

Protein Structure and the Spandrels of San Marco: Insulin's Receptor-Binding Surface Is Buttressed by an Invariant Leucine Essential for Its Stability[†]

Michael A. Weiss,^{*,‡} Satoe H. Nakagawa,[§] Wenhua Jia,[‡] Bin Xu,[‡] Qing-xin Hua,[‡] Ying-Chi Chu,^{||} Run-ying Wang,^{||} and Panayotis G. Katsoyannis^{||}

Department of Biochemistry, Case Western Reserve School of Medicine, Cleveland, Ohio 44106-4935, Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, Illinois 60637, and Department of Pharmacology and Biological Chemistry, Mount Sinai School of Medicine of New York University, New York, New York 10029

Received September 25, 2001; Revised Manuscript Received November 17, 2001

ABSTRACT: Insulin provides a model of induced fit in macromolecular recognition: the hormone's conserved core is proposed to contribute to a novel receptor-binding surface. The core's evolutionary invariance, unusual among globular proteins, presumably reflects intertwined constraints of structure and function. To probe the architectural basis of such invariance, we have investigated hydrophobic substitutions of a key internal side chain (Leu^{A16}). Although the variants exhibit perturbed structure and stability, moderate receptor-binding activities are retained. These observations suggest that the A16 side chain provides an essential structural buttress but unlike neighboring core side chains, does not itself contact the receptor. Among invertebrate insulin-like proteins, Leu^{A16} and other putative core residues are not conserved, suggesting that the vertebrate packing scheme is not a general requirement of an insulin-like fold. We propose that conservation of Leu^{A16} among vertebrate insulins and insulin-like growth factors is a side consequence of induced fit: alternative packing schemes are disallowed by lack of surrounding covariation within the hormone's hidden receptor-binding surface. An analogy is suggested between Leu^{A16} and the spandrels of San Marco, tapering triangular spaces at the intersection of the dome's arches. This celebrated metaphor of Gould and Lewontin emphasizes the role of interlocking constraints in the evolution of biological structures.

Insulin is a globular protein containing two chains, designated A (21 residues) and B (30 residues). Stored in the β cell as a Zn²⁺-stabilized hexamer, the hormone functions as a Zn²⁺-free monomer (1). The structure of a monomer in solution (2, 3) resembles the crystallographic T state (Figure 1A) as defined in diverse crystal forms (4–9). Despite its small size, insulin contains representative features of larger proteins, including canonical secondary structure and well-ordered hydrophobic core. The core, invariant among vertebrate insulins and insulin-like growth factors (8, 10), is stabilized by three disulfide bridges and close packing of aliphatic and aromatic side chains. The present study focuses on the role of a key internal side chain, Leu^{A16}, which anchors the A-chain's C-terminal α -helix (Figure 1C). Hydrophobic substitutions are introduced by chemical synthesis (11, 12). Although extensive mutagenesis studies of insulin have been conducted¹ (8), to our knowledge, substitutions at A16 have not previously been described. Our results demonstrate that Leu^{A16} provides an

essential stabilizing element but is itself unlikely to contact the insulin receptor.

Anomalies encountered in studies of structure–function relationships suggest that insulin undergoes a change in conformation on receptor binding. Two findings are of particular interest. (i) An analogue with enhanced activity (obtained by inversion of residues Thr^{B27} and Pro^{B28} in the B-chain's C-terminal β -strand; 13) exhibits unstable secondary structure in solution. Although a compact nativelike fold is retained, its α -helices exhibit attenuated ($i, i + 3$) nuclear Overhauser effects and accelerated amide proton exchange. These observations demonstrate a nonlocal coupling between the sites of mutation in the B-chain β -strand and the hormone's α -helical core. Further, that destabilizing insulin enhances its activity² suggests that native insulin incurs a hidden thermodynamic penalty on receptor binding. (ii) Conversely, an inactive single-chain analogue, mini-proinsulin, exhibits an essentially native crystal structure (14).

[†] This work was supported in part by grants from the National Institutes of Health to M.A.W. (DK40949) and P.G.K. (DK12925) and by the Diabetes Research and Training Center at The University of Chicago (S.H.N.). We dedicate this paper to the memory of Prof. Edward M. Purcell.

* To whom correspondence should be addressed: e-mail, weiss@biochemistry.cwru.edu; telephone, 216-368-5991; fax, 216-368-3419.

[‡] Case Western Reserve School of Medicine.

[§] The University of Chicago.

^{||} Mount Sinai School of Medicine of New York University.

¹ Systematic alanine scanning mutagenesis of insulin has been described (45) in the course of which an Ala^{A16} variant was attempted. Because the substitution hindered biosynthetic expression of the single-chain insulin precursor in yeast, however, no product was obtained. Remarkably, of the 21 alanine mutations constructed, only 2 sites (Val^{B12} and Ala^{A16}) were without detectable yield. The mechanism by which Ala^{A16} blocks biosynthetic expression was not characterized.

² In light of a previously proposed correlation between insulin's stability and activity (75), it is noteworthy that the Val^{A16} and Phe^{A16} analogues exhibit similar activities despite their differing stabilities. In a recent study of A8 analogues no correlation was found between stability and activity (16).

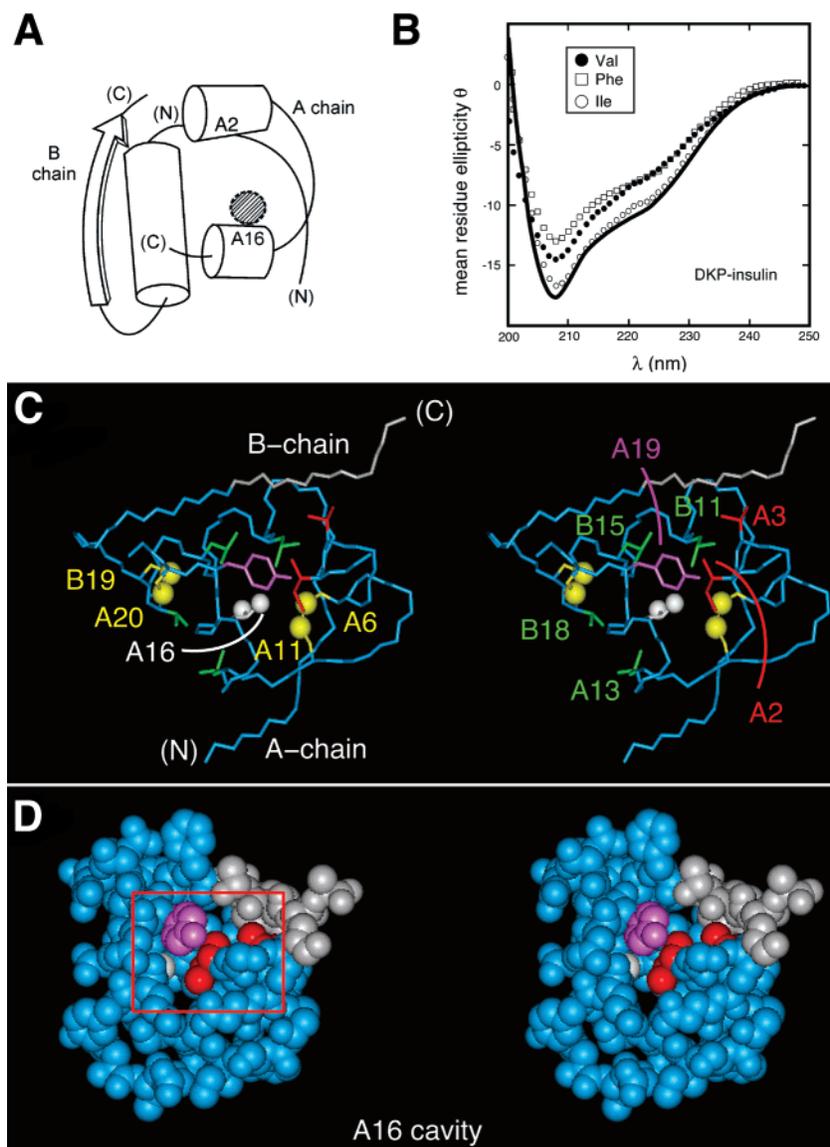


FIGURE 1: Structural overview and CD spectra. (A) Cylinder model of the native insulin T-state closed conformation. Positions of residues A2 and A16 are indicated. The shaded circle represents the cavity between chains filled by Leu^{A16}. (B) Structural perturbations indicated by far-UV CD spectra of analogues relative to the spectrum of DKP-insulin (solid line): Val^{A16} (●), Phe^{A16} (□), and Ile^{A16} (○). (C) Crystal structure of human insulin (2 Zn molecule 1; PDB accession number 4INS) highlighting the local structure in the neighborhood of Leu^{A16}. The A chain is shown in blue and residues B26–B30 (deleted in DPI; see Figure 2 and refs 24, 26, and 78) are in gray. Labeled residues are color-coded as follows: A16 (white; methyl groups are indicated by white balls), A2 and A3 (red), disulfide bridges (yellow; cysteines depicted as yellow balls), A19 (purple), and other core residues (green; A13, B11, B15, and B18). (D) Stereo representation of the hypothetical A16-associated cavity obtained on deletion of the Leu^{A16} side chain in the crystallographic protomer. The predicted cavity is 130 Å³ and is encaged by nonpolar side chains. The color scheme follows that of panel C.

Mini-proinsulin contains a peptide bond between Lys^{B29} and Gly^{A1}, which constrains the relative orientation of the B-chain's C-terminal β -strand and α -helical core. Together, these findings demonstrate a lack of correlation between insulin's classical structure and function (15, 16). Insulin's active conformation (Figure 2A) is unknown.

We and others have proposed that the B-chain's C-terminal β -strand (shown in gray in Figure 2B,C) detaches on receptor binding (14, 15, 17). This model rationalizes effects of C-terminal B-chain substitutions and deletions on receptor binding (15, 18–21). Effects of detachment of the B-chain β -strand on insulin's structure and surface have been investigated through crystallographic studies of the truncated insulin analogue (*des*-pentapeptide[B26–B30]-insulin; DPI).³ The truncated segment is shown in gray in Figures 1D, 2B,

and 2C. When amidated at B25, DPI exhibits wild-type activity (22, 23). The crystal structure of DPI (24–26) exhibits a native α -helical moiety; the absence of residues B26–B30 leads to significant exposure of core side chains Ile^{A2} and Val^{A3}. The proposed detachment of the B-chain β -strand on receptor binding would likewise expose Ile^{A2} and Val^{A3} as part of a putative “hidden” functional surface (Figure 2D). Binding of this surface is suggested by the

³ Abbreviations: CD, circular dichroism; DKP-insulin, monomeric insulin analogue containing three substitutions in the B chain (Asp^{B10}, Lys^{B28}, and Pro^{B29}); DOI, truncated insulin analogue lacking residues B23–B30; DPI, truncated insulin analogue lacking residues B26–B30; IGF-I, insulin-like growth factor I; NMR, nuclear magnetic resonance, rp-HPLC, reverse-phase high-performance liquid chromatography; UV, ultraviolet. Amino acids are designated by standard three-letter code.

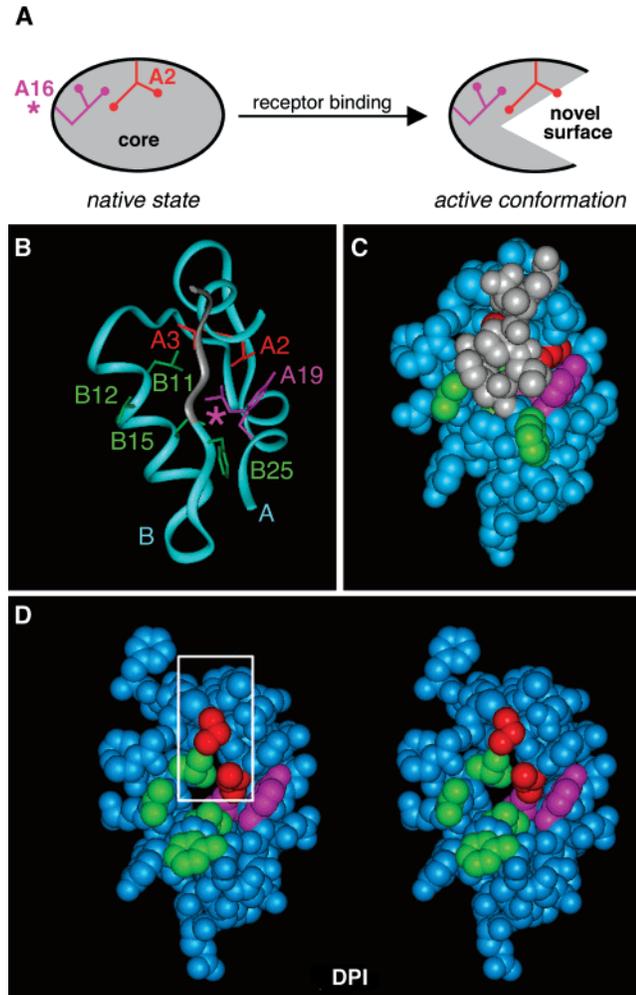


FIGURE 2: Insulin undergoes a change in conformation on receptor binding. (A) Schematic model of insulin's native state (left) and active conformation (right). On receptor binding a hidden functional surface is proposed to engage the receptor. This surface contains Ile^{A2} (red) and Val^{A3} (not shown); the present study suggests that Leu^{A16} (magenta, asterisk at left) provides a structural buttress but does not contact the receptor. (B and C) Ribbon and space-filling representations of the T-state crystallographic protomer (2 Zn molecule 1; PDB accession number 4INS) highlighting the closed position of B-chain's C-terminal segment (gray; residues B26–B30). Residues are color coded as follows: A2 and A3, red; A19, magenta; and relevant B-chain residues (B11, B12, B15, and B25), green. (D) Crystal structure of DPI lacking residues B26–B30 (PDB identifier 1PID; 24). Detachment of these residues exposes side chains of Ile^{A2} and Val^{A3} as part of hormone's proposed hidden functional surface. When amidated at B25, DPI exhibits wild-type affinity for the insulin receptor (22, 23). The color scheme is identical to that of panels B and C.

exquisite sensitivity of insulin's activity to aliphatic substitutions at A2 and A3 (Table 1; 27, 28); such effects are otherwise unexpected of conservative substitutions within a hydrophobic core (8, 15). Because additional conformational change may occur elsewhere in the hormone, this model may underestimate the actual extent of induced fit⁴ (6).

In this article we investigate the contribution of a key internal side chain, Leu^{A16}, to insulin's structure, stability, and activity. Leu^{A16}, invariant among vertebrate insulins and insulin-like growth factors (arrow in Figure 3; 8, 10), occupies a distinctive cavity between A and B chains (Figure 1D). Three nonpolar substitutions (Val, Ile, and Phe) are introduced as potential probes of insulin's hidden functional

Table 1: Prior Studies of A-Chain Analogues

analogue	activity (%)	analogue	activity (%)
(A) Residue A2 (Ile) ^a			
Ala	0.8, 0.4 ^b	Ail	4.5
Leu	5.9	Nle	1.3
(B) Residue A3 (Val)			
Ala	2.4	Leu	0.2
Ile	14.5	Nle	0.5
(C) Residue A19 (Tyr)			
Ala	0.1 ^c	Phe	13.3 ^d
Leu	0.05 ^e	<i>p</i> -F-Phe	60 ^f

^a Values (unless otherwise indicated) were obtained from ref 28 and corrected for the N^αB¹, N^εB²⁹ A-chain modification described therein. Activity was determined by receptor-binding assay using canine hepatocytes. ^b Value was obtained from ref 27. Receptor binding was assayed using rat liver plasma membranes. ^c Value was obtained from ref 45. Receptor-binding assays employed recombinant soluble human insulin receptor fusion protein. ^d Value was obtained from ref 58. Activity was determined by receptor-binding assay using human placental membrane insulin receptors as described (73). ^e Value was obtained from ref 56 using rat liver plasma membrane assays as described (27, 74). ^f Value was obtained from ref 57 by lipogenesis using rat adipocytes.

	1	*	4	8		↓	21														
human, pig	G	I	V	E	Q	C	C	T	S	I	C	S	L	Y	Q	L	E	N	Y	C	N
cattle	G	I	V	E	Q	C	C	A	S	V	C	S	L	Y	Q	L	E	N	Y	C	N
sheep	G	I	V	E	Q	C	C	A	G	V	C	S	L	Y	Q	L	E	N	Y	C	N
horse	G	I	V	E	Q	C	C	T	G	I	C	S	L	Y	Q	L	E	N	Y	C	N
whale	G	I	V	E	Q	C	C	A	S	T	C	S	L	Y	Q	L	E	N	Y	C	N
rat	G	I	V	D	Q	C	C	T	S	I	C	S	L	Y	Q	L	E	N	Y	C	N
elephant	G	I	V	E	Q	C	C	T	G	V	C	S	L	Y	Q	L	E	N	Y	C	N
guinea pig	G	I	V	D	Q	C	C	T	G	T	C	S	R	H	Q	L	Q	S	Y	C	N
Coypu hystricomorphs	G	I	V	D	Q	C	C	T	N	I	C	S	R	N	Q	L	M	S	Y	C	N
chinchilla	G	I	V	D	Q	C	C	T	S	I	C	T	L	Y	Q	L	E	N	Y	C	N
cod	G	I	V	D	Q	C	C	H	R	P	C	D	I	F	D	L	Q	N	Y	C	N
angler	G	I	V	E	Q	C	C	H	R	P	C	N	I	F	D	L	Q	N	Y	C	N
tuna	G	I	V	E	Q	C	C	H	K	P	C	N	I	F	Q	L	Q	N	Y	C	N
toadfish 2	G	I	V	E	Q	C	C	H	R	P	C	D	K	F	Q	L	Q	S	Y	C	N
toadfish 1	G	I	V	E	Q	C	C	H	R	P	C	D	I	F	D	L	Q	S	Y	C	N
bonito	G	I	H	E	Q	C	C	H	K	P	C	D	I	F	Q	L	E	N	Y	C	N
hagfish	G	I	V	E	Q	C	C	H	K	R	C	S	I	Y	N	L	Q	N	Y	C	N
silver carp	G	I	V	E	Q	C	C	H	K	P	C	S	I	F	E	L	Q	N	Y	C	N
turkey	G	I	V	E	Q	C	C	H	N	T	C	S	L	Y	Q	L	E	N	Y	C	N
duck, goose	G	I	V	E	Q	C	C	E	N	P	C	S	L	Y	Q	L	E	N	Y	C	N
snake	G	I	V	E	Q	C	C	E	N	T	C	S	L	Y	Q	L	E	N	Y	C	N
squirrel monkey	G	V	V	D	Q	C	C	T	S	I	C	S	L	Y	Q	L	E	N	Y	C	N
owl monkey	G	V	V	D	Q	C	C	T	S	I	C	S	Y	E	Q	L	Q	N	Y	C	N
IGF1	G	I	V	D	E	E	C	F	R	S	C	D	L	R	R	L	E	M	Y	C	A
IGF2	G	I	V	E	E	E	C	F	R	S	C	D	L	A	L	L	E	T	Y	C	A

FIGURE 3: Invariance of Ile^{A2}, Val^{A3}, and Leu^{A16}. Sequences of insulin A chains (upper group) and corresponding A domains of human insulin-like growth factors I and II (lower group). Residues A2, A3, and A16 are highlighted by large boxes (asterisk and arrow, respectively). In the owl monkey and squirrel monkey valine occurs at A2; the substitution Ile^{A2} → Val in human insulin causes a 3-fold decrease in receptor binding (28). The activity of owl monkey insulin is only about 20% of that of human insulin (79). Cysteines involved in disulfide bridges (A6–A11, A7–B7, and A20–B19) are shown in boldface.

surface. Does Leu^{A16}, like other core side chains, contribute to insulin's hidden functional surface, i.e., function to both stabilize insulin's structure and contact the receptor? We reason as follows. Like Ile^{A2}, Leu^{A16} is buried in native insulin but unlike Ile^{A2}, the A16 side chain is not exposed in DPI (Figure 2D). Thus, any contact between Leu^{A16} and the

⁴ Insulin's active conformation may in part recapitulate aspects of the crystallographic T → R transition, an allosteric reorganization of insulin hexamers (5, 6). This transition is remarkable for a change in the secondary structure of the B1–B8 segment from extended (T state) to α -helix (R state). Segmental reorganization is coupled to a change in configuration 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. The side chain of Leu^{A16} would remain inaccessible to solvent in an isolated R-state protomer.

receptor, should it occur as a consequence of induced fit, would require a more extensive structural reorganization than detachment of the B-chain's C-terminal β -strand. A16 substitutions may thus be regarded as probes of insulin's hidden functional surface (asterisk in Figure 2A). Experimental design highlights the unusual features of A16 analogues by comparison with prior studies of A2 and A3 substitutions by this and other laboratories (27–31).

Synthesis of A16 analogues is impeded by inefficient disulfide pairing. Once isolated, the analogues exhibit perturbed structures and very low thermodynamic stabilities. The extent of instability is greater than that observed in studies of analogues containing substitutions or deletions elsewhere in the protein. Surprisingly, the unstable A16 analogues retain substantial receptor-binding activity. Together, these results suggest that whereas Leu^{A16} is essential for protein stability, the A16 side chain does not itself contact the receptor. Interestingly, Leu^{A16} is not conserved among invertebrate insulin-like polypeptides (32), suggesting that its packing interactions are not a general requirement of an insulin-like fold. We propose that the strict conservation of Leu^{A16} among vertebrate insulins (10) is enjoined by protein stability as a *side consequence* of induced fit: lack of covariation at core sites of hormone–receptor contact has precluded exploration of alternative packing schemes. Such a derived architectural feature is reminiscent of an evolutionary metaphor, the spandrels of the great dome of San Marco in Venice (33, 34). Identification of protein “spandrels” may illuminate hidden architectural constraints in protein function and design.

EXPERIMENTAL PROCEDURES

Materials. Human insulin, [Lys^{B28},Pro^{B29}]-insulin, and Asp^{B10}-des-octapeptide[B23–B30]-insulin (Asp^{B10}-DOI) were kindly provided by Eli Lilly and Co. (Indianapolis, IN). 4-Methylbenzhydramine resin (0.63 mmol of amine/g; Star Biochemicals, Inc.) was otherwise used as a solid support for synthesis of A-chain analogues; (*N*-butoxycarbonyl-*O*-benzyl)threonine phenylacetamidomethyl (PAM) resin (0.6 mmol/g; Bachem, Inc.) was used as a solid support for synthesis of the DKP B-chain. *tert*-Butoxycarbonyl amino acids and derivatives were obtained from Bachem Inc. Amino acid analyses of synthetic chains and insulin analogues were performed after acid hydrolysis with a Hewlett-Packard Amino Quant Analyzer (model 1090). Chromatography resins CM52 and DE52 cellulose (Whatman) and Cellex E (Ecteola cellulose; Sigma) were used.

Peptide Synthesis. The protocol for solid-phase synthesis is as described (35). The C-terminal Asn in synthesis of A chains was incorporated into a solid support by coupling *tert*-butoxycarbonyl aspartic acid α -benzyl ester with 4-methylbenzhydramine resin. After the final deprotection, the Asp residue was converted to an Asn residue.

(i) **Synthesis of A-Chain S-Sulfonate Analogues.** [Ile^{A16}]-peptidyl resin (0.564 g), after deblocking and sulfitolysis, produced 172 mg of crude S-sulfonate, which on chromatography on a Cellex E column afforded 96.4 mg of purified [Ile^{A16}]-S-sulfonate. Similarly, [Val^{A16}]-peptidyl resin (0.567 g) and [Phe^{A16}]-peptidyl resin (0.86 g) yielded 116 mg of purified [Val^{A16}]-S-sulfonate and 95 mg of purified [Phe^{A16}]-S-sulfonate.

(ii) **Synthetic B-Chain S-Sulfonate.** After deblocking, sulfitolysis, and chromatographic purification 610 mg of peptidyl resin yielded ca. 125 mg of purified S-sulfonated B-chain. Amino acid analyses were in agreement with expected values.

Synthesis of Insulin Analogues. Chain combination (36) was effected by interaction of the S-sulfonated derivative of the A chain (40 mg) and the [B-DKP] chain (20 mg) in 0.1 M glycine buffer (pH 10.6, 10 mL) in the presence of dithiothreitol (7.2 mg) (2, 12). CM52 cellulose chromatography of the combination mixture enabled partial isolation of the hydrochloride form of the analogue contaminated by free B chain. Final purification was accomplished by reverse-phase high-performance liquid chromatography (rp-HPLC). For Ile^{A16}-DKP-insulin, 4 mg of crude hydrochloride was isolated, which after rp-HPLC yielded 1.5 mg of highly purified analogue. Similarly, HPLC purification yielded 0.37 mg of Val^{A16}-DKP-insulin and 0.20 mg of Phe^{A16}-DKP-insulin. Control synthesis of DKP-insulin yielded 8.0 mg of purified analogue; corresponding synthesis of human insulin yielded 7.2 mg. The analogues' predicted molecular masses were in each case verified by electrospray mass spectrometry.

Biological Assays. Receptor-binding studies were performed using a human placental membrane preparation as described (16, 37, 38). Relative activity is defined as the ratio of analogue to human insulin required to displace 50% of specifically bound ¹²⁵I-human insulin (purchased from Amersham).

Spectroscopy. Far-ultraviolet (UV) CD spectra were obtained using an Aviv spectropolarimeter equipped with an automated titration unit for guanidine denaturation studies. CD samples for wavelength spectra contained 25–50 μ M insulin analogue in 10 mM potassium phosphate (pH 7) and 50 mM KCl; samples were diluted to 5 μ M for equilibrium denaturation studies.

Thermodynamic Modeling. Guanidine denaturation data were fitted by nonlinear least squares to a two-state model as described (39). In brief, CD data $\theta(x)$, where x indicates the concentration of denaturant, were fitted by a nonlinear least-squares program according to

$$\theta(x) = \frac{\theta_A + \theta_B \exp[(-\Delta G^\circ_{\text{H}_2\text{O}} - mx)/RT]}{1 + \exp[(-\Delta G^\circ_{\text{H}_2\text{O}} - mx)/RT]}$$

where x is the concentration of guanidine hydrochloride and where θ_A and θ_B are baseline values in the native and unfolded states. These baselines were approximated by pre- and posttransition lines $\theta_A(x) = \theta_A^{\text{H}_2\text{O}} + m_A x$ and $\theta_B(x) = \theta_B^{\text{H}_2\text{O}} + m_B x$. Fitting the original CD data and baselines simultaneously circumvents artifacts associated with linear plots of ΔG as a function of denaturant according to $\Delta G^\circ(x) = \Delta G^\circ_{\text{H}_2\text{O}} + m^\circ x$ (for reviews see refs 39 and 40).

Molecular Modeling. Rigid-body models were built using the InsightII package (Biosym, Inc., San Diego, CA). The volume and location of inferred protein cavities were obtained using the program SURFNET (41).

RESULTS

Analogues were obtained by chemical synthesis in the context of DKP-insulin, an engineered monomer containing substitutions in the B-chain's dimer contact (Pro^{B28} \rightarrow Lys

and Lys^{B29} → Pro) and trimer interface (His^{B10} → Asp) (42). DKP-insulin is similar in structure to native insulin (2) and ca. twice as active (21, 42). Use of the DKP B chain in studies of A-chain analogues circumvents possible confounding effects of self-association (21). Whereas chain combination of the DKP B chain and the wild-type A chain proceeds with native efficiency to effect specific disulfide pairing, chain combination of A16 analogues is impaired by at least 5-fold.⁵ By contrast, native yields were previously observed in syntheses of analogues containing hydrophobic substitutions in the A-chain's N-terminal α -helix (2, 27, 30, 31, 43).

A16 Substitutions Perturb Structure and Stability. Ile^{A16}-DKP-insulin, Val^{A16}-DKP-insulin, and Phe^{A16}-DKP-insulin exhibit altered far-UV CD spectra relative to that of DKP-insulin (Figure 1B). The perturbation is greatest in Phe^{A16}-DKP-insulin and least in Ile^{A16}-DKP-insulin. Spectra of Phe^{A16} and Val^{A16} analogues each exhibit markedly reduced ellipticities at helix-sensitive wavelength 222 nm ($[\theta]_{222}$), consistent with loss of two turns of α -helix.⁶ These changes are associated with lower thermodynamic stabilities in studies of protein denaturation (Figure 4A). Analysis of $[\theta]_{222}$ as a function of denaturant concentration (guanidine hydrochloride) by a two-state model of protein denaturation (39) yields decrements ($\Delta\Delta G_u$) of 1.4 kcal/mol (Ile^{A16}-DKP-insulin), 2.7 kcal/mol (Val^{A16}-DKP-insulin), and 3.4 kcal/mol (Phe^{A16}-DKP-insulin).⁷ Thermodynamic parameters are given in Table 2.⁸ Remarkably, stabilities of the A16 analogues are lower than that of *des*-octapeptide(B23–B30)-Asp^{B10}-insulin (ΔG_u 4.1 kcal/mol), a truncated analogue of very low receptor-binding affinity lacking the B-chain's β -turn and C-terminal β -strand (44). A core analogue of low biological activity, *allo*-Ile^{A2}-DKP-insulin, exhibits native stability (Table 2; 31).

A16 Analogues Retain Substantial Activity. The affinity of the three A16 analogues for the human placental insulin receptor was measured by displacement of ¹²⁵I-labeled human

insulin. Surprisingly, the analogues retain substantial activity (30–65% relative to human insulin; Table 2). These values correspond to activities of ca. 40% (Ile^{A16} analogue) and ca. 20% (Val^{A16} and Phe^{A16} analogues) relative to parent DKP-insulin. Changes in free energy of binding (ΔG_b) are thus smaller (0.5 and 0.8 kcal/mol, respectively) than decrements in protein stability. Activities of A16 analogues are significantly higher than those observed in studies of corresponding substitutions at A2 or A3⁹ (such as Leu^{A2} or Leu^{A3}; see Table 1).

DISCUSSION

An overview of structure–activity relationships in insulin has been obtained by alanine scanning mutagenesis (45). This influential study employed biosynthesis of a single-chain precursor in a yeast expression system (46). Unfortunately, an Ala^{A16} analogue could not be expressed, presumably due to misfolding and/or intracellular proteolysis. Because Leu^{A16} occupies a site of special structural interest (8), we sought to obtain A16 analogues by synthetic methods (47, 48). To avoid forming a large and potentially destabilizing cavity (49) in the hydrophobic core (Figure 1D), we chose nonpolar side chains larger than alanine: Val^{A16}, Ile^{A16}, and Phe^{A16}. Although synthetic yields were limited by inefficient disulfide pairing, sufficient material was obtained to evaluate effects of these substitutions on CD-detected structure, stability, and activity.

The Size and Shape of the A16 Side Chain Are Critical to Structure and Stability. CD studies demonstrate that the size and shape of the A16 side chain make a significant contribution to insulin's structure and stability. Detailed analysis of structural perturbations [such as site(s) of main-chain distortion] will require studies by X-ray crystallography or NMR spectroscopy. Although low synthetic yields precluded such analysis here, preliminary insight may be obtained from models derived from the crystal structure of native insulin (Figure 5) (8). We consider each analogue in turn.

(i) **Ile^{A16}-DKP-insulin.** Of the three substitutions investigated, Ile^{A16} is the least perturbing. This seems reasonable since isoleucine is the most similar to leucine in size. Among mammalian species of relaxin, an insulin-like polypeptide (50), both leucine and isoleucine are observed at the A16 position (Supporting Information). Although the α -helix content of Ile^{A16}-DKP-insulin is similar to that of DKP-insulin (Figure 1B), the analogue's lower free energy indicates that packing of core side chains is less favorable. Molecular modeling suggests that, within a native structure of insulin, Ile^{A16} would be associated with a small cavity

⁵ The apparent instability of A16 analogues is greater than can be rationalized solely on the basis of the variant side chains' lower α -helical propensities (80) and tabulated hydrophobic transfer free energies (81). Such estimates, which do not consider specific features of protein structure (such as creation of destabilizing cavities), predict in each case a decrement in stability of 0.6–0.8 kcal/mol (Supporting Information).

⁶ Attenuation of helical CD bands can reflect perturbations in either structure or dynamics, i.e., enhanced fluctuations from mean helical geometry. Sites of perturbation may be segmental or distributed.

⁷ DKP-insulin is more stable than human insulin ($\Delta\Delta G_u$ 0.5 kcal/mol; Table 2) (16, 75, 82). Its enhanced stability is due to the His^{B10} → Asp substitution ($\Delta\Delta G_u$ 1.0 kcal/mol), whose stabilizing effect is in part offset by the B28–B29 switch (cf. lower stability of [Lys^{B28}, Pro^{B29}]insulin; $\Delta\Delta G_u$ –0.9 kcal/mol in Table 2).

⁸ The analogues' fitted m values are in each case lower than that of the parent unfolding curve (DKP-insulin; see Table 2). These lower m values are likely to reflect greater exposed hydrophobic surface in the absence of denaturant and/or existence of a native state ensemble containing a distribution of incompletely folded forms of differing stability. In this setting, the two-state formalism (39) may overestimate thermodynamic decrements (40). Conservative lower bounds to $\Delta\Delta G_u$ are obtained by multiplying its C_{mid} value (or $\langle C_{mid} \rangle$ if obtained as an average over an ensemble of distinct partial folds) by the parent m value (83). Respective lower bounds for Ile^{A16}, Val^{A16}, and Phe^{A16} analogues are thus 0.4, 1.7, and 2.3 kcal/mol, respectively. Global fitting of the variant unfolding transitions using the native m value leads to large errors. Because reductions in m value parallel reductions in helix content, we suggest that the m value changes are meaningful, i.e., reflect less efficient desolvation of hydrophobic side chains in the variant folded states.

⁹ Studies of engineered cavities in T4 phage lysozyme have suggested that cavities are associated on average with a free energy penalty of 0.024–0.033 kcal/Å³ (49); this relationship implies that in the absence of structural adjustments Val^{A16}-DKP-insulin would exhibit a cavity-related $\Delta\Delta G_u$ of 0.5–0.7 kcal/mol in addition to the ca. 0.8 kcal/mol penalty due to valine's lower helical propensity (80) and hydrophobicity (81). The sum of these effects, 1.3–1.5 kcal/mol, is less than the observed value of 2.7 kcal/mol. We imagine that the actual cavity size in Val^{A16}-DKP-insulin is less than 20 Å³ due to structural reorganization, including a loss of α -helix. Such reasoning omits consideration of changes in protein dynamics and solvation, which may provide entropy–enthalpy compensation (84). Large deviations from the lysozyme cavity–free energy relationship have been observed in other systems (51).

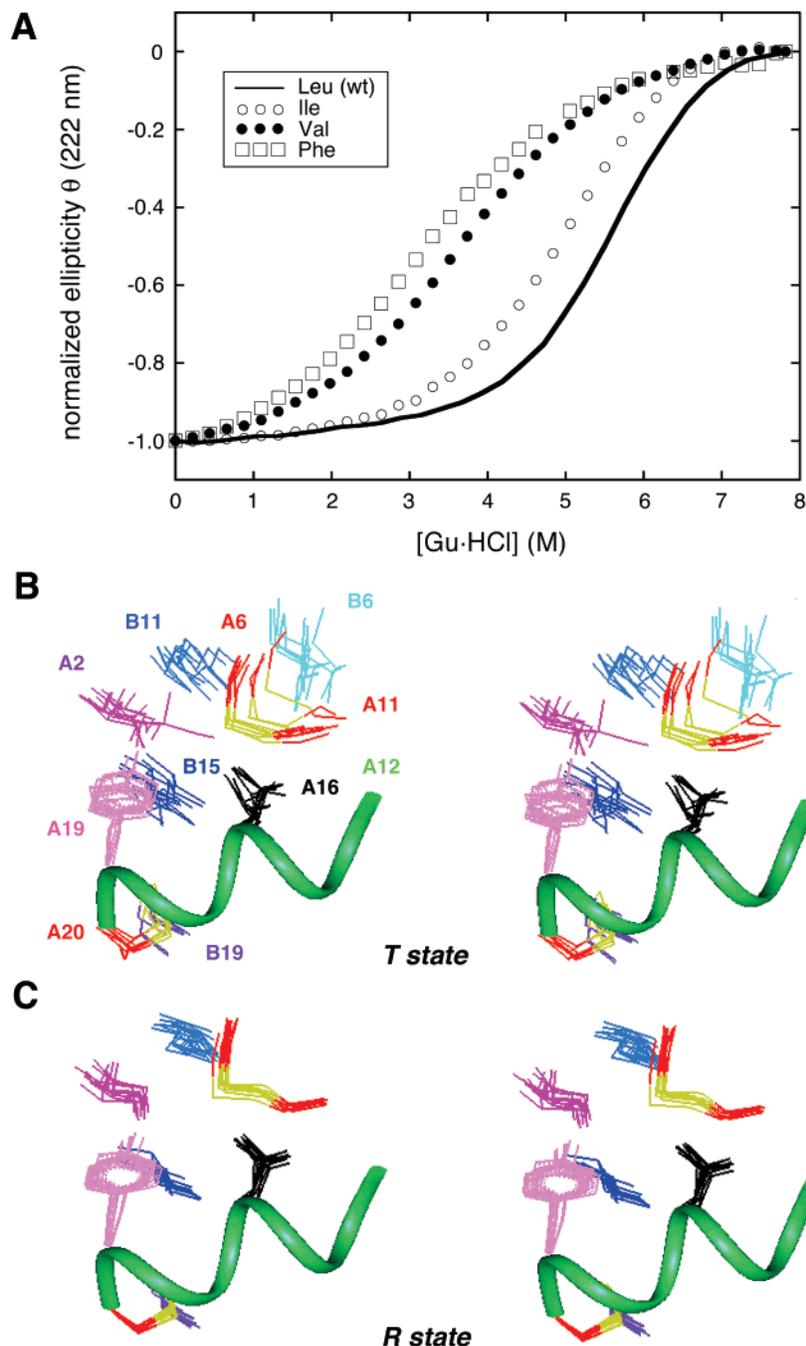


FIGURE 4: A16 substitutions impair thermodynamic stability. Guanidine unfolding transitions of DKP-insulin (Leu^{A16}; solid line) and A16 analogues: Val^{A16} (●), Phe^{A16} (□), and Ile^{A16} (○). Results of fitting to a two-state model of protein denaturation are given in Table 2. (B and C) Collection of crystal structures of the side chains relative to the A16 side chain. Structures are aligned according to main-chain atoms of residues A12–A20 (green ribbon). Side chains are color coded as follows: A16 (black), sulfur atoms (yellow), and residues involved in disulfide bonding (A6, A11, and A20; red). Other A-chain residues are shown in shades of aubergine: A19 (light purple) and A2 (purple). B-chain residues are shown in shades of blue: B6 (sky blue), B11 (blue), B15 (dark blue), and B19 (violet). T-state crystal structures (PDB identifiers 1APH, 1LPH, 1PID, 1TRZ, 1ZNI, and 4INS); R-state crystal structures (PDB identifiers 1EV3, 1TRZ, 1ZNI, 1ZNI, and 1XDA).

adjoining its γ -CH₂ and δ -CH₃ moiety (<10 Å³); the cavity is created by the absence of one of the Leu^{A16} methyl groups (central black sphere in Figure 5B). In addition, there is a potential steric clash between the γ' -CH₃ group and the side chain of B15 (intersecting green and blue spheres in Figure 5B).

(ii) *Val^{A16}-DKP-insulin*. Given the similarity between valine and isoleucine, Val^{A16}-DKP-insulin's more attenuated α -helix content and marked instability are surprising: evidently, a slightly larger packing defect at this site (ca. 20 Å³

in a rigid-body model) must be substantially destabilizing. The Val^{A16} analogue's instability strongly suggests that close packing of the A16 side chain is critical to insulin's structure and stability. Sites in other globular proteins that are unusually sensitive to packing defects have previously been described⁹ (51).

(iii) *Phe^{A16}-DKP-insulin*. Of the three analogues investigated, Phe^{A16}-DKP-insulin is the most perturbed. The aromatic ring of phenylalanine is least similar to leucine in size, shape, and functional character. Molecular modeling

Table 2: Properties of Insulin Analogues^a

analogue	activity	ΔG_u	$\Delta\Delta G_u$	C_{mid} (M)	m (kcal mol ⁻¹ M ⁻¹)
(A) engineered monomer					
human insulin	100	4.4 ± 0.1		5.3 ± 0.1	0.84 ± 0.01
[Lys ^{B29} ,Pro ^{B29}]HI	100	3.5 ± 0.1	-0.9 ± 0.2	4.9 ± 0.1	0.71 ± 0.01
Asp ^{B10} -HI	ND ^b	5.0 ± 0.1	0.6 ± 0.2	6.2 ± 0.1	0.80 ± 0.01
DKP-insulin	161 ± 19	4.9 ± 0.1		5.8 ± 0.1	0.84 ± 0.01
(B) A16 analogues					
Ile ^{A16} -DKP-insulin	64 ± 19	3.5 ± 0.1	-1.4 ± 0.2	5.4 ± 0.1	0.66 ± 0.01
Val ^{A16} -DKP-insulin	33 ± 8	2.1 ± 0.1	-2.8 ± 0.2	3.7 ± 0.1	0.57 ± 0.01
Phe ^{A16} -DKP-insulin	34 ± 11	1.5 ± 0.1	-3.4 ± 0.2	3.1 ± 0.1	0.50 ± 0.01
(C) control analogues					
Asp ^{B10} -DOI ^c	ND ^d	4.1 ± 0.1	-0.3 ± 0.2	6.4 ± 0.1	0.65 ± 0.01
<i>allo</i> -Ile ^{A2} -DKP-insulin ^e	4 ± 0.01	4.9 ± 0.1	0.0 ± 0.2	6.5 ± 0.2	0.75 ± 0.02

^a ΔG_u indicates the apparent change in free energy on denaturation in guanidine hydrochloride as extrapolated to zero denaturant concentration by a two-state model. $\Delta\Delta G_u$ indicates the difference in ΔG_u values relative to DKP-insulin. Uncertainties in two-state fitting parameters do not include possible systematic error due to possible non-two-state behavior. C_{mid} is defined as that concentration of guanidine hydrochloride at which 50% of the protein is unfolded. The m value provides the slope in plotting unfolding free energy ΔG_u ([Gu·HCl]) versus molar concentration of denaturant; this slope is proportional to the protein surface area exposed on unfolding. ^b Not determined in this study. Past values are reported to be 200% (21), 207% (75), and 435 ± 144% (76). ^c DOI denotes *des*-octapeptide[B23–B30]-insulin. ^d ND, not determined. However, past studies indicate that activities would be extremely low (<0.1%) (28, 77). ^e Values obtained from ref 31.

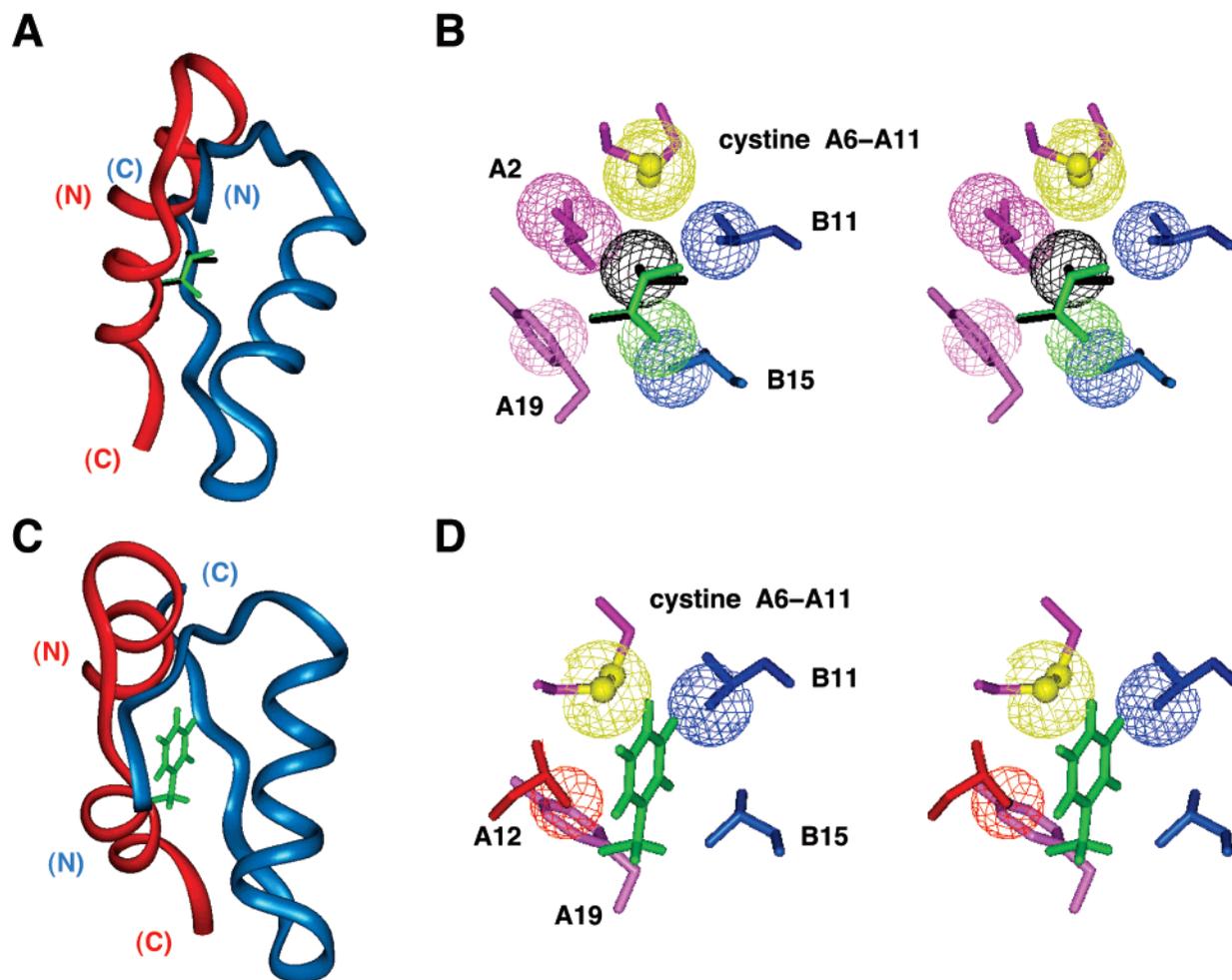


FIGURE 5: Rigid-body models of Ile^{A16}-insulin (A and B) and Phe^{A16}-insulin (C and D) based on the T₆ crystal structure (PDB identifier 4INS). Although unphysical, these models provide insight into possible packing defects in A16 analogues. (A and C) Ribbon models of overall orientation. The A chain is shown in red and the B chain in blue. (B and D) Packing of side chains. The Ile^{A16} model highlights the missing volume of the Leu^{A16} δ_2 -CH₃ group and steric clash between the γ' -CH₃ group and Leu^{B15}; the Phe^{A16} model illustrates steric clash between the aromatic ring and Leu^{B11}. Spheres indicate van de Waals radii of key atoms: A2 (C δ 1; C γ 1), A6 and A11 (sulfur atoms), A12 (oxygen; panel D), A16 (black, C δ 2; green, C γ 2), A19 (C δ 2), B11 (C δ 1), and B15 (C δ 2). Side chains are color-coded as follows: sulfur atoms (yellow), B-chain residues (B11 and B15; blue), A-chain residues (A2 and A19; purple), wild-type Leu^{A16} (black; panels A and B), variant A16 (green), and A12 (red; panel D).

suggests that within the structure of native insulin a bulky Phe^{A16} side chain would clash with residues A11, B5, B11,

and/or B15 (depending on its configuration). In the model shown the para proton of Phe^{A16} collides with the sulfur atom

of Cys^{A11} (yellow sphere in Figure 5), and one of its meta protons collides with a methyl group of Leu^{B11} (blue sphere).

The extent of thermodynamic perturbation is presumably ameliorated by structural changes (i.e., to reach the preferred ground state). The stringent requirement for leucine at A16 is surprising in light of the apparent flexibility in the environment of this site among different crystal forms. Comparison of T-state crystallographic protomers indicates significant variation in surrounding side-chain packing schemes (Figure 3B) and in the dihedral angles of Leu^{A16} (Supporting Information). R-state protomers by contrast exhibit uniform core structures (Figure 3B). Because the solution structure of DKP-insulin spans the range of T-state crystal structures (2), such variation must be insufficient to allow nativelylike packing of the variant side chains.

The attenuated helix content of Val^{A16} and Phe^{A16} analogues (Figure 1B) could in principle be due to distortion of any one (or more) of insulin's three α -helices (Figure 1A). The simplest possibility is local unraveling or instability of the A12–A18 α -helix itself. Perturbation of the central B-chain α -helix appears less likely as, of insulin's three helices, its sequence exhibits the highest helical propensity, and among crystal structures, this segment is the most highly ordered and regular in geometry (8). Further, because the A1–A8 α -helix is critical to the hormone's activity (16), the analogues' activities suggest that the A1–A8 recognition element is retained. This leaves by elimination local perturbation of the A12–A18 α -helix. Testing this proposal will require high-resolution structural studies.

Implications for Insulin's Hidden Functional Surface. The interface between insulin and the insulin receptor is proposed to contain the close apposition of nonpolar surfaces (8, 52–54). Insulin's functional surface is proposed to span Ile^{A2}, Val^{A3}, and Tyr^{A19} (8, 14, 55). Although the cocrystal structure of a hormone–receptor complex awaits determination, the activity of insulin is exquisitely sensitive to hydrophobic substitutions at these positions (Table 1; 27–29, 45, 56–58). Evidence for the direct participation of Ile^{A2} in receptor binding is provided by studies of *allo*-Ile^{A2}-DKP-insulin. The solution structure of this nonstandard analogue has recently been shown to resemble that of DKP-insulin (31), and yet its activity is reduced by 50-fold. Evidence for the direct participation of Val^{A3} is provided by studies of Leu^{A3}-insulin (insulin Wakayama; 59, 60), a mutant insulin associated with diabetes mellitus in humans (61). The activity of Leu^{A3}-insulin is essentially unperturbed (28); its ¹H NMR chemical shifts and solution structure in 20% acetic acid are similar to those of native insulin (Q. X. Hua, K. Inouye, and M. A. Weiss, unpublished observations). Evidence for a direct contact between Tyr^{A19} and the receptor is provided by the low activities of Phe^{A19} and Trp^{A19} analogues, which retain native immunoreactive structures (58, 62). Because activity is largely restored by *p*-fluoro-Phe^{A19}, a direct contact between the receptor and the *p*-OH moiety of Tyr^{A19} has been proposed (57). Steric complementarity between insulin's hidden functional surface and the insulin receptor would be expected to enforce conservation of these key side chains (10).

Given the marked effects of A2, A3, and A19 substitutions on insulin's activity (Table 1), it is surprising that misfolded A16 analogues retain such substantial receptor-binding activities. This functional tolerance suggests that (i) site(s)

of misfolding are not critical to receptor binding and (ii) Leu^{A16}, unlike Ile^{A2}, Val^{A3}, or Tyr^{A19}, does not contact the insulin receptor. Lack of engagement between the native A12–A18 α -helix and the receptor is supported by alanine scanning mutagenesis: high activities are retained following Ala substitutions at Ser^{A12}, Tyr^{A14}, Glu^{A15}, and Glu^{A17} (45). The moderate activities of A16 analogues contrast with their perturbed stabilities and have implications for models of insulin's active conformation. Unlike the side chains of Ile^{A2} and Val^{A3}, B-chain truncation in DPI does not expose Leu^{A16} (Figure 2D; 24, 25). The present results suggest that Leu^{A16} does not contribute to insulin's hidden functional surface, i.e., this conserved side chain remains buried in insulin's active conformation. This correlation does not exclude additional sites of conformational change, such as in the configuration of Gly^{B8} as observed in the hexameric T \rightarrow R transition (5, 6). Delineating sites of conformational change on receptor binding will require a cocrystal structure of the hormone–receptor complex.

Protein Spandrels and the Evolution of Insulin Sequences. The constrained environment of Leu^{A16} among vertebrate insulins and insulin-like growth factors is reminiscent of the spandrels of San Marco, an architectural feature of Renaissance engineering widely discussed as an evolutionary metaphor. The spandrel metaphor refers to the emergence of some biological structures as a side consequence of an adaptationist innovation (33). Spandrels designate the tapering triangular spaces at the intersection of the dome's arches. "The design is so elaborate, harmonious and purposeful," wrote Gould and Lewontin, "that we are tempted to view it as the starting point of any analysis, as the cause in some sense of the surrounding architecture." The triangular spaces are in fact not causes but byproducts of the dome's construction. The biological implications of this metaphor have proven contentious. Proposed as a critique of a strictly adaptationist interpretation of biological structures (33), its original scope was broad.¹⁰ Subsequent analysis has emphasized the underlying role of selection in the evolution of apparently subordinate structural features. The triangular spaces in Byzantine cathedrals, for example, provided two complementary functions: the spandrel was selected as one of several possible structural forms due to its dual role as an architectural element and opportunity of iconographic display (34). Possible architectural alternatives may have adequately served one function or the other but not both (63). This perspective highlights the adaptive importance of spandrels and by extension "subordinate" features of biological structures.

We suggest that insulin's conserved hydrophobic core also serves a dual function: both to stabilize its native state (thereby making possible specific disulfide pairing and ensuring proteolytic stability) and to contribute to the binding surface displayed to the receptor. By restricting surrounding variation, induced fit would impose a requirement for Leu^{A16} as an essential buttress. Interlocked requirements of core structure and function have in essence made any single step

¹⁰ Spandrels was originally proposed in relation to such diverse evolutionary processes as neutral drift (no adaptation and no selection), selection of one of many possible structures without adaptive advantage (evolutionary contingency), and sexual selection without adaptation. The present discussion is restricted to selection under interlocking constraints of structure and function (34, 63).

in sequence space unfavorable: insulin's fitness peak is sharply honed. Conservation of Leu^{A16} may thus be viewed as a side consequence of induced fit. The seeming complexity of induced fit may itself represent an innovation: a mechanism to hinder pathological misfolding by limiting exposure of nonpolar side chains in the native state (26, 64–67). Although unlikely to be an intrinsic requirement of an insulin-like fold, the cavity filled by Leu^{A16}, like the triangular spaces at the base of a Byzantine dome, is an obligatory feature of insulin's functional design. Rather than being a term of dismissal, the spandrel metaphor emphasizes the multiple levels of selection operative in the evolution of protein sequences.

Invertebrate insulin-like genes exhibit divergent core sequences (32). Among such sequences the A16 position contains a variety of residue types (Ala, Ile, Met, Phe, and Val; Supporting Information). Variation is likewise observed at residues A2, A6, A11, A19, B6, B11, and B15. It would be of future interest to determine the structures of such putative insulin-like sequences, i.e., to verify that such structures indeed resemble that of human insulin. Maintenance of insulin-like folds among divergent invertebrate sequences would illuminate combinatorial mechanisms of structural compensation. It would be of further interest to investigate whether these putative insulin-like proteins can bind to the vertebrate insulin receptor and if so, whether significant differences occur in the mode of binding. Surprisingly, expression of the human proinsulin gene in *Caenorhabditis elegans* has recently been shown to have biological activity, presumably by signaling through the DAF-2 insulin-like receptor¹¹ (32). It is not known whether the nematode contains noncanonical insulin-like receptors unrecognized on the basis of vertebrate sequence homology.

CONCLUDING REMARKS

The present study has demonstrated that Leu^{A16} makes an essential contribution to insulin's structure and stability. Synthesis of analogues is limited by inefficient disulfide pairing. It would be of future interest to investigate whether this block is kinetic or thermodynamic. Although crystallographic, NMR, and molecular dynamics studies have highlighted insulin's flexibility (2, 8, 68–71), the potential cavity filled by Leu^{A16} (Figure 1D) must have stringent requirements for size and shape. Improvements in structural methodologies may soon permit characterization of variant structures despite low synthetic yields. Pending advances in NMR technology (for example, use of ultrahigh magnetic fields and cryogenic probes) promise a dramatic enhancement in sensitivity. Similarly, synchrotron beam lines of high brilliance will enable diffraction to be observed from protein crystals too small for conventional X-ray analysis (72). These

advances promise to permit structural analysis of misfolded and unstable proteins and thereby illuminate general principles of protein folding.

ACKNOWLEDGMENT

We thank M. DeFelippis (Eli Lilly and Co.) for a gift of human insulin and selected analogues; N. Phillips and T. Sosnick for assistance with CD measurements and analysis; N. Narayana and D. Porubun for assistance with figures; E. Collins for preparation of the manuscript; and Prof. G. G. Dodson, G. D. Smith, and D. F. Steiner for helpful discussion. M.A.W. is indebted to the late Prof. E. M. Purcell for discussion of his (Ed's) work with S. J. Gould.

SUPPORTING INFORMATION AVAILABLE

Three tables providing dihedral angles of Leu^{A16} among inequivalent crystal structures, predicted thermodynamic properties based upon the tabulated α -helical propensities and hydrophobicities of amino acids, and core insulin-like sequences in *C. elegans*. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

1. Dodson, G., and Steiner, D. (1998) The role of assembly in insulin's biosynthesis, *Curr. Opin. Struct. Biol.* 8, 189–194.
2. 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.
3. Olsen, H. B., Ludvigsen, S., and Kaarsholm, N. C. (1996) Solution structure of an engineered insulin monomer at neutral pH, *Biochemistry* 35, 8836–8845.
4. Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A., and Vijayan, M. (1971) Atomic positions in rhombohedral 2-zinc insulin crystals, *Nature* 231, 506–511.
5. Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., and Mercola, D. (1976) Structure of insulin in 4-zinc insulin, *Nature* 261, 166–168.
6. 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.
7. Badger, J., Harris, M. R., Reynolds, C. D., Evans, A. C., Dodson, E. J., Dodson, G. G., and North, A. C. (1991) Structure of the pig insulin dimer in the cubic crystal, *Acta Crystallogr., Sect. B: Struct. Sci.* 47, 127–136.
8. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., and Reynolds, C. D. (1988) The structure of 2Zn pig insulin crystals at 1.5 Å resolution, *Philos. Trans. R. Soc. London* 319, 369–456.
9. Ciszak, E., and Smith, G. D. (1994) Crystallographic evidence for dual coordination around zinc in the T3R3 human insulin hexamer, *Biochemistry* 33, 1512–1517.
10. Steiner, D. F., and Chan, S. J. (1988) An overview of insulin evolution, *Horm. Metab. Res.* 20, 443–444.
11. Wang, S. H., Hu, S. Q., Burke, G. T., and Katsoyannis, P. G. (1991) Insulin analogues with modifications in the beta-turn of the B-chain, *J. Protein Chem.* 10, 313–324.
12. 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.
13. Keller, D., Clausen, R., Josefsen, K., and Led, J. J. (2001) Flexibility and Bioactivity of Insulin: an NMR Investigation of the Solution Structure and Folding of an Unusually Flexible

¹¹ It is not known whether insulin or proinsulin was generated in this experiment. Since *C. elegans* has genes encoding both PC2 and PC1/3 homologues (the convertases that process proinsulin to insulin), it is possible that processing of proinsulin occurs in the transgenic worm. Interestingly, the endogenous *C. elegans* ins-1, which also exhibits good activity in dauer formation, is quite similar in sequence to vertebrate insulin in having conserved A2, A16, and A19 residues. Ruvkun and colleagues speculate that human proinsulin and endogenous insulin-like molecules function in *C. elegans* as inhibitors of the putative DAF-2 tyrosine kinase rather than as activators as in vertebrate species (32). Such negative regulation appears to modulate the organism's aging and senescence.

- Human Insulin Mutant with Increased Biological Activity, *Biochemistry* 40, 10732–10740.
14. 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.
 15. 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.
 16. 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.
 17. Nakagawa, S. H., and Tager, H. S. (1989) Perturbation of insulin-receptor interactions by intramolecular hormone cross-linking. Analysis of relative movement among residues A1, B1, and B29, *J. Biol. Chem.* 264, 272–279.
 18. Kobayashi, M., Ohgaku, S., Iwasaki, M., Maegawa, H., Shigeta, Y., and Inouye, K. (1982) Supernormal insulin: [D-Phe^{B24}]-insulin with increased affinity for insulin receptors, *Biochem. Biophys. Res. Commun.* 107, 329–336.
 19. Mirmira, R. G., and Tager, H. S. (1989) Role of the phenylalanine B24 side chain in directing insulin interaction with its receptor: Importance of main chain conformation, *J. Biol. Chem.* 264, 6349–6354.
 20. Mirmira, R. G., Nakagawa, S. H., and Tager, H. S. (1991) Importance of the character and configuration of residues B24, B25, and B26 in insulin-receptor interactions, *J. Biol. Chem.* 266, 1428–1436.
 21. Shoelson, S. E., Lu, Z. X., Parlautan, L., Lynch, C. S., and Weiss, M. A. (1992) Mutations at the dimer, hexamer, and receptor-binding, surfaces of insulin independently affect insulin-insulin and insulin-receptor interactions, *Biochemistry* 31, 1757–1767.
 22. Cosmatos, A., Federigios, N., and Katsoyannis, P. G. (1979) Chemical synthesis of [*des*(tetrapeptide B27–30), Tyr(NH₂)-26-B] and [*des*(pentapeptide B26–30), Phe(NH₂)25-B] bovine insulins, *Int. J. Protein Res.* 14, 457–471.
 23. Fischer, W. H., Saunders, D., Brandenburg, D., Wollmer, A., and Zahn, H. (1985) A shortened insulin with full in vitro potency, *Biol. Chem. Hoppe Seyler* 366, 521–525.
 24. Bi, R. C., Dauter, Z., Dodson, E., Dodson, G., Giordano, F., and Reynolds, C. (1984) Insulin structure as a modified and monomeric molecule, *Biopolymers* 23, 391–395.
 25. Dai, J.-B., Lou, M.-Z., You, J.-M., and Liang, D.-C. (1987) Refinement of the structure of *des*pentapeptide (B26–B30) insulin at 1.5 Å resolution, *Sci. Sin.* 30, 55–65.
 26. 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.
 27. Kitagawa, K., Ogawa, H., Burke, G. T., Chanley, J. D., and Katsoyannis, P. G. (1984) Critical role of the A2 amino acid residue in the biological activity of insulin: [2-glycine-A]- and [2-alanine-A]insulins, *Biochemistry* 23, 1405–1413.
 28. 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.
 29. Kobayashi, M., Takata, Y., Ishibashi, O., Sasaoka, T., Iwasaki, T. M., Shigeta, Y., and Inouye, K. (1986) Receptor binding and negative cooperativity of a mutant insulin, [Leu^{A3}]-insulin, *Biochem. Biophys. Res. Commun.* 137, 250–257.
 30. 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.
 31. 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*-isoleucine^{A2} analogue, *J. Mol. Biol.* (in press).
 32. Pierce, S. B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S. A., Buchman, A. R., Ferguson, K. C., Heller, J., Platt, D. M., Pasquinelli, A. A., Liu, L. X., Doberstein, S. K., and Ruvkun, G. (2001) Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family, *Genes Dev.* 15, 672–686.
 33. Gould, S. J., and Lewontin, R. C. (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme, *Proc. R. Soc. London, Ser. B* 205, 581–598.
 34. Dennett, D. C. (1995) *Darwin's dangerous idea: evolution and the meaning of life*, Simon and Schuster, New York.
 35. Barany, G., and Merrifield, R. B. (1980) in *The Peptides* (Gross, E., and Meienhofer, J., Eds.) pp 273–284, Academic Press, New York.
 36. 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.
 37. Marshall, R. N., Underwood, L. E., Voina, S. J., Foushee, D. B., and Van Wyk, J. J. (1974) Characterization of the insulin and somatomedin-C receptors in human placental cell membranes, *J. Clin. Endocrinol. Metab.* 39, 283–292.
 38. Cara, J. F., Mirmira, R. G., Nakagawa, S. H., and Tager, H. S. (1990) An insulin-like growth factor I/insulin hybrid exhibiting high potency for interaction with the type I insulin-like growth factor and insulin receptors of placental plasma membranes, *J. Biol. Chem.* 265, 17820–17825.
 39. Sosnick, T. R., Fang, X., and Shelton, V. M. (2000) Application of circular dichroism to study RNA folding transitions, *Methods Enzymol.* 317, 393–409.
 40. Pace, C. N., and Shaw, K. L. (2000) Linear extrapolation method of analyzing solvent denaturation curves, *Proteins* (Suppl. 4), 1–7.
 41. Laskowski, R. A. (1995) SURFNET: a program for visualizing molecular surfaces, cavities and intermolecular interactions, *J. Mol. Graphics* 13, 323–330.
 42. DiMarchi, R. D., Mayer, J. P., Fan, L., Brems, D. N., Frank, B. H., Green, J. K., Hoffman, J. A., Howey, D. C., Long, H. B., Shaw, W. N., Shields, J. E., Sliker, L. J., Su, K. S. E., Sundell, K. L., and Chance, R. E. (1992) in *Peptides: Proceedings of the Twelfth American Peptide Symposium* (Smith, J. A., and Rivier, J. E., Eds.) pp 26–28, ESCOM Science Publishers B. V., Leiden, The Netherlands.
 43. Weiss, M. A., Hua, Q.-X., Jia, W., Chu, Y.-C., Wang, R.-Y., and Katsoyannis, P. G. (2000) Hierarchical protein “un-design”: insulin's intrachain disulfide bridge tethers a recognition α -helix, *Biochemistry* 39, 15429–15440.
 44. Kikuchi, K., Larner, J., Freer, R. J., Day, A. R., Morris, H., and Dell, A. (1980) Studies on the biological activity of degraded insulins and insulin fragments, *J. Biol. Chem.* 255, 9281–9288.
 45. 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.
 46. Kjeldsen, T., Brandt, J., Andersen, A. S., Egel-Mitani, M., Hach, M., Pettersson, A. F., and Vad, K. (1996) A removable spacer peptide in an alpha-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in *Saccharomyces cerevisiae*, *Gene* 170, 107–112.
 47. Katsoyannis, P. G. (1966) Synthesis of insulin, *Science* 154, 1509–1514.
 48. 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.
 49. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect, *Science* 255, 178–183.

50. Eigenbrot, C., Randal, M., Quan, C., Burnier, J., O'Connell, L., Rinderknecht, E., and Kossiakoff, A. A. (1991) X-ray structure of human relaxin at 1.5 Å. Comparison to insulin and implications for receptor binding determinants, *J. Mol. Biol.* **221**, 15–21.
51. Milla, M. E., Brown, B. M., and Sauer, R. T. (1994) Protein stability effects of a complete set of alanine substitutions in Arc repressor, *Nat. Struct. Biol.* **1**, 518–523.
52. Pullen, R. A., Lindsay, D. G., Wood, S. P., Tickle, I. J., Blundell, T. L., Wollmer, A., Krail, G., Brandenburg, D., Zahn, H., Gliemann, J., and Gammeltoft, S. (1976) Receptor-binding region of insulin, *Nature* **259**, 369–373.
53. De Meyts, P., Van Obberghen, E., and Roth, J. (1978) Mapping of the residues responsible for the negative cooperativity of the receptor-binding region of insulin, *Nature* **273**, 504–509.
54. Liang, D. C., Chang, W. R., and Wan, Z. L. (1994) A proposed interaction model of the insulin molecule with its receptor, *Biophys. Chem.* **50**, 63–71.
55. Hua, Q. X., and Weiss, M. A. (1991) Comparative 2D NMR studies of human insulin and *des*-pentapeptide insulin: sequential resonance assignment and implications for protein dynamics and receptor recognition, *Biochemistry* **30**, 5505–5515.
56. Kitagawa, K., Ogawa, H., Burke, G. T., Chanley, J. D., and Katsoyannis, P. G. (1984) Interaction between the A2 and A19 amino acid residues is of critical importance for high biological activity in insulin: [19-leucine-A]insulin, *Biochemistry* **23**, 4444–4448.
57. Stoev, S., Zakhariyev, S., Golovinsky, E., Gattner, H. G., Naithani, V. K., Wollmer, A., and Brandenburg, D. (1988) Synthesis and properties of [A19-(p-fluorophenylalanine)] insulin, *Biol. Chem. Hoppe Seyler* **369**, 1307–1315.
58. Du, X., and Tang, J. G. (1998) Hydroxyl group of insulin A19Tyr is essential for receptor binding: studies on (A19Phe)-insulin, *Biochem. Mol. Biol. Int.* **45**, 255–260.
59. 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 [LeuA3] insulin, *J. Clin. Invest.* **77**, 514–519.
60. Nanjo, K., Sanke, T., Kondo, M., Nishimura, S., Miyano, M., Linuma, J., Miyamura, K., Inouye, K., Given, B. D., and Polonsky, K. S. (1986) Mutant insulin syndrome: identification of two families with [LeuA³]insulin and determination of its biological activity, *Trans. Assoc. Am. Physicians* **99**, 132–142.
61. 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.
62. Ohta, N., Burke, G. T., and Katsoyannis, P. G. (1988) Synthesis of an insulin analogue embodying a strongly fluorescent moiety, [19-tryptophan-A]insulin, *J. Protein Chem.* **7**, 55–65.
63. Mark, R. (1996) Architecture and evolution, *Am. Sci.* **84**, 383–389.
64. Bouchard, M., Zurdo, J., Nettleton, E. J., Dobson, C. M., and Robinson, C. V. (2000) Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy, *Protein Sci.* **9**, 1960–1967.
65. Nettleton, E. J., Tito, P., Sunde, M., Bouchard, M., Dobson, C. M., and Robinson, C. V. (2000) Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry, *Biophys. J.* **79**, 1053–1065.
66. Sipe, J. D., and Cohen, A. S. (2000) Review: history of the amyloid fibril, *J. Struct. Biol.* **130**, 88–98.
67. Nielsen, L., Frokjaer, S., Carpenter, J. F., and Brange, J. (2001) Studies of the structure of insulin fibrils by Fourier transform infrared (FTIR) spectroscopy and electron microscopy, *J. Pharm. Sci.* **90**, 29–37.
68. Chothia, C., Lesk, A. M., Dodson, G. G., and Hodgkin, D. C. (1983) Transmission of conformational change in insulin, *Nature* **302**, 500–505.
69. Kruger, P., Strassburger, W., Wollmer, A., van Gunsteren, W. F., and Dodson, G. G. (1987) The simulated dynamics of the insulin monomer and their relationship to the molecule's structure, *Eur. Biophys. J.* **14**, 449–459.
70. Hua, Q. X., Kochoyan, M., and Weiss, M. A. (1992) Structure and dynamics of *des*-pentapeptide-insulin in solution: the molten-globule hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2379–2383.
71. Jacoby, E., Kruger, P., Schlitter, J., Roper, D., and Wollmer, A. (1996) Simulation of a complex protein structural change: the T ↔ R transition in the insulin hexamer, *Protein Eng.* **9**, 113–125.
72. Stevens, R. C. (2000) High-throughput protein crystallization, *Curr. Opin. Struct. Biol.* **10**, 558–563.
73. Chen, J. Q., Zhang, H. T., Hu, M. H., and Tang, J. G. (1995) Production of human insulin in an *E. coli* system with Met-Lys-human proinsulin as the expressed precursor, *Appl. Biochem. Biotech.* **55**, 5–15.
74. Burke, G. T., Chanley, J. D., Okada, Y., Cosmatos, A., Ferderigos, N., and Katsoyannis, P. G. (1980) Divergence of the in vitro biological activity and receptor binding affinity of a synthetic insulin analogue, [21-asparaginamide-A]insulin, *Biochemistry* **19**, 4547–4556.
75. Kaarsholm, N. C., Norris, K., Jorgensen, R. J., Mikkelsen, J., Ludvigsen, S., Olsen, O. H., Sorensen, A. R., and Havelund, S. (1993) Engineering stability of the insulin monomer fold with application to structure–activity relationships, *Biochemistry* **32**, 10773–10778.
76. 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**, 6408–6411.
77. Bromer, W. W., and Chance, R. E. (1967) Preparation and characterization of desoctapeptide-insulin, *Biochim. Biophys. Acta* **133**, 219–223.
78. Dai, Y., and Tang, J. G. (1996) Characteristic, activity and conformational studies of [A6-Ser, A11-Ser]-insulin, *Biochim. Biophys. Acta* **1296**, 63–68.
79. Seino, S., Steiner, D. F., and Bell, G. I. (1987) Sequence of a New World primate insulin having low biological potency and immunoreactivity, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7423–7427.
80. Padmanabhan, S., and Baldwin, R. L. (1994) Helix-stabilizing interaction between tyrosine and leucine or valine when the spacing is $i, i + 4$, *J. Mol. Biol.* **241**, 706–713.
81. Karplus, P. A. (1997) Hydrophobicity regained, *Protein. Sci.* **6**, 1302–1307.
82. Brems, D. N., Brown, P. L., Bryant, C., Chance, R. E., Green, L. K., Long, H. B., Miller, A. A., Millican, R., Shields, J. E., and Frank, B. H. (1992) Improved insulin stability through amino acid substitution, *Protein Eng.* **5**, 519–525.
83. Luo, Y., and Baldwin, R. L. (2001) How ala→gly mutations in different helices affect the stability of the apomyoglobin molten globule, *Biochemistry* **40**, 5283–5289.
84. Dunitz, J. D. (1995) Win some, lose some: enthalpy–entropy compensation in weak intermolecular interactions, *Chem. Biol.* **2**, 709–712.