A cavity-forming mutation in insulin induces segmental unfolding of a surrounding α-helix

BIN XU,[1] QING-XIN HUA,[1] SATOE H. NAKAGAWA,[2] WENHUA JIA,[1] YING-CHI CHU,[3] PANAYOTIS G. KATSOYANNIS,[3] AND MICHAEL A. WEISS[1]

[1]Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA
[2]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA
[3]Department of Pharmacology and Biological Chemistry, Mt. Sinai School of Medicine of New York University, New York, New York 10029, USA

(RECEIVED August 2, 2001; FINAL REVISION October 12, 2001; ACCEPTED October 12, 2001)

Abstract

To investigate the cooperativity of insulin's structure, a cavity-forming substitution was introduced within the hydrophobic core of an engineered monomer. The substitution, Ile^{A2}→Ala in the A1–A8 α-helix, does not impair disulfide pairing between chains. In accord with past studies of cavity-forming mutations in globular proteins, a decrement was observed in thermodynamic stability ($\Delta\Delta G_u$ 0.4–1.2 kcal/mole). Unexpectedly, CD studies indicate an attenuated α-helix content, which is assigned by NMR spectroscopy to selective destabilization of the A1–A8 segment. The analog's solution structure is otherwise similar to that of native insulin, including the B chain's supersecondary structure and a major portion of the hydrophobic core. Our results show that (1) a cavity-forming mutation in a globular protein can lead to segmental unfolding, (2) tertiary packing of Ile^{A2}, a residue of low helical propensity, stabilizes the A1–A8 α-helix, and (3) folding of this segment is not required for native disulfide pairing or overall structure. We discuss these results in relation to a hierarchical pathway of protein folding and misfolding. The Ala^{A2} analog's low biological activity (0.5% relative to the parent monomer) highlights the importance of the A1–A8 α-helix in receptor recognition.

Keywords: Protein unfolding; insulin; cooperativity; protein structure; hormone insulin receptor; NMR spectroscopy

Supplemental material: See www.proteinscience.org.

Insulin is a small globular protein containing two chains, designated A (21 residues) and B (30 residues). The hormone is stored in the pancreatic β cell as a Zn^{2+}-stabilized hexamer and functions in the bloodstream as a Zn^{2+}-free monomer (Dodson and Steiner 1998). The structure of a monomer in solution resembles the crystallographic T-state (Fig. 1A), as defined in a variety of crystal forms (Blundell et al. 1971; Bentley et al. 1976; Bi et al. 1984; Dai et al. 1987; Baker et al. 1988; Badger et al. 1991; Ciszak and Smith 1994; Bao et al. 1997). Despite its small size, insulin contains representative features of larger proteins, including canonical elements of secondary structure and a well-ordered hydrophobic core. The present study focuses on the structural role of Ile^{A2}, an invariant side chain in the core (Blundell et al. 1971; Bentley et al. 1976; Derewenda et al. 1989; Badger et al. 1991; Kitagawa et al. 1984a,b; Nakagawa and Tager 1992). Adjoining side chains Ile^{A2} and Cys^{A6} (part of the internal A6–A11 disulfide bridge) anchor the A1–A8 α-helix through a network of long-range contacts involving Leu^{A16}, Tyr^{A19}, Leu^{B6}, Leu^{B11}, and Leu^{B15} (Fig. 1C). Reorganization of the B chain on receptor binding is proposed to expose Ile^{A2} to engage the insulin receptor

Reprint requests to: Michael Weiss, Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA; email: weiss@biochemistry.cwru.edu; fax: (216) 368-3419.

Abbreviations: CD, circular dichroism; DG/SA, distance-geometry simulated annealing; DKP-insulin, analog containing three substitutions in the B chain (see Table 1); DQF-COSY, double-quantum filtered correlated spectroscopy; HPLC, high-performance liquid chromatography; IGF-I, single-chain insulin-like growth factor; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; RMS, root mean square; RMSD, RMS deviations; rp-HPLC, reverse-phase HPLC; Amino acids are represented by standard one- and three-letter codes. "Native" elements of structure designate feature of crystal structures and may not correspond to the functional conformation in a receptor complex.

Article and publication are at http://www.proteinscience.org/cgi/doi/10.1101/ps.32102.

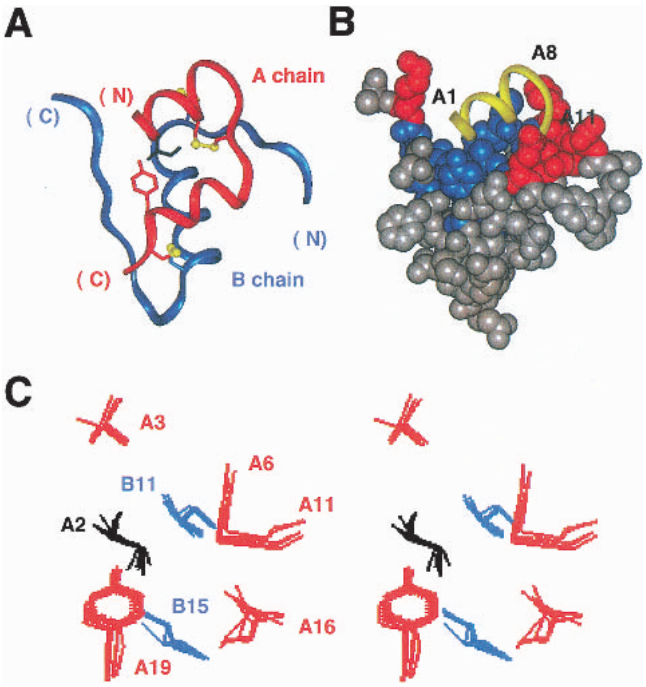


Fig. 1. Structure of insulin. (*A*) Ribbon representation of crystal structure of native 2-Zn porcine insulin (T-state protomer; PDB entry 4INS). The A chain is shown in red and B chain in blue; disulfide bridges are shown in yellow (balls and sticks). The solution structure of DKP-insulin is similar to the crystallographic T-state (Hua et al. 1996). (*B*) Surface representation of T-state protomer (2-Zn molecule 1) showing "internal template" in contact with N-terminal A-chain α-helix (yellow ribbon; residues A1–A11). Polar or charged side chains are shown in red; nonpolar or aromatic side chains in blue. Images were generated using InsightII (Biosym, Inc.). (*C*) Stereo representation of side-chain environment of Ile^{A2} in multiple crystal structures (T-state protomers; PDB entries 1APH, 1LPH, 1TRZ, 1ZNI, 2INS, 4INS). Structures are aligned with respect to main-chain atoms of residues A2–A8, A13–A20, and B9–B19.

(Derewenda et al. 1991; Hua et al. 1991; Nakagawa and Tager 1992). Conservation of Ile^{A2} presumably reflects its dual role in structure and function. In this article we investigate the consequences of a cavity-forming mutation, Ala^{A2}. Previous studies have shown that an Ala^{A2} substitution impairs insulin's activity by 200-fold (Kitagawa et al. 1984a; Nakagawa and Tager 1992) but have not addressed its structural consequences.

Protein structures ordinarily show cooperativity and robustness to point mutations (Eriksson et al. 1992). Whereas cavity-forming mutations in globular proteins typically result in decreased thermodynamic stability (Sandberg and Terwilliger 1989; Milla et al. 1994; Xu et al. 1998, 2001), native-like structures are maintained through local conformational adjustments (Eriksson et al. 1992; Xu et al., 1998). The present study exploits a cavity-forming mutation in insulin to investigate the stability and modularity of the A1–A8 α-helix. Experimental design is influenced by an analogy proposed between the dynamics of the insulin monomer and a molten globule (MG; Hua et al. 1992, 1993). Mutagenesis studies of α-helices within equilibrium MG models have highlighted the limited cooperativity of individual α-helices (Pfeil 1981; Dolgikh et al. 1985; Pfeil et al. 1986; Griko et al. 1988; Schulman and Kim 1996; Schulman et al. 1997). Given the insulin monomer's molten features, would one or more of its helical segments lack cooperativity? Folding of insulin and insulin-like growth factors is characterized by stepwise stabilization of native structural elements with successive disulfide bond formation (Narhi et al. 1993; Hua et al. 1996b; Hober et al. 1997). In one branch (Hua et al. 1996a,b; Weiss et al. 2000; Qiao et al. 2001), the A1–A8 α-helix occurs as a late event, coupled to formation of the final disulfide bridge (cystine A6–A11) (Weiss et al. 2000). The preceding intermediate contains a native-like subdomain, which is proposed to provide a specific template to stabilize folding of the A1–A8 segment as an amphipathic α-helix (Hua et al. 1996a; Weiss et al. 2000).

In this article we describe the synthesis and characterization of an Ile^{A2}→Ala analog. This large-to-small substitution occurs within an amphipathic α-helix unusually rich in β-branched amino acids (sequence GIVEQCCT). The substitution is constructed with an engineered monomer (DKP-insulin; Weiss et al. 1991; DiMarchi et al. 1992; Shoelson et al. 1992; Ciszak et al. 1995) to facilitate spectroscopic studies. DKP-insulin contains three substitutions in the B chain (Table 1) that impair formation of dimers, trimers, and higher-order oligomers (DiMarchi et al. 1992, Shoelson et al. 1992). The engineered monomer, which shows enhanced potency (Shoelson et al. 1992) and thermodynamic stability (Weiss et al. 2001), provides a convenient template for analysis of substitutions in the A chain (Hua et al. 1996a; Weiss et al. 2000). A combination of CD and NMR is used to show that an Ala^{A2} substitution in DKP-insulin causes segmental destabilization of the A1–A8 segment. An otherwise native subdomain is retained. The thermodynamic stability of Ala^{A2}-DKP-insulin is lower than that of DKP-

Table 1. *Design of insulin analogues*

Substitution		Location	Purpose
B10	His → Asp	surface of α-helix	destabilize trimer
B28	Pro → Lys	surface of β-strand	destabilize dimer
B29	Lys → Pro	surface of β-strand	destabilize dimer
A2	Ile → Ala	hydrophobic core	cavity-forming mutation

DPK-insulin contains the three B-chain substitutions (Weiss et al. 1991; Shoelson et al. 1992) whereas Ala^{A2}-DKP-insulin contains four substitutions. The B28-B29 inversion was motivated by homology to IGF-I to destabilize insulin's classical dimer interface (Brems et al. 1992; DiMarchi et al. 1992). The crystal structure of [Lys^{B28}, Pro^{B29}]-insulin (KP-insulin) has been determined as a T_3R_{f3} hexamer (Ciszak et al. 1995); the solution structure of DKP-insulin has also been determined (Hua et al. 1996a). In the present study, the affinity of DKP-insulin for a human placental insulin-receptor preparation was 160% relative to native human insulin.

insulin. Despite these perturbations, the variant A-chain pairs with the DKP-B chain with native fidelity and yield. The structure of the mutant insulin shows that a cavity-forming mutation in the core of a protein can lead to segmental unfolding in accord with a hierarchical perspective (Hober et al. 1992; Miller et al. 1993; Narhi et al. 1993; Hua et al. 1996a; Weiss et al. 2000). In particular, folding of the A1–A8 segment is not integral to insulin's overall structure or mechanism of disulfide pairing. To our knowledge, this is the first structure of an A2 insulin analog to be described. We speculate that insulin's limited cooperativity may facilitate segmental conformational change during the protein's complex "life cycle" of biosynthesis, assembly, and action (Derewenda et al. 1991; Hua et al. 1991; Dodson and Steiner 1998; Lipkind and Steiner 1999).

Results

Ala^{A2}-DKP-insulin was prepared by total chemical synthesis (see Materials and Methods). In contrast to studies of certain B-chain substitutions (Wang et al. 1991; Hu et al. 1993), the yield of insulin chain combination was similar to that obtained in the synthesis of native or DKP-insulin (Hua et al. 1996a). The analog's affinity for the human insulin receptor is 150-fold lower than that of native insulin and 250-fold lower than that of DKP-insulin; such a marked impairment is in accord with previous studies (Kitagawa et al. 1984a,b; Nakagawa and Tager 1992). CD spectra indicate a reduction in α-helix content relative to DKP-insulin (Fig. 2B and Table 2); the value of mean residue ellipticity at 222 nm is −7510 deg·cm·mole^{-1}, significantly less negative than the parent value of −9580 deg·cm·mole^{-1}. Ala^{A2}-DKP-insulin also shows decreased thermal and thermodynamic stabilities (asterisks in Fig. 2, C and D). Analysis of native and variant guanidine-denaturation transitions by a two-state model (Brems et al. 1990; Sosnick et al. 2000) implies free energies of unfolding (ΔG_u) of 4.9 kcal/mole (DKP-insulin) and 3.7 kcal/mole (Ala^{A2}-DKP-insulin) at 4°C. Because the variant *m* value is also lower than the native *m* value, such fitting may underestimate the variant's stability and so overestimate the thermodynamic perturbation (Pace and Shaw 2000). An upper bound on the analog's ΔG_u value of 4.5 kcal/mole may be calculated using the native *m* value and variant C_{mid} value (see Materials and

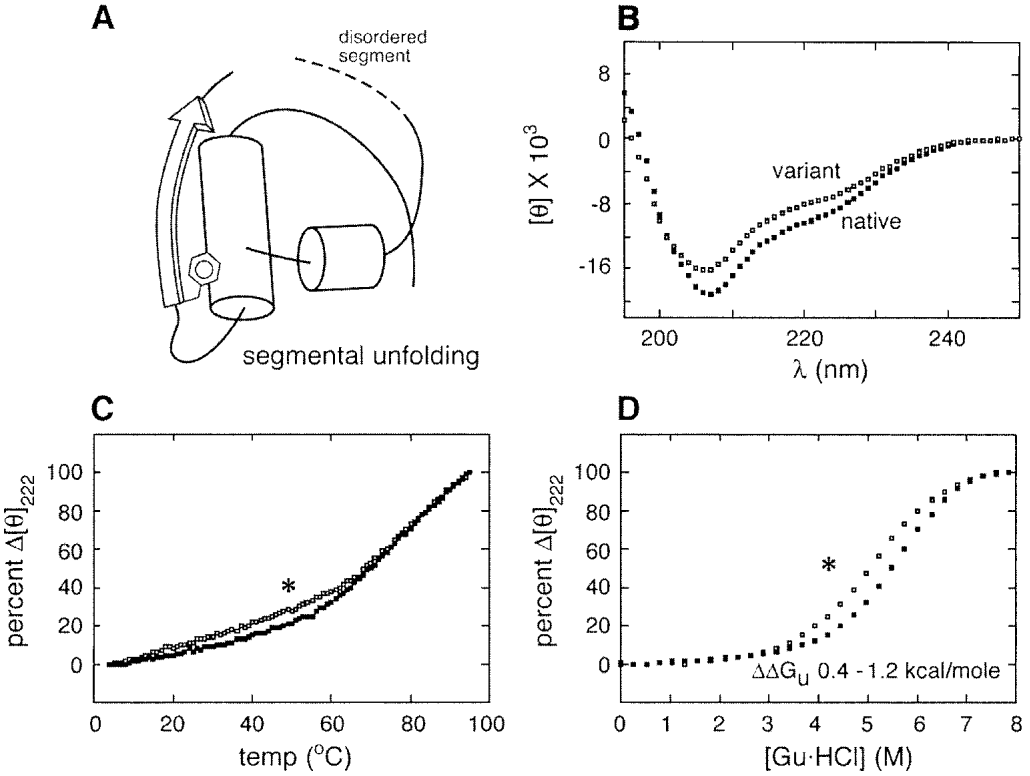


Fig. 2. Model of partial insulin fold (*A*) and CD studies (*B–D*). (*A*) Cylinder model of segmental unfolding of Ala^{A2}-DKP-insulin. (*B*) Far-UV spectra of DKP-insulin [native (filled squares)], and Ala^{A2}-DKP-insulin [variant (open squares)] indicate attenuated helix content in variant. (*C*) Thermal unfolding of Ala^{A2}-DKP-insulin (open squares), monitored by ellipticity at 222 nm, precedes (asterisk) that of DKP-insulin (filled squares). (*D*) Guanidine unfolding transitions of Ala^{A2}-DKP-insulin (open squares) and DKP-insulin (filled squares) show that the variant shows decreased thermodynamic stability as inferred by a two-state model ($\Delta\Delta G_u$ 0.4–1.2 kcal/mole; see Table 2 for ΔG_u values).

Table 2. *Spectroscopic and thermodynamic studies*

Mean residue ellipticities[a] Protein	195 nm	208 nm	222 nm
DKP-insulin	5.80	−18.7	−9.58
Ala^{A2}-DKP-insulin	2.26	−15.6	−7.51

Thermodynamic studies[b] Protein	ΔG_u	$\Delta\Delta G_u$	C_{mid}[b]	m (kcal mol^{-1} M^{-1})
DKP-insulin	4.9 ± 0.04	—	5.8 ± 0.10	0.84 ± 0.01
Ala^{A2}-DKP-insulin	3.7 ± 0.08[c]	1.2 ± 0.12[c]	5.2 ± 0.03	0.71 ± 0.02

[a] Values are shown as $[\theta] \times 10^3$.
[b] C_{mid} is defined as the denaturant concentration at which 50% of the protein is unfolded. The m values is the slope $d(\Delta G_u)/d(M)$.
[c] The analogue's 15% decrease in m value raises the possibility of an underestimate of ΔG_u and hence overestimate of $\Delta\Delta G_u$ (see text). An alternative estimate based on linear extrapolation of the native m value yields values of 4.5 kcal/mole and 0.4 kcal/mole, respectively.

Methods for justification; Luo and Baldwin 2001). The analog's decrement in stability ($\Delta\Delta G_u$) is thus estimated to lie between 0.4 and 1.2 kcal/mole (Table 2).

^{1}H-NMR studies of Ala^{A2}-DKP-insulin were conducted at neutral pH in aqueous solution and in 20% deuterioacetic acid (Weiss et al. 1991), as described in previous studies of DKP-insulin (see Materials and Methods; Hua et al. 1996a). Complete sequential assignment was in each case obtained; patterns of NOEs and chemical shifts are similar under the two conditions (electronic supplemental material). Differences in chemical shifts (relative to those of DKP-insulin) are in general small and not localized (Table 3). Analysis of secondary shifts (defined as differences between observed chemical shifts and tabulated random-coil values) indicates an overall correlation (Fig. 3 and electronic supplemental material). Perturbations in A1–A5 H_α resonances trend downfield, that is, away from helix-related values and toward random-coil values. The magnitude of such changes is small. Baseline chemical shift dispersion in the native A1–A8 α-helix of DKP-insulin is less marked than would be predicted based on crystal structures (Hua et al. 1996a; Jacoby et al. 1996), presumably because of protein dynamics[1].

Table 3. *Chemical-shift differences (pH 7 and 25°C)*[a]

Residue	H_N	$C_\alpha H$	$C_\beta H$	Other
A3 Val		−0.14		
A4 Glu	+0.15			$C_\gamma H_2$ −0.17
A5 Gln		+0.10		
A6 Cys			+0.13	
A10 Ile		−0.10		
A11 Cys		−0.16		
A12 Ser	−0.11			
A15 Gln			−0.13, −0.18	
A16 Leu	+0.16			
A17 Glu			+0.12	
A19 Tyr	−0.11	+0.13	+0.16	
B2 Val	−0.18			
B4 Gln	−0.12			
B12 Val	−0.12			
B13 Glu		+0.13	−0.10	
B15 Leu			−0.13	
B17 Leu	−0.14		−0.14	$C_\gamma H$ −0.14
B24 Phe			−0.12	
B27 Thr		−0.10		

[a] The site of substitution (residue A2) is not included. Positive (negative) values represent chemical shifts that are upfield (downfield) in the parent spectrum relative to the variant spectrum.

Comparison of native and variant NOESY spectra indicates segmental perturbations in the A1–A8 region; native-like spectra features are otherwise retained. Whereas NOEs between amide protons (d_{NN} contacts) in the A12–A19 and B9–B19 segments are similar to those observed in DKP-insulin, d_{NN} contacts in the A1–A8 segment are weak or absent (Fig. 4A and Table 4). Analysis of helix-related (i, i + 3) contacts ($d_{\alpha N(i,i+3)}$ and $d_{\alpha\beta(i,i+3)}$; shown in Wüthrich format in Fig. 4C) shows retention of native-like A12–A19 and B9–B19 helices; such contacts are absent in the A1–A8 segment. NMR-defined secondary structure is thus in accord with attenuated CD ellipticity at 222 nm as a consequence of segmental destabilization of the A1–A8 α-helix (Fig. 2A). Native-like contacts between side chains in the B9–B19 α-helix (Leu^{B11}, Val^{B12}, and Leu^{B15}) and the B-chain's C-terminal β-strand (Phe^{B24} and Tyr^{B26}) are retained (Fig. 4B), indicating maintenance of the B chain's supersecondary structure. Whereas Ile^{A2} in DKP-insulin shows prominent NOEs to the aromatic resonances of Tyr^{A19}, however, no long-range NOEs are observed involv-

[1]Of insulin's three helices, the A1–A8 α-helix is least regular among crystal structures. In NMR studies, an A2–A8 α-helix is well defined, but largely lacks protected amide resonances in D_2O solution (Hua and Weiss 1991; Hua et al. 1996a). Chemical shifts of the Ile^{A2} side chain are less upfield than would be predicted by a ring-current model of Tyr^{A19} (Jacoby et al. 1996), presumably because of variability in the details of A2–A19 packing.

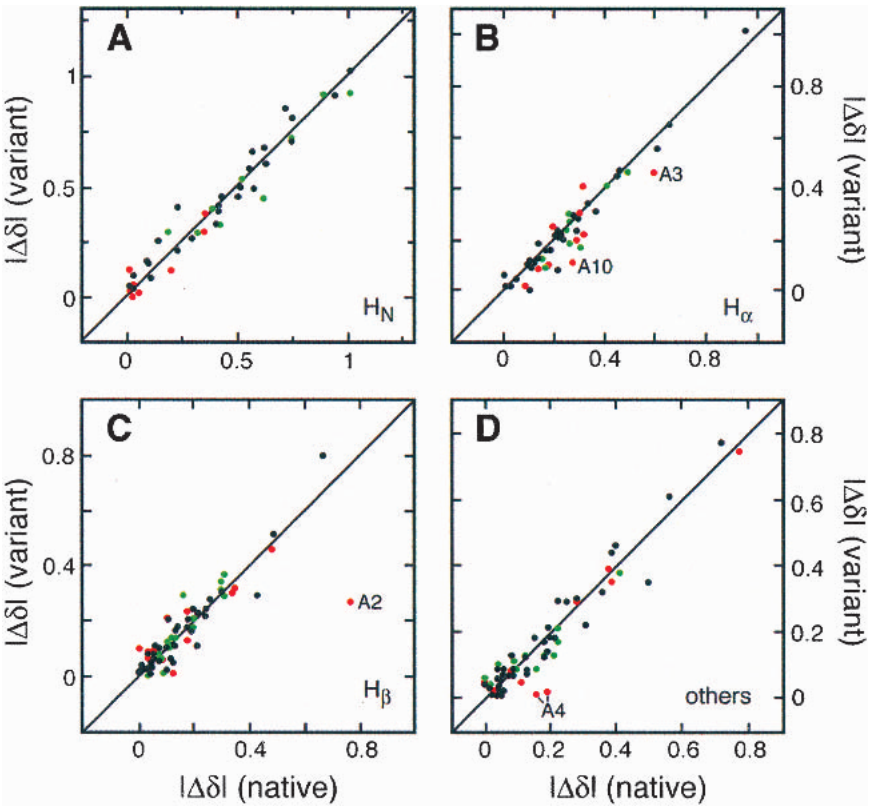


Fig. 3. Correlation of ^{1}H-NMR secondary shifts (|Δδ|) between Ala^{A2}-DKP-insulin (ordinate) and DKP-insulin (abscissa) by proton class. Such differences reflect sites of change in proton magnetic environments: (*A*) amide, (*B*) H_α, (*C*) H_β, and (*D*) aromatic and methyl resonances. Shifts were observed at pH 7.0 and 25°C. The A1–A11 segment is shown in red, the remainder of the A chain in green, and the B chain in black. Outlying points are assigned as shown. Tables of chemical shifts and secondary shifts are provided as supplemental material (Table S1 in electronic supplemental material). Secondary shifts are defined as difference between observed chemical shifts and tabulated random-coil values (Wüthrich 1986).

ing the methyl resonance of Ala^{A2} (dashed line in Fig. 4B; see also Table 4). Native-like long-range NOEs elsewhere in the molecule are retained. These include contacts between the side chains of Ile^{A10} and His^{B5} (not shown), Leu^{A13} and Phe^{B1}, and Tyr^{A19} and Leu^{B15} (Fig. 4B and Table 4).

The solution structure of Ala^{A2}-DKP-insulin was calculated based on 416 NOE-derived distances, dihedral and hydrogen-bond restraints (Table 5). DG/SA ensembles of DKP-insulin and Ala^{A2}-DKP-insulin are shown in Figure 5, A and B, respectively. Whereas the variant protein retains the overall insulin motif, the A1–A8 helix is not well defined. Disorder in the A1–A3 segment is associated with imprecision[2] in the B26–B30 segment as a result of loss of interchain contacts. Excepting the variable position of Ala^{A2} and the imprecise convergence of the location of the internal A6–A11 disulfide bridge[3], the remainder of the analog's hydrophobic core (Fig. 5D) is similar to that of DKP-insulin (Fig. 5C). The side chains of cystine A6–A11, Leu^{A16}, Tyr^{A19}, Leu^{B11}, and Leu^{B15} remain largely inaccessible to solvent in the Ala^{A2} structure. It is possible that fluctuations in the analog's structure can lead to transient exposure of these core side chains. The Tyr^{A19} side chain has similar average solvent accessibility value in the Ala^{A2}-DKP ensemble as in the DKP ensemble, but is more variable within the ensemble of Ala^{A2}-DKP (20% ± 7% versus 22% ± 3%). No residual cavity is consistently observed in the ensemble. Imprecision of the A1–A8 segment of Ala^{A2}-DKP-insulin and adjoining regions of the B chain reflects absence of

[2]Imprecision is defined by RMS deviations between members of an ensemble of models. Following alignment according to the main-chain atoms of residues A12–A20 and B9–B19, the A1–A8 segment shows a main-chain RMSD of 0.71 Å in DKP-insulin versus 1.51 Å in Ala^{A2}-DKP-insulin.

[3]Cys^{A11} is somewhat better defined than Cys^{A6}. RMSD values of these side chains are variant ensemble, A6 1.93 Å and A11 1.39 Å; parent ensemble, A6 0.55 Å and A11 0.61 Å.

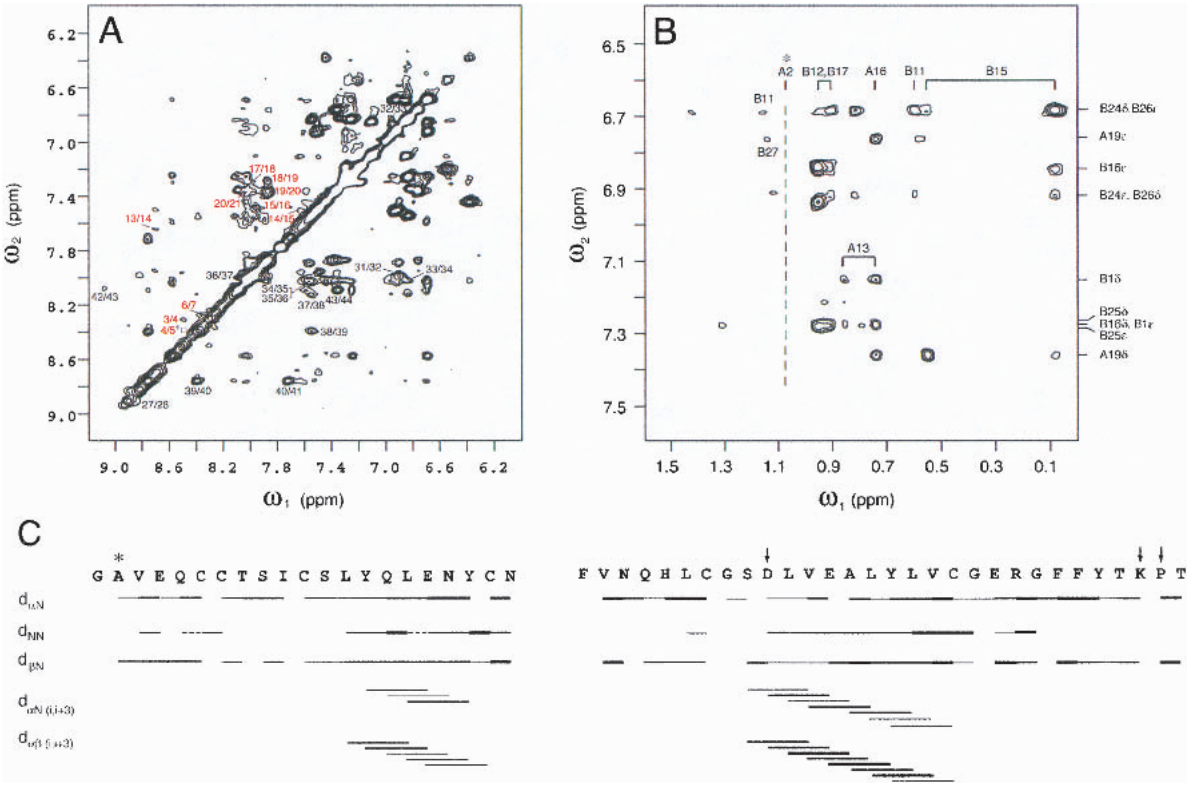


Fig. 4. NOESY spectra of Ala^{A2}-DKP-insulin (*A* and *B*) and summary of sequential assignments (*C*). (*A*) Amide and aromatic region in H_2O shows d_{NN} connectivities in A chains (red labels) and B chains (black labels). Location of key unobserved peaks are marked by asterisks at predicted locations (open squares). (*B*) NOEs between aromatic and aliphatic resonances in D_2O include long-range NOEs involving side chains of Tyr^{A19}, Phe^{B1}, Phe^{B24}, and Tyr^{B26} and methyl groups of Leu^{B15}, Leu^{B11}, Leu^{A16}, Val^{B12}, Leu^{B17}, and Thr^{B27}. (*C*) Sequential and medium-range NOEs in Ala^{A2}-DKP-insulin are summarized in Wüthrich format. Relative strength of NOEs is schematically represented by thickness of line. Dashed lines indicate possible cross peaks obscured by resonance overlap.

restraints. In 20% acetic acid, motional narrowing[4] is observed among resonances in the A1–A8 segment, which indicates that the variant segment is in fast exchange among a range of configurations. Motional narrowing is not observed in aqueous solution at neutral pH. Rather, several resonances in the A1–A8 segment are broader than those of analogous spin systems in the well-ordered subdomain. It is likely that line widths at neutral pH are influenced by conformational exchange on a millisecond time scale, leading to differential broadening.

Discussion

The present study used a combination of CD and NMR spectroscopy to characterize the response of an engineered insulin monomer to a potential cavity in its hydrophobic core. Whereas CD provides an overall estimate of the extent of order in an ensemble, NMR enables specific sites of order and disorder to be distinguished. These complementary probes in part circumvent the "absence of evidence" dilemma posed by the variable precision of polypeptide segments in a distance-geometry model (Liu et al. 1992; Clore et al. 1993; Tjandra et al. 2000). Thus, whereas imprecision of the A1–A8 segment in the DG/SA ensemble of Ala^{A2}-DKP-insulin reflects the absence of selected NOE restraints ("absence of evidence"; Table 4), its physical reality is corroborated by attenuation of the analog's helix-specific CD signature ("evidence of absence"; Table 2). In the future, it would be of interest to characterize the dynamics of Ala^{A2}-DKP-insulin on multiple time scales by heteronuclear relaxation methods (Barbato et al. 1992; Peng and Wagner 1992; Mulder et al. 2000) and analysis of residual dipolar couplings (Tjandra and Bax 1997; Tjandra et al. 1997; Tjandra et al. 2000; Goto et al. 2001). The extent and time scale of structural fluctuations in Ala^{A2}-DKP-insulin are presently not well characterized.

[4]Motional narrowing is defined as the general narrowing of NMR resonances as a result of molecular motions in the laboratory frame. In the present context, motional narrowing refers to differential narrowing because of local or segmental mobility relative to baseline line widths determined by macromolecular tumbling (τ_c).

Table 4. *Segmental and long-range NOEs*

Class	Atoms	Structure
Retained A1-A9 NOEs in Ala^{A2} analogue		
Segmental	Glu^{A4} $H_{\beta1,2}$–Cys^{A7} H_N	helix-related
Long-range	Val^{A3} H_α–Leu^{B11} $H_{\delta1\text{-}CH3}$	helix-helix
	Cys^{A6} $H_{\beta2}$–Leu^{B6} $H_{\delta1,2\text{-}CH3}$	helix-extended strand
	Cys^{A6} $H_{\beta1,2}$–$Leu^{B11}H_{\delta1,2\text{-}CH3}$	helix-helix
	Cys^{A7} H_α–His^{B5} $H_{\beta2}$	helix-extended strand
	Cys^{A7} H_α–Leu^{B6} H_N	helix-extended strand
	Cys^{A7} $H_{\beta1}$–Cys^{B7} H_α	disulphide-related
Absent A1-A9 NOEs in Ala^{A2} Analogue (relative to DKP-insulin)		
Segmental	Gly^{A1} H_α–Glu^{A4} H_N	helix-related
	Ile^{A2} H_N–Gln^{A5} $H_{\beta2}$	helix-related
	Ile^{A2} H_α–Gln^{A5} H_N	helix-related
	Ile^{A2} H_α–Gln^{A5} $H_{\beta1,2}$	helix-related
	Ile^{A2} $H_{\delta\text{-}CH3}$–Gln^{A5} $H_{\beta2}$	helix-related
	Val^{A3} H_α–Cys^{A6} H_N	helix-related
	Val^{A3} H_α–Cys^{A7} $H_{\beta1}$	helix-related
	Val^{A3} $H_{\gamma1\text{-}CH3}$–Cys^{A7} $H_{\beta1,2}$	helix-related
	Glu^{A4} H_α–Cys^{A7} H_N	helix-related
	Glu^{A4} H_α–Cys^{A7} $H_{\beta1,2}$	helix-related
	Glu^{A4} H_α–Thr^{A8} H_β	helix-related
	Gln^{A5} $H_{\beta1,2}$–Ser^{A9} $H_{\beta2}$	helix-turn
	Cys^{A6} H_α–Ile^{A10} $H_{\gamma1\text{-}CH2}$	helix-turn
	Cys^{A6} H_α–$Cys^{A11(a)}$ H_N	disulphide-related
	Cys^{A6} H_α–$Cys^{A11(a)}$ $H_{\beta1,2}$	disulphide-related
Long-range	Gly^{A1} H_α–Tyr^{A19} H_ϵ	supersecondary
	Ile^{A2} H_β–Tyr^{A19} H_δ	supersecondary
	Ile^{A2} H_β–Tyr^{A19} H_ϵ	supersecondary
	Ile^{A2} $H_{\gamma'\text{-}CH3}$–Tyr^{A19} H_δ	supersecondary
	Ile^{A2} $H_{\gamma'\text{-}CH3}$–Tyr^{A19} H_ϵ	supersecondary
	Ile^{A2} $H_{\delta\text{-}CH3}$–Tyr^{A19} H_δ	supersecondary
	Ile^{A2} $H_{\delta\text{-}CH3}$–Tyr^{A19} H_ϵ	supersecondary
	Ile^{A2} $H_{\gamma'\text{-}CH3}$–Tyr^{B26} H_α	helix-β-strand
	Ile^{A2} $H_{\gamma'\text{-}CH3}$–Tyr^{B26} H_δ	helix-β-strand
	Ile^{A2} $H_{\gamma'\text{-}CH3}$–Tyr^{B26} H_ϵ	helix-β-strand
	Ile^{A2} $H_{\gamma\text{-}CH2}$–Leu^{A16} H_α	supersecondary
	Ile^{A2} $H_{\gamma'\text{-}CH3}$–Thr^{B27} H_N	helix-β-strand
	Val^{A3} $H_{\gamma1,2\text{-}CH2}$–Tyr^{B26} H_δ	helix-β-strand
	Val^{A3} $H_{\gamma1,2\text{-}CH2}$–Tyr^{B26} H_ϵ	helix-β-strand
	Ser^{A9} H_α–His^{B5} $H_{\epsilon2}$	turn-extended strand

[a] NOE crosspeaks involving Cys^{A11} protons are not observed in NOESY spectra of Ala^{A2}-DKP-insulin, presumably due to conformational broadening. This confounds interpretation of unobserved A11-related NOEs (not listed in table).

To our knowledge, segmental unfolding of portions of a protein surrounding a putative cavity has not previously been observed as a response to a large-to-small substitution (for review, see Matthews 1995). Its occurrence here may reflect the small size of the insulin monomer and the key role played by the side chain of Ile^{A2} in tethering the A1–A8 α-helix to the hydrophobic core. Our results are in accord with a prescient early study of A2 and A19 analogs that proposed that van der Waals interactions between the side chains of Ile^{A2} and Tyr^{A19} are essential to insulin's structure and function (Kitagawa et al. 1984b). The economy of insulin's design leaves little "redundancy" within its network of tertiary contacts. The likelihood of structural changes among mutant insulins highlights the need for structural studies when interpreting relative binding activities. Insulin's putative receptor-binding surface as defined by alanine scanning mutagenesis (Kristensen et al. 1997) does not resolve structural effects from the role of individual side chains in receptor recognition. The very low activity of Ala^{A2}-DKP-insulin, for example, is likely to reflect both segmental unfolding and loss of a specific contact between the Ile^{A2} and the insulin receptor. Packing of the Ile^{A2} side chain in the hormone-receptor interface (Pullen et al. 1976; Baker et al. 1988) is proposed to require a conformational change in the B chain (Derewenda et al. 1991; Hua et al. 1991; Nakagawa and Tager 1992). Mutations in the B chain may affect activity by hindering or facilitating this putative change (Kobayashi et al. 1982; Mirmira and Tager 1989;

Table 5. *DG/SA restraints and statistical parameters*

	Ala^{A2}-DKP-insulin	DKP-insulin
NOE restraints		
total[a]	367	613
sequential	119	200
medium	97	125
long-range	101	113
intra-residue	50	175
Dihedral angle restraints	32	42
ϕ-angles	21	27
χ_1-angles	11	15
Hydrogen-bond	17	23
Main chain (RMSD)		
α-helices[b]	0.31 Å	0.35 Å
A-chain (2–8)	1.51 Å	0.71 Å
A-chain (9–20)	0.46 Å	0.51 Å
B-chain[c]	0.46 Å	0.42 Å
Average restrain violations		
NOE violation	0.042 Å	0.035 Å
dihedral angle violation	0.95°	0.57°
Deviations from ideal covalent geometry		
bond length	0.006 Å	0.007 Å
bond angle	0.98°	0.99°
Empirical energy function[d] (kcal/mole)		
NOE restraint energy	36 ± 6	39 ± 6
van der Waals	18 ± 4	21 ± 2
improper dihedral angles	8 ± 1	11 ± 2
constrainted dihedral angles	2 ± 1	1 ± 1
covalent bond lengths	30 ± 1	31 ± 1
bond angles	73 ± 6	79 ± 5

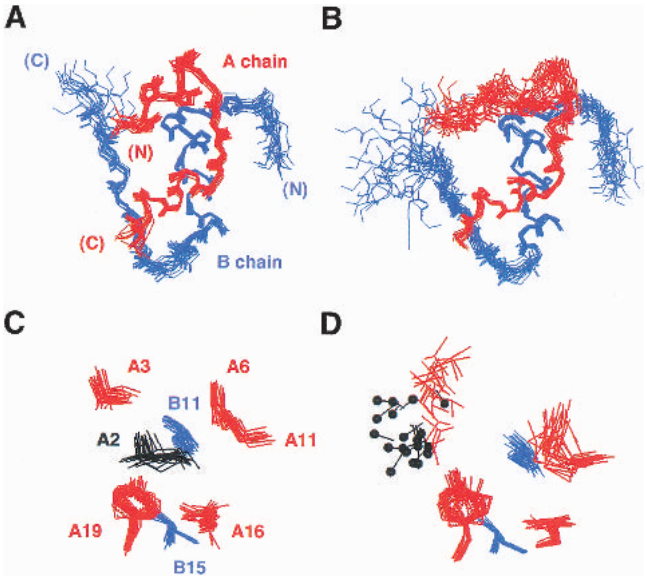
[a] The solution structure of DKP insulin is as previously described (Hua et al. 1996a); the restraint list has been refined in the present study based on additional data at pH 7.6 and 32°C. Difference in total reflects primarily a difference in use of intra-residue restraints of low information content.
[b] Helices: A13-20 and B9-19.
[c] RMSD calculated according to C_α of B4-24.
[d] NOE and dihedral force constants were 40 kcal $Å^{-2}$ and 40 kcal $radian^{-2}$, respectively (1 kcal = 4.18 kJ).

Hua et al. 1991; Mirmira et al. 1991; Kurapkat et al. 1997). In contrast, insulin's A1–A8 α-helix is thought to function as a preformed recognition element (Weiss et al. 2000). Because insulin's active structure is not well understood, detailed interpretation of relative activities will require crystallographic analysis of a hormone-receptor complex.

Cavity-forming mutations in globular proteins

Loss-of-volume substitutions in cores of globular proteins often create novel cavities or crevices (Eriksson et al. 1992, 1993). Although the size of packing defects depends on the extent of local conformational adjustments, impaired packing interactions are associated with decreased thermodynamic stability (Eriksson et al. 1992). The extent of destabilization varies from case to case. Structural studies of T4 phage lysozyme variants have indicated a correlation between cavity size and thermodynamic impairment (Eriksson et al. 1992; Xu et al. 1998). Extensive thermodynamic studies of alanine substitutions have been described in T4 phage lysozyme (Xu et al. 1998), Arc repressor (Milla et al. 1994), gene 5 protein (Terwilliger 1995), and human growth hormone (Wells 1994). In relation to such studies, Ala^{A2}-DKP-insulin shows a typical thermodynamic decrement (0.4–1.2 kcal/mole relative to DKP-insulin; Table 2). Because the partial fold of Ala^{A2}-DKP-insulin presumably represents its ground state, however, creation of a crevice in a hypothetical native-like structure (estimated to be 75 $Å^3$ if a native-like A1–A8 α-helix would have been maintained) must impose a larger thermodynamic penalty than that associated with partial unfolding. The relationship between cavity volume and stability proposed by Matthews and coworkers (Eriksson et al. 1992) would have predicted a corresponding free-energy decrement of 1.8–2.5 kcal/mole.

The physical origins of the Ala^{A2} analog's thermodynamic decrement are likely to be complex, the net consequence of simultaneous and compensating changes in configurational entropy and enthalpy on the one hand, and solvation entropy and enthalpy on the other. Unfavorable contributions to the observed $\Delta\Delta G_u$ presumably include



Fig, 5. Structures of DKP-insulin and Ala^{A2}-DKP-insulin. (*A*) Solution structure of DKP-insulin (Hua et al. 1996a) is similar to the crystallographic T-state (Baker et al. 1988). (*B*) Solution structure of Ala^{A2}-DKP-insulin has disordered N-terminal A-chain segment (A1–A8) and more divergent C-terminal B-chain tail (B26–B30), whereas remainder of molecule is essentially unchanged. (*C*) Environment of Ile^{A2} in DKP-insulin ensemble is similar to that in collection of crystal structures (see Fig. 1C). (*D*) Environment of Ala^{A2} in Ala^{A2}-DKP-insulin. Ala^{A2} and Val^{A3} are disordered, whereas the A6–A11 disulfide bridge is less-well converged; the remainder of the hydrophobic core is similar to that of DKP-insulin. β-Carbons of Ala^{A2} side chain are shown as black balls. Structures in panels *A–D* are aligned with respect to main-chain atoms of residues A13–A20 and B9–B19.

loss of van der Waals interactions in the perturbed hydrophobic core. Favorable contributions presumably arise from changes in configurational entropy and from the substitution of an amino acid of low intrinsic α-helical propensity (the native isoleucine) by an amino acid of high intrinsic α-helical propensity (alanine; O'Neil and DeGrado 1990; Chakrabartty et al. 1994). Changes in solvation free energy (Eisenberg and McLachlan 1986) are difficult to estimate in the context of segmental instability.

Studies of model globular proteins have illuminated how surrounding structure can adjust to large-to-small substitutions. In general, such changes are modest and localized to the neighborhood of the substitution. The extent of change appears to depend on how rigid or extensive is the network of surrounding interactions (Xu et al. 1998). The largest structural change characterized in a series of T4 lysozyme variants (main-chain RMSD 0.3 Å) involves a hinge-bending displacement leading to closure of the predicted cavity (Xu et al. 1998). This substitution (L84A) occurs at a side chain that, like Ile^{A2} in insulin, is close to the protein surface and engaged in a tertiary contact of limited size. It is possible that crystallization of such mutant proteins can in itself enhance order, as indicated in principle by laser Raman spectroscopy: line widths of globular proteins in solution are broader than in a single crystal[5] (Altose et al. 2001).

Possible implications for folding, misfolding, and function

Studies of the oxidative refolding of a single-chain insulin precursor (analogous to proinsulin; Qiao et al. 2001) and IGF-Is (Hober et al. 1992; Miller et al. 1993; Hua et al. 1996b) indicate a nonrandom pathway of disulfide pairing. Analysis of equilibrium models obtained by pairwise substitution of cystines by alanine or serine has indicated that successive disulfide pairing is accompanied by stepwise stabilization of native-like structural elements (Narhi et al. 1993; Hua et al. 1996b; Hober et al. 1997; Weiss et al. 2000). Because insulin's isolated A and B chains contain sufficient information to specify the folding of proinsulin (Wang and Tsou 1991), structures of two-chain analogs have been investigated as peptide models of protein-folding intermediates (Oas and Kim 1988). Of particular interest are the partial folds of insulin analogs lacking the A6–A11 disulfide bridge (*des*-[A6–A11]^{Ser}-DKP-insulin and *des*-[A6–A11]^{Ala}-DKP-insulin; Hua et al. 1996b; Weiss et al. 2000). These analogs are remarkable for segmental unfolding of the A1–A11 segment. Although the extent of disorder and thermodynamic instability are more marked than those observed here, in each case the response of the molecule is segmental and characterized by detachment of the A2 side chain from the hydrophobic core. Structures of the disulfide analogs indicate that one face of the corresponding intermediate's hydrophobic core provides an internal template for the A1–A8 coil→helix transition (Hua et al. 1996a), that is, as an "internal template" to align the A6 and A11 thiolates for specific disulfide pairing (Fig. 1B). The structure of this template, as inferred from the crystal structure of the native T state, is shown in Figure 6. The A1–A4 portion of the α-helix (yellow ribbon in Fig. 6A) occupies a deep groove in the surface of the monomer. The floor of the groove is occupied by invariant side chains Tyr^{A19} and

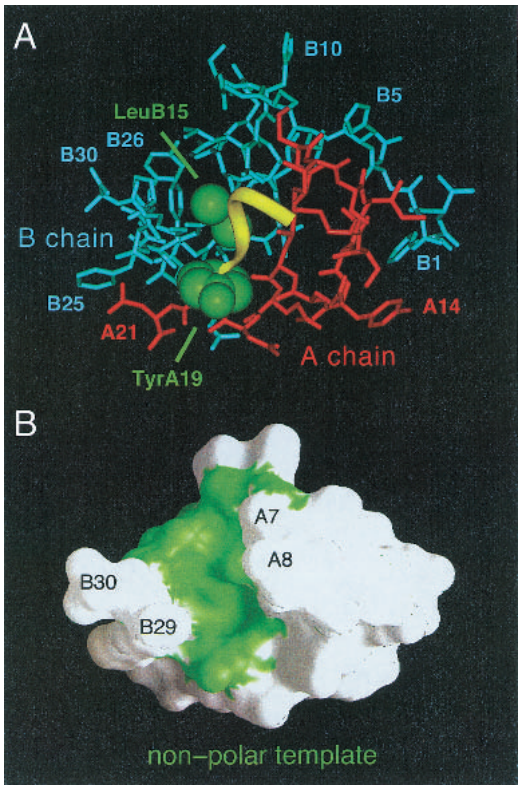


Fig. 6. Structure of the A1–A8 α-helix with respect to its nonpolar template. (*A*) Stick representation of crystallographic T-state protomer (Baker et al. 1988) showing A1–A5 (yellow ribbon) and underlying nonpolar template, largely defined by the side chains of Tyr^{A19} and Leu^{B15} (green, shown in space-filling representation) as well as of Leu^{B11} and Tyr^{B26}. Image was generated using InsightII (Biosym, Inc.). (*B*) Surface representation of template (green) generated by GRASP (Nicholls et al. 1991). For clarity, residues A1–A5 were removed to highlight the underlying nonpolar groove. In oxidative folding of proinsulin and insulin-related growth factors, this groove is proposed to provide a template to direct folding of the A1–A8 α-helix as a late event (Hua et al. 1996a,b; Weiss et al. 2000).

[5]This issue has been addressed in comparative crystallographic (Eriksson et al. 1993) and NMR studies (Mulder et al. 2000) of the L99A cavity mutant of T4 lysozyme (Eriksson et al. 1992, 1993). The L99A substitution is of particular interest because it creates a large internal cavity capable of binding ligands. The variant crystal structure is similar to that of wild type, including in its pattern of thermal B factors. In solution, however, studies of protein dynamics by heteronuclear NMR relaxation methods show that, whereas no change occurs on the picosecond-to-nanosecond time scales, conformational exchange is accentuated on the microsecond-to-millisecond time scale. Such cavity-associated motions involve an extensive portion of the protein surrounding the cavity (Mulder et al. 2000).

Leu^{B15}. The internal-template hypothesis raised the possibility that—even with the three native cystines intact—the A1–A8 α-helix is not integral to the structure of the insulin monomer (Weiss et al. 2000). The structure of Ala^{A2}-DKP-insulin supports this hypothesis.

Modularity of the A1–A8 segment implies that, regardless of its local state of organization, native-like structure is elsewhere maintained. This finding extends to the A-chain observations well-established in studies of the B chain. Crystallographic studies of insulin hexamers have shown, for example, that insulin's overall structure is compatible with alternative B1–B8 configurations: either extended strand (as in the T-state; Baker et al. 1988) or α-helix[6] (the R/R_f state; Bentley et al. 1976; Smith et al. 1984). Further, the monomer's globular α-helical domain is unaffected by truncation or detachment of the B chain's C-terminal β-strand (B24-B30; Bao et al. 1997). A degree of modularity of the A1–A8 α-helix was previously inferred from comparison of crystal structures (Chothia et al. 1983). The T→R transition is accompanied by an approximately 20° rotation of this helix, which in turn requires multiple adjustments in core packing. The concerted series of conformational changes among insulin hexamers, including variable packing of the A1–A8 α-helix, has been analyzed as a model for the transmission of conformational change in proteins (Chothia et al. 1983).

The biological "life cycle" of insulin is likely to require a series of conformational changes. Although the A1–A8 segment appears to function in receptor binding as a recognition α-helix (Pullen et al. 1976; Baker et al. 1988; Hua et al. 1996a; Weiss et al. 2000), nonhelical configurations may be important in prohormone processing. Cleavage of the junction of the C-peptide-portion (Arg^{A0}) of proinsulin and Gly^{A1} (Duguay et al. 1997) is effected by a subtilisin-related converting enzyme conserved among neuroendocrine processing proteases (Lipkind and Steiner 1999). Such enzymes are thought to accept extended strands—but not α-helices—into their active sites (Lipkind and Steiner 1999). In such a complex, the side chain of Ile^{A2} is itself proposed to dock within a conserved pocket of the converting enzyme rather than within proinsulin's hydrophobic core (Lipkind and Steiner 1999). Accordingly, we suggest that the structure of Ala^{A2}-DKP-insulin provides a model of how fluctuations in the structure of proinsulin may enable its docking with the converting enzyme. Selective unfolding of the A1–A8 segment may also facilitate partial exposure of hydrophobic surfaces, leading to aberrant aggregation as intermediates in insulin fibrillation (Brange et al. 1997; Kelly et al. 1997; Bouchard et al. 2000; Nettleton et al. 2000; Sipe and Cohen 2000; Nielsen et al. 2001). This process would be analogous to the ordered assembly of β-sheet-rich structures encountered in pathological amyloidogenesis (Orpiszewski and Benson 1999). Because insulin fibrillation is accelerated by partial thermal unfolding, it is of future interest to investigate the structure and dynamics of the A1–A8 segment in native insulin at elevated temperatures and effects of the Ile^{A2}→Ala substitution on rates of fibrillation.

[6]The R_f state, observed in T_3R_{f3} hexamers, differs from the R state at residues B1–B3. Whereas R_6 hexamers contain a B1–B19 α-helix, the R_f state contains a B3–B19 or B4–B19 α-helix with frayed N-termini.

Materials and methods

Materials

4-Methylbenzhydrylamine resin (0.6 mmole of amine/g; Star Biochemicals, Inc.) was used as solid support for synthesis of the A-chain analog; (N-butoxy-carbonyl, O-benzyl)-threonine-PAM resin, (0.56 mmole/g; Bachem, Inc.) was used as solid support for synthesis of the B-chain analog. *tert*-Butoxycarbonyl-amino acids and derivatives were obtained from Bachem and Peninsula Laboratories; N,N′-dicyclohexylcarbodiimide and N-hydroxybenzotriazole (recrystallized from 95% ethanol) from Fluka. Chromatography resins were preswollen microgranular carboxymethylcellulose (CM-cellulose; Whatman CM52), DE53 cellulose (Whatman) and Cellex E (Ecteola cellulose; Sigma); solvents were HPLC grade.

Peptide synthesis

The general protocol for solid-phase synthesis is as described (Barany and Merrifield 1980). The C-terminal Asn in the synthesis of the A chain was incorporated into solid support by coupling *tert*-butoxycarbonyl aspartic acid α-benzyl ester with 4-methylbenzhydrylamine resin. After the final deprotection, the Asp residue was converted to an Asn residue.

(i) Synthetic A-chain S-sulfonate

Peptidyl resin (0.82 g), after deblocking, sulfitolysis, and chromatographic purification (Hu et al. 1993; Hua et al. 1996a), yielded ~244 mg of purified S-sulfonated Ala^{A2} A-chain variant.

(ii) Synthetic B-chain S-sulfonate

After deblocking, sulfitolysis, and chromatographic purification, 610 mg of peptidyl resin yielded ~125 mg of purified S-sulfonated B chain. Amino-acid analyses were in agreement with expected values.

Peptide purification

Crude S-sulfonated A chain was purified by chromatography on a Cellex E column (1.5 × 47 cm) as described (Hu et al. 1993; Hua et al. 1996a), dialyzed against distilled water, and lyophilized to yield the purified Ala^{A2} A-chain S-sulfonate. Crude S-sulfonated [Asp^{B10}, Lys^{B28}, Pro^{B29}] B chain was likewise purified on a cellulose DE53 column (1.5 × 47 cm), dialyzed, and lyophilized to yield the DKP B-chain S-sulfonate.

Chain recombination

Chain recombination used S-sulfonated A- and B chains (approximately 2:1 by weight) in 0.1 M glycine (pH 10.6) in the presence of dithiothreitol (Chance et al. 1981). The DKP-insulin analog was isolated from the combination mixture as described (Hu et al. 1993; Hua et al. 1996a) and purified on a 0.9 × 23 cm CM-cellulose chromatography and rp-HPLC on a Vydac 218 TP column (0.46 × 25 cm); the latter used a flow rate of 0.5 mL/min with 20&–80% linear gradient of 80% aqueous acetonitrile containing 0.1% trifluoroacetic acid over 80 min. Rechromatography of this material on reverse-phase HPLC (rp-HPLC) under the same conditions gave a single sharp peak. Amino acid analysis and mass spectrometry gave expected values.

Biological assays

For insulin receptor-binding studies, plasma membranes were partially purified from human placenta. [^{125}I]-insulin was purchased from Dupont NEN. Assays were performed at 4°C as described (Marshall et al. 1974; Cara et al. 1990; Weiss et al. 2001) with relative activity defined as the ratio of human insulin to analog required to displace 50% of specifically bound ^{125}I-insulin.

Spectroscopy

^{1}H-NMR spectra were obtained at 600 MHz at 25°C in 50 mM potassium phosphate (pH 7) and in 20% deuteroacetic acid (pH 1.9) as described (Hua and Weiss 1991; Hua et al. 1996a); the protein concentration was 1.5 mM. No significant differences were observed in the pattern of interresidue NOEs and chemical shifts at pH 7 or pH 1.9 for Ala^{A2}-DKP, comparing with those for DKP-insulin (Hua et al. 1996a). The range of pH conditions enables resonances overlapping in one spectrum to be resolved in another. Acidic pH facilitates analysis of amide resonances, some of which are attenuated at pH 7 as a result of base-catalyzed solvent exchange. Resonance assignment was based on 2D NOESY (mixing times 100 and 200 ms), total correlated spectroscopy (TOCSY) (mixing times 30 and 55 ms), and double-quantum filtered correlated spectroscopy (DQF-COSY) spectra. Spectra in H_2O were obtained using pulse-field gradients and laminar shaped pulses (Hua et al. 1996a). CD spectra were obtained using an Aviv spectropolarimeter equipped with thermistor temperature control and automated titration unit for guanidine denaturation studies. CD samples contained 25–50 μM insulin or analog in 50 mM potassium phosphate (pH 7); samples were diluted to 5 μM for equilibrium denaturation studies.

Thermodynamic modeling

Guanidine denaturation data were fitted by nonlinear least squares to a two-state model as described (Sosnick et al. 2000; Weiss et al. 2000). In brief, CD data $\theta(x)$, where x indicates the concentration of denaturant, were fitted by a nonlinear least-squares program according to

$$\theta(x) = \frac{\theta_A + \theta_B e^{(-\Delta G^o_{H_2O} - mx)/RT}}{1 + e^{-(\Delta G^o_{H_2O} - mx)/RT}}$$

where x is the concentration of guanidine and where θ_A and θ_B are baseline values in the native and unfolded states as approximated by pre- and post-transition lines $\theta_A(x) = \theta_A{}^{H_2}O + m_A x$ and $\theta_B(x)$ &equals $\theta_B{}^{H_2}O + m_B x$. Fitting the original CD data and base lines simultaneously circumvents artifacts associated with linear plots of ΔG as a function of denaturant according to $\Delta G^O(x) = \Delta G_{H_2}{}^O + m^O x$ (for review, see Sosnick et al. 2000). Nonetheless, the m value obtained in fitting the variant unfolding curve is significantly lower than the m value obtained in fitting the wild-type unfolding curve (DKP-insulin; see Table 2). This situation can be associated with an underestimate of the Ala^{A2} analog's stability. The analog's lower m value may reflect its greater exposed hydrophobic surface in the absence of denaturant or existence of a native-state ensemble containing a distribution of incompletely folded forms of differing stability. Analysis of unfolding curves in this setting has recently been considered by Luo and Baldwin in a study of an equilibrium MG (apomyoglobin; Luo and Baldwin 2001). In brief, the apparent C_{mid} is interpreted as a mean value

$$\langle C_m \rangle = \sum \alpha_i C_{mi}$$

where each species in the ensemble makes a fractional contribution to the weight-averaged unfolding curve. If the distribution of stabilities in the mutant ensemble is small relative to its mean decrement relative to the wild type, then the variant's stability may be estimated in a linear-extrapolation model by the product of the wild-type slope and $\langle C_m \rangle$. This approximation fails if only a small fraction of the variant ensemble contributes to the experimental probe of foldedness. Because insulin does not show a detectable change in tyrosine fluorescence on unfolding, concordance of fluorescent and codeleted unfolding curves could not be evaluated as a criterion of apparent two-state MG unfolding as proposed by Luo and Baldwin (2001).

Molecular modeling

The volume and location of inferred protein cavities were obtained using the program SURFNET (Laskowski 1995). Analysis of polar and nonpolar protein surfaces was obtained using the Connolly surface feature implemented in InsightII (Biosym, Inc.).

Electronic supplemental material

Supplemental information contains one figure depicting changes in ^{1}H-NMR secondary shifts between Ala^{A2}-DKP-insulin and DKP-insulin by proton class in 20% acetic acid at pH 1.9 and 25°. Tables S1 and S2 provide chemical shifts of assigned ^{1}H NMR resonances of Ala^{A2}-DKP for conditions corresponding to Figures 3 and S1, respectively. Table S3 is a list of DG/SA restraints for Ala^{A2}-DKP. The atomic coordinates for Ala^{A2}-DKP-insulin have been deposited in the RCSB Protein Data Bank (PDB number 1K3M).

Acknowledgments

This work was supported in part by grants from the National Institute of Health to P.G.K. and M.A.W. and by the Diabetes Research and Training Center at the University of Chicago. We thank P. DeMeyts, G.G. Dodson, E. Dodson, D.F. Steiner, J. Whittingham, and the late H.S. Tager for helpful discussion.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby

marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Altose, M.D., Zheng, Y., Dong, J., Palfey, B.A., and Carey, P.R. 2001. Comparing protein-ligand interactions in solution and single crystals by Raman spectroscopy. *Proc. Natl. Acad. Sci.* **98:** 3006–3011.

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 Crystallographica Sec. B, Struct. Sci.* **47:** 127–136.

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. *Phil. Trans. Royal Soc. London Ser.* **319:** 369–456.

Bao, S.J., Xie, D.L., Zhang, J.P., Chang, W.R., and Liang, D.C. 1997. Crystal structure of *des*heptapeptide(B24-B30)insulin at 1.6 Å resolution: Implications for receptor binding. *Proc. Natl. Acad. Sci.* **94:** 2975–2980.

Barany, G. and Merrifield, R.B. 1980. In *The peptides* (eds. E. Gross and J. Meienhofer), pp. 273–284. Academic Press, New York.

Barbato, G., Ikura, M., Kay, L.E., Pastor, R.W., and Bax, A. 1992. Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible. *Biochemistry* **31:** 5269–5278.

Bentley, G., Dodson, E., Dodson, G., Hodgkin, D., and Mercola, D. 1976. Structure of insulin in 4-zinc insulin. *Nature* **261:** 166–168.

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.

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.

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.

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.

Brems, D.N., Brown, P.L., Heckenlaible, L.A., and Frank, B.H. 1990. Equilibrium denaturation of insulin and proinsulin. *Biochemistry* **29:** 9289–9293.

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.

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.

Chakrabartty, A., Kortemme, T., and Baldwin, R.L. 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* **3:** 843–852.

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* (eds. D.H. Rich and E. Gross), pp. 721–728. Pierce Chemical Co., Rockford, IL.

Chothia, C., Lesk, A.M., Dodson, G.G., and Hodgkin, D.C. 1983. Transmission of conformational change in insulin. *Nature* **302:** 500–505.

Ciszak, E. and Smith, G.D. 1994. Crystallographic evidence for dual coordination around zinc in the T3R3 human insulin hexamer. *Biochemistry* **33:** 1512–1517.

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 $Lys^{B28}Pro^{B29}$-human insulin. *Structure* **3:** 615–622.

Clore, G.M., Robien, M.A., and Gronenborn, A.M. 1993. Exploring the limits of precision and accuracy of protein structures determined by nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **231:** 82–102.

Dai, J.-B., Lou, M.-Z., You, J.-M., and Liang, D.-C. 1987. Refinement of the structure of *des*pentapetide (B26-B30) insulin at 1.5 Å resolution. *Sci. Sin.* **30:** 55–65.

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.

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.

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., Slieker, L.J., Su, K.S.E., Sundell, K.L., and Chance, R.E. 1992. In *Peptides: Proceedings of the Twelfth American Peptide Symposium* (eds. J.A. Smith and J.E. Rivier), pp. 26–28. ESCOM Science Publishers B.V., Leiden, The Netherlands.

Dodson, G. and Steiner, D. 1998. The role of assembly in insulin's biosynthesis. *Curr. Opin. Struct. Biol.* **8:** 189–194.

Dolgikh, D.A., Abaturov, L.V., Bolotina, I.A., Brazhnikov, E.V., Bychkova, V.E., Gilmanshin, R.I., Lebedev Yu, O., Semisotnov, G.V., Tiktopulo, E.I., and Ptitsyn, O.B. 1985. Compact state of a protein molecule with pronounced small-scale mobility: Bovine α-lactalbumin. *Eur. Biophys. J.* **13:** 109–121.

Duguay, S.J., Milewski, W.M., Young, B.D., Nakayama, K., and Steiner, D.F. 1997. Processing of wild-type and mutant proinsulin-like growth factor-IA by subtilisin-related proprotein convertases. *J. Biol. Chem.* **272:** 6663–6670.

Eisenberg, D. and McLachlan, A.D. 1986. Solvation energy in protein folding and binding. *Nature* **319:** 199–203.

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.

Eriksson, A.E., Baase, W.A., and Matthews, B.W. 1993. Similar hydrophobic replacements of Leu99 and Phe153 within the core of T4 lysozyme have different structural and thermodynamic consequences. *J. Mol. Biol.* **229:** 747–769.

Goto, N.K., Skrynnikov, N.R., Dahlquist, F.W., and Kay, L.E. 2001. What is the average conformation of bacteriophage T4 lysozyme in solution? A domain orientation study using dipolar couplings measured by solution NMR. *J. Mol. Biol.* **308:** 745–764.

Griko, Y.V., Privalov, P.L., Venyaminov, S.Y., and Kutyshenko, V.P. 1988. Thermodynamic study of the apomyoglobin structure. *J. Mol. Biol.* **202:** 127–138.

Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. 1992. Disulfide exchange folding of insulin-like growth factor I. *Biochemistry* **31:** 1749–1756.

Hober, S., Uhlen, M., and Nilsson, B. 1997. Disulfide exchange folding of disulfide mutants of insulin-like growth factor I in vitro. *Biochemistry* **36:** 4616–4622.

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.

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.

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.

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.* **89:** 2379–2383.

Hua, Q.X., Ladbury, J.E., and Weiss, M.A. 1993. Dynamics of a monomeric insulin analogue: Testing the molten-globule hypothesis. *Biochemistry* **32:** 1433–1442.

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. 1996a. Mapping the functional surface of insulin by design: Structure and function of a novel A-chain analogue. *J. Mol. Biol.* **264:** 390–403.

Hua, Q.X., Narhi, L., Jia, W., Arakawa, T., Rosenfeld, R., Hawkins, N., Miller, J.A., and Weiss, M.A. 1996b. 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.

Jacoby, E., Hua, Q.X., Stern, A.S., Frank, B.H., and Weiss, M.A. 1996. Structure and dynamics of a protein assembly. 1H-NMR studies of the 36 kDa R6 insulin hexamer. *J. Mol. Biol.* **258:** 136–157.

Kelly, J.W., Colon, W., Lai, Z., Lashuel, H.A., McCulloch, J., McCutchen, S.L., Miroy, G.J., and Peterson, S.A. 1997. Transthyretin quaternary and tertiary structural changes facilitate misassembly into amyloid. *Adv. Protein Chem.* **50:** 161–181.

Kitagawa, K., Ogawa, H., Burke, G.T., Chanley, J.D., and Katsoyannis, P.G. 1984a. 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.

Kitagawa, K., Ogawa, H., Burke, G.T., Chanley, J.D., and Katsoyannis, P.G. 1984b. 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.

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.

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.

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.

Laskowski, R.A. 1995. SURFNET: A program for visualizing molecular surfaces, cavities and intermolecular interactions. *J. Molec. Graphics* **13:** 323–330.

Lipkind, G. and Steiner, D.F. 1999. Predicted structural alterations in proinsulin during its interactions with prohormone convertases. *Biochemistry* **38:** 890–896.

Liu, Y., Zhao, D., Altman, R., and Jardetzky, O. 1992. A systematic comparison of three structure determination methods from NMR data: Dependence upon quality and quantity of data. *J. Biomol. NMR* **2:** 373–388.

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.

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.

Matthews, B.W. 1995. Studies on protein stability with T4 lysozyme. *Adv. Protein Chem.* **46:** 249–278.

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.

Miller, J.A., Narhi, L.O., Hua, Q.X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K.S., Tsai, L., and Weiss, M.A. 1993. Oxidative refolding of insulin-like growth factor 1 yields two products of similar thermodynamic stability: a bifurcating protein-folding pathway. *Biochemistry* **32:** 5203–5213.

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.

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.

Mulder, F.A., Hon, B., Muhandiram, D.R., Dahlquist, F.W., and Kay, L.E. 2000. Flexibility and ligand exchange in a buried cavity mutant of T4 lysozyme studied by multinuclear NMR. *Biochemistry* **39:** 12614–12622.

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.

Narhi, L.O., Hua, Q.X., Arakawa, T., Fox, G.M., Tsai, L., Rosenfeld, R., Holst, P., Miller, J.A., and Weiss, M.A. 1993. Role of native disulfide bonds in the structure and activity of insulin-like growth factor 1: Genetic models of protein-folding intermediates. *Biochemistry* **32:** 5214–5221.

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.

Nicholls, A., Sharp, K.A., and Honig, B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11:** 281–296.

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.

Oas, T.G. and Kim, P.S. 1988. A peptide model of a protein folding intermediate. *Nature* **336:** 42–48.

O'Neil, K.T. and DeGrado, W.F. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* **250:** 646–651.

Orpiszewski, J. and Benson, M.D. 1999. Induction of β-sheet structure in amyloidogenic peptides by neutralization of aspartate: A model for amyloid nucleation. *J. Mol. Biol.* **289:** 413–428.

Pace, C.N. and Shaw, K.L. 2000. Linear extrapolation method of analyzing solvent denaturation curves. *Proteins* (Suppl.) **4:** 1–7.

Peng, J.W. and Wagner, G. 1992. Mapping of the spectral densities of N-H bond motions in eglin c using heteronuclear relaxation experiments. *Biochemistry* **31:** 8571–8586.

Pfeil, W. 1981. Thermodynamics of α-lactalbumin unfolding. *Biophys. Chem.* **13:** 181–186.

Pfeil, W., Bychkova, V.E., and Ptitsyn, O.B. 1986. Physical nature of the phase transition in globular proteins. Calorimetric study of human α-lactalbumin. *FEBS Lett.* **198:** 287–291.

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.

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.

Sandberg, W.S. and Terwilliger, T.C. 1989. Influence of interior packing and hydrophobicity on the stability of a protein. *Science* **245:** 54–57.

Schulman, B.A. and Kim, P.S. 1996. Proline scanning mutagenesis of a molten globule reveals non-cooperative formation of a protein's overall topology. *Nat. Struc. Biol.* **3:** 682–687.

Schulman, B.A., Kim, P.S., Dobson, C.M., and Redfield, C. 1997. A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat. Struct. Biol.* **4:** 630–634.

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.

Sipe, J.D. and Cohen, A.S. 2000. Review: History of the amyloid fibril. *J. Struct. Biol.* **130:** 88–98.

Smith, G.D., Swenson, D.C., Dodson, E.J., Dodson, G.G., and Reynolds, C.D. 1984. Structural stability in the 4-zinc human insulin hexamer. *Proc. Natl. Acad. Sci.* **81:** 7093–7097.

Sosnick, T.R., Fang, X., and Shelton, V.M. 2000. Application of circular dichroism to study RNA folding transitions. *Meth. Enzymol.* **317:** 393–409.

Terwilliger, T.C. 1995. Engineering the stability and function of gene V protein. *Adv. Protein. Chem.* **46:** 177–215.

Tjandra, N. and Bax, A. 1997. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278:** 1111–1114.

Tjandra, N., Omichinski, J.G., Gronenborn, A.M., Clore, G.M., and Bax, A. 1997. Use of dipolar 1H-15N and 1H-13C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat. Struct. Biol.* **4:** 732–738.

Tjandra, N., Marquardt, J., and Clore, G.M. 2000. Direct refinement against proton-proton dipolar couplings in NMR structure determination of macromolecules. *J. Magn. Reson.* **142:** 393–396.

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.

Wang, S.H., Hu, S.Q., Burke, G.T., and Katsoyannis, P.G. 1991. Insulin analogues with modifications in the β-turn of the B-chain. *J. Protein Chem.* **10:** 313–324.

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 ^{2}H and ^{13}C labeling with application to protein design. *Biochemistry* **30:** 7373–7389.

Weiss, M.A., Hua, Q.-X., Jia, W., Chu, Y.-C., Wang, R.-Y., and Katsoyannis, P.G. 2000. Hierarchiacal protein 'un-design': Insulin's intrachain disulfide bridge tethers a recognition α-helix. *Biochemistry* **39:** 15429–15440.

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 analogs at position A8 are uncorrelated with their thermodynamic stabilities. *J. Biol. Chem.* **276:** 40018–40024.

Wells, J.A. 1994. Structural and functional basis for hormone binding and receptor oligomerization. *Curr. Opin. Cell Biol.* **6:** 163–173.

Wüthrich, K. 1986. *NMR of proteins and nucleic acids.* John Wiley and Sons, New York.

Xu, J., Baase, W.A., Baldwin, E., and Matthews, B.W. 1998. The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. *Protein Sci.* **7:** 158–177.

Xu, J., Baase, W.A., Quillin, M.L., Baldwin, E.P., and Matthews, B.W. 2001. Structural and thermodynamic analysis of the binding of solvent at internal sites in T4 lysozyme. *Protein Sci.* **10:** 1067–1078.