

Chiral Mutagenesis of Insulin. Foldability and Function Are Inversely Regulated by a Stereospecific Switch in the B Chain^{†,‡}

Satoe H. Nakagawa,[§] Ming Zhao,^{§,||} Qing-xin Hua,⁺ Shi-Quan Hu,^{#,®} Zhu-li Wan,⁺ Wenhua Jia,⁺ and Michael A. Weiss^{*,+}

Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637,

Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106-4935, and Department of Pharmacology and Biological Chemistry, Mt. Sinai School of Medicine of New York University, New York, New York 10029

Received September 13, 2004; Revised Manuscript Received November 23, 2004

ABSTRACT: How insulin binds to its receptor is unknown despite decades of investigation. Here, we employ chiral mutagenesis—comparison of corresponding D and L amino acid substitutions in the hormone—to define a structural switch between folding-competent and active conformations. Our strategy is motivated by the T → R transition, an allosteric feature of zinc-hexamers in which an invariant glycine in the B chain changes conformations. In the classical T state, Gly^{B8} lies within a β -turn and exhibits a positive ϕ angle (like a D amino acid); in the alternative R state, Gly^{B8} is part of an α -helix and exhibits a negative ϕ angle (like an L amino acid). Respective B chain libraries containing mixtures of D or L substitutions at B8 exhibit a stereospecific perturbation of insulin chain combination: L amino acids impede native disulfide pairing, whereas diverse D substitutions are well-tolerated. Strikingly, D substitutions at B8 enhance both synthetic yield and thermodynamic stability but markedly impair biological activity. The NMR structure of such an inactive analogue (as an engineered T-like monomer) is essentially identical to that of native insulin. By contrast, L analogues exhibit impaired folding and stability. Although synthetic yields are very low, such analogues can be highly active. Despite the profound differences between the foldabilities of D and L analogues, crystallization trials suggest that on protein assembly substitutions of either class can be accommodated within classical T or R states. Comparison between such diastereomeric analogues thus implies that the T state represents an inactive but folding-competent conformation. We propose that within folding intermediates the sign of the B8 ϕ angle exerts kinetic control in a rugged landscape to distinguish between trajectories associated with productive disulfide pairing (positive T-like values) or off-pathway events (negative R-like values). We further propose that the crystallographic T → R transition in part recapitulates how the conformation of an insulin monomer changes on receptor binding. At the very least the ostensibly unrelated processes of disulfide pairing, allosteric assembly, and receptor binding appear to utilize the same residue as a structural switch; an “ambidextrous” glycine unhindered by the chiral restrictions of the Ramachandran plane. We speculate that this switch operates to protect insulin—and the β -cell—from protein misfolding.

Insulin, a small globular protein containing three disulfide bridges, plays a central role in the regulation of vertebrate

metabolism. The hormone is stored in the pancreatic β cell as a zinc hexamer and functions in the bloodstream as a zinc-free monomer. The functional surface of insulin has long been the object of speculation. Despite decades of investigation by mutagenesis and X-ray crystallography (1, 2), structures of insulin and insulin analogues do not consistently predict relative potencies (3–5). Such anomalies suggest that a change in structure occurs on receptor binding. Here, chiral mutagenesis—comparison of corresponding D and L amino acid substitutions of an invariant glycine—is employed to investigate the interrelation of structure and function in the B chain.¹ Experimental design is motivated by the classical T → R transition, an allosteric feature of zinc-insulin hexamers (6–8).² Remarkably, chiral stabilization of the native T state (the predominant conformation in solution;

[†] This work was supported in part by Diabetes Research and Training Center at the University of Chicago (S.H.N. and M.Z.) and by grants from the National Institutes of Health to P. G. Katsoyannis (DK56673) and M.A.W. (DK409049 and DK069764).

[‡] The coordinates of D-Ala^{B8}-DKP-insulin will be deposited on acceptance in the RCSB Protein Data Bank (accession number pending).

^{*} To whom correspondence should be addressed. E-mail: michael.weiss@case.edu; telephone: (216) 368-5991; fax: (216) 368-3419.

[§] University of Chicago.

^{||} Present address: Glycopep Company, 2201 W. Campbell Park Dr., Chicago, IL 60612.

⁺ Case Western Reserve University.

[#] Mt. Sinai School of Medicine of New York University.

[®] Present address: Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, OH 44106-4935.

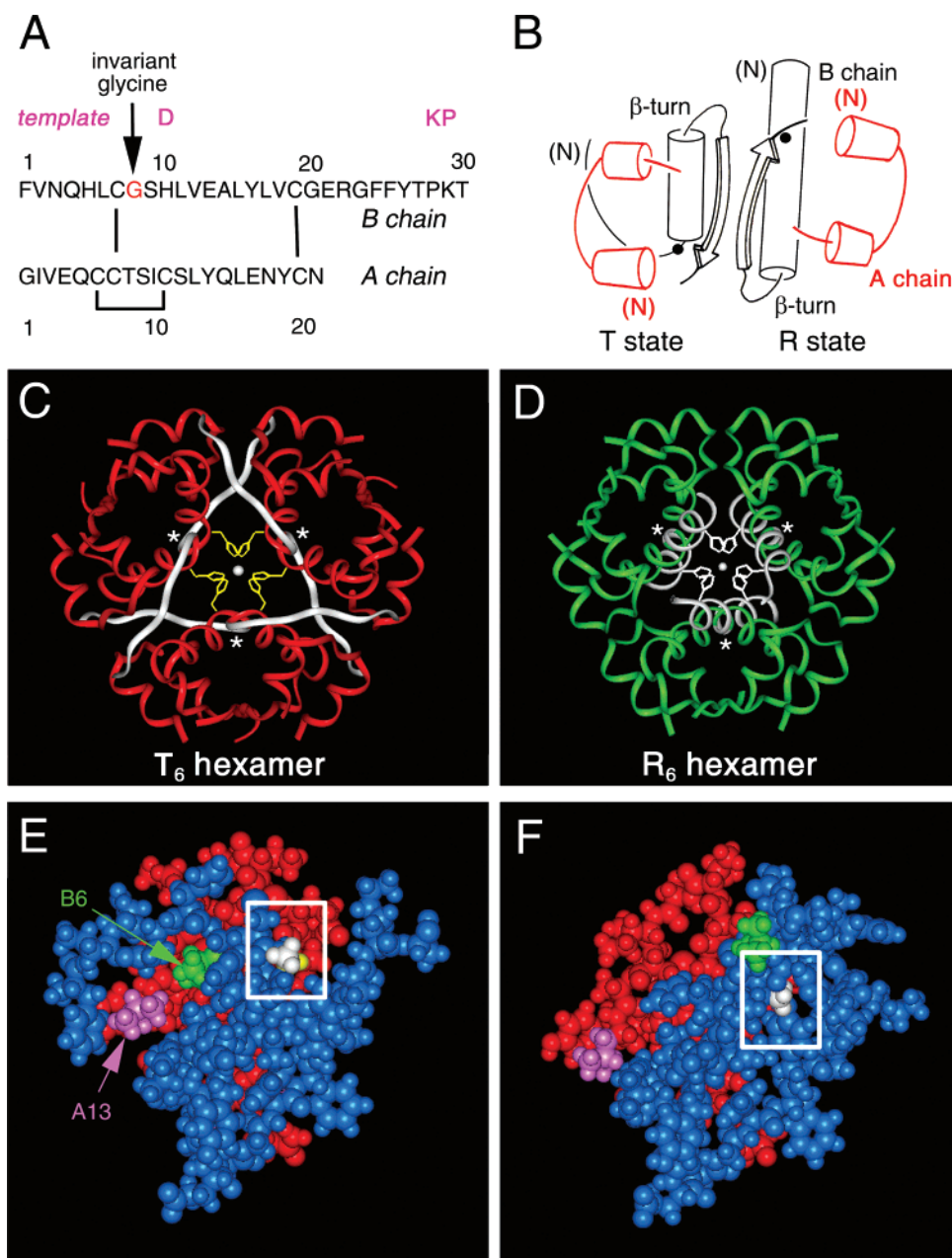


FIGURE 1: Overview of insulin structure. (A) Sequence of human insulin indicating invariant glycine in B chain (Gly^{B8}; arrow). Substitutions in monomeric DKP-insulin template are shown in magenta. Disulfide bridges are indicated by lines. (B) Cylinder model of insulin showing the T state (left) and R state (right) in the crystallographic dimer. Black balls indicate positions of Gly^{B8}. (C and D) Ribbon models of T₆ (C) and R₆ (D) zinc hexamers. Asterisks indicate positions of Gly^{B8}; the variable B1–B8 segment is highlighted in silver. The central zinc ions are coordinating side chains of His^{B10} and are also shown. (E and F) Molecular surfaces of T state (E) and R state (F) protomers indicating B8 site (boxes) and model of D-Ala^{B8} substitution (methyl group in white and H_α in yellow). Also shown are Leu^{A13} (magenta) and Leu^{B6} (green). The A and B chains are otherwise shown in red and blue, respectively. 4INS (1) and 1EV3 (39) were used in panels C and D.

refs 9 and 10) markedly impairs its binding to the insulin receptor (IR).³ Because the extent of impairment exceeds that ordinarily encountered in studies of mutant insulins (1, 2, 11), we suggest that the D substitution impairs an R-like conformational switch required for high-affinity receptor binding.

¹ Chiral mutagenesis of an invariant side chain in the A chain (Ile^{A2}) has previously been employed to investigate a hidden functional surface (30, 43). An *allo*-Ile^{A2} analogue (in which the chirality of the β carbon is inverted) exhibits native structure but low activity (30, 33). The present study extends this approach to the polypeptide main chain.

We focus on an invariant glycine in the B chain (Gly^{B8}; arrow in Figure 1A). This glycine follows a motif-specific cysteine (Cys^{B7}) and is broadly conserved among vertebrate insulin-like polypeptides (Figure 2). The environment of Gly^{B8} differs between the two classes of crystallographic protomers, T and R (Figure 1B). Whereas residues B7–B10 form a type II' β turn in the T state (Figure 1C; ref 1), the same residues lie within an extended α helix in the alternative R state (Figure 1D; refs 6 and 7). Gly^{B8} lies at the junction of this chameleon segment (12) and the central α helix (B9–B19). Although the B8 junction is exposed on the surface

		*																														
Insulins																																
Human, Pig		F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	T	(A)
Cattle		F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	A	
Sheep		F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	A	
Horse		F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	A	
Sei Whale		F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	A	
Rat		F	V	K	Q	H	L	C	G	P	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	S	
Elephant		F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	T	
Guinea Pig		F	V	S	R	H	L	C	G	S	N	L	V	E	A	T	Y	S	V	C	E	D	D	G	F	F	Y	I	P	K	D	
Coypu Hystricomorphs		Y	V	S	Q	R	L	C	G	S	Q	L	V	D	T	L	Y	S	V	C	R	H	R	G	F	F	Y	R	P	N	D	
Chinchilla		F	V	N	K	H	L	C	G	S	H	L	V	D	A	L	Y	L	V	C	G	D	R	G	F	F	Y	N	P	K	-	
Cod		A	P	P	Q	H	L	C	G	S	H	L	V	D	A	L	Y	L	V	C	G	D	R	G	F	F	Y	N	P	K	-	
Angler	V	A	P	A	Q	H	L	C	G	S	H	L	V	D	A	L	Y	L	V	C	G	D	R	G	F	F	Y	N	P	K	-	
Tuna	V	A	P	P	Q	H	L	C	G	S	H	L	V	D	A	L	Y	L	V	C	G	D	R	G	F	F	Y	N	P	K	-	
Toadfish 2		A	P	P	Q	H	L	C	G	S	H	L	V	D	A	L	Y	L	V	C	G	D	R	G	F	F	Y	N	S	-	-	
Toadfish 1		A	P	P	Q	H	L	C	G	S	H	L	V	D	A	L	Y	L	V	C	G	E	R	G	F	F	Y	N	P	K	-	
Bonito		A	A	N	P	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	Q	P	K	T	
Hagfish		R	T	T	G	H	L	C	G	K	D	L	V	D	A	L	Y	I	A	C	G	V	R	G	F	F	Y	D	P	T	K	M
Silver Carp		G	A	P	Q	H	L	C	G	S	H	L	V	D	A	L	Y	L	V	C	G	D	R	G	F	F	Y	T	P	K	T	
Turkey		A	A	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	S	P	T	T	
Duck, Goose		A	A	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	S	P	K	T	
Rattlesnake		A	P	N	Q	R	L	C	G	S	H	L	V	E	A	L	F	L	I	C	G	E	R	G	F	F	Y	S	P	R	T	
IGF-1																																
Human, Pig		A	G	P	E	T	L	C	G	A	E	L	V	D	A	L	Q	F	V	C	G	D	R	G	F	Y	F	N	K	P	T	
Chicken		A	G	P	E	T	L	C	G	A	E	L	V	D	A	L	Q	F	V	C	G	D	R	G	F	Y	F	S	K	P	T	
Frog		A	G	P	E	T	L	C	G	A	E	L	V	D	T	L	Q	F	V	C	G	D	R	G	F	Y	F	S	K	P	T	
Carp		A	G	P	E	T	L	C	G	A	E	L	V	D	T	L	Q	F	V	C	G	D	R	G	F	Y	F	S	K	P	T	
Catfish		A	G	P	Y	T	L	C	R	A	E	L	V	D	S	L	Q	P	V	C	G	Q	R	G	F	Y	F	S	R	P	M	
Shark		A	S	P	E	T	L	C	G	A	E	L	V	D	A	L	Q	F	V	C	G	D	R	G	F	Y	F	N	K	P	A	
IGF-2																																
Human		R	P	S	E	T	L	C	G	G	E	L	V	D	T	L	Q	F	V	C	G	D	R	G	F	Y	F	S	R	P	A	
Chicken		G	T	A	E	T	L	C	G	G	E	L	V	D	T	L	Q	F	V	C	G	D	R	G	F	Y	F	S	R	P	V	
Salmon		A	S	A	E	T	L	C	G	G	E	L	V	D	A	L	Q	F	V	C	E	D	R	G	F	Y	F	S	R	P	T	
Puffer fish		T	S	I	E	T	L	C	G	G	E	L	V	D	A	L	Q	F	V	C	E	D	R	G	F	Y	F	S	R	P	T	
Eel		G	S	G	E	T	L	C	G	G	E	L	V	D	A	L	Q	F	V	C	E	D	R	G	F	Y	F	S	R	P	T	
Hagfish IGF		-	L	S	E	T	L	C	G	S	E	L	V	D	T	L	Q	F	V	C	D	D	R	G	F	F	F	V	R	Q	H	
Lamprey IGF		-	A	S	E	T	L	C	G	S	E	L	V	D	A	L	O	F	V	C	D	D	R	G	F	Y	F	V	R	-	H	

FIGURE 2: Glycine is invariant at position B8 among insulin sequences and also conserved among insulin-like growth factors (IGF-I and IGF-II). Comparison of representative B chain or B domain sequences is shown; B8 residues are boxed (asterisk).

of the T and R state protomers, local packing schemes thus differ among crystal forms (Figure 3A,B). These contrasting environments are associated with different B8 ϕ dihedral angles (Figure 3C,D): positive in the T state (like a D amino acid; $56.4^\circ \pm 4.1$ among multiple crystal structures) and negative in the R state (like an L amino acid; $-63.0^\circ \pm 3.2$) (see Supporting Information). Such Ramachandran relationships motivate investigation of whether D and L amino acid substitutions would induce stereospecific perturbations of folding or function. In particular, because the structure of the insulin monomer in solution closely resembles the crystallographic T state (4, 10, 13), would D substitutions at B8 stabilize the β turn, and if so, would such a chiral "lock" enhance or impair biological activity? Conversely, what

would be the effects of L substitutions? Molecular modeling suggests that in a T-like protomer D or L side chains would project into solvent (Figure 1E); in a putative R-like protomer, L substituents would be exposed, whereas D side chains would be buried (Figure 1F).

To address these questions, we first employ combinatorial peptide chemistry to compare effects of L and D amino acid substitutions on insulin foldability. A novel in vitro selection is designed based on chain combination (14); mass spectrometry (MS) is used to distinguish between allowed and disallowed B chain sequences (15). Respective peptide libraries containing mixtures of D or L substitutions at B8 exhibit a stereospecific perturbation of insulin chain com-

² The crystallographic T \rightarrow R transition is characterized by a change in the secondary structure of the B1–B8 segment from extended (T state) to α -helix (R state). This reorganization is coupled to a change in the conformation of Gly^{B8} and handedness of cystine A7–B7. The sulfur atoms of the latter are exposed in the T state but buried in a nonpolar crevice in the R state. Whereas the structure of an isolated insulin in solution resembles the T state (10, 13), R state features have been observed only in zinc (and cobalt-substituted) hexamers. In particular, minor helix-related NOEs in the B1–B8 segment diagnostic of a subpopulation of R state conformers have not been detected in an insulin monomer in solution (10, 13, 33). The relevance of the T \rightarrow R transition to the function of insulin as a monomer is unknown (7).

³ Abbreviations: CD, circular dichroism; DG, distance geometry; DKP-insulin, insulin analogue containing three substitutions in B chain (Asp^{B10}, Lys^{B28}, and Pro^{B29}); DQF–COSY, double-quantum-filtered correlated spectroscopy; HPLC, high-performance liquid chromatography; IGF–I, insulin-like growth factor I; IR, insulin receptor; kDa, kilo-Dalton of mass; MALDI, matrix assisted laser desorption ionization; NMR, nuclear magnetic resonance; MS, mass spectrometry; NOEs, nuclear Overhauser enhancements; NOESY, NOE spectroscopy; RMD, restrained molecular dynamics; rp-HPLC, reverse-phase HPLC; SA, simulated annealing; RMSD, root-mean-square deviation; rpm, revolutions per minute; TOCSY, total correlation spectroscopy; TOF, time-of-flight; and UV, ultra-violet. Amino acids are designated by standard one- and three-letter codes.

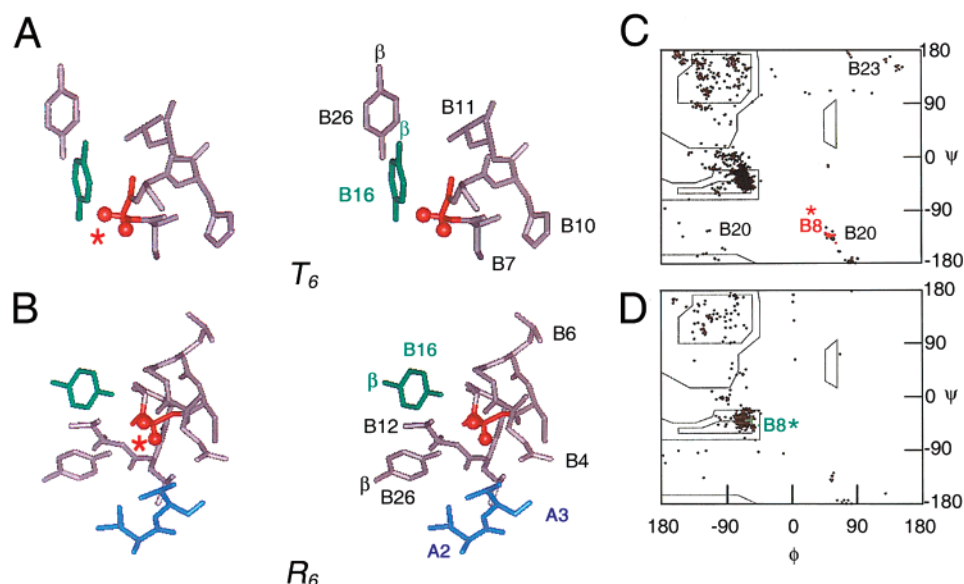


FIGURE 3: Structural relationships in crystal structures of insulin surrounding Gly^{B8}. (A and B) In T state protomers Gly^{B8} participates in a β turn (red asterisk in A) whereas in R state protomers, B8 is part of an α -helix (red asterisk in B). In the T state protomer, B8 is near Tyr^{B26} of same molecule (B8 C α –O ϵ B26 distance 5.3 Å) and dimer-related Tyr^{B16} (green); in the R state protomer, B8 is near these residues and in addition Val^{B12} (an (*i*, *i* + 4) contact in α -helix) and A chain of adjoining molecule in hexamer (Ile^{A2} and Val^{A3}; blue). For clarity, C β atoms of aromatic side chains are as indicated. Coordinates were obtained from T₃R₃^f zinc hexamer (Protein Databank 1TYL; ref 75). (C and D) Ramachandran maps derived from crystal structures (Protein Databank 4INS (1), 1TYL (75), and 1TRZ (36)). Whereas in T state protomers Gly^{B8} has ϕ dihedral angles (red; C), in the R state protomers, Gly^{B8} belongs to an extended α -helix and so exhibits negative ϕ values (green; D).

bination: L amino acids impede native disulfide pairing, whereas diverse D substitutions are well-tolerated. We then extend these findings to characterize representative diastereomeric pairs of B8 analogues. D Substitutions enhance the thermodynamic stability of insulin but markedly impair its biological activity, whereas unstable L analogues can be highly active.⁴ Remarkably, a single D methyl group attached to the surface of insulin is shown to stabilize the native structure of insulin in solution ($\Delta\Delta G_u > 1$ kcal/mol) but impair its receptor binding by 1000-fold. Because this decrement far exceeds effects of mutations on the protein surface at neighboring sites (1, 11), these observations suggest that the canonical T state of insulin represents an inactive conformation: the introduced D methyl group acts as a spanner in the works to block an R-like change in B8 conformation on receptor binding. We propose that this invariant glycine in the B chain functions as a Ramachandran switch between folding-competent and active conformations, mirroring aspects of the classical T \rightarrow R allosteric transition (6, 7). We thus envisage that the flexibility of Gly^{B8} both protects the β -cell from toxic protein misfolding and enables high-affinity hormone–receptor recognition.

MATERIALS AND METHODS

Materials. Human insulin was provided by Eli Lilly and Co. (Indianapolis, IN). Isolated A chain *tetra*-S-sulfonate derivatives were obtained by oxidative sulfitolysis of insulin followed by separation of S-sulfonated A and B chains as described (16). B chain analogues were prepared (16) using

solid-phase synthesis (17, 18). Chain combination (19) was effected by interaction of the S-sulfonated derivative of the A chain (1.25 mM) and B chain (0.5 mM) in 0.1 M glycine buffer (pH 10.6, 4 mL) in the presence of dithiothreitol (4.5 mM) (15, 18). Each analogue was purified by a combination of size-exclusion chromatography (Bio-Gel P-4 in 3 M acetic acid) and preparative rp-HPLC (C18). DKP analogues each contain three DKP substitutions to prevent self-association of insulin (His^{B10} \rightarrow Asp, Pro^{B28} \rightarrow Lys, and Lys^{B29} \rightarrow Pro; refs 13 and 20). Nonstandard analogue D-*p*-amino-Phe^{B8}-B1-biotinyl-amidocaproyl-DKP-insulin (21) was kindly provided by P. G. Katsoyannis. Two-disulfide analogues of insulin containing pairwise substitution of cystine A7–B7 by serine were likewise synthesized as described (22). Fidelity of synthesis was in each case verified by MS.

Peptide Libraries. Mixtures of Fmoc-protected D or L amino acids were prepared by mixing equimolar solutions of each residue type except cysteine, glycine, and proline. Respective B chain libraries were synthesized using Fmoc chemistry by dividing, coupling, and recombining peptide resins to ensure equimolarity of the 17 B chain analogues within each peptide pool. The predicted diversity of the libraries was verified by MALDI-TOF MS. Respective libraries were combined with the native A chain (in S-sulfonate form); chain combination was initiated on addition of dithiothreitol and monitored by MALDI-TOF MS (Perkin-Elmer/Applied Biosystems). Similar efficiencies of MS detection were verified in control studies of representative D and L analogues obtained by conventional synthesis (D-Ala^{B8}-insulin and L-Ala^{B8}-insulin). Proper disulfide pairing in a representative D analogue (D-Ala^{B8}-DKP-insulin) was verified by NMR spectroscopy and in a representative L analogue (L-Ser^{B8}-insulin) by X-ray crystallography (J. Whittingham and G. G. Dodson, personal communication).

⁴ A preliminary description of D and L B8 libraries was presented at the 1997 American Peptide Symposium (77). L-Alanine has also been introduced at B8 by biosynthetic expression of a single-chain precursor in yeast (11, 86). L-Ala^{B8} was associated with lower yield and impaired receptor binding by ~ 30 -fold in accord with the present results.

Receptor-Binding Assays. Relative activity is defined as the ratio of analogue to wild-type human insulin required to displace 50 % of specifically bound ^{125}I -human insulin. A human placental membrane preparation containing the IR was employed as described (23). In all assays, the percentage of tracer bound in the absence of competing ligand was less than 15% to avoid ligand-depletion artifacts.

Ultracentrifugation. Sedimentation equilibrium and sedimentation velocity measurements were conducted at 22 °C using a Beckman analytical ultracentrifuge (model XLI) as described (24). The proteins were made 0.4 mM in 10 mM potassium phosphate (pH 7.4) and 50 mM KCl. Sedimentation equilibrium studies were conducted at 16 000 and 30 000 rpm. Molecular-weight values were calculated using a partial specific volume of 0.73 mL/g and density 1 g/cm³. Sedimentation velocity studies were conducted at 60 000 rpm, and results were fitted with the program DCDT Plus. The dimerization-defective analogue KP-insulin (Pro^{B28} → Lys and Lys^{B29} → Pro; Humalog (24)) was employed as a control.

Visible Absorption Spectroscopy. To probe the TR transition of Co²⁺-substituted insulin hexamers, the d-d optical absorption bands of Co²⁺ (a characteristic feature of a tetrahedral complex) were monitored as described (25–27). Solutions contained 0.2 mM insulin or insulin analogue in a buffer consisting of 0.07 mM CoCl₂ and 50 mM phenol in 50 mM Tris-HCl (pH 8). Spectra were obtained in the presence and absence of 0.8 M NaSCN.

CD Spectroscopy. Circular dichroism (CD) spectra were obtained using a thermister-controlled Aviv spectropolarimeter equipped with an automated titration unit for guanidine denaturation studies. CD samples for wavelength spectra contained 25–50 μM insulin analogue in 10 mM potassium phosphate (pH 7.4) and 50 mM KCl. For equilibrium denaturation studies, samples were diluted to 5 μM; guanidine-HCl was employed as denaturant as described (28). Data were obtained at 4 °C. Guanidine denaturation data were fitted by nonlinear least squares to a two-state model (29).

NMR Spectroscopy. High-resolution studies of insulin as a monomer in aqueous solution have been made possible by substitutions that destabilize dimer- and hexamer-forming surfaces but do not hinder receptor binding (10, 13, 20, 30). DKP-insulin is employed here as a template for study of B8 analogues (substitutions are shown in magenta in Figure 1A). Spectra were obtained at 600 MHz as described (13, 28, 30) at pH 7.0 and 25 °C; pH 8.0 and 32 °C; and in 20% v/v deuterioacetic acid, pH 1.9 and 25 °C. Resonance assignment was based on homonuclear 2-D NOESY (mixing times 80 and 200 ms), total correlation spectroscopy (TOCSY; mixing time 55 ms), and double-quantum filtered correlated spectroscopy (DQF-COSY) spectra. Helix-related hydrogen bonds were inferred from the pattern of protected amide resonances as observed in D₂O solution containing 20% deuterioacetic acid (31).

Crystallization Trials. Crystals of variant zinc-insulin hexamers containing D-Ala^{B8} substitutions were sought by hanging-drop vapor diffusion using well-established T₆ and ostensible R₆ conditions (7, 32) as follows.

(i) **T₆ Conditions.** Under a series of conditions permitting growth of well-ordered single crystals of native insulin over two weeks at room temperature, the D-Ala^{B8} variant exhibited unusually rapid growth of multiple small crystals leading in

2 days to clusters unsuitable for X-ray diffraction studies. For these trials, the proteins were made 10 mg/mL in 0.02 M HCl and mixed with an equal volume of reservoir solution consisting of 100 mM sodium citrate (pH 7.4–8.9), 10–15% acetone, and 0.08–0.12% zinc acetate.

(ii) **Ostensible R₆-Like Conditions.** Crystals of D-Ala^{B8}-insulin suitable for X-ray diffraction analysis were obtained under conditions resembling those that ordinarily lead to growth of wild-type crystals containing R₆ hexamers. The crystals were grown in the presence of a 1:2.5 ratio of Zn²⁺ to protein monomer and a 3.7:1 ratio of phenol to protein monomer in Tris-HCl buffer as described (33). Drops consisted of 1 μL of protein solution (10 mg/mL in 0.02 M HCl) mixed with 1 μL of reservoir solution (0.02 M Tris-HCl, 0.05 M sodium citrate, 5% acetone, 0.03% phenol, and 0.01% zinc acetate at pH 8.8). Each drop was suspended over 1 mL of reservoir solution. Crystals were obtained at room temperature after 2 weeks. Preliminary diffraction data from 17.4 to 2.5 Å resolution were collected at room temperature using an Raxis II area detector with CuKα radiation, located in the Department of Biochemistry, University of Chicago. Data were processed with programs DENZO (version 1.9.6) and SCALEPACK (version 1.9.6). Crystals belong to space group R3 with equivalent hexagonal unit-cell dimensions $a = b = 80.84$ Å and $c = 38.84$ Å, corresponding to one dimer per asymmetric unit. These dimensions are characteristic of a T₃R₃^f hexamer (see Results). Control crystallization trials of wild-type human insulin under the same conditions yielded crystals belonging to the same space group but containing R₆ hexamers as expected. Crystals of L-Ala^{B8}-insulin were not sought due to low synthetic yields.

Molecular Modeling. Distance-geometry/simulated annealing calculations (DG/SA) were performed using the program DG-II as described (5); restrained molecular dynamics (RMD) calculations were performed using X-PLOR (34).

RESULTS

Our results are presented in three parts. The use of D- and L B chain libraries is described in part I and extended to characterize representative (D, L) B8 analogues. Their structures, stabilities, and receptor-binding activities are investigated in part II. Pairwise substitution of a key cystine by serine is employed in part III to explore the coupling between B8 chirality and disulfide pairing. In addition to the B8 peptide libraries, the present study employs ten B8 analogues obtained by individual syntheses (Table 1).

I. Combinatorial Chain Combination. Chain combination provides a functional selection for foldable B chain variants as outlined in Figure 4A. Allowed substitutions enable pairing with wild-type A chains to yield a two-chain product with characteristic molecular mass (ca. 5.8–6.0 kDa); disallowed substitutions cannot. Two libraries of B chains were synthesized, one containing diverse D amino acids at B8 and the other, L amino acids. Each contains 17 variants. Cysteine was excluded to avoid confounding disulfide bridges; proline was excluded due to its anomalous conformation properties. Neither library contained glycine, the wild-type residue. The peptides were protected by S-sulfonate modification; their diversity was verified by MS. Following

Table 1: Receptor–Binding Affinities of B8 Analogues^a

analogue	rel. affinity (%)	analogue	rel. affinity (%)
human insulin (HI)	100 (7)	DKP-insulin (DKP-Ins)	160 ± 15 (2)
L-Ala ^{B8} -HI ^b	1.0 ± 0.4 (3)	L-Ser ^{B8} -DKP-Ins ^c	90 ± 6 (3)
D-Ala ^{B8} -HI	0.11 ± 0.02 (4)	D-Ser ^{B8} -DKP-Ins	1.1 ± 0.1 (3)
L-Ala ^{B8} -DKP-Ins	3.6 ± 0.6 (4)	D-Pmp ^{B8} -DKP-Ins ^d	0.52 ± 0.11 (3)
D-Ala ^{B8} -DKP-Ins	0.17 ± 0.04 (5)	D-Arg ^{B8} -DKP-Ins	0.054 ± 0.002 (3)
[Ser ^{A7} , Ser ^{B7}]-DKP-Ins ^e	<0.01	D-Ala ^{B8} -[Ser ^{A7} , Ser ^{B7}]-DKP-Ins	<0.01
L-Ala ^{B8} -[Ser ^{A7} , Ser ^{B7}]-DKP-Ins	<0.01		

^a The dissociation constant for native human insulin is 0.4–0.5 nM. Numbers in parentheses indicate number of replicates. ^b Kristensen and colleagues reported that Ala^{B8}-insulin binds to the isolated receptor ectodomain with relative affinity 3 ± 2%. ^c Similar but lower values have recently been obtained by Feng and colleagues in studies of a L-Ala^{B8}-HI analogue (35, 86). ^d Pmp designates para-amino-Phe. The analogue also contains an N-terminal biotinyl-amidocaproyl modification at B1, which is associated itself with a small decrement in activity (21). Correction for this effect predicts a relative affinity of 0.63 ± 0.13% for the B8 modification. ^e Value obtained from ref 22.

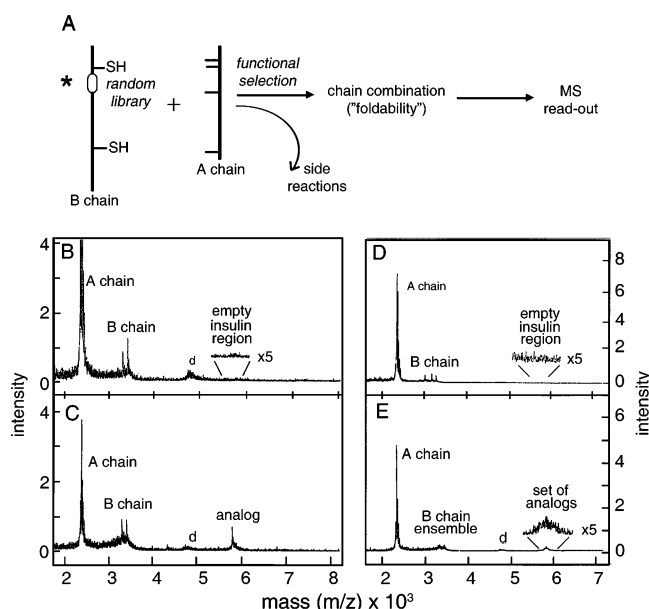


FIGURE 4: Foldability as probed by combinatorial peptide synthesis with MS read out. (A) Functional-selection scheme based on insulin chain combination. B chain analogue or library of analogues is mixed with wild-type A chain to initiate specific disulfide pairing. Products are detected by mass. (B–E) MALDI-TOF MS spectra of the chain-combination reactions following 24 h at 4 °C. (B and C) Control reactions of L-Ala^{B8} (B) and D-Ala^{B8} (C) B chains. (D and E) Combinatorial reactions of L and D libraries at B8, respectively. Whereas L analogues undergo inefficient chain combination, productive pairing is observed to yield multiple diverse D substitutions at B8. MS spectra also detect oxidized A and B chains and A chain dimers (d).

addition of a 2:1 molar excess of the wild-type *S*-sulfonate-protected A chain, the chain combination was initiated by the addition of dithiothreitol and allowed to proceed under air oxidation for 24 h at 4 °C. The resulting D and L reaction mixtures were analyzed by MALDI-TOF MS (Figure 4D,E) in relation to a control reaction employing the wild-type B chain. Whereas the B8 L amino acid library yielded only a trace amount of insulin analogues (empty insulin region in Figure 4D), the D library by contrast resulted in native or enhanced foldability (Figure 4E). The D reaction mixture was subjected to gel filtration and partial rp-HPLC purification. All possible D-amino acid substitutions were detected by MS (Supporting Information). Among these substitutions, the polar side chain at B8 (such as D-Lys and D-Arg) was obtained at higher yield than nonpolar or aromatic side chains (such as D-Ala and D-Trp). All fractions exhibit very low

affinity for the IR (<0.5% relative to insulin). Stereospecific interference with chain combination was verified in single syntheses of D-Ala^{B8} and L-Ala^{B8} insulin analogues (Figure 4B,C). The D-Ala^{B8} substitution is associated with a 3-fold increase in the efficiency of chain combination. By contrast, following prolonged reaction times (48–72 h), the L-Ala^{B8} analogue was obtained and purified with a yield >30-fold lower than that of the D analogue. Native disulfide pairing was verified in part by peptide mapping following *Staphylococcus aureus* V8 protease digestion.⁵ Native disulfide pairing in a representative D analogue (D-Ala^{B8}-DKP-insulin) was explicitly demonstrated by 2-D NMR spectroscopy. MALDI-TOF MS spectra of the purified D-Ala^{B8} and L-Ala^{B8} analogues indicated similar efficiencies of detection, suggesting that the original MS characterization of the D and L libraries faithfully reflects relative yields.

II. Biochemical Characterization of Representative Analogues. Both D and L Ala^{B8}-insulin exhibit low receptor-binding activities (0.11 ± 0.02 and 1.0 ± 0.4%, respectively; Figure 5A and Table 1). The low activity of L-Ala^{B8}-insulin is consistent with prior studies (11, 35). Both D and L-Ala^{B8}-insulin are able to undergo the T → R transition in phenol-stabilized Co²⁺-insulin hexamers as probed by observation of phenol-dependent d–d transitions characteristic of the R state-specific tetrahedral metal-binding site (26, 27). The concentration of phenol was made 50 mM to favor formation of the R state (Materials and Methods). Although the spectral line shape of the d–d transitions in the D-Ala^{B8} hexamer is similar to that of wild-type Co²⁺-substituted R₆ hexamers, their intensity is nevertheless reduced by about 2-fold (Supporting Information), suggesting predominant formation of T₃R₃^f hexamers as the octahedral metal-binding site of a T₃ trimer would not appreciably contribute to the d–d band. In accord with this observation, preliminary crystallographic studies of phenol-stabilized Zn²⁺-coordinated D-Ala^{B8}-insulin hexamers grown under such conditions, although expected to form R₆ hexamers, instead exhibit unit cell dimensions similar to those of T₃R₃^f hexamers in the same crystal form (Supporting Information). In a given space group, these dimensions provide an empirical fingerprint of the hexamer type. The *a* and *b* dimensions (80.84 Å), for example, are similar to those of T₃R₃^f PDB 1TRZ (80.64 Å; ref 36) and 1QJ0 (80.94 Å; ref 37), whereas rhombohedral crystals

⁵ This protocol verifies the presence of cystine A20-B19 but does not distinguish between potential bridges A6–A11 or A7–A11. The MS read-out (Figure 4A) does not distinguish between native pairing and possible disulfide isomers.

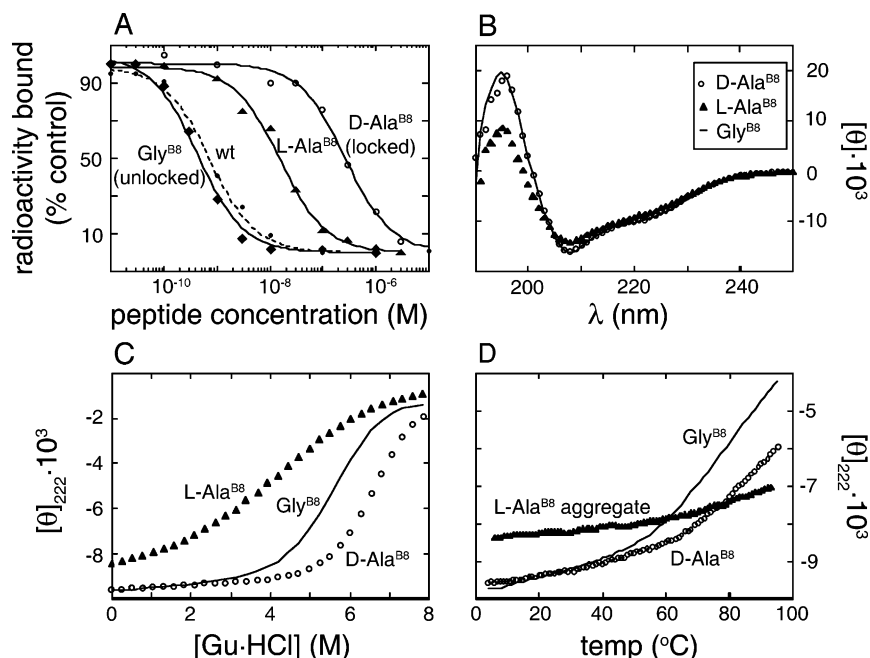


FIGURE 5: Activity and stability of B8 analogues. (A) Receptor-binding titrations of DKP-insulin analogues relative to native human insulin (dashed line): Gly^{B8} (diamonds), L-Ala^{B8} (triangles), and D-Ala^{B8} (open circles). Relative affinities of these and other analogues are given in Table 1. (B) Far-UV CD spectra of DKP insulin analogues: Gly^{B8} (solid line), L-Ala^{B8} (triangles), and D-Ala^{B8} (open circles). (C and D) Guanidine (C) and thermal (D) denaturation studies of DKP-insulin analogues (symbols as in panel B). L-Ala^{B8} aggregates exhibit an anomalous thermal signature. Extent of folding was monitored by ellipticity at 222 nm. Receptor-binding and CD studies were conducted at 4 °C.

containing R₆ hexamers exhibit distinct unit-cell dimensions (PDB 1ZEG, 1ZEH, and 1EV3; refs 38 and 39). As expected, preliminary analysis of the present crystals by molecular replacement is consistent with formation of a nativelike T₃R₃^f containing three bound phenol molecules (Z.-I. Wan and M. A. Weiss, unpublished results). Together, these spectroscopic and crystallographic observations suggest that the D substitution partially impairs the TR transition, arresting the hexamer at an intermediate stage of structural organization. Our results do not exclude formation of D-Ala^{B8} R₆ hexamers at higher concentrations of phenol.

Wild-type insulin in solution undergoes self-association to yield dimers, tetramers, hexamers, and higher-order oligomers, which may be differently affected by D- or L substitutions at B8. To avoid such potentially confounding effects and to facilitate spectroscopic studies, D-Ala^{B8} and L-Ala^{B8} analogues were resynthesized in the context of DKP-insulin, an engineered monomer of high activity (13, 20, 30). The previous stereospecific difference in efficiency of chain combination was unaffected by the DKP substitutions, leading to enhanced yield of D-Ala^{B8}-DKP-insulin and more than tenfold lower yield of L-Ala^{B8}-DKP-insulin. CD spectroscopy demonstrates that the D analogue retains native secondary structure (circles in Figure 5B), whereas the L analogue is perturbed: helical features at 222, 208, and 195 nm are attenuated (triangles in Figure 5B). Since Gly^{B8} does not reside in an α -helix in the native (T-like) structure in solution, the attenuated helix content of L-Ala^{B8}-insulin implies nonlocal changes in conformation or dynamics. The D and L substitutions effect opposite changes in protein stability: L-Ala^{B8}-insulin is more sensitive to denaturant-induced unfolding, whereas D-Ala^{B8}-insulin is more robust (Figure 5C). Respective free energies of unfolding are given in Table 2. The thermodynamic changes are marked:

Table 2: Thermodynamic Stabilities of B8 Analogues

analogue	ΔG_u (kcal/mol) ^a	C_{mid} (M) ^b	m (kcal/mol/M) ^c
Insulin	4.4 ± 0.1	5.3 ± 0.1	0.84 ± 0.01
DKP-insulin	4.9 ± 0.1	5.8 ± 0.1	0.84 ± 0.01
L-Ala ^{B8} -DKP-Ins ^d	2.3 ± 0.2	3.5 ± 0.3	0.95 ± 0.02
D-Ala ^{B8} -DKP-Ins	6.4 ± 0.1	6.8 ± 0.1	0.66 ± 0.05
L-Ser ^{B8} -DKP-Ins	1.9 ± 0.1	3.7 ± 0.2	0.89 ± 0.02
D-Ser ^{B8} -DKP-Ins	5.8 ± 0.1	6.5 ± 0.1	0.51 ± 0.01

^a ΔG_u indicates apparent change in free energy on denaturation in guanidine-HCl as extrapolated to zero denaturant concentration by a two-state model (29). ^b C_{mid} is defined as that concentration of guanidine-HCl at which 50% of the protein is unfolded. ^c The m value provides the slope in plotting unfolding free energy ΔG_u versus molar concentration of denaturant; this slope is proportional to the protein surface area exposed on unfolding. ^d Abbreviation for DKP-insulin template.

respective $\Delta\Delta G_u$ values are 1.5 ± 0.2 kcal/mol (D-Ala^{B8}; “locked” in the native state) and -2.1 ± 0.3 kcal/mol (L-Ala^{B8}). These $\Delta\Delta G_u$ values are larger than those ordinarily encountered in studies of single amino acid substitutions in insulin (5, 23, 40, 41). Inversion of C α chirality at B8 (i.e., comparison of D-Ala^{B8} and L-Ala^{B8} analogues) is thus associated with an extraordinary 4.1 ± 0.3 kcal/mol difference in stability. D-Ala^{B8}-DKP-insulin also exhibits greater thermal stability than DKP-insulin (Figure 5D). Comparison of D-Ser^{B8}- and L-Ser^{B8}-DKP-insulin analogues yields similar trends in efficiency of chain combination, helix content, and stability (Table 2 and Supporting Information). D- and L-Ser^{B8} analogues differ in stability by 3.9 ± 0.2 kcal/mol.

D-Ala^{B8}- and D-Ser^{B8}-DKP-insulin each exhibit low activities (0.17 ± 0.04 and $1.1 \pm 0.1\%$, respectively; Table 1). Because of the favorable His^{B10} \rightarrow Asp substitution in the DKP template (20), the affinity of D-Ala^{B8}-DKP-insulin is almost 2-fold higher than that of D-Ala^{B8}-insulin ($0.11 \pm$

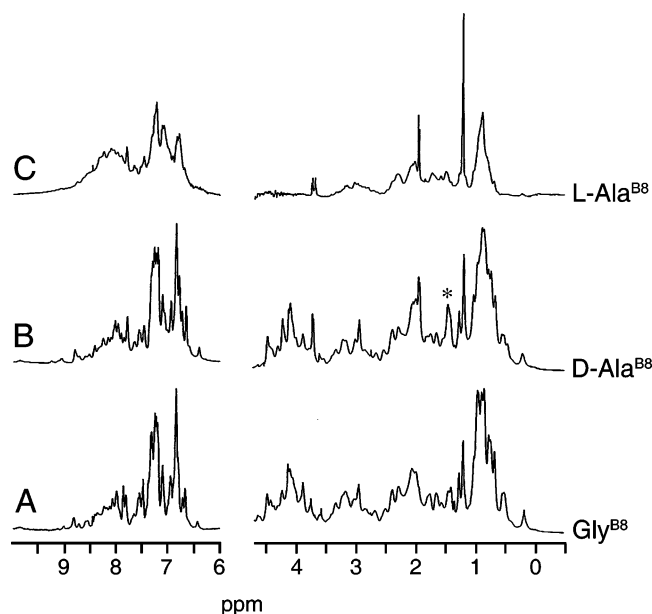


FIGURE 6: ^1H NMR studies of DKP-insulin analogues at 600 MHz at 25 $^{\circ}\text{C}$ (pH 7.4). (A) Gly^{B8}, (B) D-Ala^{B8}, and (C) L-Ala^{B8}. Asterisk in panel B indicates position of D-Ala^{B8} methyl resonance (see Figure 7). The spectrum of the Gly^{B8} and D-Ala^{B8} analogues are similar whereas the L-Ala analogue undergoes aberrant aggregation.

0.02%). Surprisingly, L-Ser^{B8}-DKP-insulin exhibits substantial receptor-binding activity (90% relative to human insulin), suggesting that the low activity of L-Ala^{B8} analogues is not a general property of L amino acid substitutions. The ^1H NMR spectrum of L-Ala^{B8}-DKP-insulin (Figure 6C) exhibits resonance broadening relative to spectra of DKP-insulin or D-Ala^{B8}-DKP-insulin (Figure 6A,B). Studies of protein aggregation by equilibrium ultracentrifugation and sedimentation velocity indicate that (like DKP-insulin) the D-Ala^{B8} analogue is monomeric, whereas L-Ala^{B8}-DKP-insulin forms predominantly dimers and higher-order oligomers at a protein concentration of 75 μM . Because self-association occurs

despite the presence of the DKP substitutions (20), it is presumably mediated by non-native protein surfaces (see Discussion). By contrast, the quality of the ^1H NMR spectrum of L-Ser^{B8}-DKP-insulin is similar to that of DKP-insulin and its D analogues (Supporting Information), indicating an initial absence of aberrant self-association. (Unlike DKP-insulin and D-Ala^{B8}-DKP-insulin, whose NMR spectra remain unchanged after two weeks in solution, L-Ser^{B8}-DKP-insulin undergoes slow but progressive aggregation over several days at 25 $^{\circ}\text{C}$, leading likewise to resonance broadening.) These findings suggest that the low activity of L-Ala^{B8} analogues is due to confounding aggregation and not directly related to B8 stereochemistry. Aggregation of the L-Ala^{B8} analogue is also associated with an anomalous thermal CD signature (Figure 5D).

The ^1H NMR spectrum of D-Ala^{B8}-DKP-insulin is tractable by homonuclear NMR methods (42), permitting complete resonance assignment (Supporting Information). Chemical shifts are essentially identical to those observed in DKP-insulin; significant changes (magnitude >0.1 ppm) are observed only at neighboring residues Cys^{B7} and Ser^{B9} (Supporting Information). The novel D-Ala^{B8} spin system is well-resolved (Figure 7). The pattern of interresidue NOEs is likewise similar to that of DKP-insulin, in each case consistent with structures of T-state crystallographic protomers.⁶ Maintenance of T state-specific long-range interactions by the N-terminal arm of the B chain is demonstrated by the retention of interchain NOEs between the B1–A13 and B5–A10 side chains. The orientation of the B9–B19 α -helix relative to the A chain, as probed by a network of long-range NOEs in the hydrophobic core (Supporting Information), is also unaffected by the D-Ala^{B8} substitution. NOEs involving the β protons of the six cysteines are essentially identical to those observed in the spectrum of DKP-insulin, indicating maintenance of native disulfide pairing. As expected based on native models (Figures 1E and 3A), NOEs are observed from the β -CH₃ group of

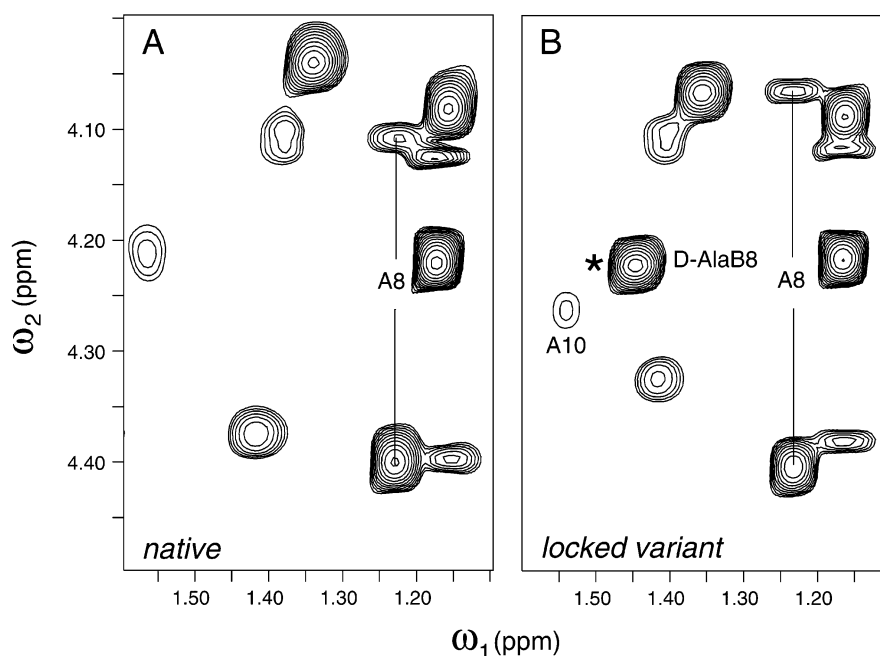


FIGURE 7: ^1H NMR TOCSY spectra of (A) DKP-insulin and (B) D-Ala^{B8}-DKP-insulin highlighting novel D-Ala methyl resonance in analogue (asterisk in panel B). Spectra were acquired at 600 MHz at 32 $^{\circ}\text{C}$ and pD 7.6 (direct meter reading) with a TOCSY mixing time of 55 ms.

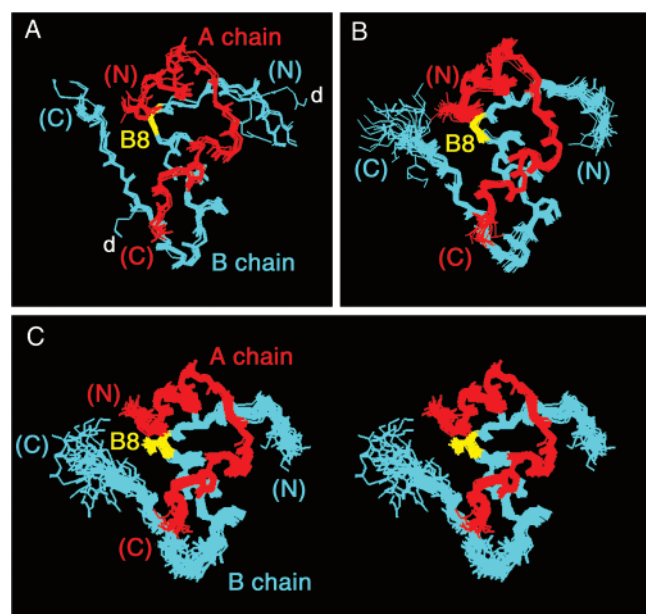


FIGURE 8: D-Ala^{B8}-DKP-insulin exhibits nativelike solution structure. (A) Collection of crystal structures showing superimposition of T state protomers (Protein Databank 4INS (1), 1TYL (75), 1PID (76), 1TRZ (36), and 1TYM (75)). The position of Gly^{B8} is highlighted in yellow. The B chain is otherwise shown in blue, and the A chain is red. Structure of truncated analogue des-pentapeptide[B26–B30]-insulin (DPI) is indicated by *d*'s at positions B1 and B25. (B) Solution structure of DKP-insulin (PDB 1LNP; ref 28). (C) Ensemble of DG/RMD structures of D-Ala^{B8}-DKP-insulin showing position of D-Ala side chain (yellow). The coloring scheme in panels B and C corresponds to that in panel A.

D-Ala^{B8} to the amide protons of B11 and B12, one methyl resonance of Val^{B12}, and the H_ε aromatic protons of Tyr^{B26}. The methyl group of D-Ala^{B8} is also likely to be near the methyl groups of Val^{A3} and Leu^{B11}, but such cross-peaks are difficult to resolve in homonuclear 2-D NOESY spectra at 600 MHz.

DG/RMD calculations verify that the solution structure of D-Ala^{B8}-DKP-insulin (Figure 8C) is similar to the native T state (Figure 8A) and essentially identical to that of DKP-insulin (Figure 8B) with the exception of the D-methyl group at B8 (yellow). The structure was calculated on the basis of 389 NOEs, 42 dihedral restraints, and 36 hydrogen-bond-related restraints (Supporting Information). The B8 Ramachandran angles in the DG/RMD ensemble are $\phi = 22.1 \pm 5.2$ and $\psi = -103.5 \pm 4.6$; these values are similar but not identical to those characteristic of Gly^{B8} in crystallographic T state protomers. The positive ϕ value for D-Ala^{B8} stands in contrast to the classical α -helical conformation of L-Ala^{B14} in the NMR ensemble ($\phi -67.7 \pm 3.6$, $\psi -39.0 \pm 3.0$). D-Ala^{B8} is shielded in part by the side chains of Val^{A3} (nearby despite the absence of explicit A3–B8 constraints), Val^{B12}, and Tyr^{B26}. The solvent accessibility of the B8 methyl group ($28 \pm 13\%$) is lower than what would be predicted based on the wild-type crystallographic T state (ca. 70%)

⁶ The NOESY spectrum of D-Ala^{B8}-DKP-insulin contains some additional interresidue contacts that are very weak or not observed in DKP-insulin but consistent with crystal structures of wild-type insulin (Supporting Information). The increased intensity of these NOEs may reflect damping of fluctuations by the D side chain. Stabilization of the protein is not accompanied by accentuated dispersion of ¹H NMR chemical shifts.

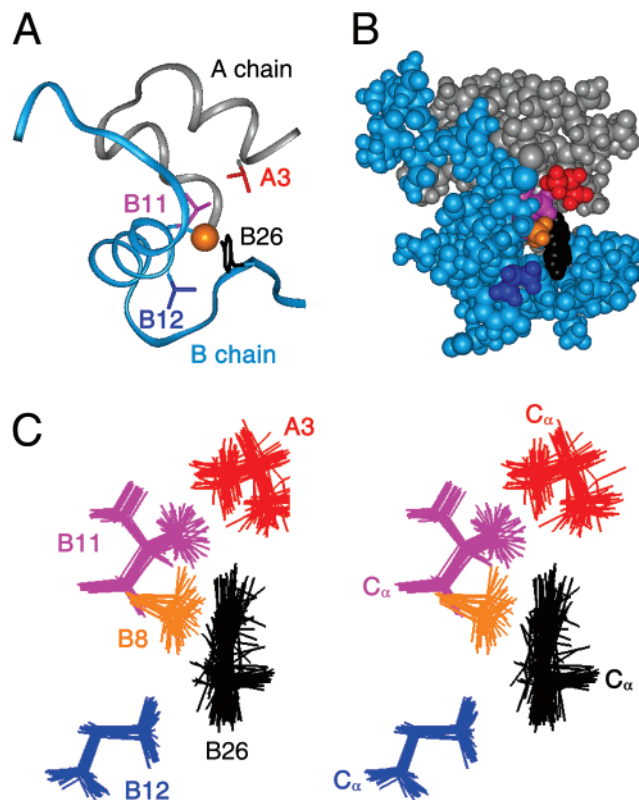


FIGURE 9: Environment of D-Ala^{B8} side chain at protein surface. (A and B) Ribbon and space-filling models of representative structure of D-Ala^{B8}-DKP-insulin showing position of D-Ala^{B8} methyl group (tawny; sphere in panel A at 0.7 van der Waals radius) relative to side chains of Val^{A3} (red), Val^{B12} (dark blue), and Tyr^{B26} (black). Leu^{B11} (magenta) lies behind B8 in core. The A chain is otherwise shown in gray and the B chain otherwise in blue. (C) Stereopair showing spatial relationships among residues A3, B8, B11, B12, and B26. The ensemble was aligned on these side chains to illustrate B8 pocket. Color scheme is the same as in panels A and B.

due to a small adjustment in the position of Tyr^{B26}. Because the C-terminal segment of the B chain appears to reorganize on receptor binding (see Discussion), this adjustment is unlikely to be of functional significance. The conserved side chains of Val^{A3}, Leu^{B11}, Val^{B12}, and Tyr^{B26} create a local nonpolar cup that partially surrounds D-Ala^{B8} (Figure 9). B11 is a critical component of the hydrophobic core (15), whereas the side chains of A3, B12, and B26 contribute to receptor binding (1, 18, 43); their spatial relationships in this or wild-type structures may be reorganized on receptor binding (2). Restraint information and statistical parameters are given in the Supporting Information. The nativelike structure of D-Ala^{B8}-DKP-insulin is in accord with its efficient folding (as probed by the increased yield of chain combination) and enhanced stability but stands in contrast to its 1000-fold decrement in receptor-binding activity relative to DKP-insulin (Table 1).

III. Importance of Cystine A7–B7. It is possible that the very low yields encountered in syntheses of L-Ala^{B8}-insulin and other L variants are due to either (a) the low thermodynamic stabilities of the final products (Table 2) or (b) a kinetic block to disulfide pairing. In an effort to distinguish between these mechanisms, we investigated the role of neighboring cystine A7–B7. Pairwise substitution of A7 and B7 by serine has previously been shown to lead to a molten

globule of marginal stability ($\Delta\Delta G_u < 2$ kcal/mol) but not impairment of specific pairing of the remaining four cysteines (22). Because the stability of the two-disulfide analogue is lower than that of L-Ala^{B8}-DKP-insulin and yet efficient disulfide pairing is maintained, relative yields are unlikely to be under thermodynamic control. Substitution of Gly^{B8} by D or L-Ala in the context of [Ser^{A7}, Ser^{B7}]-DKP-insulin likewise enables analogous disulfide pairing without significant differences in yield. Specific pairing of cysteines A20–B19 and A6–A11 was achieved in each case with yields 70–80% of that observed in wild-type chain combination. The absence of a stereospecific block to disulfide pairing indicates that the folding defect associated with L-Ala^{B8} is overcome by removal of the A7–B7 disulfide bridge. Conversely, the absence of stereospecific augmentation of disulfide pairing indicates that D-Ala^{B8} does not facilitate formation of noncontiguous cysteines A6–A11 and A20–B19. Together, these findings suggest that in the wild-type pathway of disulfide pairing the B8 dihedral angle is critical in allowing or hindering formation of the A7–B7 disulfide bridge.⁷ In accord with past studies (22), the two disulfide analogues are essentially devoid of biological activity (relative affinities <0.01%; Table 1).

DISCUSSION

Glycine can be conserved at key sites in proteins because of its small size. The interior of the collagen triple helix, for example, can accommodate only glycine, rationalizing the conserved Gly–X–Y triplet repeat (44). Similarly, in myoglobin and hemoglobin, the close approach of the B and E helices leads to broad conservation of glycine at helix position B6 (45). The absence of a side chain also enables glycine to adopt either positive or negative ϕ dihedral angles and thus occupy regions of the Ramachandran plot unfavorable to the other amino acids. Insulin and insulin-like growth factors contain an invariant glycine in the B chain (Gly^{B8}; asterisk in Figure 2) whose conformation switches as part of the T → R transition (Figure 3; refs 6 and 7). Although the physiological significance of this transition is unknown (as the R state has only been observed in zinc insulin hexamers and not among structures of IGFs or engineered insulin monomers; ref 8), we propose that this glycine is conserved to make possible a structural switch between folding-competent and active conformations of the hormone. Thus, impairment of the TR transition in a variant insulin hexamer by substitution of Gly^{B8} by D-Ala also impairs binding to the insulin receptor. The substitution nonetheless preserves the structure and enhances the stability of an engineered insulin monomer. We discuss these results from the complementary perspectives of stability, folding and function.

Stereospecific Stabilization of a Protein. Respective D and L substitutions at B8 give rise to large and opposite changes in thermodynamic stability ($\Delta\Delta G_u$; see Table 2). Why do D or L substitutions so markedly enhance or perturb stability? D-Ala^{B8} and D-Ser^{B8} enhance the stability of DKP-insulin by 1.5 ± 0.2 and 0.9 ± 0.2 kcal/mol, respectively. We

suggest that this gain in free energy has four sources. First, the D substitution may decrease the entropy of the unfolded state and hence the entropic penalty of folding. The extent of stabilization is greater than $RT \ln(2)$ (~ 0.5 kcal/mol), the expected entropic effect of restricting a single glycine's ϕ angle to half of the accessible Ramachandran plane. It is possible that a D substitution at B8 stabilizes a nascent β turn in the unfolded state, biasing the values of multiple dihedral angles and so increasing the resulting entropic benefit. Second, local interactions of the D side chain in the folded state may be favorable (or unfavorable). van der Waal contacts between D-Ala^{B8} and the neighboring side chains of Val^{A3}, Leu^{B11}, Val^{B12}, and Tyr^{B26}, for example (Figure 9), may contribute to stability; this environment may be less favorable in the polar Ser^{B8} analogue and so account for its lower stability. Third, possible damping of fluctuations elsewhere in the protein, presumed to be coupled to changes in the B8 dihedral angles, may strengthen multiple weak interactions; such an effect may be offset by an entropic penalty. Finally, a change in the topography and polarity of the B8-associated protein surface may alter its solvation properties. Of these four potential mechanisms, only the first—chiral restriction of the unfolded state—would be expected to be independent of the details of the folded state and to circumvent entropy–enthalpy compensation. Accordingly, it would be of interest in the future to test whether chiral substitution of a β turn in a globular protein may provide a general strategy for protein stabilization, and if so, whether the mechanism is primarily entropic. Such a strategy would extend established principles of medicinal chemistry widely employed in the stabilization of cyclic peptides.

The decrement in stability associated with L-Ala^{B8} or L-Ser^{B8} substitutions ($\Delta\Delta G_u$ 2.6 ± 0.3 and 3.0 ± 0.3 kcal/mol, respectively) is also likely to have complex origins. We imagine that a distorted local structure of the B7–B10 β turn is accompanied by transmitted perturbations in the helical structure of the protein. It is not clear from CD where the perturbations occur in the protein (i.e., in the neighboring B chain α -helix or in the A chain) and whether the attenuated CD signature reflects a discrete structural change (foreshortening of an α -helix) or a distributed dynamic destabilization without a change in helical endpoints. The CD spectrum of KP-insulin (containing substitutions Pro^{B28} → Lys and Lys^{B29} → Pro; ref 24), an analogue in clinical use as a rapid-acting insulin formulation (46), also exhibits an apparent attenuation of helix content even though crystallographic and NMR studies have demonstrated nativelike secondary structure (13, 47, 48). A molecular understanding of L-Ser^{B8} analogues will require high-resolution NMR or crystallographic analysis of these distorted structures. It is noteworthy that the unfavorable L-amino acids are not accommodated at B8 (as might be imagined) simply by propagation of the B9–B19 α helix into the N-terminal segment (i.e., adoption of an extended R-like conformation). This possibility is excluded by the CD perturbations implying attenuation (rather than enhancement) of helix-related features. It would be intriguing if L-Ser^{B8} is nevertheless accommodated in a high-affinity hormone–receptor complex by an R-like conformational change.

Protein Stability and Stereospecific Modulation of the TR Transition. Although the present study has focused on the insulin monomer, the biologically active species, our ex-

⁷ The kinetic hypothesis is in accord with prior studies indicating that chain combination is in part under kinetic control (28) and that pairing of cysteine A7–B7 plays a critical role in the oxidative folding pathway of insulin-like polypeptides (78, 79).

perimental design is stimulated by the extensive foundation of crystallographic studies of zinc-insulin hexamers. Such assemblies are remarkable for a ligand-dependent equilibrium among T_6 , $T_3R_3^f$, and R_6 hexamers (8). The TR transition is characterized by long-range conformational changes associated with a switch in the Ramachandran position of Gly^{B8}. D Substitutions at B8 significantly stabilize the T-like conformation of an insulin monomer and would hence be expected to shift the equilibrium among T_6 , $T_3R_3^f$, and R_6 hexamers toward the T state. Such a shift is in accord with our observation of attenuated Co^{2+} d–d transitions, which are diagnostic of the R state-specific tetrahedral metal-binding site in phenol-stabilized hexameric complexes (25–27). These results suggest that—although D-Ala^{B8}-insulin can indeed form an R state—the variant R state is less stable than the variant T state under conditions (high phenol concentration) in which wild-type insulin is almost entirely in the R_6 conformation. The D-Ala^{B8} substitution thus impedes (but does not entirely prevent) both binding of an insulin monomer to the IR and the hexameric TR transition.

Crystallization trials of D-Ala^{B8}-insulin zinc hexamers under conditions that ordinarily give rise to a wild-type R_6 lattice instead yield $T_3R_3^f$ crystals. This result rationalizes the ca. 2-fold attenuation of the previous Co^{2+} d–d transitions and demonstrates that the TR transition has been interrupted at an intermediate stage. Although the molecular details of this structure have not been determined, the overall ability of D-Ala^{B8} to be accommodated within an R^f state-specific α helix would be consistent with studies of model helicogenic peptides by W. DeGrado and co-workers in which a single D-Ala was shown to be energetically well-tolerated within an α helix (49). Although local conformational adjustments may occur, such accommodation incurs only a modest cost in free energy ($\Delta\Delta G_u$ 1 kcal/mol relative to L-Ala and 0.2 kcal/mol relative to glycine at the same peptide position). The ability of a canonical α -helix to accommodate one or more D amino acids has been elegantly demonstrated by Karle and co-workers in crystallographic studies of model peptides in helicogenic solvents (84). Impairment of the TR transition by D-Ala^{B8} is thus likely to have two mechanisms: marked chiral stabilization of the T state and modest chiral destabilization of the R state.

The crystal structure of D-Ala^{B8}-insulin as a $T_3R_3^f$ hexamer promises to extend prior peptide studies (84) to define how a D amino acid is accommodated in an α helix within a large protein assembly. Substitution of Gly^{B8} by L-Ser is likewise compatible with either T or R^f conformation in crystallographic hexamers (J. Whittingham and G. G. Dodson, personal communication). We would expect such an L substitution to shift the equilibrium among T_6 , $T_3R_3^f$, and R_6 hexamers toward the R state. Although the Ramachandran plot is often interpreted in terms of forbidden regions, these considerations emphasize that D or L substitutions influence the relative thermodynamic stabilities of competing conformations rather than provide covalent locks to tether or exclude specific structures.

Because the solution structure of D-Ala^{B8}-DKP-insulin closely resembles the parent monomer and in turn the crystallographic T state, we initially sought to obtain crystals of D-Ala^{B8}-insulin under a range of conditions suitable for growth of wild-type T_6 crystals. Surprisingly, the analogue exhibited crystallization properties different from those of

wild-type insulin: anomalously rapid growth of clusters of small crystals. Suitable T_6 crystals of the analogue were therefore not obtained despite extensive trials. Accelerated crystallization, although a practical impediment to obtaining highly ordered single crystals, suggests that the D substitution enhances lattice growth either by stabilizing intermolecular contacts or damping structural fluctuations that otherwise retard the growth of wild-type crystals. It is possible that in the future suitable single T_6 crystals of D-Ala^{B8}-insulin can be obtained under conditions that are less favorable for wild-type crystallization, retarding the rate of crystallization to permit orderly growth of single crystals. Because the wild-type T_6 crystal form has been found to diffract to atomic resolution (1, 50), such crystals would be of interest to enable comparative analysis of anisotropic B values and alternative side-chain conformations as probes of protein dynamics. It is also possible that at higher concentrations of phenol, D-Ala^{B8}-insulin can be driven into the R_6 state as the free energy of ligand binding compensates for chiral destabilization of the R-state-specific α -helix.

Insulin Chain Combination and the Folding of Proinsulin. In vivo insulin is the product of a larger single-chain precursor, designated proinsulin (86 residues; ref 51). Native disulfide pairing (A6–A11, A7–B7, and A20–B19) occurs in the endoplasmic reticulum (ER). In vitro studies suggest that such pairing is directed exclusively by sequences in the A and B domains (14); the connecting peptide (35 residues) provides a passive tether. Combination of the isolated A and B chains has thus been studied as a peptide model of proinsulin-folding intermediates (28, 52). The present study of chain-combination efficiency exploits peptide libraries to demonstrate a stereospecific preference for D amino acids at B8. In fact, D-Ala^{B8} enhances the yield of chain combination by 3-fold; to our knowledge, such enhancement is without precedent in the literature of insulin synthesis. It is likely that even higher yields would be conferred by other members of the D library. Equally striking, the corresponding L library yields no productive sequences. Use of libraries containing all possible D or L side chains (exclusive of cysteine and proline) thus indicates that the key feature is the chirality of the α carbon (and hence preferred sign of the ϕ dihedral angle) rather than the size or chemical features of individual side chains. Representative L analogues can nonetheless be prepared in very low yield. The relevance of the peptide libraries to proinsulin is suggested by studies of the expression of a foreshortened single-chain analogue (porcine insulin precursor; 52 residues) in *Saccharomyces cerevisiae*: substitution of Gly^{B8} by L-Ala markedly impairs productive folding and secretion (35). It would be of future interest to extend these results to the folding and secretion of human proinsulin analogues by a transfected mammalian neuroendocrine cell line.

Together, our results suggest that the B8 conformation plays a critical role in the mechanism of disulfide pairing. Because D substitutions stabilize a native T-like fold (the structure of an isolated insulin monomer; refs 10 and 13), exclusion of L substitutions suggests that local nativelylike structures are populated in the relevant protein-folding intermediates and serve to guide disulfide pairing. Although these intermediates are not well-characterized (53), pairwise substitution of cystine A7–B7 by serine leads to rescue of chain combination and essentially complete attenuation of

stereospecific interference by L-Ala^{B8}. This finding suggests that the B8 conformation is coupled to productive orientation of the adjoining Cys^{B7}. We speculate that a kinetic block to formation of the A7–B7 disulfide bridge leads to competing off-pathway events. The yield of insulin chain combination (like that of proinsulin folding; ref 54) is well-known to be limited by aberrant aggregation of partially folded species (28). Analogous partial unfolding, even in the presence of the three native disulfide bridges, may underlie the unexpected self-association of L-Ala^{B8}-DKP-insulin and the slow aggregation of the L-Ser^{B8}-DKP-insulin NMR sample. Such aggregation is ordinarily mediated by the B chain (or B domain) and can lead to irreversible cross- β assembly of fibrils. It would be of future interest to investigate whether B8 L substitutions in proinsulin lead to formation of pathological electron-dense deposits in the endoplasmic reticulum in cell culture or accelerated aggregation or fibrillation in vitro. Such features are characteristic of a mutant proinsulin in a murine model of diabetes mellitus (the Akita mouse; ref 55).

The marked enhancement in the efficiency of chain combination associated with the D-Ala^{B8} substitution is unlikely simply to reflect the greater stability of the end product (the insulin analogue). Past syntheses of unrelated insulin variants of enhanced stability (such as Asp^{B10}-insulin and His^{A8}-insulin; 60, 68, 85) have exhibited only modest increases in yield (<20%). The magnitude of the increase in yields of Ala^{B8}-insulin and Ala^{B8}-DKP-insulin (3-fold) therefore suggests favorable kinetic control of the reaction. It is possible that the D substitution biases the conformation of a protein-folding intermediate to facilitate A7–B7 disulfide pairing (an example of on-pathway control). It is also possible that the D substitution impairs competing side reactions, such as formation of cyclic B chains, covalent B-chain dimers, or B-chain fibrils (off-pathway control). We speculate that both of these mechanisms are operative. On the one hand, because D substitutions stabilize disulfide pairing in the native state, it is reasonable to suppose that they also stabilize the productive orientation of Cys^{A7} and Cys^{B7} in a transition state. On the other hand, it is also plausible that D substitutions at B8 impede cross- β assembly of B chains. Such a unique combination of on- and off-pathway effects may account for the extraordinary stereospecific augmentation of yields observed in the present D-Ala^{B8} syntheses.

Induced Fit and the Active Conformation of Insulin. How insulin binds to the insulin receptor has long been the subject of speculation (1, 2). That classical structures represent inactive conformers has been suggested by studies of an inactive single-chain analogue (mini-proinsulin) in which a peptide bond tethers the C-terminal β strand of the B chain (Lys^{B29}) to the N-terminal α helix of the A chain (Gly^{A1}; refs 3 and 56).⁸ On receptor binding, the β strand presumably detaches from the core to contact the receptor (4) and expose an otherwise hidden functional surface in the A chain (30,

33). The present study exploits chiral stabilization of the B7–B10 β turn by D-amino acid substitutions at B8 to provide a complementary example of an inactive analogue with natively like structure. Substitution of Gly^{B8} by D-Ala or D-Ser enhances the thermodynamic stability of insulin but markedly impairs its binding to the insulin receptor (Table 2). As in mini-proinsulin (56), the extent of impairment is highly unusual among mutant insulins (1, 11). The solution structure of D-Ala^{B8}-DKP-insulin resembles wild-type insulin (and the parent DKP-insulin monomer; ref 13) with the exception of the protruding methyl group at B8. Although we cannot exclude that D side chains at B8 themselves introduce a steric clash at the receptor interface (see below), we propose that the severe decrement in binding reflects a chiral impediment to induced fit: the analogues are locked into an inactive T-like conformation. This model thus envisages that both the N- and the C-terminal segments of the B chain reorganize on receptor binding.⁹ Such reorganization would in each case expose nonpolar side chains that are otherwise inaccessible in the T state and so make possible a more extensive hormone–receptor interface. Whereas displacement of the C-terminal segment of the B chain would expose Ile^{A2} and Val^{A3} (3, 4, 30, 33, 43, 65), reorganization of the B1–B8 segment may likewise enable Ile^{A10}, Leu^{A13}, and Leu^{B6} to contact the receptor (Figure 1E,F; 2, 27, 69).

Why is steric clash unlikely to account for the inactivity of B8 D analogues? This model seems implausible for several reasons. First, single amino acid substitutions on the surface of insulin generally affect receptor binding by less than 100-fold. Alanine-scanning mutagenesis of the protein surface, for example, yields a range of relative receptor-binding affinities between 30% (Leu^{A13} \rightarrow Ala) and 405% (Arg^{B22} \rightarrow Ala) (11). Even at presumed sites of van der Waals contact between insulin and the IR, enlarging the size of a side chain by one methyl group typically has only modest effects. Examples are provided by critical conserved valines in α helices: Val^{A3} and Val^{B12}. Substitution of valine by isoleucine or *allo*-isoleucine leads to a one-methyl protrusion of one branch or the other. In each case, the respective receptor binding affinities are decreased by less than 10-fold (see Supporting Information; refs 18 and 43). Similarly, substitution of Val^{A3} or Val^{B12} by *t*-leucine (which substitutes H β by a third methyl group) perturbs receptor binding by less than tenfold (18, 43).¹⁰ Accommodation of additional methyl substituents at B12 is of particular interest in light of its proximity to B8 (in both T and R states; Supporting Information). Other sites near B8 also tolerate marked changes in side-chain volume as follows. (i) Only modest changes in affinity are observed on substitution of Ser^{B9} by

⁹ The T \rightarrow R transition is also characterized by a small displacement of the C-terminal B chain β strand away from the A chain, breaking the B25 NH \cdots O=C A19 hydrogen bond observed in the T state (1). The two modes of induced fit proposed to occur in the N- and C-terminal segments of the B chain, although spatially separated in the protein, may in part be correlated (G. G. Dodson, personal communication).

¹⁰ Whereas substitution of Val^{B12} by Leu likewise impairs receptor binding by only less than 20-fold (16), Leu^{A3}-insulin exhibits an unusual 500-fold decrease (43, 81), indicating limited flexibility around the presumed A3-binding pocket (82). A site of clinical mutation causing diabetes mellitus (83), Val^{A3} is largely buried in a crevice between chains that is distant from the B8-associated surface. The crystal structure of Leu^{A3}-insulin (as an R₆ zinc hexamer) is essentially identical to that of wild-type insulin (82).

⁸ Mini-proinsulin refolds more efficiently than does proinsulin to achieve native disulfide pairing (80). Its enhanced foldability mirrors a loss of biological activity (64): the tethered native state is presumably constrained from reorganizing on receptor binding (3). Analogous (but less stringent) tethering has been described in an analogue containing a chemical cross-link between the ϵ -amino group of Lys^{B29} and the α -amino group of Gly^{A1} (63).

the larger side chains of Asn or Asp (57). (ii) Substitution of His^{B5} (near B8 in the structure of DPI; refs 58 and 59) by Asn or Ala is similarly well-tolerated. (iii) Substitution of His^{B10} by Asn or Leu leads to 2- and 3-fold decrements in receptor binding; Asp^{B10} enhances binding whereas Lys^{B10} impairs binding by 7-fold, presumably due to favorable or unfavorable long-range electrostatic interactions (60, 61). Given these quantitative trends, the 1000-fold decrement in activity caused by D-Ala^{B8} seems anomalous and unlikely to be due to rigidity of packing near this part of the insulin surface. Indeed, whereas the steric-clash model envisages a snug fit between Gly^{B8} and the insulin receptor, we instead imagine that this site adjoins a cavity between respective protein surfaces. This model would rationalize the relative affinities of D-Ser^{B8}-DKP-insulin: although serine is larger than alanine, the residual binding of the D-Ser^{B8} analogue is 10-fold higher than that of D-Ala^{B8}-DKP-insulin. Such enhancement, difficult to reconcile with the steric-clash model, might arise from the introduction of a new interaction between the serine side chain and an otherwise uncontacted wall of the cavity, partially offsetting the D-associated chiral impairment. Similarly, the bulky side chain of D-*p*-amino-Phe (Pmp) is accommodated at B8 with less impairment than that caused by D-Ala^{B8} (Table 1), presumably due to favorable interactions in the putative B8-associated cavity.

Insulin analogues of extremely low activity ordinarily fall into two classes. The first are substitutions that lead to transmitted structural perturbations, presumably impairing both local and nonlocal hormone–receptor contacts. Examples include substitution of α helical contact residues Val^{A3} or Val^{B12} by glycine (relative affinities 0.1–0.2 and 0.32%; refs 18, 43, and 62), expected to destabilize the helical segment (40, 62). Substitution of a framework aromatic side chain in the hydrophobic core (Tyr^{A19} \rightarrow Ala), expected to disrupt tertiary structure, likewise impairs receptor binding by 1000-fold (11). Similarly, pairwise substitution Cys^{A7} and Cys^{B7} by serine (hence removing the A7–B7 disulfide bridge) leads to a molten partial fold whose activity is reduced by at least 5000-fold (13). The present NMR studies demonstrate that D-Ala^{B8} analogues do not fall into this class as the solution structure of D-Ala^{B8}-DKP-insulin is essentially identical to that of the parent monomer (13). The observed confluence of native structure and very low activity is reminiscent of single-chain insulin analogues in which the C-terminus of the B chain is connected to the N-terminus of the A chain directly or by a short linker (3, 56, 63, 64). This topological constraint is envisaged to hinder induced fit rather than impose a local steric clash. We propose an analogy between such a topological restraint and chiral restriction of the flexibility of Gly^{B8} by a D substitution. Consistent with the previous discussion of the thermodynamics of the TR transition, also impaired by D-Ala^{B8}, we imagine that the 1000-fold decrement in binding of D-Ala^{B8} analogues arises in two ways: from chiral stabilization of the initial inactive T state and from the cost of accommodating D-Ala^{B8} in an induced R-like α helix (49). In essence, induced fit is impaired because the initial state is harder to leave and the final state is less favorable to enter.

The induced-fit hypothesis at the C-terminal region of the B chain is supported by the high activities of analogues that exhibit structural perturbations of the B chain. Substitution of Phe^{B24} (which anchors the C-terminal β strand to the core;

ref 1) by glycine leads to an altered segmental structure with near-native activity (4, 65).¹¹ Further, the activity of insulin is enhanced by chiral substitution of Phe^{B24} by D-Phe or D-Ala (20, 66); these substitutions appear to be incompatible with native structure. The present study has yielded analogous findings at B8, which suggest a second site of conformational change in the N-terminal region of the B chain. Whereas the nativelylike D-Ala^{B8}-stabilized T state is essentially inactive, the L-Ser^{B8}-distorted structure exhibits high activity despite its attenuated α helix content and lower stability. We imagine that structural distortions associated with D substitutions at B24 or L substitutions at B8 facilitate induced fit of the respective C- or N-terminal segments of the B chain and so are compatible with high activity. It would be of future interest to investigate the solution structures of these analogues.

Limitations of Conventional Mutagenesis. The present study illustrates a limitation of alanine scanning mutagenesis. The low activity of L-Ala^{B8}-insulin (ref 11 and present results) is likely to reflect aberrant aggregation rather than local structural changes. Because L-Ser^{B8}-DKP-insulin is initially monomeric but exhibits a similar decrement in stability and CD-detected helix content, rapid aggregation is a particular feature of the Ala^{B8} analogue. Such aggregation was unanticipated as the DKP substitutions disallow classical self-association: inversion of B28–B29 destabilizes the dimer-related β sheet, whereas the substitution His^{B10} \rightarrow Asp destabilizes the hexamer-forming interface (20, 24, 48, 67). That these substitutions do not prevent self-association of Ala^{B8}-DKP-insulin implies that aggregation is mediated by distorted protein surfaces. Interestingly, L-Ser^{B8} is compatible with nativelylike assembly of zinc-insulin hexamers and crystallization (J. Whittingham and G. G. Dodson, personal communication). Native zinc-mediated assembly presumably protects L-Ser^{B8}-insulin from the slow aggregation observed in NMR samples of L-Ser^{B8}-DKP-insulin. With respect to the different properties of L-Ala and L-Ser analogues, it may be relevant that generally among protein crystal structures serine is more likely than alanine to exhibit positive ϕ angles.

Confounding effects of structural distortion have previously been encountered in alanine scanning mutagenesis of insulin (11). At some sites, Ala variants could not be expressed, presumably due to misfolding and degradation in the yeast expression system (11). Kinetic barriers to disulfide pairing in vitro can also block synthetic chain combination (16, 68). Of particular interest, substitution of invariant side chain Ile^{A2} by Ala permits robust expression (11) but leads to segmental unfolding of the A1–A8 α helix (5). In this case, substitution by *allo*-Ile^{A2} enabled the specific contribution of this side chain to be resolved (30, 33). Detachment or reorganization of the C-terminal segment of the B chain, as envisaged previously, would expose Ile^{A2} to engage the receptor. Analogous reorganization of the B1–B8 segment would likewise expose the central surface of the A chain (A10–A13) to provide an additional contact surface (2, 69). The total chemical synthesis of insulin

¹¹ Two structural models of Gly^{B24} analogues have been proposed in which the C-terminal segment is either detached from the core (4) or aberrantly attached via non-native contacts by Phe^{B25} (65). Reassessment of these models suggests that the unanchored segment is primarily disordered but able to make transient contacts with the core (Q.-x. Hua and M. A. Weiss, manuscript in preparation).

permits use of nonstandard amino acids to extend the ordinary repertoire of site-directed mutagenesis. The use of stereoisomers such as *allo*-Ile (30, 43) and D amino acids, as exemplified here, may be of general utility in studies of synthetic proteins (70).

CONCLUSION

Important insight into the informational content of protein sequences has been obtained through genetic analysis of allowed and disallowed substitutions (71, 72). Application of random cassette mutagenesis to peptide libraries, pioneered by Kent and co-workers (15), provides a chemical approach to dissecting determinants of foldability. In the present study, the ability of B chain variants to combine with the wild-type A chain provides a functional selection. Advances in both synthetic methodologies and MS sequencing methods may enable this approach to be extended to multiple sites in the A and B chains. Such studies are motivated by the hypothesis that the evolution of insulin sequences is constrained in part by foldability: complementary requirements for specific disulfide pairing and against toxic misfolding. The present study has established that L and D substitutions at position B8 exhibit striking stereospecific modulation of synthetic yields and thermodynamic stabilities. We speculate that the flexibility of Gly^{B8} serves to protect proinsulin from misfolding and aberrant aggregation in the β -cell.

The functional implications of the present study transcend our initial goals pertaining to combinatorial chain combination and protein folding. Ala^{B8}-DKP-insulin provides a novel example of an insulin analogue with native structure but very low activity. Previous examples include mini-proinsulin (whose structure is tethered by a peptide bond between the A and B chains; ref 3), a nonstandard main-chain analogue containing an ester linkage at B25 (26, 73), and the chiral analogue *allo*-Ile^{A2}-insulin (30, 33). Each of these structures provides evidence for induced fit in receptor binding. We propose that Gly^{B8} functions as a structural switch to enable both specific disulfide pairing and high-affinity receptor binding. The proposed switch recapitulates a key aspect of the classical T \rightarrow R transition (8): a ligand-dependent reorganization of the insulin hexamer remarkable for a change in the ϕ angle of Gly^{B8} from positive (as preferred by D amino acids) to negative (as preferred by L amino acids). It would be of future interest to address whether other features of the R state (such as its extended B1–B19 α helix or change in conformation of cystine A7–B7) would also be induced on receptor binding. Although the receptor-bound conformation of insulin is likely to be neither completely T-like nor completely R-like, we anticipate that similar sites of conformation change will be utilized. If so, future cocrystal structures of insulin–receptor complexes will validate the prescient vision of Dorothy Hodgkin and colleagues, whose studies of zinc insulin crystals provided a model of conformational change (74). The pioneering perspective of the Hodgkin school (1) thus suggests that D amino acid substitutions at B8 act as a chiral spanner in the works of receptor recognition.

ACKNOWLEDGMENT

We thank P. G. Katsoyannis for support and advice regarding peptide synthesis; M. DeFelippis (Eli Lilly and

Co.) for ultracentrifuge studies and the gift of human insulin; B. H. Frank, Y.-M. Feng, and J. Whittingham for communication of results prior to publication; and P. Arvan, T. L. Blundell, P. De Meyts, G. G. Dodson, A. Fersht and D. F. Steiner for helpful discussions. M.A.W. thanks C. M. Dobson and M. Karplus for encouragement. This article is a contribution from the Cleveland Center for Structural Biology. These studies were stimulated in part by historical remarks of Prof. Dorothy C. Hodgkin on the 20th anniversary of the determination of the crystal structure of insulin (The University of York, York, England, 1989). We dedicate this article to her memory.

SUPPORTING INFORMATION AVAILABLE

Nine figures illustrating disulfide pairing and structural relationships in insulin crystals, visible absorption spectra of cobalt-substituted hexamers, additional CD and NMR spectra, diagonal plot of NOEs, and summary of NMR sequential assignment. Nine tables providing B8 dihedral angles, summary of mutations at sites neighboring B8, crystallographic unit-cell dimensions, NMR resonance assignments, statistical information pertaining to DG/RMD ensemble, and restraints. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

1. 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. Royal Soc. London Ser. 319*, 369–456.
2. De Meyts, P., and Whittaker, J. (2002) Structural biology of insulin and IGF-I receptors: implications for drug design, *Nat. Rev. Drug Discovery* 1, 769–783.
3. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., and Markussen, J. (1991) X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue, *J. Mol. Biol.* 220, 425–433.
4. Hua, Q. X., Shoelson, S. E., Kochoyan, M., and Weiss, M. A. (1991) Receptor binding redefined by a structural switch in a mutant human insulin, *Nature* 354, 238–241.
5. Xu, B., Hua, Q. X., Nakagawa, S. H., Jia, W., Chu, Y. C., Katsoyannis, P. G., and Weiss, M. A. (2002) A cavity-forming mutation in insulin induces segmental unfolding of a surrounding α -helix, *Protein Sci.* 11, 104–116.
6. Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., and Mercola, D. (1976) Structure of insulin in 4-zinc insulin, *Nature* 261, 166–168.
7. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., and Swenson, D. (1989) Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer, *Nature* 338, 594–596.
8. Brader, M. L., and Dunn, M. F. (1991) Insulin hexamers: new conformations and applications, *Trends Biochem. Sci.* 16, 341–345.
9. Hua, Q. X., Narhi, L., Jia, W., Arakawa, T., Rosenfeld, R., Hawkins, N., Miller, J. A., and Weiss, M. A. (1996) Native and non-native structure in a protein-folding intermediate: spectroscopic studies of partially reduced IGF-I and an engineered alanine model, *J. Mol. Biol.* 259, 297–313.
10. Olsen, H. B., Ludvigsen, S., and Kaarsholm, N. C. (1996) Solution structure of an engineered insulin monomer at neutral pH, *Biochemistry* 35, 8836–8845.
11. Kristensen, C., Kjeldsen, T., Wiberg, F. C., Schaffer, L., Hach, M., Havelund, S., Bass, J., Steiner, D. F., and Andersen, A. S. (1997) Alanine scanning mutagenesis of insulin, *J. Biol. Chem.* 272, 12978–12983.
12. Minor, D. L., Jr., and Kim, P. S. (1994) Measurement of the β -sheet-forming propensities of amino acids, *Nature* 367, 660–663.

13. Hua, Q. X., Hu, S. Q., Frank, B. H., Jia, W., Chu, Y. C., Wang, S. H., Burke, G. T., Katsoyannis, P. G., and Weiss, M. A. (1996) Mapping the functional surface of insulin by design: structure and function of a novel A chain analogue, *J. Mol. Biol.* **264**, 390–403.
14. Katsoyannis, P. G., and Tometsko, A. (1966) Insulin synthesis by recombination of A and B chains: a highly efficient method, *Proc. Natl. Acad. Sci. U.S.A.* **55**, 1554–1561.
15. Muir, T. W., Dawson, P. E., Fitzgerald, M. C., and Kent, S. B. H. (1996) Probing the chemical basis of binding activity in an SH3 domain by protein signature analysis, *Chem. Biol.* **3**, 817–825.
16. Hu, S. Q., Burke, G. T., Schwartz, G. P., Ferderigos, N., Ross, J. B., and Katsoyannis, P. G. (1993) Steric requirements at position B12 for high biological activity in insulin, *Biochemistry* **32**, 2631–2635.
17. Barany, G., and Merrifield, R. B. (1980) in *The Peptides* (Gross, E., and Meienhofer, J., Eds.) pp 273–284, Academic Press, New York.
18. Nakagawa, S. H., Tager, H. S., and Steiner, D. F. (2000) Mutational analysis of invariant valine B12 in insulin: implications for receptor binding, *Biochemistry* **39**, 15826–15835.
19. Chance, R. E., Hoffman, J. A., Kroeff, E. P., Johnson, M. G., Schirmer, W. E., and Bormer, W. W. (1981) in *Peptides: Synthesis, Structure, and Function; Proceedings of the Seventh American Peptide Symposium* (Rich, D. H., and Gross, E., Eds.) pp 721–728, Pierce Chemical Co., Rockford, IL.
20. Shoelson, S. E., Lu, Z. X., Parlaatan, 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.
21. Xu, B., Hu, S. Q., Chu, Y. C., 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.
22. Hua, Q.-X., Nakagawa, S. H., Jia, W., Hu, S. Q., Chu, Y.-C., Katsoyannis, P. G., and Weiss, M. A. (2001) Hierarchical protein folding: asymmetric unfolding of an insulin analogue lacking the A7–B7 interchain disulfide bridge, *Biochemistry* **40**, 12299–12311.
23. Weiss, M. A., Hua, Q.-X., Jia, W., Nakagawa, S. H., Chu, Y.-C., Hu, S.-Q., and Katsoyannis, P. G. (2001) Activities of monomeric insulin analogues at position A8 are uncorrelated with their thermodynamic stabilities, *J. Biol. Chem.* **276**, 40018–40024.
24. DeFelippis, M. R., Bakaysa, D. L., Bell, M. A., Heady, M. A., Li, S., Pye, S., Youngman, K. M., Radziuk, J., and Frank, B. H. (1998) Preparation and characterization of a cocrystalline suspension of [Lys^{B28}, Pro^{B29}]-human insulin analogue, *J. Pharmaceut. Sci.* **87**, 170–176.
25. 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.
26. Wollmer, A., Gilge, G., Brandenburg, D., and Gattner, H. G. (1994) An insulin with the native sequence but virtually no activity, *Biol. Chem. Hoppe-Seyler* **375**, 219–222.
27. Nakagawa, S. H., and Tager, H. S. (1991) Implications of invariant residue Leu^{B6} in insulin–receptor interactions, *J. Biol. Chem.* **266**, 11502–11509.
28. Hua, Q.-X., Chu, Y.-C., Jia, W., Phillips, N. F. B., Wang, R.-Y., Katsoyannis, P. G., and Weiss, M. A. (2002) Mechanism of insulin chain combination. Asymmetric roles of A-chain α -helices in disulfide pairing, *J. Biol. Chem.* **277**, 43443–43453.
29. Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) Application of circular dichroism to study RNA folding transitions, *Methods Enzymol.* **317**, 393–409.
30. Xu, B., Hua, Q. X., Nakagawa, S. H., Jia, W., Chu, Y. C., Katsoyannis, P. G., and Weiss, M. A. (2002) Chiral mutagenesis of insulin's hidden receptor-binding surface: structure of an *allo*-isoleucineA2 analogue, *J. Mol. Biol.* **316**, 435–441.
31. Hua, Q. X., Jia, W., Frank, B. H., and Weiss, M. A. (1993) Comparison of the dynamics of an engineered insulin monomer and dimer by acid-quenched amide proton exchange. Nonlocal stabilization of interchain hydrogen bonds by dimerization, *J. Mol. Biol.* **230**, 387–394.
32. Harding, M. M., Hodgkin, D. C., Kennedy, A. F., O'Connor, A., and Weitzmann, P. D. (1966) The crystal structure of insulin. II. An investigation of rhombohedral zinc insulin crystals and a report of other crystalline forms, *J. Mol. Biol.* **16**, 212–226.
33. Wan, Z., Xu, B., Chu, Y. C., Katsoyannis, P. G., and Weiss, M. A. (2003) Crystal structure of *allo*-Ile^{A2}-insulin, an inactive chiral analogue: implications for the mechanism of receptor binding, *Biochemistry* **42**, 12770–12783.
34. Brunger, A. T. (1993) *XPLOR Manual, Version 3.1*, Yale University Press, New Haven, CT.
35. Guo, Z. Y., Tang, Y. H., Wang, S., and Feng, Y. M. (2003) Contribution of the absolutely conserved B8Gly to the foldability of insulin, *Biol. Chem.* **384**, 805–809.
36. Ciszak, E., and Smith, G. D. (1994) Crystallographic evidence for dual coordination around zinc in the T₃R₃ human insulin hexamer, *Biochemistry* **33**, 1512–1517.
37. Tang, L., Whittingham, J. L., Verma, C. S., Caves, L. S. D., and Dodson, G. G. (1999) Structural consequences of the B5 histidine tyrosine mutation in human insulin characterized by X-ray crystallography and conformational analysis, *Biochemistry* **38**, 12041–12051.
38. Whittingham, J. L., Edwards, D. J., Antson, A. A., Clarkson, J. M., and Dodson, G. G. (1998) Interactions of phenol and *m*-cresol in the insulin hexamer and their effect on the association properties of B28 Pro → Asp insulin analogues, *Biochemistry* **37**, 11516–11523.
39. Smith, G. D., Ciszak, E., Magrum, L. A., Pangborn, W. A., and Blessing, R. H. (2000) R₆ hexameric insulin complexed with *m*-cresol or resorcinol, *Acta Crystallogr. D56*, 1541–1548.
40. Weiss, M. A., Wan, Z., Zhao, M., Chu, Y. C., Nakagawa, S. H., Burke, G. T., Jia, W., Hellmich, R., and Katsoyannis, P. G. (2002) Nonstandard insulin design: structure–activity relationships at the periphery of the insulin receptor, *J. Mol. Biol.* **315**, 103–111.
41. Huang, K., Xu, B., Hu, S. Q., Chu, Y. C., Hua, Q. X., Qu, Y., Li, B., S., W., Y., W. R., 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.
42. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
43. Nakagawa, S. H., and Tager, H. S. (1992) Importance of aliphatic side-chain structure at positions 2 and 3 of the insulin A chain in insulin–receptor interactions, *Biochemistry* **31**, 3204–3214.
44. Mayo, K. H. (1996) NMR and X-ray studies of collagen model peptides, *Biopolymers* **40**, 359–370.
45. Perutz, M. F. (1969) Structure and function of hemoglobin, *Harvey Lect.* **63**, 213–261.
46. Vajdos, F. F., Ultsch, M., Schaffer, M. L., Deshayes, K. D., Liu, J., Skelton, N. J., and de Vos, A. M. (2001) Crystal structure of human insulin-like growth factor-1: detergent binding inhibits binding protein interactions, *Biochemistry* **40**, 11022–11029.
47. Weiss, M. A., Hua, Q. X., Lynch, C. S., Frank, B. H., and Shoelson, S. E. (1991) Heteronuclear 2D NMR studies of an engineered insulin monomer: assignment and characterization of the receptor-binding surface by selective ²H and ¹³C labeling with application to protein design, *Biochemistry* **30**, 7373–7389.
48. Ciszak, E., Beals, J. M., Frank, B. H., Baker, J. C., Carter, N. D., and Smith, G. D. (1995) Role of C-terminal B chain residues in insulin assembly: the structure of hexameric LysB28ProB29-human insulin, *Structure* **3**, 615–622.
49. Fairman, R., Anthony-Cahill, S. J., and DeGrado, W. F. (1992) The helix-forming propensity of d-alanine in a right-handed α -helix, *J. Am. Chem. Soc.* **114**, 5458–5459.
50. Sakabe, N., Skabe, K., and Sakaki, K. (1981) in *Structural Studies On Molecules of Biological Interest* (Dodson, G., Glusker, J. P., and Sayre, D., Eds.) pp 509–526, Oxford University Press, Oxford.
51. Dodson, G., and Steiner, D. (1998) The role of assembly in insulin's biosynthesis, *Curr. Opin. Struct. Biol.* **8**, 189–194.
52. Wang, C. C., and Tsou, C. L. (1991) The insulin A and B chains contain sufficient structural information to form the native molecule, *Trends Biochem. Sci.* **16**, 279–281.
53. Min, C. Y., Qiao, Z. S., and Feng, Y. M. (2004) Unfolding of human proinsulin. Intermediates and possible role of its C-peptide in folding/unfolding, *Eur. J. Biochem.* **271**, 1737–1747.
54. Winter, J., Klappa, P., Freedman, R. B., Lilie, H., and Rudolph, R. (2002) Catalytic activity and chaperone function of human protein-disulfide isomerase are required for the efficient refolding of proinsulin, *J. Biol. Chem.* **277**, 310–317.
55. Ron, D. (2002) Proteotoxicity in the endoplasmic reticulum: lessons from the Akita diabetic mouse, *J. Clin. Invest.* **109**, 443–445.

56. Hua, Q.-X., Hu, S. Q., Jia, W., Chu, Y.-C., Burke, G. T., Wang, S. H., Wang, R. Y., Katsoyannis, P. G., and Weiss, M. A. (1998) Mini-proinsulin and mini-IGF-I: homologous protein sequences encoding nonhomologous structures, *J. Mol. Biol.* 277, 103–118.
57. 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.
58. Bi, R. C., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Giorgino, F., Reynolds, C. D., and Tolley, S. P. (1983) Molecular-replacement studies on crystal forms of despentapeptide insulin, *Acta Crystallogr.* B39, 90–98.
59. Dai, J.-B., Lou, M.-Z., You, J.-M., and Liang, D.-C. (1987) Refinement of the structure of despentapeptide (B26–B30) insulin at 1.5 Å resolution, *Sci. Sin.* 30, 55–65.
60. Burke, G. T., Schwartz, G. P., and Katsoyannis, P. G. (1984) Nature of the B10 amino acid residue. Requirements for high biological activity of insulin, *Int. J. Pept. Protein Res.* 23, 394–401.
61. Schwartz, G. P., Burke, G. T., and Katsoyannis, P. G. (1987) A superactive insulin: [B10-aspartic acid]insulin(human), *Proc. Natl. Acad. Sci. U.S.A.* 84, 6408–11.
62. Olsen, H. B., Ludvigsen, S., and Kaarsholm, N. C. (1998) The relationship between insulin bioactivity and structure in the NH₂-terminal A chain helix, *J. Mol. Biol.* 284, 477–488.
63. Cutfield, J., Cutfield, S., Dodson, E., Dodson, G., Hodgkin, D., and Reynolds, C. (1981) Evidence concerning insulin activity from the structure of a cross-linked derivative, *Hoppe-Seyler's Z. Physiol. Chem.* 362, 755–761.
64. Markussen, J., Jorgensen, K. H., Sorensen, A. R., and Thim, L. (1985) Single chain *des*-(B30) insulin. Intramolecular cross-linking of insulin by trypsin catalyzed transpeptidation, *Int. J. Peptide Protein Res.* 26, 70–77.
65. Ludvigsen, S., Olsen, H. B., and Kaarsholm, N. C. (1998) A structural switch in a mutant insulin exposes key residues for receptor binding, *J. Mol. Biol.* 279, 1–7.
66. Kobayashi, M., Ohgaku, S., Iwasaki, M., Maegawa, H., Shigeta, Y., and Inouye, K. (1982) Supernormal insulin: [D-PheB24]-insulin with increased affinity for insulin receptors, *Biochem. Biophys. Res. Commun.* 107, 329–336.
67. 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.
68. Weiss, M. A., Nakagawa, S. H., Jia, W., Xu, B., Hua, Q. X., Chu, Y. C., Wang, R. Y., and Katsoyannis, P. G. (2002) Protein structure and the spandrels of San Marco: insulin's receptor-binding surface is buttressed by an invariant leucine essential for protein stability, *Biochemistry* 41, 809–819.
69. Schäffer, L. (1994) A model for insulin binding to the insulin receptor, *Eur. J. Biochem.* 221, 1127–1132.
70. Dawson, P. E., and Kent, S. B. H. (2000) Synthesis of native proteins by chemical ligation, *Annu. Rev. Biochem.* 69, 925–962.
71. Reidhaar-Olson, J. F., and Sauer, R. T. (1988) Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences, *Science* 241, 53–57.
72. Scott, J. K., and Smith, G. P. (1990) Searching for peptide ligands with an epitope library, *Science* 249, 386–390.
73. Kurapkat, G., De Wolf, E., Grotzinger, J., and Wollmer, A. (1997) Inactive conformation of an insulin despite its wild-type sequence, *Protein Sci.* 6, 580–587.
74. Chothia, C., Lesk, A. M., Dodson, G. G., and Hodgkin, D. C. (1983) Transmission of conformational change in insulin, *Nature* 302, 500–505.
75. Smith, G. D., and Ciszak, E. (1994) The structure of a complex of hexameric insulin and 4'-hydroxyacetanilide, *Proc. Natl. Acad. Sci. U.S.A.* 91, 8851–8855.
76. Brange, J., Dodson, G. G., Edwards, D. J., Holden, P. H., and Whittingham, J. L. (1997) A model of insulin fibrils derived from the X-ray crystal structure of a monomeric insulin (*des*-pentapeptide insulin), *Proteins* 27, 507–516.
77. Zhao, M., Nakagawa, S. H., Hua, Q. X., and Weiss, M. A. (1998) Exploring the foldability and function of insulin by combinatorial peptide chemistry, in *Proceedings of the American Peptide Symposium*, pp 369–371.
78. Qiao, Z. S., Guo, Z. Y., and Feng, Y. M. (2001) Putative disulfide-forming pathway of porcine insulin precursor during its refolding in vitro, *Biochemistry* 40, 2662–2668.
79. Qiao, Z. S., Min, C. Y., Hua, Q. X., Weiss, M. A., and Feng, Y. M. (2003) In vitro refolding of human proinsulin. Kinetic intermediates, putative disulfide-forming pathway, folding initiation site, and potential role of C-peptide in folding process, *J. Biol. Chem.* 278, 17800–17809.
80. Markussen, J. (1985) Comparative reduction/oxidation studies with single chain *des*-(B30) insulin and porcine proinsulin, *Int. J. Pept. Protein Res.* 25, 431–434.
81. Nanjo, K., Sanke, T., Miyano, M., Okai, K., Sowa, R., Kondo, M., Nishimura, S., Iwo, K., Miyamura, K., Given, B. D. 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, 514–519.
82. Wan, Z., Huang, K., Xu, B., Hu, S. Q., Wang, S., Chu, Y. C., Katsoyannis, P. G., and Weiss, M. A. (2005) Diabetes-associated mutations in human insulin: crystal structure and photocross-linking studies of A chain variant insulin *Wakayama*, *Biochemistry* (in press).
83. Steiner, D. F., Tager, H. S., Chan, S. J., Nanjo, K., Sanke, T., and Rubenstein, A. H. (1990) Lessons learned from molecular biology of insulin—gene mutations, *Diabetes Care* 13, 600–609.
84. Karle, I. L., Gopi, H. N., and Balam, P. (2003) Crystal structure of hydrophobic 19-residue peptide helix containing three centrally located D amino acids, *Proc. Natl. Acad. Sci. U.S.A.* 100, 13946–13951.
85. Wan, Z., Xu, B., Huang, K., Chu, Y., Li, B., Nakagawa, S., Qu, Y., Hu, S., Katsoyannis, P. G., and Weiss, M. A. (2004) Enhancing the activity of insulin at the receptor interface: Crystal structure and photo-cross-linking of A8 analogues, *Biochemistry* 43, 16119–16133.
86. Zhang, H., Tang, Y. H., and Feng, Y. M. (2000) Possible role of B8Gly in insulin structural motif, *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* 32, 480–484.

BI0480250