopt nonrandom conformations.

Crystallographic Resolution of the Bound B22–B27 Segment. The above motivated further crystallographic analysis. We collected diffraction data from >40 crystals of the prior insulin μIR complex (11); although these data were mostly of lower resolution than originally described, combination of the best three datasets enabled technical improvement on merging of intensities to 3.5-Å

Table 1. Summary of insulin analogs

Analog	Modification	Templates*	Rationale
1	D-Ala^{B20}, D-Ala^{B23}	Insulin; KP-insulin	Locked β-turn
2	ΔPhe^{B25}	KP-insulin; DKP-insulin	β-breaker at B25
3	ΔPhe^{B24}	KP-insulin; DKP-insulin	β-breaker at B24
4	Gly^{B24}	KP-insulin; DKP-insulin	Destabilized hinge

*All templates use the human insulin sequence, with KP-insulin ("lispro") having substitutions ProB28Lys and LysB29Pro and DKP-insulin having the additional substitution HisB10Asp.

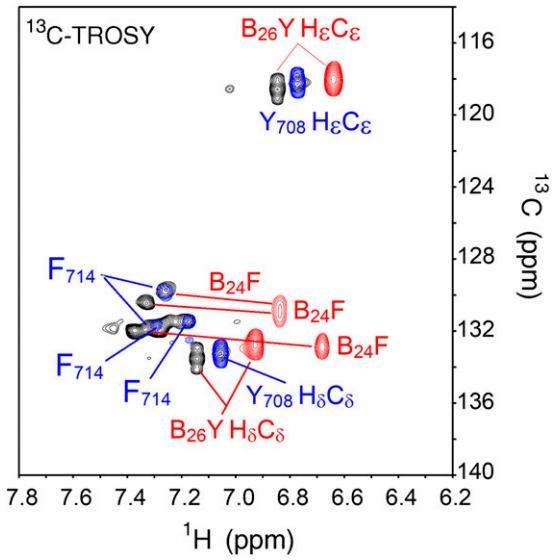

Fig. 3. TROSY NMR analysis of unlabeled L1–CR fragment of the receptor α-subunit (construct IR310.T) complexed with labeled KP-porcine insulin and labeled αCT peptide (complex 1). [^{1}H-^{13}C]-TROSY spectrum of selected aromatic ring resonances in the μIR complex (black) is shown relative to [^{1}H-^{13}C] HSQC spectra of KP-porcine insulin (red) and free αCT (blue). Labeled aromatic sites in the insulin analog were Phe^{B24} and Tyr^{B26}; labeled aromatic sites in αCT were Tyr708 and Phe714 as indicated. The characteristic upfield secondary chemical shifts of Phe^{B24} and Tyr^{B26} in free insulin are attenuated in the complex. For aliphatic [^{1}H-^{13}C] HSQC spectra, see *SI Appendix*, Fig. S3.

resolution as assessed by $\langle I/\sigma(I)\rangle$ and $CC_{1/2}$ criteria (18). Refinement against the combined diffraction dataset yielded interpretable density segments adjacent to insulin B21 and to αCT Val715, respectively (Fig. 1*C*). The first of these segments was successfully modeled and crystallographically refined (R_{work}/R_{free} = 0.264/0.284) as B22–B27 docked largely within a crevice between the C-terminal region of αCT and the L1 surface. The second of these refined segments encompassed αCT residues 716–719. Full details regarding X-ray data processing and crystallographic refinement are provided in the *SI Appendix*, with statistics in *SI Appendix*, Table S3.

Corresponding to well-defined electron density (Fig. 1*D*), the side chain of Phe^{B24} is anchored in a hydrophobic pocket bounded by B-chain residues Val^{B12}, Leu^{B15}, Tyr^{B16}, and Cys^{B19} and receptor residues Asn15, Leu37, Phe39, and Phe714 (Fig. 4 *A* and *B*)[†]. Modeling connected Phe^{B24} through density to Cys^{B19} via a type 1 β-turn with positive ϕ angles at Gly^{B20} and Gly^{B23}—similar to those of WT insulin (2). The B20–B23 β-turn has minimal interaction with the μIR. The side chain of Phe^{B25} (in less well-defined density than that of Phe^{B24}; Fig. 1*D*) projects away from L1–β_2 to insert between αCT residues Val715 and Pro718 (Fig. 4*C*); contacts are also made to the αCT main chain at 715–718. The Phe^{B25} main chain adjoins the side chain of Arg14, but no contacts occur between the aromatic ring and either L1 or other insulin side chains. The side chain of Tyr^{B26} lies within a shallow depression defined by L1 side chains Asp12 and Arg14; B26 side-chain density is less well defined than either Phe^{B24} or Phe^{B25} (Fig. 1*D*). Although the B26 side chain contacts only L1, its main chain abuts both Arg14 and Val715. The orientation of Thr^{B27} is unclear (Fig. 1*D*), with little apparent μIR interaction (Fig. 4*B*). Electron density extending C-terminal to αCT residue 715 was readily modeled (Fig. 1*E*). Residues

[†]In the original map, Phe^{B24} side-chain density was overlain in part by Asn^{A21}; the latter was found here to project away from the axis of the A-chain C-terminal helix (Fig. 4*C*).

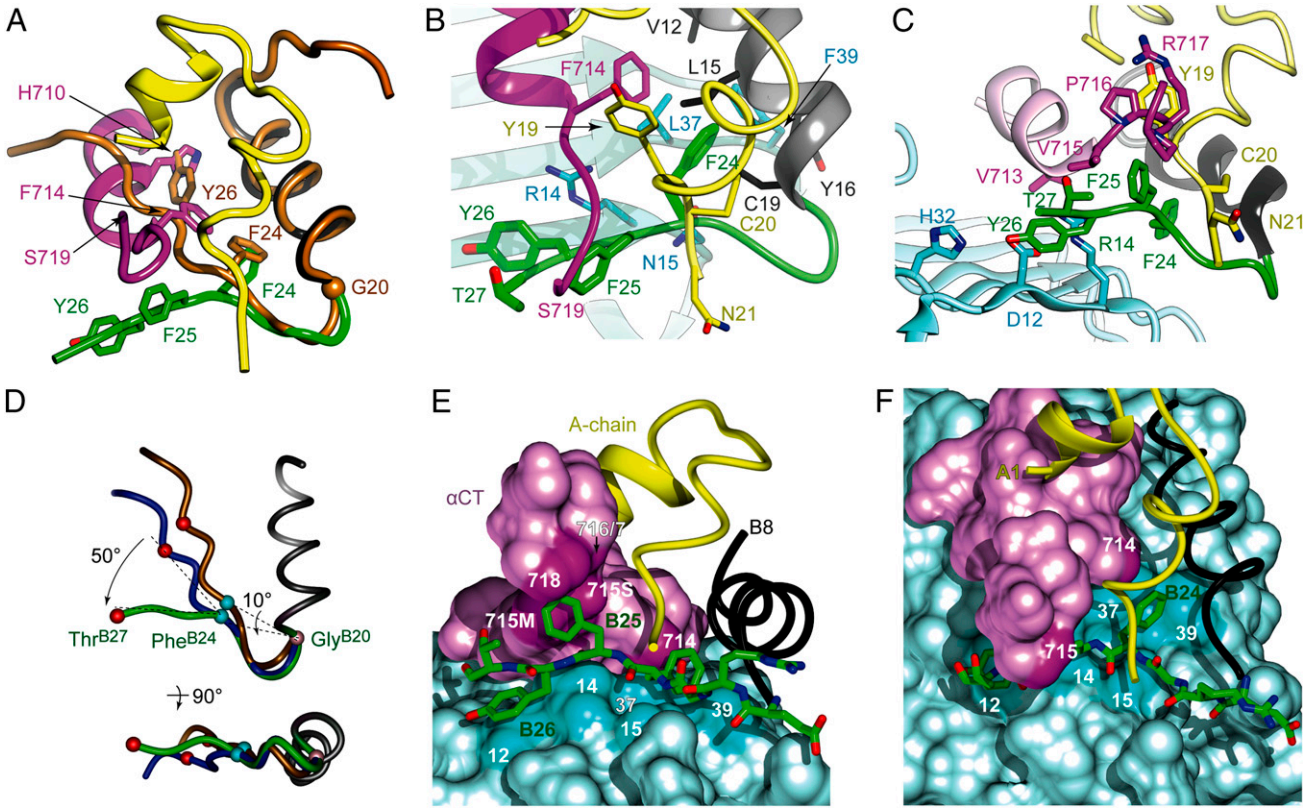

Fig. 4. Structural change in the B-chain C-terminal segment on μIR engagement. (*A*) Overlay of the B-chain of free insulin onto the μIR complex. Brown, B1–B30 conformation in free insulin; black, B7–B19 conformation of μIR-bound insulin; green, B20–B27 conformation of μIR-bound insulin; yellow, A-chain conformation of μIR-bound insulin; magenta, αCT segment. In the μIR complex, αCT side chains His710 and Phe714 fill a volume occupied by Tyr^{B26} in free insulin; the location of the Phe^{B24} side chain is approximately maintained by rotameric change. (*B* and *C*) Environments of Phe^{B24} (*B*) and residues B25–B27 (*C*) in the μIR complex. Colors are as in *A*, with the L1 domain in cyan. (*D*) Dissection of the hinge-like opening of residues B20–B30 into a ∼10° rotation of B20–B30 about B20 followed by a ∼50° rotation of B24–B30 about B24. Brown, B20–B30 conformation in free insulin; black/green, B-chain conformation in μIR complex; blue, intermediate conformation illustrating ∼10° rotation of the B20–B23 β-turn. (*E*) Packing of residues B20–B27 between αCT and L1–β_2 sheet surfaces (C, green; N, blue; O, red). Black, residues B8–B19; yellow, insulin A chain. Contact surfaces of αCT with B24–B26 are highlighted in magenta, and those of L1 with B24–B26 in cyan; surfaces not abutting B24–B26 are in lighter shades. Main- and side-chain surfaces of Val715 are labeled M and S. (*F*) Orthogonal view to *E*, showing interactions of Phe^{B24} side chain with nonpolar surface of L1–β_2 sheet and those of conserved residues A1–A3 with αCT.

716–719 are directed away from L1 in an irregular conformation (Fig. 4 *A*–*C*); the helical portion of αCT terminates at Phe714. Bulges in density correspond to Pro716 and Pro718 (Fig. 1*E*), whereas the Arg717 side chain appears disordered. Although detailed conformations are uncertain, it is clear that the displaced B24–B26 segment inserts between the first strand of the L1–β_2 sheet and αCT residues 715–718 (Fig. 4*C*); contacts with Phe^{B25} (above) stabilize residues 716–719.

Detachment of the B-chain C-terminal strand from the hormone core can be factored into two rotations that result in its repositioning between αCT and L1. The first is rotation of the B20–B27 segment by ∼10° about B20, and the second is a further rotation of B24–B27 by ∼50° about B24 (Fig. 4*D*). Although the first rotation effects maximal displacement at Phe^{B24}, the latter side chain undergoes a compensating rotameric change (Fig. 4*B*), preserving contacts with the B-chain helix. The second rotation positions B25–B27 antiparallel to the first strand of L1–β_2 and nearly perpendicular to the B-chain α-helix (Fig. 4 *A* and *B*). This series of conformational changes couples detachment of the insulin B-chain C-terminal strand with insertion of the αCT helix (residues 705–714) into the volume occupied by residues B25–B30 in the free hormone (11). Such detachment motivated design of the following insulin analogs to enable interpretation in the context of the holoreceptor.

Analog 1: Stabilization of a β-Turn. [D-Ala^{B20}, D-Ala^{B23}]-insulin (analog 1) contains pairwise D-Ala substitutions, designed to reinforce the native positive ϕ dihedral angles of Gly^{B20} and Gly^{B23} in the B20–B23 β-turn (19, 20). This chiral "lock" enhanced thermodynamic stability ($\Delta\Delta G_u = 1.0 \pm 0.2$ kcal/mol; *SI Appendix*, Fig. S4*A* and Table S4). Whereas in WT insulin crystal structures the β-turn exhibits unusually high thermal B-factors with incomplete side-chain electron density (2), the crystal structure of receptor-free analog 1 (determined at 1.40-Å resolution; *SI Appendix* and *SI Appendix*, Table S5) exhibits a well-ordered β-turn with the D-methyl groups projecting into solvent (Fig. 5); the overall structure closely resembles free insulin (*SI Appendix*, Fig. S5*A*). Molecular modeling suggests that the D-methyl groups can readily be accommodated within the μIR complex (*SI Appendix*, Fig. S5*A*).

To probe dynamics, we used ^{1}H-NMR, using (here and below) a monomeric insulin template[‡] to circumvent self-association of the WT hormone (21). Aromatic ^{1}H-NMR resonances at sites flanking the stabilized β-turn (Tyr^{B16}, Phe^{B24}, and Tyr^{B26}) exhibited enhanced chemical-shift dispersion (*SI Appendix*, Fig. S6 *A* and *B*) and a denser network of interresidue nuclear

[‡]Monomeric insulin analogs contained substitutions Pro^{B28}→Lys and Lys^{B29}→Pro (i.e., parent KP-insulin; PDB ID code 2KJJ) or these with addition of His^{B10}→Asp (DKP-insulin; PDB ID code 2JMN) (Table 1).

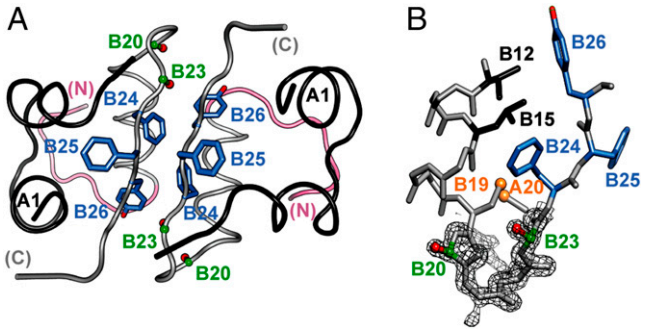

Fig. 5. Structure of analog 1. (*A*) Ribbon model of zinc-free dimer showing residues D-Ala^{B20} and D-Ala^{B23} in each protomer in red (side chain methyl groups) and green (C_α atoms). Side chains of Phe^{B24}, Phe^{B25}, and Tyr^{B26} (and their dimer-related mates) are shown in blue. B-chain ribbons are pink (residues B1–B8) or light gray (B9–B30); A chains are black. (*B*) Expanded view of residues B12–B26, overlaid with weighted $2F_{obs}$-F_{calc} difference electron density. The sulfur atoms of cystine B19-A20 are shown in gold; the coloring scheme is otherwise as in *A*.

Overhauser enhancements (NOEs) than were observed in the parent template (*SI Appendix*, Fig. S6 *C* and *D*), indicating that conformational fluctuations within the B-chain supersecondary structure are damped by the chiral modifications. These NOEs are in accordance with the analog's crystal structure (Fig. 5). The activity of analog 1 as an engineered monomer, evaluated by receptor-binding affinity in vitro and glycemic potency in a rat model of DM (22), was at least as high as that of WT insulin (*SI Appendix*, Figs. S4*B* and 7). Together, these data suggest that the B20–B23 β-turn, as visualized in the refined μIR complex, is maintained as a distinct structural element on holoreceptor binding and activation.

Analog 2: β-Breaker at Phe^{B25}. We next probed the importance of the B-chain hinge at B24–B25 via analog 2, which contains a rigid kink at B25 introduced by ΔPhe (*Z* isomer; Fig. 2*B*). Preserving side-chain size and aromaticity, ΔPhe contains trigonal C_α/C_β configurations, leading to extended conjugation of π electrons (from aromatic ring to formal double bonds C_α=C_β and C=O) (23). Such delocalization enforces near planarity, associated in turn with a break in β-strand–associated main-chain dihedral angles (23). Introduction of ΔPhe at position B25 is well tolerated in the free hormone as monitored by circular dichroism (CD; *SI Appendix*, Fig. S8*A*); thermodynamic stability was indistinguishable from that of its monomeric parent (*SI Appendix*, Fig. S8*B* and Table S4).

^{1}H-NMR spectra of engineered insulin monomers containing ΔPhe^{B25} are similar to those of the native templates (Fig. 6 *A* and *C*). The solution structure of analog 2 (Fig. 6*B* and *D*), determined by heteronuclear multidimensional NMR (24) (*SI Appendix*, Fig. S9 *A–D*; restraints and statistical parameters are given in *SI Appendix*, Table S6), closely resembles WT insulin. Native-like positioning of the B24–B25–B26 aromatic rings is maintained relative to the α-helical domain with a single family of main-chain dihedral angles at B25 (Fig. 6*D*). Control ensembles imposing alternative main-chain conformations [based on peptide structures (23); *SI Appendix*] exhibited similar overall features, including positioning of Tyr^{B26}. Although recapitulating a native-like fold, ΔPhe^{B25} exhibited severe deficits in receptor binding (*SI Appendix*, Fig. S4*B*), μIR assembly (*SI Appendix*, Fig. S4*C*), downstream signaling in cell culture (*SI Appendix*, Fig. S10 *A–C*), and biological activity in vivo (*SI Appendix*, Fig. S10 *F* and *G*). These impairments were more severe than ordinarily encountered among insulin analogs (2). Such marked inactivity suggests that a planar B25 constraint limits the main-chain flexibility required for detachment of the B24–B27 segment and subsequent engagement with the receptor. Indeed, modeling predicts steric

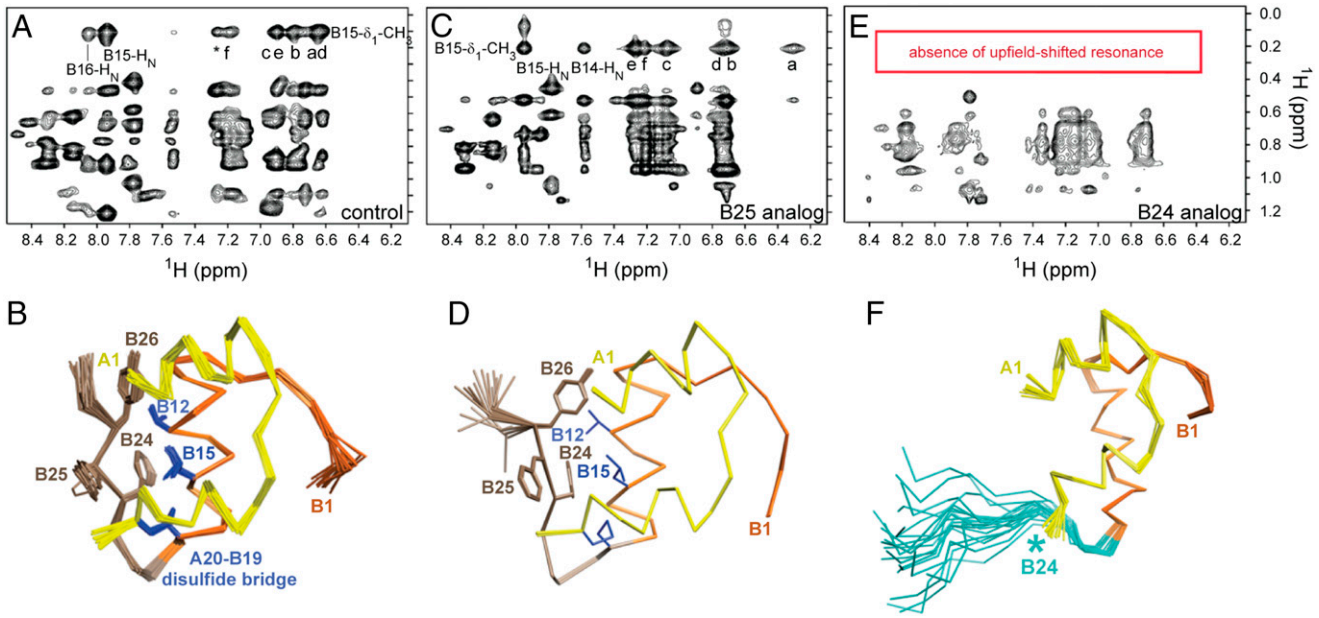

Fig. 6. NMR studies of ΔPhe insulin analogs. (*A*) Baseline NOESY spectrum of KP-insulin. Cross-peak (ω_1, ω_2) assignments are (*a*) Leu^{B15} δ$_1$-CH$_3$/Phe^{B24} H$_\delta$, (*b*) Leu^{B15} δ$_1$-CH$_3$/Phe^{B24} H$_\varepsilon$, (*c*) Leu^{B15} δ$_1$-CH$_3$/Phe^{B24} H$_\varepsilon$, (*d*) Leu^{B15} δ$_1$-CH$_3$/Tyr^{B26} H$_\varepsilon$, (*e*) Leu^{B15} δ$_1$-CH$_3$/Tyr^{B26} H$_\delta$, and (*f*) Leu^{B15} δ$_1$-CH$_3$/Tyr^{A19} H$_\delta$. *B15-δ$_1$-CH$_3$/B25 H$_\delta$. (*B*) Solution structure of parent DKP-insulin (PDB ID code 2JMN). (*C*) NOESY spectrum of ΔPhe^{B25}-KP-insulin. Peak assignments are as labeled in *A*. (*D*) Superposition of 20 DG/RMD structures of ΔPhe^{B25}-DKP-insulin. In each case the A chain is shown in yellow, residues B1–B20 in orange, and B-chain segment B21–B30 in dark brown. Selected side chains are labeled. (*E*) NOESY spectrum of ΔPhe^{B24}-KP-insulin. Red box indicates upfield region of NOESY spectrum in which contacts from the native A19, B24, and B26 aromatic rings to the methyl resonances of Leu^{B15} δ$_{1,2}$-CH$_3$ are ordinarily observed due to ring-current effects; such upfield-shifted methyl resonances are absent in spectra of the ΔPhe^{B24} analog. (*F*) Models of ΔPhe^{B24}-DKP-insulin. NOESY spectra in *A*–*C* were acquired at 25 °C with a mixing time of 200 ms.

clashes between αCT and both ΔPhe^{B25} and Tyr^{B26} and also between L1 and the B27–B30 segment, despite the latter's flexibility (Fig. 7 and *SI Appendix*, Fig. S5*B*). Resolution of these clashes presumably requires conformational distortion of either αCT or the variant B chain.

Analog 3: β-Breaker at Phe^{B24}. Analog 3, containing a ΔPhe-induced kink at position B24, retained high holoreceptor affinity (*SI Appendix*, Fig. S10*A*) and TK activation (*SI Appendix*, Fig. S10 *B* and *C*) with substantial potency in vivo (*SI Appendix*, Fig. S10 *D* and *E*). Its CD spectrum by contrast exhibited attenuated α-helix content (*SI Appendix*, Fig. S8*A*), and stability was markedly reduced (*SI Appendix*, Fig. S8*B*): $\Delta\Delta G_u = -0.8 \pm 0.2$ kcal/mol at 25 °C with respect to the monomeric template (baseline $\Delta G_u = -3.0 \pm 0.1$ kcal/mol; *SI Appendix*, Table S4). Further, ^{1}H-NMR spectra of analog 3 were broadened with reduced chemical shift dispersion (Fig. 6*E* and *SI Appendix*, Fig. S8*F*). Broadening worsened over hours at 25 °C, suggesting progressive aggregation despite the analog's intended monomeric design (17). Similar trends were observed under a variety of conditions (temperature, buffer composition, and pH). Although these features precluded detailed NMR analysis, modeling suggested that the analog populates an ensemble of partial folds encompassing the μIR-bound state of insulin (Fig. 6*F*). Indeed, under a variety of structural assumptions (*SI Appendix*), the near-planar residue at B24 is found to direct the C-terminal B-chain segment away from the helical core of the hormone, rationalizing its productive engagement with the insulin receptor (*SI Appendix*, Fig. S5*C*).

Together, the contrasting properties of analogs 2 and 3 elucidate key structural requirements of the B-chain hinge, with immobilization (via ΔPhe^{B25}) preserving native structure and assembly at the price of receptor binding and forced opening (via ΔPhe^{B24}) preserving activity at the price of protein instability and nonnative aggregation.

Analog 4: Gly^{B24} Detachment and Fibrillation. Insulin fibrillation, a central concern in pharmaceutical formulation (25), provides a model of aggregation-coupled misfolding (26). Whereas protective substitutions are rare among standard variants (22, 27), analog 1 was resistant to fibrillation. On gentle agitation at 37 °C (22), KP-insulin (made 60 μM in PBS at pH 7.4) exhibited a lag time of 3 ± 0.3 d ($n = 20$); under the same conditions, [D-Ala^{B20}, D-Ala^{B23}]-KP-insulin exhibited lag times of 15 and 18 d ($n = 2$), presumably due to its greater dynamic and thermodynamic stability. Analogs 2 and 3 also exhibited longer lag times,[§] but interpretation of these findings is confounded by potential effects of the planar ΔPhe conformation on cross–β-assembly (28, 29).

To test the relationship between C-terminal B-chain detachment (as seen in WT insulin at elevated temperatures that promote insulin fibrillation) (30) and misfolding, we prepared an analog in which Phe^{B24} was substituted by glycine. Analog 4 (Gly^{B24}-KP-insulin) thus lacks a nonpolar side chain to anchor the B24–B28 segment to the α-helical hormone core. Consistent with past studies (4, 5, 31), this analog retains native potency (*SI Appendix*, Fig. S11*A*), despite its decreased thermodynamic stability ($\Delta\Delta G_u = -0.7 \pm 0.2$ kcal/mol at 25 °C relative to KP-insulin) and despite its seeming lack of a conserved receptor contact (*Discussion*). ^{1}H-NMR studies (at neutral pH) demonstrated marked flexibility of the B20–B30 segment without precise packing of Phe^{B25} or Tyr^{B26} against the central B-chain α-helix (*SI Appendix*, Fig. S11 *B–D* and Table S7). The fibrillation lag time of analog 4 is foreshortened (1.0, 1.6, and 1.6 d; $n = 3$). We envisage that the analog's anomalous activity and susceptibility to fibrillation have a common origin: enhanced detachability of the B-chain hinge. Such a shared mechanism is consistent with the accelerated fibrillation of analogs containing C-terminal truncated B chains (25) and the forestalled fibrillation of single-chain insulins (SCIs) whose topology favors the closed hinge conformation (32).

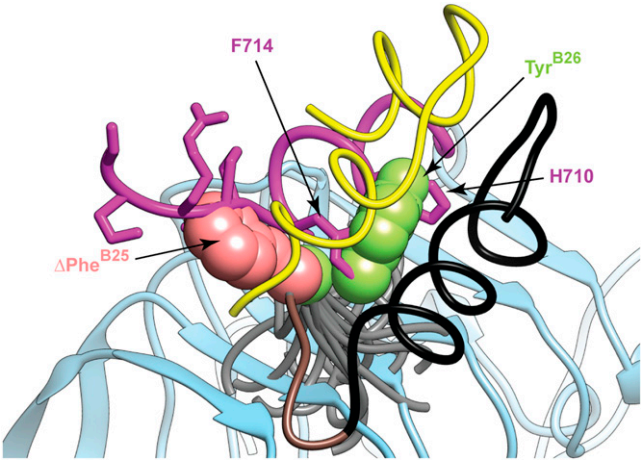

Fig. 7. Solution structure of analog 2 superposed on the μIR complex. Yellow, insulin A chain; black, insulin residues B1–B20; brown, B21–B24; gray, B27–B30; magenta, αCT; cyan, L1 domain. Side-chain atoms are shown for αCT residues His-710 and Phe-714 (magenta), ΔPhe^{B25} (dark pink, space-filling), and Tyr^{B26} (light green, space-filling). Conformational restriction by ΔPhe^{B25} predicts steric clashes (see main text).

Discussion

The classical structure of insulin, determined as a zinc-coordinated hexamer, is remarkable for an aromatic-rich dimerization element. This interface contains six invariant residues (Phe^{B24}-Phe^{B25}-Tyr^{B26} and their dimer-related partners) within an antiparallel β-sheet. The present study defines how the elements of this sheet reorganize to engage conserved receptor surfaces.

Concordance with Prior Biochemical Data. The salient feature of the B24–B27 segment in the μIR complex is its alternating pattern of side-chain contacts with L1 and αCT: Phe^{B24} and Tyr^{B26} are primarily directed toward L1, Phe^{B25}, and Thr^{B27} toward αCT. Such alternation was foreshadowed by photo–cross-linking studies of the insulin-holoreceptor complex (8), supporting the relevance of the μIR model.

Conservation of both Phe^{B24} and its cognate binding pocket among vertebrate sequences is striking. Ala scanning mutagenesis of the pocket-lining residues Asn15, Leu37, Phe39, and Phe714 impairs high-affinity hormone binding to the holoreceptor (33) or ectodomain (34).

The B25-related surface is partly accessible to solvent. Lined by Pro716 and Pro718 in IR-A (Pro716 and Lys718 in IR-B) and multiple main-chain atoms, this open surface presumably contributes to the rigid requirement for an aromatic side chain at B25 (2, 4, 35). In the holoreceptor, substitution of Val715 by Ala eliminates detectable hormone binding (33). It is unclear, however, why substitution of Phe^{B25} by Ala, deleterious in full-length insulin (13), is by contrast well tolerated in truncated analog *des*-pentapeptide[B26-B30]-insulin-amide (see below) (35). It is possible that native Phe^{B25}–receptor interactions contribute to

[§]Analog 2 exhibited lag times of 9, 10, and 10 d ($N = 3$); analog 3 exhibited lag times of 5, 5, and 6 d ($N = 3$). Whereas the latter may have been influenced by both the generic properties of ΔPhe and the initial state of ΔPhe^{B24}-KP-insulin (i.e., protective noncanonical α-helical aggregates), the more marked delay characteristic of monomeric ΔPhe^{B25}-KP-insulin is likely to reflect the dynamic and thermodynamic stabilization of its closed conformation (Fig. 6*D*). Negative-stain transmission electron microscopic studies demonstrated fibrillar structures in all end point samples except those formed by ΔPhe^{B24} -KP-insulin, which appeared amorphous.

coordinated displacement of the B26–B30 segment on receptor binding and are not required by the truncated analog.

The interaction of Tyr^{B26} with site 1 is sparse and restricted to Asp12 and Arg14. This observation rationalizes (*i*) the high activity of an Ala^{B26} analog (36), (*ii*) the dispensability of Tyr^{B26} in truncated analogs (37), and (*iii*) that whereas holoreceptor substitution Arg14Ala (near both Tyr^{B26} and the B25 main chain) impairs insulin binding $>10^3$-fold, Ala substitution of Asp12 (in contact only with Tyr^{B26}) impairs hormone binding by only 6-fold (33). We suggest that the conservation of Tyr^{B26} (2) arises not from its role in receptor binding but instead from its contribution to proinsulin folding (38) and insulin self-assembly (2, 39). Indeed, in the free hormone, Tyr^{B26} inserts within a conserved nonpolar interchain crevice (2); on μIR binding, this crevice is occupied by key αCT side chains His710 and Phe714 (11).

The present structure rationalizes the reduced affinity of clinical variant Leu^{B25}-insulin *Chicago* (7) and essential inactivity of Leu^{A3}-insulin *Wakayama* (7, 40). Receptor engagement at these sites reflects mutual induced fit. Leu^{A3}, predicted to clash with αCT, exhibits a residual receptor-binding affinity (0.1% relative to WT insulin) similar to that of ΔPhe^{B25}. The μIR structure further accounts for effects of C-terminal truncation or modification of the B chain[¶] (41). Truncation at B25 (with C-terminal amide), B26, or B27 yields analogs with complete activity, in accordance with our observations that, beyond Phe^{B24}, successive C-terminal residues exhibit progressively more limited (B25–B27) or absent (B28–B30) interactions with the μIR. Modification of the latter residues can modulate pharmacokinetic properties without loss of activity as exploited in clinical formulations (41).

Analog 4 exploits the intrinsic flexibility of glycine. Its high activity (like that of related analogs) (4, 5, 31, 42) poses a seeming paradox: how can the hormone-receptor interface tolerate the absence of an invariant side chain—especially an aromatic ring anchored within a canonical binding pocket? This question is made more pointed by the enhanced receptor-binding affinity conferred by D-Ala^{B24} and other D substitutions at B24 (6, 42, 43). On the one hand, destabilization of the B20–B30 segment (Gly^{B24}) or its predetachment (D-Ala^{B24}) might mitigate the cost of induced fit and so enhance binding affinity. Such enhancement might be reinforced by an entropic advantage gained by desolvation of nonpolar surfaces—anomalously exposed due to the B24 substitutions—releasing interfacial water molecules into bulk solvent. On the other hand, the entropic cost of immobilizing a flexible B-chain segment on receptor binding would weaken affinity. Further, were such analogs to bind in the same mode as WT insulin, the variant complex would incur a cavity penalty due to an unoccupied B24-related pocket (44). It is possible that these effects cancel each other, rationalizing native activities despite complex perturbations to underlying thermodynamic drivers as a striking example of entropy-enthalpy compensation (EEC).

An alternative and simpler explanation posits a one-residue shift in register between the C-terminal B-chain β-strand and site 1 surface (45). This model envisages that Phe^{B25} occupies the erstwhile B24-binding pocket (and likewise Tyr^{B26} occupies the B25 pocket), leaving a noncanonical five-residue loop between this FY motif and the B-chain α-helix. This model is consistent with the structural compatibility of a register-shifted Thr^{B27} with the exposed B26-related depression in the μIR and lack of interactions by residues B28–B30. Although in their free states Gly^{B24}-DKP-insulin and D-Ala^{B24}-DKP-insulin (6) lack ^{1}H-NMR features of a register shift (at least not as a stable structural element), we imagine that such a reorganized interface may be more favorable than a binding mode that leaves the B24-related pocket empty. The register-shifted and EEC models may be distinguished through crystallographic studies of variant μIR complexes.

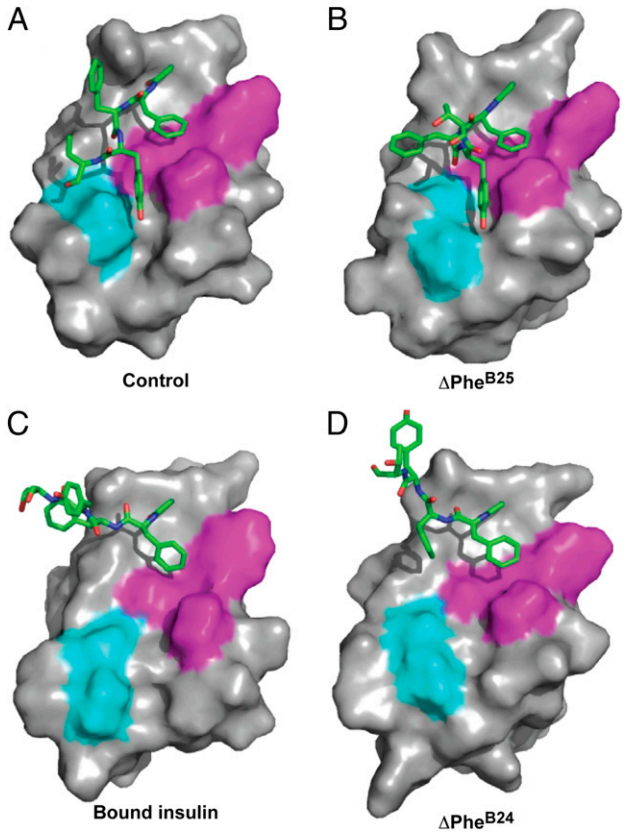

Fig. 8. Orientation of B23–B27 segments in free insulin analogs. The closed B-chain conformation in DKP-insulin (*A*) is recapitulated in the inactive constrained conformation of ΔPhe^{B25}-DKP-insulin (*B*). The open conformation observed in the μIR complex (*C*) is in accordance with molecular models of the active but less stable ΔPhe^{B24} analog (*D*). Green, B23–B27 segment; magenta, the insulin internal contact surface of Phe^{B24} (which includes Val^{B12}, Leu^{B15}, Tyr^{B16}, and Cys^{B19}; Tyr^{B26} also packs against Val^{B12}); cyan, additional contact surface of Tyr^{B26} with A-chain residues Ile^{A2} and Val^{A3}. The latter nonpolar side chains are exposed in *C* and *D* and so poised to engage αCT on receptor binding.

Induced Fit and Biological Signaling. The function of the B-chain hinge in cellular signaling was probed through comparative studies of ΔPhe analogs. Remarkably, although ΔPhe^{B24}-KP-insulin (analog 3) is unstable and prone to aggregation, native receptor-binding affinity was maintained. Further, analog 3 was able to direct reduction of blood glucose concentration in vivo. Modeling, based on qualitative CD and NMR features, suggests that the ΔPhe^{B24} analog populates an ensemble of partial folds consistent with displacement of the B20–B27 segment in the μIR complex. Unlike ΔPhe^{B24}-KP-insulin, the free ΔPhe^{B25} analog exhibited unperturbed stability with the B26–B30 segment directed toward the hormone's helical core as in WT insulin. The contrasting conformations of the B-chain C-terminal segments in the ΔPhe analogs are shown in Fig. 8 in relation to the free structure of insulin and its μIR-bound conformation. ΔPhe^{B25}-DKP-insulin thus recapitulates classical structural relationships (1, 2). Strikingly, however, this analog has essentially no biological activity, mirroring Leu^{A3}-insulin (31, 46) and inactive SCIs (3). Loss of activity presumably reflects steric clash between the B25–B30 segment and αCT. Thus, the μIR model—despite its

[¶]Nonstandard analogs directing B26 (and, where present, residues beyond) away from the insulin core exhibit high affinity presumably due to general predisplacement of B26–B30 rather than specific mimicry of the receptor-bound state as originally envisioned (69).

intrinsic limitations as a small fragment of the holoreceptor complex—may rationalize the contrasting properties of the constrained analogs and broadly inform classical structure–activity relationships as discussed above.

Insulin-Like Growth Factor System. Concordant with their sequence homology, insulin and the two insulin-like growth factors (IGF-I and IGF-II) are each capable of binding to both the insulin receptor (isoforms A and B) and IGF-1R (47). In structures of free IGFs (48, 49), the corresponding B22–B26 segment has an almost identical disposition (with respect to the α-helical core) to that in insulin; the L1 structure and αCT sequence are also conserved between the insulin receptor and IGF-1R (11). We thus expect that IGF B domains undergo a similar hinge-like detachment on binding to site 1 of IGF-1R. We note in particular that (*i*) the B20–B23 β-turn (sequence GERG in insulin) is conserved among IGFs both in sequence (GDRG) and structure (48, 49)—suggesting that the homologous turn is also maintained as IGFs engage IGF-1R; and (*ii*) the aromatic triplet in insulin (Phe^{B24}-Phe^{B25}-Tyr^{B26}) is conserved in IGFs (as Phe-Tyr-Phe) as are its cognate binding residues in IGF-1R (*SI Appendix*, Fig. S12 *A* and *B*).

Determinants of Foldability. A monogenic form of DM is due to dominant mutations in the insulin gene leading to misfolding of proinsulin (50, 51). Age of onset reflects the severity of the folding defect and extent of interference with WT insulin biosynthesis (38). Of particular interest is Gly^{B23}→Val (50), predicted (as an L-amino acid) to destabilize a positive ϕ dihedral angle at B23 (20). We suggest that Val^{B23} impairs disulfide pairing in the variant proinsulin by perturbing the orientation of Phe^{B24} and adjoining C-terminal B-domain segment. This model is supported by observations that this segment facilitates classical insulin chain combination (37) and that Phe^{B24} stabilizes nascent structure in peptide models of proinsulin folding intermediates (38).

Substitution of Phe^{B24} by serine (insulin *Los Angeles*) leads to variable onset of DM in early adulthood (46). In both the free and μIR-bound hormone, the side chain of Phe^{B24} abuts cystine B19–A20. Although secretion of Ser^{B24}-insulin was observed in a patient (46), Ser^{B24} was found in a β-cell line to perturb disulfide pairing (52). Ser^{B24}-associated DM may thus have dual origins—impaired receptor binding by the mutant insulin (46) and β-cell dysfunction resulting from chronic endoplasmic reticular (ER) stress (52). The extent of impaired binding (14-fold) is less marked than that of insulins *Chicago* or *Wakayama*: Phe^{B25}→Leu (50-fold) or Val^{A3}→Leu (1000-fold) (31, 46). Because Ser^{B24} introduces segmental B-chain destabilization similar to that of Gly^{B24} (53), its partial activity may be rationalized by the above register-shifted or EEC models.

Fibrillation. Chiral stabilization of the B20–B23 β-turn (analog 1) delays insulin fibrillation, whereas its destabilization (analog 4) promotes fibrillation. These findings may be rationalized in relation to a general scheme of insulin fibrillation (Fig. 9) based on conformational distortion of a susceptible monomer (27). Whereas the native hormone is protected within storage hexamers (Fig. 9, *Left*), nonnative aggregates can form amyloidogenic nuclei leading to cross–β-assembly in protofilaments and fibrils (Fig. 9, *Right*). We envisage hinge opening both enables receptor binding and underlies the susceptibility of the monomeric hormone to fibrillation. Like populated partial folds in amyloidogenic proteins in general (28, 29), conformational distortion of insulin represents a breakdown of the cooperativity of protein folding.

Lack of correlation between fibrillation lag times of insulin analogs and native state stability (22, 27) reflects a key aspect of fibrillation: the unfolded-state ensemble is off-pathway (Fig. 9, *Top*). Whereas measurements of stability ordinarily probe free-energy differences (ΔG_u) relative to this ensemble, susceptibility to fibrillation may be influenced by the relative stabilities of partial folds, intervening kinetic barriers, and mechanisms of secondary nucleation (29). That hinge opening enhances susceptibility to fibrillation is in accordance with the stereospecific effects of L- and D-Ala substitutions at B24: although these diastereomers exhibit similar thermodynamic stabilities, the lag time of the (active) D-Ala^{B24} analog (with unstructured B20-B30 segment) is threefold shorter than that of the (inactive) L-Ala^{B24} analog, whose hinge is partially closed (6). It would be of future interest to dissect which individual kinetic steps in the mechanism of insulin fibrillation (54, 55) are influenced by substitutions at position B24 in relation to the dynamics of hinge opening. Such analogs provide a model of segmental disorder, a general feature of mutant proteins associated with diverse diseases of toxic protein misfolding and amyloid deposition (29).

Aggregation-coupled misfolding may have evolutionary consequences. In the caviomorph rodent *Octodon degus*, for example, endogenous fibrillation of a divergent insulin can lead to selective islet amyloidosis associated with β-cell loss and DM (56, 57). The degu B chain contains multiple anomalous substitutions predicted to impair dimerization (Tyr^{B26}→Arg, Thr^{B27}→Pro, and Pro^{B28}→His) and the zinc-stabilized hexamer assembly (His^{B10}→Asn). The marked susceptibility of such rodents to β-cell dysfunction when fed a diet high in carbohydrates (especially free sugars, which are almost absent in its natural plant-based diet) may reflect both extracellular- and intracellular proteotoxicity, respectively, due to amyloidogenesis and impaired folding of the divergent proinsulin—the latter leading to ER stress as in human neonatal DM (52). The anomalous molecular features of degu insulin may thus honor in the breach our suggestion that the closed conformation of insulin and its stable self-assembly in secretory vesicles (2, 39) evolved to protect the β cell from proteotoxicity.

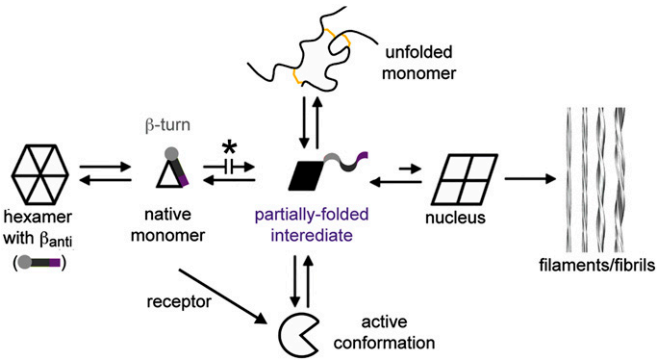

Fig. 9. Mechanism of insulin fibrillation via partial unfolding. The native state is protected by classical self-assembly (far left), mediated in part by an anti-parallel β-sheet (β_{anti}) at the dimer interfaces of the hexamer. C-terminal segment of B-chain is represented by light gray circle (B20–B23 β-turn), gray bar (B24–B28 β-strand), and purple bar (less ordered C-terminal residues B29 and B30) (1, 2). Disassembly leads to equilibrium between native- and partially folded monomers (open triangle and black trapezoid). Detachment of the C-terminal B-chain segment within a partial fold (5, 6) may lead to the off-pathway, active conformation (open circle) or to an aggregated nucleus en route to a protofilament assembly (far right). Asterisk highlights protective modifications stabilizing the closed conformation: D-Ala–locked β-turn (analog 1), ΔPhe^{B25} (analog 2 relative to ΔPhe^{B24}), and SCIs (25, 32, 70, 71). The unfolded state, constrained by native disulfide bridges (gold), shown in schematic form at the top, is off-pathway. Reproduced with permission from ref. 26.

Concluding Remarks. The present study illustrates the power of nonstandard structural constraints (as exemplified by chiral β-turn stabilization and by main-chain planarity enforced through extended π conjugation) to interrogate a functional hinge in a globular protein. Our results have extended the repertoire of insulin structures through multicrystal refinement (18) of the conformation of the hormone bound to the L1 and αCT domains of the receptor α-subunit (11). In this μIR complex, the C-terminal β-turn (residues B20–B23) and β-strand (B24–B27) rotate away from the α-helical core of insulin to enable its conserved aromatic motif (Phe^{B24}-Phe^{B25}-Tyr^{B26}) to dock between L1 and αCT. The B chain thus contains a hinge whose closed setting functions in β-cell biosynthesis—from the nascent folding of proinsulin (38) to zinc-mediated self-assembly (2, 39)—and whose open setting mediates biological activity.

The μIR model provides a simplified view of the structural complexity of the intact insulin receptor. Indeed, asymmetric binding of insulin to site 1, as visualized here, defines only the first step of what is likely to be a complex choreography of conformational change leading to activation of the receptor tyrosine kinase domain and its multiple signaling outputs. Clues regarding such subsequent steps have been provided by analysis of mutations at the edge of the present site 1-binding surface (58) or at sites in contact with site 2 of the holoreceptor (9). These analogs distinguish between the thermodynamics of receptor binding and determinants of biological signaling. The μIR surface in contact with the conserved aromatic side chains at insulin positions B24–B26 is intimately connected to putative signaling surfaces of αCT and insulin. In particular, the deep and distinctive B24-binding pocket may provide an attractive target for engineering insulin analogs with enhanced potency or selective activities as biased agonists. Such molecular engineering defines a key emerging frontier of pharmacology.

Materials and Methods

Detailed materials and methods are available in *SI Appendix*.

Preparation of Insulin Analogs. [^{13}C, ^{15}N]-labeled peptides were prepared by solid-phase peptide synthesis, purified by HPLC, and assessed by MS. Unlabeled B23–B30 octapeptides containing Ala, Gly, or ΔPhe substitutions were likewise prepared; ΔPhe-containing octapeptides were also synthesized as described (12). Labeled or variant insulins were prepared by semisynthesis (42). Analog 1 was prepared by chain combination (15). Biosynthetic isotopic labeling of the α-helical domain of analogs 2 and 3 (residues B1–B22 and A1–A21) was accomplished in *Pichia pastoris* (14).

μIR Stability and Holoreceptor Binding. Relative stabilities of variant μIR complexes were assessed in a competitive binding assay using polyethylene glycol (PEG) precipitation (59); holoreceptor binding affinities were obtained as previously described (60).

TK Activation Assays. Relative activities of insulin analogs were evaluated in an in vitro assay of hormone-stimulated receptor autophosphorylation (33).

Biological Studies. Insulin-stimulated receptor autophosphorylation and protein kinase B phosphorylation were assessed in an IGF-1R–deficient mouse embryo fibroblast cell line expressing the human insulin receptor (isoform B) (61). Potency of insulin analogs was tested in male Sprague–Dawley rats rendered diabetic by streptozotocin (22). Statistical significance was assessed using a Student *t* test.

Fibrillation Assays. Fibrils were induced at 37 °C by gentle agitation of a solution of insulin analog at 60 μM in PBS (pH 7.4) and monitored by thioflavin T fluorescence and transmission electron microscopy (EM) (22).

Crystallographic Analysis of the μIR Complex. Crystal growth and data collection protocols were as previously described (11). Six datasets displayed diffraction to higher than 4.0-Å resolution—two successfully comerged to 3.5 Å with that originally reported. The original structure [Protein Data Bank (PDB) ID code 3W11 (11)] was then refined against the merged set. Inspection of the $2F_{obs}$-F_{calc} difference electron density map revealed two polypeptide-like segments: one running approximately parallel to the first strand of L1–β_2 and the other extending onward from IR 715 (the last modeled residue of the αCT); these segments could be successfully mapped to insulin B22–B27 and αCT 716–719, respectively. These additional residues were then included in the model, and the entire model further refined at 3.5-Å resolution. Data processing and refinement statistics are in *SI Appendix*, Table S3.

Crystallographic Analysis of Analog 1. Zinc-free crystals of analog 1 were obtained during attempts to cocrystallize with the μIR, Fab 83-7, and αCT peptide; conditions were subsequently refined to 0.7 M trisodium citrate and 0.1 M imidazole-HCl (pH 8.0). Diffraction data were collected at beamline MX2 (Australian Synchrotron). The structure was refined (following molecular replacement) to a resolution of 1.40 Å; data processing and refinement statistics are in *SI Appendix*, Table S5.

Circular Dichroism and Stabilities. CD spectra acquisition, assays of guanidine denaturation, and thermodynamic modeling were undertaken as previously described (22).

NMR Spectroscopy and Structure Calculations. Spectra were acquired at 600, 700, and 800 MHz; [^{1}H, ^{13}C]-TROSY spectra were acquired at 900 MHz (16). Data were processed using NMRPipe (62). Structures of analogs 1 and 4 were determined as previously described (6, 15); interpretation of long-range NOEs was aided by use of multiple mixing times and model-based back-calculation of predicted NOESY spectra and aromatic ring-current shifts. Heteronuclear studies of analog 2 were performed as previously described (22). In brief, assignments were based on 3D HNCACB, CBCA(CO)NH, C(CO)NH, H(CCO)NH, and HCCH-TOCSY spectra and extended by analyses of 3D ^{13}C- and ^{15}N-separated NOESY-HSQC spectra. Distance restraints in the labeled domain were derived from 4D-NOESY spectra. The solution structure of analog 2 was calculated using X-PLOR-NIH. Models of analog 3 were similarly calculated based on structural assumptions.

Molecular Modeling of μIR Complexes. Modeling was performed using MODELLER (63) with μIR fragments from PDB ID code 3W11 (11) and native insulin from PDB ID code 2G4M (64). Molecular dynamics calculations were performed with GROMACS (v4.5.5) (65) using the OPLS-aa force field (66, 67). Partial atomic charges for ΔPhe were obtained from the electrostatic potential calculated with GAUSSIAN-09 (68).

ACKNOWLEDGMENTS. We thank Q. X. Hua and K. Huang for initial studies; M. Gupta for assistance with ΔPhe-related synthesis; P. G. Katsoyannis, S. B. Kent, and S. H. Nakagawa for advice; K. Siddle for supplying the 83-7 hybridomas; and L. Lu and the fermentation group at the Commonwealth Scientific and Industrial Research Organization (CSIRO) Materials Science and Engineering laboratory (Parkville, Australia). This work was supported by Australian National Health and Medical Research Council (NHMRC) Project Grants 1005896 and 1058233 and the Hazel and Pip Appel Fund (to M.C.L.); National Institutes of Health (NIH) Grants DK04949 and DK079233 and The Leona M. and Harry B. Helmsley Charitable Trust (to M.A.W.); Grant 7-13-IN-31 from the American Diabetes Association (to N.B.P.); and NHMRC Independent Research Institutes Infrastructure Support Scheme Grant 361646 and a Victorian State Government Operational Infrastructure support grant (to the Walter and Eliza Hall Institute). S.J.C. and D.F.S. received support from National Institutes of Health Grants DK013914 and UC DRTC DK020595. V.P. received support from Institutional NIH Medical Scientist Training Program Grant T32 GM007250 and NIH Fellowship F30 DK094685-04. Structural data were obtained at the Australian Synchrotron (beam line MX2) and Cleveland Center for Membrane and Structural Biology.

1. Adams MJ, et al. (1969) Structure of rhombohedral 2 zinc insulin crystals. *Nature* 224(5218):491–495.
2. Baker EN, et al. (1988) The structure of 2Zn pig insulin crystals at 1.5 A resolution. *Philos Trans R Soc Lond B Biol Sci* 319(1195):369–456.
3. Derewenda U, et al. (1991) X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue. *J Mol Biol* 220(2):425–433.
4. Mirmira RG, Nakagawa SH, Tager HS (1991) Importance of the character and configuration of residues B24, B25, and B26 in insulin-receptor interactions. *J Biol Chem* 266(3):1428–1436.
5. Hua QX, Shoelson SE, Kochoyan M, Weiss MA (1991) Receptor binding redefined by a structural switch in a mutant human insulin. *Nature* 354(6350):238–241.
6. Hua QX, et al. (2009) Enhancing the activity of a protein by stereospecific unfolding: Conformational life cycle of insulin and its evolutionary origins. *J Biol Chem* 284(21):14586–14596.

7. Shoelson S, et al. (1983) Three mutant insulins in man. *Nature* 302(5908):540–543.
8. Xu B, et al. (2009) Decoding the cryptic active conformation of a protein by synthetic photoscanning: Insulin inserts a detachable arm between receptor domains. *J Biol Chem* 284(21):14597–14608.
9. De Meyts P, Whittaker J (2002) Structural biology of insulin and IGF1 receptors: implications for drug design. *Nat Rev Drug Discov* 1(10):769–783.
10. Smith BJ, et al. (2010) Structural resolution of a tandem hormone-binding element in the insulin receptor and its implications for design of peptide agonists. *Proc Natl Acad Sci USA* 107(15):6771–6776.
11. Menting JG, et al. (2013) How insulin engages its primary binding site on the insulin receptor. *Nature* 493(7431):241–245.
12. Gupta M, et al. (2008) Dehydrophenylalanine (DeltaPhe) as a β breaker: Extended structure terminated by a DeltaPhe-induced turn in the pentapeptide Boc-Phe1-Ala2-Ile3-DeltaPhe4-Ala5-OMe. *ChemBioChem* 9(9):1375–1378.
13. Kobayashi M, et al. (1984) Changes in receptor binding, biological activity and immunoreactivity of insulin caused by replacing the residues B23-B26 with alanine. *Biomed Res* 5(3):267–272.
14. Huang Y, Liang Z, Feng Y (2001) The relationship between the connecting peptide of recombined single chain insulin and its biological function. *Sci China C Life Sci* 44(6): 593–600.
15. Huang K, et al. (2004) How insulin binds: The B-chain α-helix contacts the L1 β-helix of the insulin receptor. *J Mol Biol* 341(2):529–550.
16. Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci USA* 94(23):12366–12371.
17. Hua QX, Jia W, Weiss MA (2011) Conformational dynamics of insulin. *Front Endocrinol (Lausanne)* 2:48.
18. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084):1030–1033.
19. Anil B, Song B, Tang Y, Raleigh DP (2004) Exploiting the right side of the Ramachandran plot: Substitution of glycines by D-alanine can significantly increase protein stability. *J Am Chem Soc* 126(41):13194–13195.
20. Nakagawa SH, et al. (2006) Chiral mutagenesis of insulin. Contribution of the B20-B23 β-turn to activity and stability. *J Biol Chem* 281(31):22386–22396.
21. Weiss MA, Hua QX, Lynch CS, Frank BH, Shoelson SE (1991) Heteronuclear 2D NMR studies of an engineered insulin monomer: Assignment and characterization of the receptor-binding surface by selective ^{2}H and ^{13}C labeling with application to protein design. *Biochemistry* 30(30):7373–7389.
22. Yang Y, et al. (2010) An Achilles' heel in an amyloidogenic protein and its repair: Insulin fibrillation and therapeutic design. *J Biol Chem* 285(14):10806–10821.
23. Gupta M, Chauhan VS (2011) De novo design of α,β-didehydrophenylalanine containing peptides: From models to applications. *Biopolymers* 95(3):161–173.
24. Yang Y, et al. (2010) Solution structure of proinsulin: Connecting domain flexibility and prohormone processing. *J Biol Chem* 285(11):7847–7851.
25. Brange J, Andersen L, Laursen ED, Meyn G, Rasmussen E (1997) Toward understanding insulin fibrillation. *J Pharm Sci* 86(5):517–525.
26. Jiménez JL, et al. (2002) The protofilament structure of insulin amyloid fibrils. *Proc Natl Acad Sci USA* 99(14):9196–9201.
27. Nielsen L, Frokjaer S, Brange J, Uversky VN, Fink AL (2001) Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry* 40(28):8397–8409.
28. Booth DR, et al. (1997) Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* 385(6619):787–793.
29. Dobson CM (2003) Protein folding and misfolding. *Nature* 426(6968):884–890.
30. Hua QX, Weiss MA (2004) Mechanism of insulin fibrillation: The structure of insulin under amyloidogenic conditions resembles a protein-folding intermediate. *J Biol Chem* 279(20):21449–21460.
31. Shoelson SE, Lu ZX, Parlautan L, Lynch CS, Weiss MA (1992) Mutations at the dimer, hexamer, and receptor-binding surfaces of insulin independently affect insulin-insulin and insulin-receptor interactions. *Biochemistry* 31(6):1757–1767.
32. Huang K, Dong J, Phillips NB, Carey PR, Weiss MA (2005) Proinsulin is refractory to protein fibrillation: Topological protection of a precursor protein from cross-β assembly. *J Biol Chem* 280(51):42345–42355.
33. Whittaker J, Whittaker L (2005) Characterization of the functional insulin binding epitopes of the full-length insulin receptor. *J Biol Chem* 280(22):20932–20936.
34. Williams PF, Mynarcik DC, Yu GQ, Whittaker J (1995) Mapping of an NH2-terminal ligand binding site of the insulin receptor by alanine scanning mutagenesis. *J Biol Chem* 270(7):3012–3016.
35. Mirmira RG, Tager HS (1991) Disposition of the phenylalanine B25 side chain during insulin-receptor and insulin-insulin interactions. *Biochemistry* 30(33):8222–8229.
36. Kristensen C, et al. (1997) Alanine scanning mutagenesis of insulin. *J Biol Chem* 272(20):12978–12983.
37. Cosmatos A, Ferderigos N, Katsoyannis PG (1979) Chemical synthesis of [*des*(tetrapeptide B^{27-30}), Tyr$(NH_2)^{26}$-B] and [*des*(pentapeptide B^{26-30}), Phe$(NH_2)^{25}$-B] bovine insulins. *Int J Pept Protein Res* 14(5):457–471.
38. Weiss MA (2013) Diabetes mellitus due to the toxic misfolding of proinsulin variants. *FEBS Lett* 587(13):1942–1950.
39. Dodson G, Steiner D (1998) The role of assembly in insulin's biosynthesis. *Curr Opin Struct Biol* 8(2):189–194.
40. Nanjo K, et al. (1986) Diabetes due to secretion of a structurally abnormal insulin (insulin Wakayama). Clinical and functional characteristics of [Leu^{A3}] insulin. *J Clin Invest* 77(2):514–519.
41. Mayer JP, Zhang F, DiMarchi RD (2007) Insulin structure and function. *Biopolymers* 88(5):687–713.
42. Mirmira RG, Tager HS (1989) Role of the phenylalanine B24 side chain in directing insulin interaction with its receptor. Importance of main chain conformation. *J Biol Chem* 264(11):6349–6354.
43. Žáková L, et al. (2013) Structural integrity of the B24 site in human insulin is important for hormone functionality. *J Biol Chem* 288(15):10230–10240.
44. Matthews BW (1995) Studies on protein stability with T4 lysozyme. *Adv Protein Chem* 46:249–278.
45. Ludvigsen S, Olsen HB, Kaarsholm NC (1998) A structural switch in a mutant insulin exposes key residues for receptor binding. *J Mol Biol* 279(1):1–7.
46. Shoelson SE, Polonsky KS, Zeidler A, Rubenstein AH, Tager HS (1984) Human insulin B24 (Phe–Ser). Secretion and metabolic clearance of the abnormal insulin in man and in a dog model. *J Clin Invest* 73(5):1351–1358.
47. De Meyts P, Whittaker J (2002) Structural biology of insulin and IGF1 receptors: Implications for drug design. *Nat Rev Drug Discov* 1(10):769–783.
48. Vajdos FF, et al. (2001) Crystal structure of human insulin-like growth factor-1: Detergent binding inhibits binding protein interactions. *Biochemistry* 40(37): 11022–11029.
49. Brzozowski AM, et al. (2002) Structural origins of the functional divergence of human insulin-like growth factor-I and insulin. *Biochemistry* 41(30):9389–9397.
50. Støy J, et al.; Neonatal Diabetes International Collaborative Group (2007) Insulin gene mutations as a cause of permanent neonatal diabetes. *Proc Natl Acad Sci USA* 104(38): 15040–15044.
51. Colombo C, et al.; Early Onset Diabetes Study Group of the Italian Society of Pediatric Endocrinology and Diabetes (SIEDP) (2008) Seven mutations in the human insulin gene linked to permanent neonatal/infancy-onset diabetes mellitus. *J Clin Invest* 118(6):2148–2156.
52. Liu M, et al. (2010) Mutant INS-gene induced diabetes of youth: Proinsulin cysteine residues impose dominant-negative inhibition on wild-type proinsulin transport. *PLoS ONE* 5(10):e13333.
53. Hua QX, Shoelson SE, Inouye K, Weiss MA (1993) Paradoxical structure and function in a mutant human insulin associated with diabetes mellitus. *Proc Natl Acad Sci USA* 90(2):582–586.
54. Knowles TPJ, et al. (2009) An analytical solution to the kinetics of breakable filament assembly. *Science* 326(5959):1533–1537.
55. Buell AK, et al. (2013) Electrostatic effects in filamentous protein aggregation. *Biophys J* 104(5):1116–1126.
56. Nishi M, Steiner DF (1990) Cloning of complementary DNAs encoding islet amyloid polypeptide, insulin, and glucagon precursors from a New World rodent, the degu, *Octodon degus*. *Mol Endocrinol* 4(8):1192–1198.
57. Hellman U, et al. (1990) Amino acid sequence from degu islet amyloid-derived insulin shows unique sequence characteristics. *Biochem Biophys Res Commun* 169(2): 571–577.
58. Whittaker J, et al. (2012) α-Helical element at the hormone-binding surface of the insulin receptor functions as a signaling element to activate its tyrosine kinase. *Proc Natl Acad Sci USA* 109(28):11166–11171.
59. Kristensen C, Andersen AS, Ostergaard S, Hansen PH, Brandt J (2002) Functional reconstitution of insulin receptor binding site from non-binding receptor fragments. *J Biol Chem* 277(21):18340–18345.
60. Whittaker L, Hao C, Fu W, Whittaker J (2008) High-affinity insulin binding: Insulin interacts with two receptor ligand binding sites. *Biochemistry* 47(48):12900–12909.
61. Denley A, et al. (2007) Differential activation of insulin receptor substrates 1 and 2 by insulin-like growth factor-activated insulin receptors. *Mol Cell Biol* 27(10):3569–3577.
62. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6(3):277–293.
63. Fiser A, Sali A (2003) Modeller: Generation and refinement of homology-based protein structure models. *Methods Enzymol* 374:461–491.
64. Mueller-Dieckmann C, et al. (2007) On the routine use of soft X-rays in macromolecular crystallography. Part IV. Efficient determination of anomalous substructures in biomacromolecules using longer X-ray wavelengths. *Acta Crystallogr D Biol Crystallogr* 63(Pt 3):366–380.
65. Hess B, Kutzner C, Van Der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4(3):435–447.
66. Jorgensen WL, Tirado-Rives J (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* 110:1657–1666.
67. Damm W, Frontera A, Tirado–Rives J, Jorgensen WL (1997) OPLS all-atom force field for carbohydrates. *J Comput Chem* 18(16):1955–1970.
68. Frisch M, et al. (2009) *Gaussian 09, Revision A. 02* (Gaussian, Wallingford, CT).
69. Jirácek J, et al. (2010) Implications for the active form of human insulin based on the structural convergence of highly active hormone analogues. *Proc Natl Acad Sci USA* 107(5):1966–1970.
70. Hua QX, et al. (2008) Design of an active ultrastable single-chain insulin analog: Synthesis, structure, and therapeutic implications. *J Biol Chem* 283(21):14703–14716.
71. Phillips NB, Whittaker J, Ismail-Beigi F, Weiss MA (2012) Insulin fibrillation and protein design: Topological resistance of single-chain analogs to thermal degradation with application to a pump reservoir. *J Diabetes Sci Tech* 6(2):277–288.