Hierarchical Protein "Un-Design": Insulin's Intrachain Disulfide Bridge Tethers a Recognition α-Helix†

Michael A. Weiss,*,‡ Qing-Xin Hua,‡ Wenhua Jia,‡ Ying-Chi Chu,§ Run-Ying Wang,§ and Panayotis G. Katsoyannis*,§

Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44016, and Department of Biochemistry & Molecular Biology of New York University, Mount Sinai School of Medicine, New York, New York 10029

Received August 11, 2000; Revised Manuscript Received October 20, 2000

ABSTRACT: A hierarchical pathway of protein folding can enable segmental unfolding by design. A monomeric insulin analogue containing pairwise substitution of internal A6−A11 cystine with serine [[Ser^{A6},Ser^{A11},Asp^{B10},Lys^{B28},Pro^{B29}]insulin (DKP[A6−A11]^{Ser})] was previously investigated as a model of an oxidative protein-folding intermediate [Hua, Q. X., et al. (1996) *J. Mol. Biol. 264*, 390−403]. Its structure exhibits local unfolding of an adjoining amphipathic α-helix (residues A1−A8), leading to a 2000-fold reduction in activity. Such severe loss of function, unusual among mutant insulins, is proposed to reflect the cost of induced fit: receptor-directed restoration of the α-helix and its engagement in the hormone's hydrophobic core. To test this hypothesis, we have synthesized and characterized the corresponding alanine analogue [[Ala^{A6},Ala^{A11},Asp^{B10},Lys^{B28},Pro^{B29}]insulin (DKP[A6−A11]^{Ala})]. Untethering the A6−A11 disulfide bridge by either amino acid causes similar perturbations in structure and dynamics as probed by circular dichroism and ^{1}H NMR spectroscopy. The analogues also exhibit similar decrements in thermodynamic stability relative to that of the parent monomer as probed by equilibrium denaturation studies ($\Delta\Delta G_u = 3.0 \pm 0.5$ kcal/mol). Despite such similarities, the alanine analogue is 50 times more active than the serine analogue. Enhanced receptor binding ($\Delta\Delta G = 2.2$ kcal/mol) is in accord with alanine's greater helical propensity and more favorable hydrophobic-transfer free energy. The success of an induced-fit model highlights the applicability of general folding principles to a complex binding process. Comparison of DKP[A6−A11]^{Ser} and DKP[A6−A11]^{Ala} supports the hypothesis that the native A1−A8 α-helix functions as a preformed recognition element tethered by insulin's intrachain disulfide bridge. Segmental unfolding by design provides a novel approach to dissecting structure−activity relationships.

Protein folding and recognition define central problems in biochemistry and biophysics (*1*). Important insights have emerged from classical studies of amino acids and their conformational propensities (*2*−*4*). Application of such principles to actual systems is often made difficult, however, by the diversity and complexity of protein structures (*5*). An example of broad interest is rational design of ligands for cellular receptors, a problem that ordinarily requires detailed consideration of molecular structure, dynamics, and solvation (*6*). Such features are relevant since the stability of an interface (like that of a globular protein; *7*) reflects differences between large and opposing thermodynamic driving forces (*8*). In this article, we suggest that two general properties of amino acids, empirical thermodynamic scales of hydrophobicity (*9*, *10*) and conformational propensity (*11*−*15*), can rationalize relative activities of a special class of insulin analogues. These analogues, designed by analogy to protein-folding intermediates (*16*, *17*), lack one of insulin's three disulfide bridges (*18*, *19*). A combination of total protein synthesis (*20*−*22*) and biophysical characterization is used to test the hypothesis that biological activities mirror the thermodynamics of induced fit. Our results illuminate the function and evolution of insulin's intrachain disulfide bridge.

Insulin is a small globular protein containing two chains, designated A (21 residues) and B (30 residues) (*23*). 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 (*24*). The structure of such a monomer in solution (*19*, *25*, *26*) resembles the crystallographic T state (Figure 1A; *27*), which has been defined at high resolution in a variety of crystal forms (*28*, *29*). Despite its small size, insulin contains representative features of larger proteins, including canonical elements of secondary structure and a well-ordered hydrophobic core. The native protein contains three disulfide bridges, one within the A chain (A6−A11) and the others between chains (A7−B7 and A20−B19; Figure 1A). Our study focuses on the A6−A11 disulfide bridge, which is situated in the hydrophobic core between aliphatic side chains of both chains (Figure 1B). These neighboring residues (Ile^{A2}, Val^{A3}, Leu^{B6}, and Leu^{B11}) are invariant among vertebrate insulins and insulin-like growth

† This work was supported by grants from the extramural diabetes program of the National Institutes of Health to M.A.W. and P.G.K.

* To whom correspondence should be addressed at Case Western Reserve University School of Medicine (M.A.W.) or the Mount Sinai School of Medicine (P.G.K.).

‡ Case Western Reserve University School of Medicine.

§ Mount Sinai School of Medicine.

10.1021/bi001905s CCC: $19.00
Published on Web 11/23/2000

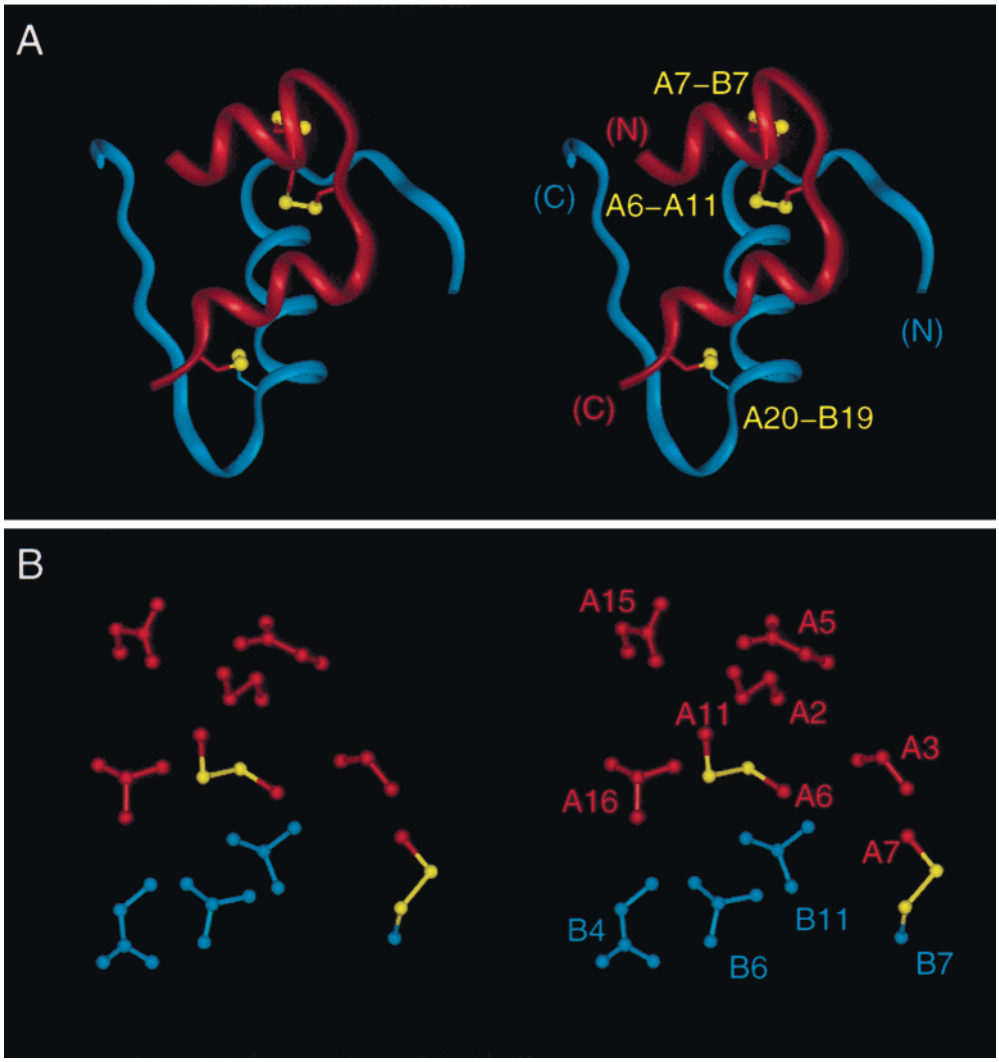

FIGURE 1: Structure of insulin. (A) Stereo ribbon representation of the T state protomer of porcine insulin in the T_6 crystal structure (two-Zn molecule 1 in Chinese nomenclature; PDB entry 4INS). The A chain is shown in red and the B chain in blue, and disulfide bridges are shown in yellow (balls and sticks). Sulfur atoms are shown with radii that are 40% of the van der Waals radius. The solution structure of DKP-insulin is similar to the crystallographic T state (*27*). (B) Ball-and-stick stereo model of side chains illustrating the nonpolar environment of the A6−A11 cystine in the hydrophobic core of the T state. The A7−B7 cystine exposed at the protein surface is also shown. Coloring scheme as described for panel A.

factors (*23*). The A20−B19 disulfide bridge is likewise buried, whereas the A7−B7 cystine lies on the protein surface. Oxidative refolding of proinsulin and insulin chain combination in each case lead to native disulfide pairing (*22*, *30*−*32*).

Insulin's receptor-binding surface has been extensively investigated by mutagenesis (*27*, *33*−*41*). Although residues in both chains are required for high activity, such analyses highlight the particular importance of residues in the N-terminal region of the A chain (Gly^{A1}, Ile^{A2}, and Val^{A3}; *33*, *34*, *40*). The $Val^{A3} \rightarrow$ Leu substitution, for example, reduces the level of receptor binding by 500-fold, a decrement larger than that of analogous mutations in the B chain.[1] This mutation (designated insulin *Wakayama*) causes a rare monogenic form of human diabetes mellitus with hyperinsulinemia (*42*). Although conformational changes are proposed to occur in the B chain on receptor binding (*25*, *27*, *32*, *36*, *38*, *39*, *43*), residues A1−A3 appear to function within the native α-helix (shaded in Figure 2A) (*19*, *44*, *45*).[1] This model is supported by characterization of a two-disulfide analogue lacking the A6−A11 cystine (*19*). Pairwise substitution of Cys^{A6} and Cys^{A11} with serine in a monomeric analogue ($DKP[A6-A11]^{Ser}$; see Table 1) leads to segmental unfolding of the A1−A8 α-helix in an otherwise native subdomain (Figure 2A). The analogue exhibits a 2000-fold reduction

Table 1: Design of Insulin Analogues[a]

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
A6 Cys → Ala or Ser	internal SS bridge	destabilize A1−A8 helix
A11 Cys → Ala or Ser	internal SS bridge	destabilize A1−A8 helix

[a] DKP-insulin contains the three B chain substitutions (*56*), whereas $DKP[A6-A11]^{Ala}$ and $DKP[A6-A11]^{Ser}$ contain five substitutions; the latter analogues retain native A7−B7 and A20−B19 disulfide bridges (*19*). The B28−B29 inversion was motivated by homology to IGF-I to destabilize insulin's classical dimer interface (*54*).

[1] Examples of anomalous low-affinity analogues are substitution of Ile^{A2} with norleucine (*33*) or *allo*-isoleucine (*40*). Although their structures have not been described, it is expected that insulin's hydrophobic core could accommodate such similar aliphatic side chains.

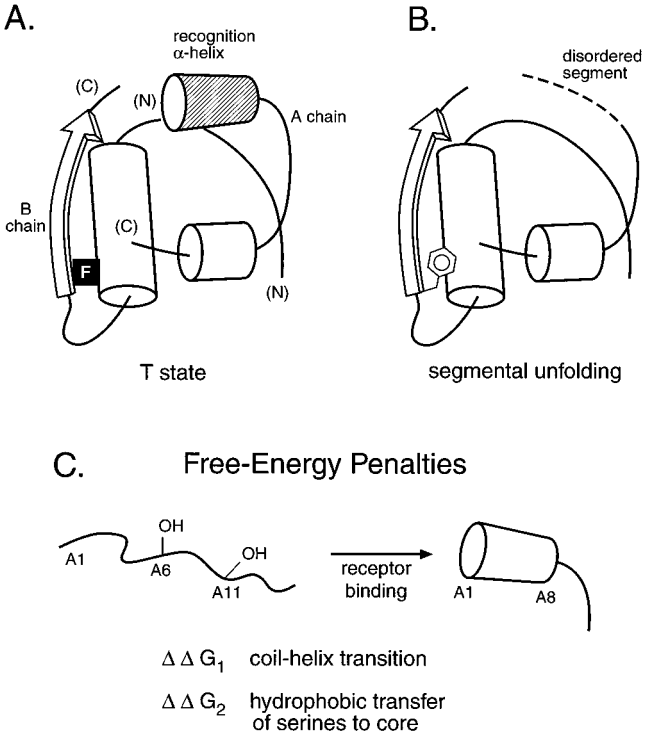

FIGURE 2: Outline of the experimental design. (A and B) Cylinder models of the engineered insulin monomer (DKP-insulin, panel A) and the two-disulfide analogue (DKP[A6−A11]^{Ser}, panel B) (*19*). The NMR structure of DKP-insulin (*19*) is similar to that of the crystallographic T state (*27*). The black square in panel A indicates the position of the Phe^{B24} side chain, drawn explicitly in panel B. The effects of detachment of the A1−A8 fragment (dashed line in panel B) contrast with the effects of substitutions at Phe^{B24}, which anchors the B chain β-strand against the central α-helix. Its direct involvement in receptor binding is unlikely as substitution with D-Phe and D-Ala enhances activity to equal extents (*36*). (C) Induced-fit model proposes receptor-directed α-helical folding of the disordered A1−A8 segment of two-disulfide analogues and its re-engagement within the hormone's hydrophobic core. Such induced fit would in principle incur costs in free energy associated with the coil → helix transition ($\Delta\Delta G_1$) and changes in solvation ($\Delta\Delta G_2$; hydrophobic-transfer free energies). The values of these residue specific penalties are estimated in parts A and B of Table 2, respectively.

in biological activity (*19*), a decrement more severe than that ordinarily associated with point mutation (*27*). Because A1−A3 and other putative receptor contacts are retained in the analogue's primary structure, its inactivity was ascribed to the free energy penalty of refolding the A1−A8 α-helix on receptor binding. The internal disulfide bridge was therefore proposed to tether the native α-helix as a preformed recognition element. It was not established, however, whether segmental unfolding in DKP[A6−A11]^{Ser} is a consequence of untethering the disulfide bridge or of the polar serine substitutions employed. The importance of this issue has been emphasized in analogous studies of bovine pancreatic trypsin inhibitor (BPTI).[2] The choice of serine or alanine in pairwise substitution of internal cystines leads in BPTI to marked differences in the variant protein's structure, dynamics, and stability (*46*−*48*). Because native structure is retained in two-disulfide BPTI analogues containing pairwise substitutions in the 30−51 cystine, for example, the substituted side chains lie in hydrophobic environments. As a consequence, the serine analogue exhibits a greater decrement in thermodynamic stability than does the alanine analogue (*48*).

The study presented here seeks to compare the effects of alanine and serine substitutions in insulin's A6−A11 disulfide bridge. Would DKP[A6−A11]^{Ala} also undergo segmental unfolding, and if so, would the two partial folds exhibit similar decrements in thermodynamic stability? Comparison of DKP[A6−A11]^{Ser} and DKP[A6−A11]^{Ala} promises to extend the BPTI paradigm and in addition to provide an opportunity to investigate a second problem: the relationship between peptide folding and induced fit in hormone−receptor recognition. Is the marked loss of activity exhibited by DKP[A6−A11]^{Ser} a necessary consequence of its segmental unfolding or due to the particular properties of serine? Because insulin's native structure is presumably regained in the variant hormone−receptor complex, we propose that the thermodynamics of binding reflect the thermodynamics of folding of an amphipathic α-helix. Accordingly, we hypothesize that the free energy penalty of induced fit has two components (Figure 2B), that due to an induced coil → helix transition ($\Delta\Delta G_1$) and that due to repacking of the serine side chains in the hormone's hydrophobic core ($\Delta\Delta G_2$). This hypothesis predicts that the corresponding alanine analogue would, if otherwise similar to the serine analogue in structure and stability, exhibit enhanced biological activity. Further, the extent of enhancement would be determined by intrinsic differences between serine and alanine in helical propensity (*11*−*13*) and hydrophobic-transfer free energy (*2*−*4*, *10*). The magnitude of these differences may readily be estimated by empirical thermodynamic scales.

To test these predictions, we have synthesized DKP[A6−A11]^{Ala}, probed its structure by circular dichroism (CD) and 2D ^{1}H NMR, and assessed its thermodynamic stability by equilibrium-denaturation studies. Our results indicate that the efficiency of chain combination (*41*), the distinctive pattern of chemical shifts (*19*), and the decrement in stability are each similar to those of DKP[A6−A11]^{Ser}. Such correspondence, which stands in contrast to the disparate properties of BPTI disulfide analogues (*46*−*48*), suggests that segmental unfolding of the A1−A8 α-helix is due to untethering of the disulfide bridge rather than to the polarity of the serine side chains. Although protein folding is in general a cooperative process, in this case a segmental coil → helix transition is coupled to formation of an adjoining disulfide bridge and uncoupled from formation of native structure elsewhere in the molecule. Further and in striking accord with an induced-fit model of receptor recognition, DKP[A6−A11]^{Ala} is 50-fold more active than DKP[A6−A11]^{Ser} in both receptor binding and metabolic signaling. Such enhanced function (i) highlights the applicability of general principles of protein stability to the formation of a

[2] Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; DG, distance geometry; DKP-insulin, monomeric insulin analogue containing three substitutions in the B chain (Asp^{B10}, Lys^{B28}, and Pro^{B29}; see Table 1); DKP[A6−A11]^{Ala}, insulin analogue containing two substitutions in the A chain (Ala^{A6} and Ala^{A11}) and DKP substitutions in the B chain; DKP[A6−A11]^{Ser}, insulin analogue containing two substitutions in the A chain (Ser^{A6} and Ser^{A11}) and DKP substitutions in the B chain; DQF-COSY, double-quantum-filtered correlated spectroscopy; IGF-I, insulin-like growth factor I; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; rmsd, root-mean-square difference; RP-HPLC, reverse-phase high-performance liquid chromatography; SA, simulated annealing; TOCSY, total correlated spectroscopy; 2D NMR, two-dimensional NMR.

complex hormone−receptor interface (*6*−*8*), (ii) rationalizes the otherwise surprising nonconservation of Cys^{A6} and Cys^{A11} among invertebrate insulin-like sequences in *Caenorhabditis elegans* (*49*, *50*), and (iii) supports the hypothesis that among vertebrate insulins the A6−A11 cystine serves to tether a preformed recognition α-helix. Segmental unfolding of insulin thus exploits its hierarchic pathway of disulfide pairing to dissect a relationship between structure and activity.

MATERIALS AND METHODS

Materials. 4-Methylbenzhydrylamine resin (0.6 mmol of amine/g; Star Biochemicals, Inc.) was used as a solid support for synthesis of the A chain analogue; (*N*-butoxycarbonyl,*O*-benzyl)threonine-PAM resin (0.56 mmol/g; Bachem, Inc.) was used as a solid support for synthesis of the B chain analogue. *tert*-Butoxycarbonyl amino acids and derivatives were obtained from Bachem and Peninsula Laboratories; *N*,*N*′-dicyclohexylcarbodiimide and *N*-hydroxybenzotriazole (recrystallized from 95% ethanol) were from Fluka. Amino acid analyses of synthetic chains and insulin analogues were performed after acid hydrolysis; protein determinations were carried out by the Lowry method using native insulin as the standard. 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 in ref *51*. The C-terminal Asn in synthesis of the A chain was incorporated into the solid support by coupling *tert*-butoxycarbonylaspartic acid 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 (1.23 g), after deblocking, sulfitolysis, and chromatographic purification (*19*, *41*), yielded ca. 186 mg of purified S-sulfonated A chain (Ala^{A6} and Ala^{A11}). *(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.

Peptide Purification. Crude S-sulfonated A chain was purified by chromatography on a Cellex E column (1.5 cm × 47 cm) as described previously (*19*, *41*), dialyzed against distilled water, and lyophilized to yield the purified [Ala^{A6},Ala^{A11}]A chain S-sulfonate. Crude S-sulfonated B chain was likewise purified on a cellulose DE53 column (1.5 cm × 47 cm), dialyzed, and lyophilized to yield the [Asp^{B10},Lys^{B28},Pro^{B29}]B chain S-sulfonate.

Chain Recombination. Chain recombination employed S-sulfonated A and B chains (approximately 2:1 by weight) in 0.1 M glycine (pH 10.6) in the presence of dithiothreitol (*30*). The insulin analogue was isolated from the combination mixture as described previously (*19*, *41*) and purified via 0.9 cm × 23 cm CM-cellulose chromatography and RP-HPLC on a Vydac 218 TP column (0.46 cm × 25 cm); the latter used a flow rate of 0.5 mL/min with a 20 to 80% linear gradient of 80% aqueous acetonitrile containing 0.1% trifluoroacetic acid over the course of 80 min. Rechromatography of this material on RP-HPLC under the same conditions gave a single sharp peak. Amino acid analysis and mass spectrometry gave the expected values.

Biological Assays. For binding studies using rat insulin receptors, plasma membranes were purified from rat liver. [^{125}I]Insulin and [3-^{3}H]glucose were purchased from Dupont NEN. Assays were performed as described previously (*19*, *41*) with relative activity defined as the ratio of porcine insulin to analogue required to displace 50% of the specifically bound [^{125}I]insulin. Relative activity in the lipogenesis assay is defined as ratio of porcine insulin to analogue required to achieve 50% of the maximal conversion of [3-^{3}H]-glucose to organic-extractable material (glycogen).

Spectroscopy. ^{1}H NMR spectra were obtained at 600 MHz at 25 °C in 50 mM potassium phosphate (pH 7) and in 20% deuterioacetic acid (pH 1.9) as described previously (*41*); the protein concentration was 1.5 mM. Resonance assignment was based on 2D NOESY (mixing times of 100 and 200 ms), TOCSY (mixing times of 30 and 55 ms), and DQF-COSY spectra. Spectra in H_2O were obtained using pulse-field gradients and laminar-shaped pulses (*19*). CD spectra were obtained using an Aviv spectropolarimeter equipped with a thermister temperature control and automated titration unit for guanidine denaturation studies. CD samples contained 25−50 μM insulin or analogue in 50 mM potassium phosphate (pH 7); samples were diluted to 5 μM for equilibrium denaturation studies (Figure 3C). Data were fitted by nonlinear least squares to a two-state model as described previously (*52*).

RESULTS

DKP[A6−A11]^{Ala} was prepared by total synthesis (see Materials and Methods). The analogue contains two substitutions in the A chain and three in the B chain (Table 1). The protein design is based on an engineered insulin monomer (*53*−*55*). Respective substitutions in the B chain destabilize insulin's classical dimer interface (Pro^{B28} → Lys and Lys^{B29} → Pro) and hexamer contact (His^{B10} → Asp) and so provide a monomeric template with enhanced biological activity (DKP-insulin; *56*). Use of a monomeric template simplifies biophysical analysis by preventing confounding effects of insulin self-association (*19*). Chain combination yielded native disulfide pairing (A7−B7 and A20−B19 cystines) with an efficiency similar to that observed in the total synthesis of DKP[A6−A11]^{Ser} and native insulin. Non-native disulfide isomers (*22*) were not observed. Such efficiency and fidelity of disulfide pairing contrast with the >10-fold loss of yield reported in syntheses of insulin analogues containing other substitutions elsewhere in the hydrophobic core, such as Leu^{B11} and Val^{B12} (*41*).

Alanine and Serine Analogues Exhibit Similar Partial Folds. CD spectra of the two-disulfide analogues at 4 °C exhibit similar attenuations of helix context (Figure 3A) relative to the parent DKP-insulin monomer (●). Similar results are obtained at 25 °C (not shown). The extent of attenuation is consistent with a helix → coil transition affecting A1−A8 as reported in the solution structure of DKP[A6−A11]^{Ser} at 25 °C (*19*). The CD results exclude a native folding transition at low temperatures.

The ^{1}H NMR spectrum of the alanine analogue at pH 7.2 and 25 °C is similar to that of the serine analogue (Figure 4B,C); each exhibits incomplete but nonrandom dispersion of chemical shifts relative to the spectrum of DKP-insulin (Figure 4A). Despite their similarities in chemical shift,

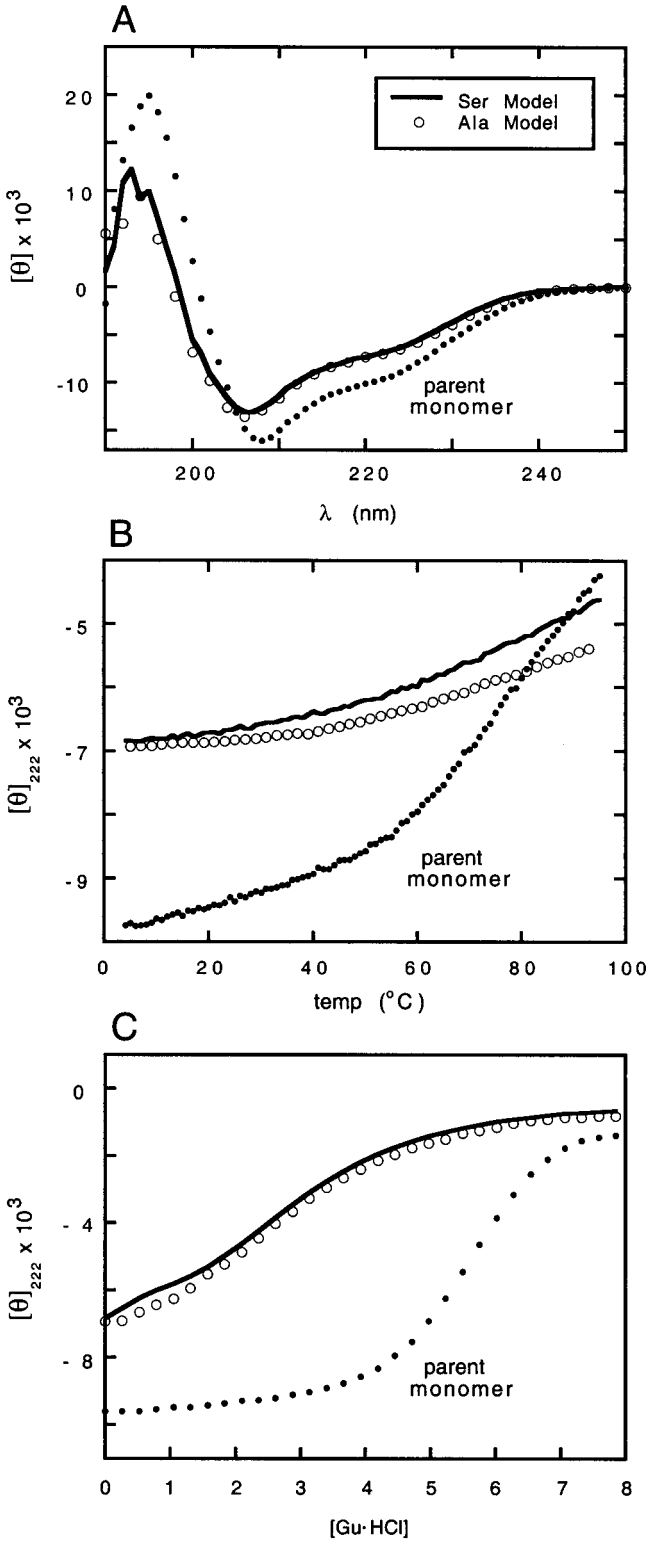

FIGURE 3: CD spectra and stabilities of insulin analogues. (A) Far-UV spectra of DKP-insulin ["parent monomer" (●)], DKP[A6−A11]^{Ser} (—), and DKP[A6−A11]^{Ala} (○) demonstrate similar attenuation in helix context in either two-disulfide analogue. For clarity, the spectrum of the alanine analogue is represented by every other point. (B) Thermal unfolding of the serine analogue (—) and the alanine analogue (○), monitored by ellipticity at 222 nm, is less cooperative than that of DKP-insulin [parent monomer (●)]. $[\theta]_{222}$ values of DKP[A6−A11]^{Ser} and DKP[A6−A11]^{Ala} are similar in the range of 4−25 °C. (C) The guanidine unfolding transition of serine and alanine analogues (— and ○) and DKP-insulin [parent monomer (●)] demonstrates that untethered analogues exhibit similarly marked destabilization (see the text for ΔG_u values).

resonances are in general broader in the spectrum of DKP-[A6−A11]^{Ala} than in the spectrum of DKP[A6−A11]^{Ser}, suggesting partial self-association. 2D ^{1}H NMR spectra in each case contain 51 unique spin systems. Detailed analysis of the spectrum of DKP[A6−A11]^{Ala} is as described for the serine analogue (*19*; see the Supporting Information). As in the spectrum of DKP[A6−A11]^{Ser}, resonances from the native-like domain exhibit chemical shifts similar to those in the parent monomer, whereas the majority of resonances in or near the A1−A11 segment exhibit attenuated secondary shifts (i.e., perturbations toward random coil). Examples are provided by the upfield aliphatic resonances of Ile^{A2}, Ile^{A10}, Leu^{B11}, and Leu^{B15} (right-hand panel of Figure 4). The B chain methyl resonances retain high-field chemical shifts (reflecting aromatic ring currents in the native-like subdomain), whereas the A chain methyl resonances regress toward random-coil values (near 0.9 ppm).

An overview of chemical shift differences between DKP-[A6−A11]^{Ala} and DKP-insulin is shown in Figure 5A; chain- and region specific perturbations are coded by color. The predominant sites of perturbation are in the A1−A11 segment (red). Conversely, the structural similarity between alanine and serine analogues is indicated by corresponding patterns of chemical shifts and secondary chemical shifts (Figure 5B). Such detailed correspondence between the ^{1}H NMR spectra of DKP[A6−A11]^{Ala} and DKP[A6−A11]^{Ser} provides a fingerprint of structural similarity. Tables of chemical shifts and changes in chemical shifts are provided as Supporting Information. As expected, complementary analysis of d_{NN} and other diagnostic nuclear Overhauser enhancements (NOEs) corroborates the fact that the variant A chains in each case lack helix-related NOEs in their N-terminal regions (A1−A8) but otherwise retain native secondary structure (Supporting Information).

Analogues Exhibit Similar Thermodynamic Perturbations. The native DKP monomer exhibits a cooperative thermal unfolding transition as monitored by ellipticity at the helix-sensitive wavelength of 222 nm [Figure 3B (●)]; the apparent midpoint (T_m) is near 60 °C. In contrast, the two-disulfide analogues each exhibit incremental and progressive attenuation of ellipticity with increasing temperatures. Thermal unfolding curves of DKP[A6−A11]^{Ala} and DKP[A6−A11]^{Ser} (○ and —, respectively) are distinguishable but exhibit similar overall features (Figure 3B). An analogous trend is observed on analysis of protein unfolding in the denaturant guanidine-HCl (Figure 3C). Whereas DKP-insulin exhibits a cooperative and apparent two-state transition with a ΔG_u of 4.3 ± 0.3 kcal/mol, the two-disulfide analogues exhibit a marked thermodynamic destabilization (Figure 3C). Unfolding of DKP[A6−A11]^{Ala} (○) closely follows the unfolding transition of DKP[A6−A11]^{Ser} (—). Inferred values of ΔG_u are not statistically different (1.1 ± 0.4 and 1.3 ± 0.2 kcal/mol, respectively). The decrement in free energies of unfolding between the untethered and native monomers is 3.0 ± 0.5 kcal/mol. This represents a profound destabilization relative to the parent value.

The Alanine Analogue Is More Potent Than the Serine Analogue. Biological activity of the two-disulfide analogues was measured relative to porcine insulin by two complementary assays. Specific binding to the insulin receptor was assessed in a rat liver membrane fraction; signaling through the insulin receptor was assessed by the ability of an analogue

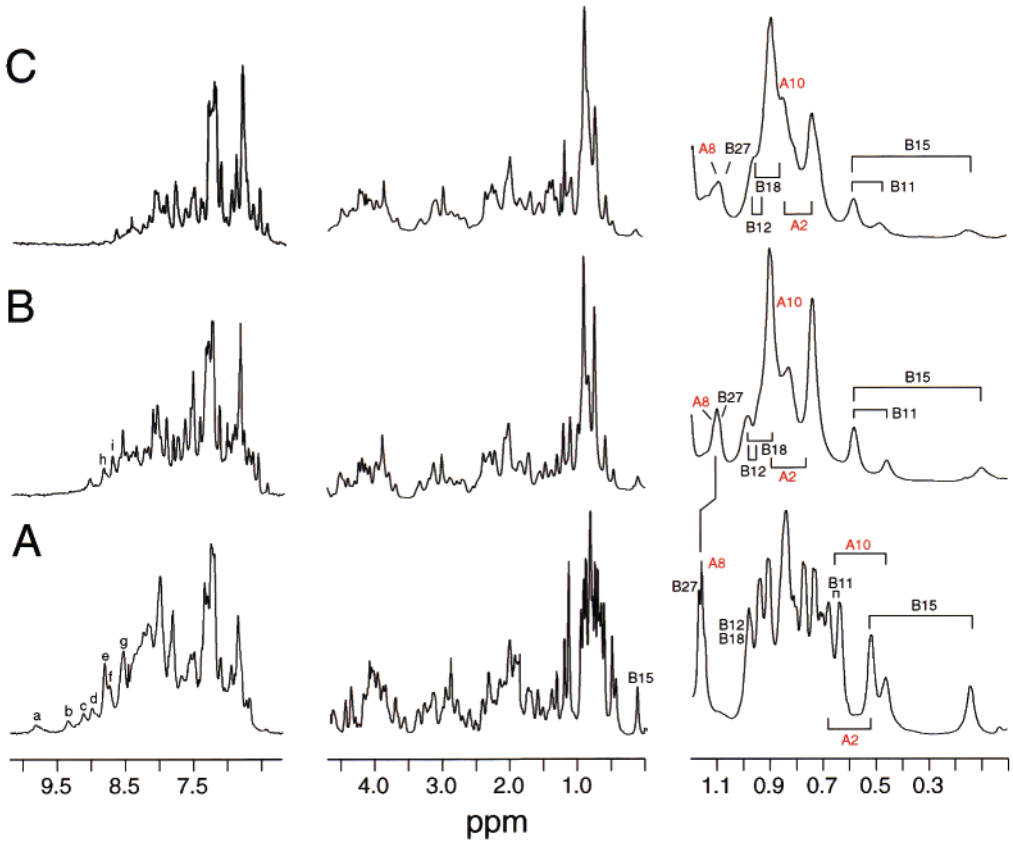

FIGURE 4: Structural similarity of two-disulfide analogues. ^{1}H NMR spectra of (A) parent DKP-insulin monomer, (B) DKP[A6−A11]^{Ser}, and (C) DKP[A6−A11]^{Ala} demonstrate that the two-disulfide analogues exhibit similar dispersions and patterns of chemical shift perturbations. Assignments of selected aliphatic spin systems are shown at the right. Aliphatic spin systems are shown in expanded form at the right; A chain assignments are highlighted in red and B chain assignments in black. Assignment of selected downfield amide resonances of (a) A11, (b) B8, (c) B9, (d) B6, (e) B7, (f) B19, (g) B5, (h) B7, and (i) B19. Spectra A and B were obtained at 25 °C and pH 7.0 (pD 6.6), whereas spectrum C was obtained at obtained at 32 °C and pH 8.0 (pD 7.6). The spectrum of DKP[A6−A11]^{Ala} is somewhat broader than that of DKP[A6−A11]^{Ser}, presumably due to partial self-association mediated by the more apolar disordered A1−A11 segment. Spectra in 90% H_2O were obtained by selective excitation without solvent presaturation.

to stimulate lipogenesis in isolated rat adipocytes. Whereas DKP[A6−A11]^{Ser} was previously reported to exhibit relative activities of 0.1% in each assay, DKP[A6−A11]^{Ala} exhibits relative activities of 5% (Figure 6A,B). At saturating concentrations, stimulation of lipogenesis is in each case identical to that of native porcine insulin (Figure 6B). This indicates that when the hormone−receptor complex is formed, it achieves native potency in metabolic signaling.

DISCUSSION

The importance of insulin's three cystines (Figure 1A) is demonstrated by the inactivity of the hormone on reduction and by the low activities of synthetic analogues containing elongated, rearranged, or broken disulfide bridges (*21*, *22*). The particular role of the A6−A11 cystine has been investigated through the synthesis and characterization of an analogue lacking this disulfide bridge (*19*). The analogue contained pairwise substitution of a polar side chain in an engineered monomer (DKP[A6−A11]^{Ser}). The analogue's structure at 25 °C exhibits segmental unfolding of the N-terminal A chain α-helix (- - - in Figure 2B). Because chain combination occurred with native efficiency and fidelity, the native A1−A8 α-helix is not required for specification of the remaining disulfide pairing (A20−B19 and A7−B7). The structure is otherwise native-like, further suggesting that the A1−A8 segment is also not required to nucleate the hydrophobic core or specify the orientation between chains (*19*). The very low biological activity of the analogue was interpreted as evidence that the native A1−A8 segment functions as a preformed recognition α-helix. This conclusion is of interest in light of models proposing reorganization of the B chain in the hormone−receptor complex (*25*, *27*, *32*, *36*, *38*, *39*, *43*).

Our original study (*19*) left open the possibility that segmental unfolding of DKP[A6−A11]^{Ser} is particular to serine and not a general consequence of untethering the disulfide bridge. This possibility was suggested by the polarity of the serine side, which would be expected to be less compatible with a hydrophobic environment (Figure 1B) than would aliphatic substitutions. Our concern was accentuated by analogous studies of analogues of BPTI (*46*−*48*). Like insulin, BPTI is a small globular protein containing three disulfide bridges (Figure 7A). Pairwise substitutions of cystines in BPTI with serine were found to be more destabilizing than substitutions with alanines (*46*−*48*). Characterization of the divergent properties of BPTI analogues[3] has motivated the present reinvestigation of the A6−A11 cystine in human insulin using pairwise alanine substitutions. Given that the previous study of DKP[A6−A11]^{Ser} was restricted to 25 °C, how robust is its finding of segmental unfolding? Surprisingly, we have observed that DKP[A6−A11]^{Ser} and DKP[A6−A11]^{Ala} exhibit essentially

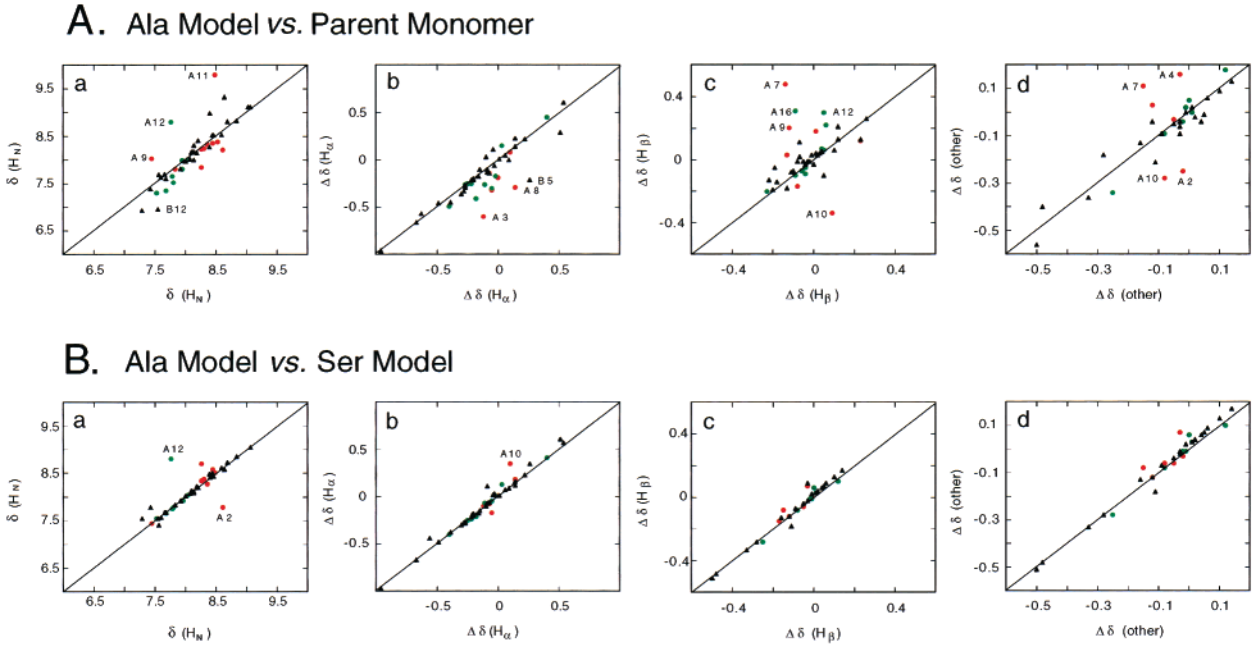

FIGURE 5: NMR footprints. (A) Large differences in ^{1}H NMR chemical shifts between (A) DKP[A6−A11]^{Ala} and DKP-insulin reflect the site and magnitude of structural changes in alanine analogue relative to parent monomer. (B) Small differences in ^{1}H NMR chemical shifts between DKP[A6−A11]^{Ala} and DKP[A6−A11]^{Ser} (A) demonstrate similarities between alanine and serine analogues in structure and dynamics. Each comparison has four parts: (a) chemical shift amide resonances and secondary shifts of H_α (b), H_β methylene (c), and aromatic and methyl (d) resonances. Secondary shifts are calculated as the difference between observed chemical shifts and tabulated random-coil values. 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 Supporting Information.

identical physical properties and further that partial unfolding occurs even at a low temperature (4 °C). Nevertheless, although the two analogues each exhibit reduced levels of receptor binding (Figure 6), DKP[A6−A11]^{Ala} does not exhibit the severe decrement in activity described previously. The alanine analogue's higher relative affinity and potency have important implications as discussed in turn below.

The A6−A11 Cystine Tethers a Preformed Recognition α*-Helix*. Comparison of CD and ^{1}H NMR spectra demonstrates that the structure of DKP[A6−A11]^{Ala} is similar to that of DKP[A6−A11]^{Ser}. Segmental unfolding of the A1−A8 α-helix in the two-disulfide analogue is thus not a consequence of polarity per se. Its occurrence in an alanine analogue (a nonpolar side chain with high helical propensity) suggests that the A6−A11 cystine functions to tether this α-helix and to ensure its integration within the hydrophobic core. Pairwise substitution with alanine or serine also leads to similarly profound decrements in overall stability (Figure 3B,C). We imagine that the similar physical properties of DKP[A6−A11]^{Ala} and DKP[A6−A11]^{Ser} reflect the low intrinsic stability of the A1−A8 α-helix and hence a requirement for covalent tethering. This α-helix contains a high proportion of *β*-branched amino acids (Ile^{A2}, Val^{A3}, and Thr^{A8}), which have low intrinsic helical propensities (*11*−*13*, *57*). Further, the A1−A8 helix is least regular in structure among different insulin crystal forms (*27*−*29*), and (unlike the other helices in insulin, B9−B19 and A12−A18; Figure 2A) fails to exhibit protected amide protons in D_2O solution (*19*, *25*, *26*). BPTI in contrast exhibits significantly greater intrinsic physical and thermodynamic stability (*48*). The extent of thermodynamic destabilization of DKP[A6−A11]^{Ala} and DKP[A6−A11]^{Ser} is surprising given that a predominance of native-like structure is in each case retained. Detailed interpretation of the present ΔG_u and $\Delta\Delta G_u$ values is limited by the inequivalence of denatured states (i.e., containing two or three disulfide bridges and different residues at A6 and A11) and by the unavailability of ΔH and ΔS values.

DKP[A6−A11]^{Ser} and DKP[A6−A11]^{Ala} each exhibit reduced biological activity (Figure 6). Although the extent of perturbed function differs (see below), their shared impairment supports the hypothesis that the A1−A8 segment functions as a preformed recognition α-helix (*19*, *44*, *45*). Considerable mutational evidence has accumulated that the side chains of A1−A3 are critical to receptor binding whereas A4, A5, A7, and A8 are not (*33*, *34*, *40*, *58*). In themselves such studies do not distinguish whether the substituted side chain contributes directly or indirectly to bioactivity, i.e., as a receptor contact or overall structural element. Nor do such studies address the structural context in which Ile^{A2} and Val^{A3} function, i.e., within the native α-helix or as part of a novel secondary structure induced on receptor binding. These issues are pertinent as the hormone's active structure is unknown.

[3] An example is provided by single-disulfide BPTI analogues containing the 5−55 disulfide bridge, an internal cystine that can be formed early in one branch of the folding pathway. Whereas the alanine analogue exhibits only native structure at temperatures of <30 °C and under these conditions functions to inhibit trypsin stoichiometrically (*47*), the serine analogue is partially folded at 4 °C and unfolded at 25 °C, the temperature at which the corresponding intermediate accumulates during folding (*46*). Similarly, an extensive series of thermodynamic measurements was performed comparing pairwise substitutions of the 30−51 cystine (*48*). The C30A/C51A variant was found to be 2.8 kcal/mol more stable than the C30S/C51S variant as probed by equilibrium denaturation. Such differences in structure and stability are ascribed to the polarity of serine, which unlike alanine would resist native folding events that bury cysteine side chains at neutral pH (*47*).

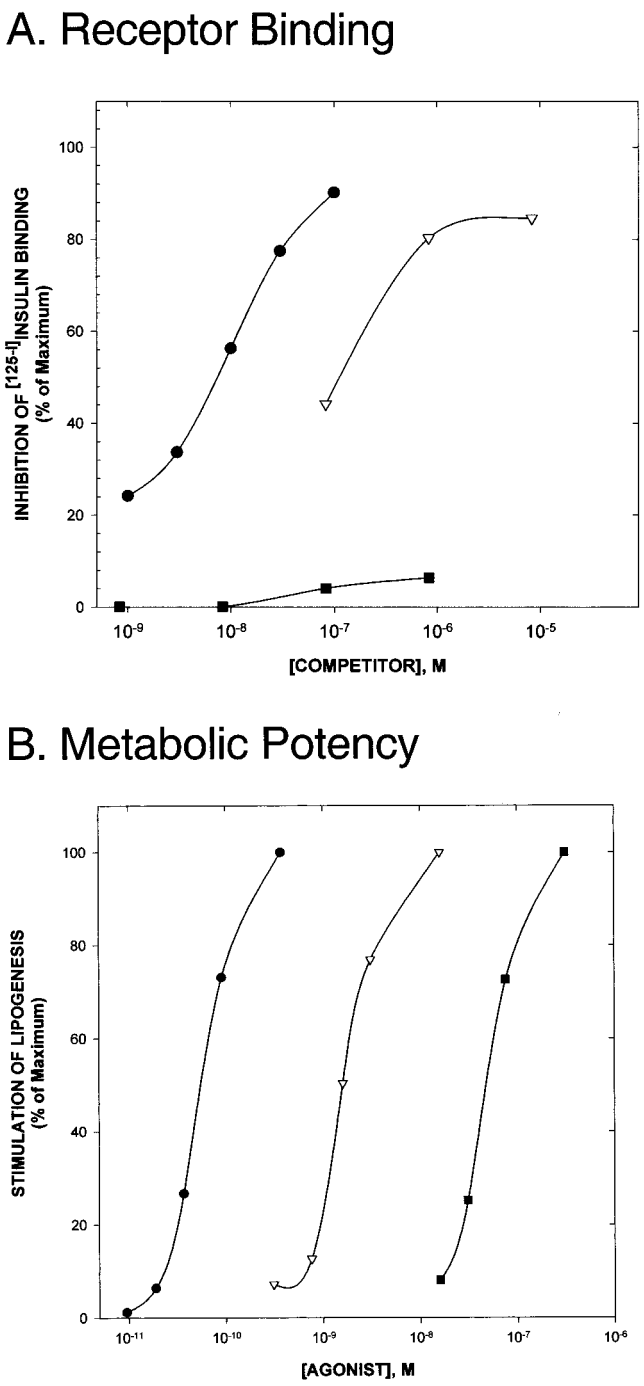

FIGURE 6: Biological activities of insulin analogues. (A) Stimulation of lipogenesis expressed as a percentage of maximum incorporation of tritium from [3-^{3}H]glucose in organic-extractable material (glycogen) as measured in adipocytes. (B) Receptor binding is assessed relative to displacement of ^{125}I-labeled porcine insulin. In each case, data for porcine insulin are represented by black circles (●), DKP[A6−A11]^{Ala} by white triangles (▽), and DKP[A6−A11]^{Ser} by black squares (■). The serine analogue's severe impairment of receptor binding precluded saturation of receptor occupancy in experiment A.

The long-range structural reorganization seen in crystal structures (allosteric transitions among T_6, T_3R_3, and R_6 hexamers; *27*−*29*) highlights the potential for conformational change in the B chain. Detachment or reorganization of the B chain's C-terminal β-strand on receptor binding is further suggested by the anomalous properties of molecules modified in this region (*36*, *38*, *39*, *43*, *59*).

Unlike the B chain (*28*, *29*), the A chain exhibits a single canonical structure containing two helices (Figures 1A and 2A). Although some variability can occur in its positioning relative to the B chain and in the details of side chain packing in the hydrophobic core (*60*), the A chain's secondary structure and overall orientation relative to the B chain remain essentially unchanged. This orientation ensures that the side chains of Ile^{A2} and Val^{A3} contact the nonpolar surface of the B chain's C-terminal β-strand. Although such contacts would ordinarily be expected to render Ile^{A2} and Val^{A3} inaccessible to the insulin receptor, the side chains are exposed in an active truncated analogue [*des*-pentapeptide-(B26−B30)insulin; *61*−*63*] and shielded in an inactive single-chain analogue (in which a peptide bond joins Lys^{B29} and Gly^{A1}; *43*). Together with mutational studies, structures of these modified molecules have led to a model in which detachment of the B chain's C-terminal β-strand on receptor binding exposes Ile^{A2} and Val^{A3} for interaction with the insulin receptor (*25*). This model rationalizes the exquisite sensitivity of receptor binding to seemingly conservative substitutions at A2 and A3. The extent of the proposed detachment is unknown.

The low bioactivities of DKP[A6−A11]^{Ser} and DKP[A6−A11]^{Ala} presumably reflect local unfolding of the A1−A8 segment rather than altered receptor contacts as the A6−A11 cystine is deeply buried in the native hormone's hydrophobic core (Figure 1B). (Even on detachment of the B chain's C-terminal β-strand in a putative receptor complex, the A6 and A11 side chains would be inaccessible to the receptor.) The analogue's primary structure otherwise retains Gly^{A1}, Ile^{A2}, Val^{A3}, and other receptor-binding residues in the A and B chains. We imagine that the helical structure of the A1−A8 segment is restored on receptor binding. In this model, the decrements in activity would reflect the free energy cost of refolding the A1−A8 segment to restore the native configuration of receptor contacts. An alternative model (analogous to that proposed for the B chain's C-terminal β-strand; *25*) posits that on receptor binding the A1−A8 segment detaches from the core of the protein and/or changes secondary structures (*32*). In this alternative scenario, destabilization of the A1−A8 segment might have no effect on activity or even result in enhanced activity. The latter indeed occurs in the B chain wherein substitution of Phe^{B24} (an invariant side chain in the hydrophobic core; Figure 2A) with D-Phe or D-Ala enhances activity (*36*, *59*) despite destabilizing the B chain's C-terminal β-strand (Q.-X. Hua et al., unpublished results). Like substitutions of the A6−A11 cystine, D-amino acid substitutions at B24 cause only segmental perturbations, leaving the preponderance of the structure native-like. The opposite functional effects of such destabizing mutations in the A and B chains suggest that the native A1−A8 segment (unlike the B24−B28 β-strand) indeed functions as a preformed recognition element.

Principles of Protein Stability Rationalize Biological Activities. The qualitative model outlined above has quantitative implications. We propose an analogy between the low peptide folding activities of DKP[A6−A11]^{Ser} and DKP[A6−A11]^{Ala}: *The enhanced relative activity of the alanine analogue arises exclusively from its enhanced compatibility with the bound structure of the hormone.* This hypothesis assumes that (i) native and variant receptor−hormone

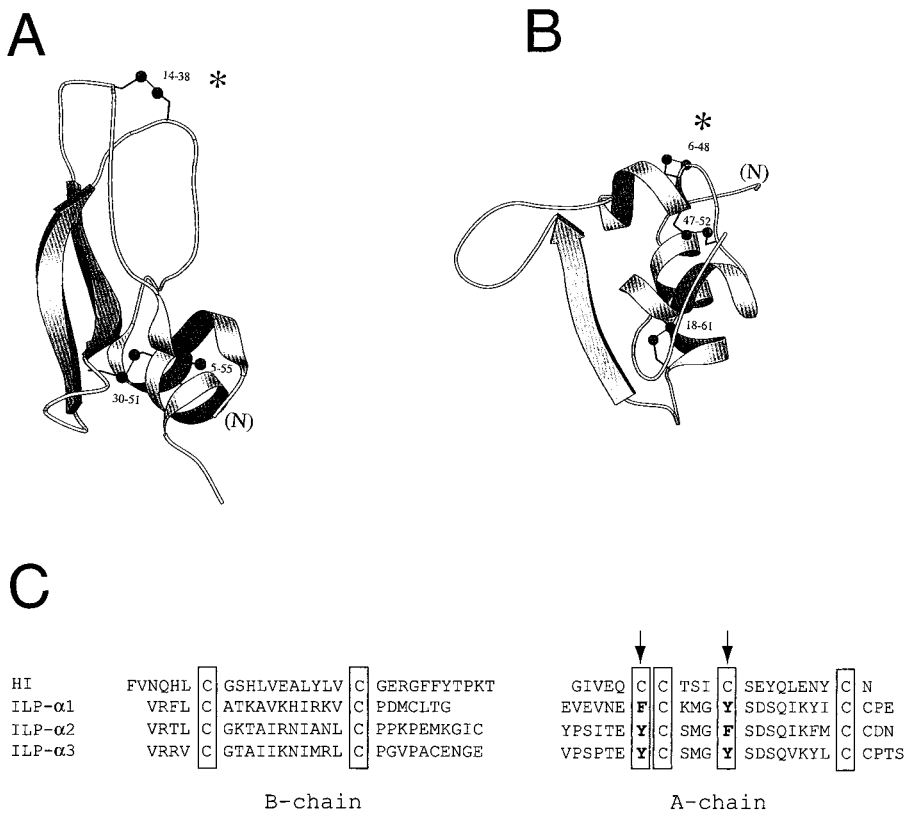

FIGURE 7: Ribbon models of (A) BPTI and (B) IGF-I (PDB entries 1BPI and 2GF1, respectively). Asterisks indicated the respective solvent-exposed disulfide bridges (14−38 in BPTI and 6−68 in IGF−I; the latter corresponds to A7−B7 in insulin), which play contrasting roles in respective oxidative refolding pathways (*65*−*70*). Further, whereas alanine and serine substitutions at the internal 30−51 cystine in BPTI lead to marked differences in thermodynamic stability (*48*), analogous substitutions in insulin's internal A6−A11 cystine (47−52 in IGF-I) lead to similar stabilities. Images were generated using Molscript. (C) Comparison of the sequence of human insulin with putative insulin-like peptides encoded in the genome of *C. elegans* (*49*, *50*). Cysteine residues are boxed; arrows indicate Cys^{A6} and Cys^{A11} in vertebrate insulin. Variant residues in nematode insulin-like sequences are indicated in bold. Disulfide pairing schemes and structures of worm peptides have not been established.

complexes are essentially identical in structure, (ii) segmental unfolding of the A1−A8 segment in the two analogues affects in the same way other functional surfaces of insulin, and (iii) the induced fit of the A1−A8 segment is independent of binding processes elsewhere in the hormone−receptor interface. By focusing on the *differences* between the analogues, the model also enables shared and poorly characterized features of receptor binding (such as possible conformational changes in the B chain) to be neglected. These assumptions simplify the complex process of hormone−receptor recognition to a more tractable and well-studied problem, the thermodynamics of a coil → helix transition (*11*−*13*; Figure 2A and Table 2A) and its engagement in a hydrophobic core (Figure 1B and Table 2B).

A simple calculation builds on the fundamental distinguishing properties of amino acids, hydrophobicity (*2*−*4*, *9*, *10*) and intrinsic conformational propensity (*11*−*13*). Kauzmann originally proposed that hydrophobic interactions between nonpolar side chains make a predominant contribution to the stability of globular proteins and protein interfaces (*2*). A hydrophobicity scale was developed by Tanford and co-workers that was based on the free energy of transfer of amino acids from organic solvents to water (*3*). Although the thermodynamics of such partitioning depends on the solvent used and in any case fails to capture the heterogeneity of actual protein interiors (*9*), a wealth of experimental and theoretical studies support Kauzmann's proposal (*10*) and the empirical utility of a hydrophobicity scale. A critical assessment of such scales has been described by P. A. Karplus (*10*) and has led to the definition of a "pure" hydrophobicity scale that captures the essence of Tanford's concept (see the footnote of Table 2B). We imagine that the present analogues, on receptor binding, transfer the A6 and A11 side chains from a solvated environment to the bound hormone's hydrophobic core. Hydrophobic-transfer free energies thus provide an estimate of this component of the differential free energy of the induced fit between $DKP[A6-A11]^{Ser}$ and $DKP[A6-A11]^{Ala}$ as given in Table 2B. The relevant values are provided by the side chain specific estimates given in the bottom line of the table. Because two serine or alanine side chains are proposed to transfer from solvated to apolar environments on receptor binding, the estimated "purely hydrophobic" contribution to relative receptor affinity ($\Delta\Delta G_2$ in the scheme outlined in Figure 2B) is given by 2Δ or 1.6 kcal/mol.

Protein stability is also influenced by conformational preferences of particular amino acids for α-helix, β-sheet, or β-strand (*11*−*15*). Empirical thermodynamic scales (relative to the random-coil state) have been defined on the basis of the relative stabilities of variant peptides containing a "guest" amino acid in a well-defined "host", such as an alanine-rich α-helical peptide (*11*, *13*) or coiled-coil dimer (*12*). Representative values are given in Table 2A. We imagine that the present analogues, on receptor binding,

Table 2: Predicted Thermodynamic Differences between Alanine and Serine

(A) Helical Propensities (kcal/mol)[a]			
Ala	Ser	Δ	model
−1.88	−1.10	0.78	peptide monomer[b]
−0.77	−0.35	0.42	coiled-coil dimer[c]

(B) Hydrophobic-Transfer Free Energies (kcal/mol)[d,e]				
Ala	Ser	Δ	2Δ	model
−0.42	0.05	0.47	0.94	$H_2O \rightarrow$ octanol (ΔG^o_{trans})[f]
−2.15	−1.40	0.75	1.5	$H_2O \rightarrow$ octanol (residue burial)[g]
−1.0	−0.20	0.80	1.6	$H_2O \rightarrow$ octanol (side chain burial)[h]

[a] Differences in α-helical propensities estimate the free energy cost of an induced A1−A8 α-helix on receptor binding. Of the sites of substitution, only residue A6 participates in the putative α-helix; the configuration of Cys^{A11} lies in the β region (*27*). [b] Values were obtained by Baldwin and co-workers (*13*) at 273 °C using a modified Lifson−Roig theory applied to peptides of the form Ac-YGGKAAAAKAXAAKAAAAK-$CONH_2$ wherein the guest amino acid was placed at position X (underlined). [c] Values were obtained by O'Neil and DeGrado (*12*) on the basis of urea-induced equilibrium-denaturation studies (ΔG_u measurements) of the coiled-coil Ac-EWALEKKLAALEXKLQALEKKLEALEHG wherein the guest amino acid was placed at position X (underlined). [d] A "pure" hydrophobicity scale was proposed by P. A. Karplus (*10*) to resolve the contribution of hydrophobicity from confounding effects such as hydrogen bonding and local details of protein structure. [e] 2Δ is calculated as twice the difference in values between alanine and serine; the factor of 2 arises from the pairwise substitutions (A6 and A11). [f] Values were taken from Fauchere and Pliska (*9*); these represent raw data in solvent-transfer studies of amino acids without correction for apolar surface area. [g] Values include correction for apolar surface area (*10*). [h] Values were obtained from previous line by subtraction of value of glycine (*10*).

transfer the A6 side chain from a random-coil state to an α-helical configuration. We neglect the cost of transfer of the A11 side chain to the β region of the Ramachandran plot.[4] The empirical helical propensities of serine and alanine in monomeric and dimeric peptides thus provide upper and lower estimates of this component of the differential free energy of the induced fit ($\Delta\Delta G_1$ in the scheme outlined in Figure 2B). These values are 0.78 and 0.43 kcal/mol, respectively (Table 2A, lines 1 and 2, column Δ).

The sum of predicted hydrophicity- and helix-related terms thus predicts that the alanine analogue would exhibit higher affinity for the insulin receptor and that the extent of enhancement would lie between 2.0 and 2.6 kcal/mol. The 50-fold tighter binding of DKP[A6−A11]^{Ala} relative to DKP-[A6−A11]^{Ser} corresponds to a free energy difference of 2.2 kcal/mol, which is in the middle of this range. In the absence of additional data, it is possible that this agreement is fortuitous. Given the simplicity of our model and the likely complexity of the insulin−receptor complex, however, the agreement seems remarkable. In any case, the observed level of activity of DKP[A6−A11]^{Ala} validates a critical prediction of the local induced-fit model and its implication for native insulin: *that the A1−A8 segment ordinarily functions as a preformed recognition element.* It is noteworthy that relative activities correlate with derived residue-specific local thermodynamic parameters and not with changes in the analogue's global thermodynamic stability relative to the parent insulin monomer ($\Delta\Delta G_u \approx 3$ kcal/mol in each case).

Peptide Models of Protein-Folding Intermediates. Because the isolated A and B chains contain sufficient information to specify the folding of proinsulin (*21*, *31*, *32*, *61*), the present partial folds suggest an analogy to peptide models of protein-folding intermediates (*64*). BPTI intermediates that are well-populated at neutral pH exhibit native structure (*47*, *68*−*70*). Large kinetic barriers among such folded intermediates lead to a preferred final step: formation of an *external* disulfide bridge between solvent-exposed loops (asterisk in Figure 7A). Although proinsulin's oxidative refolding pathway is not well-characterized, that of a homologous protein [insulin-like growth factor I (IGF-I)] has been studied in detail (*65*−*67*).[5] Insulin's disulfide bridges are conserved in IGF-I (Figure 7B); the A6−A11 cystine corresponds to the 47−52 cystine in IGF-I (arrow in Figure 7B). Unlike BPTI, IGF-I exhibits a hierarchical pathway in which successive intermediates exhibit stepwise increases in structural organization. Because the 47−52 cystine is the last disulfide bridge in IGF-I to form (*65*−*67*), the partial folds of DKP[A6−A11]^{Ser} and DKP[A6−A11]^{Ala} suggest that an exposed face of the corresponding intermediate's hydrophobic core provides an internal template for the A1−A8 coil → helix transition (*19*), i.e., as an "internal template" to align the A6 and A11 thiolates for specific disulfide pairing.[6]

Evolution and Divergence of Insulin-like Sequences. In native insulin, side chains from the A1−A11 segment form an integral part of the hydrophobic core (Figure 1B). Indeed, long-range contacts are observed from this segment to each of the remaining elements of secondary structure. Segmental unfolding of the two-disulfide analogues nevertheless suggests that the A1−A8 helix is *modular*; regardless of its local state of organization, native-like structure is elsewhere maintained. Comparison of R and T state protomers in insulin hexamers likewise suggests that the segmental structure of residues B1−B8 is likewise modular (*27*−*29*). Unlike the A1−A8 α-helix, the variable B1−B9 segment (strand in the T state and helix in the R state) is extrinsic to the hydrophobic core of the monomer.

The surprising modularity of the A1−A8 α-helix and the relative increase in the activity of DKP[A6−A11]^{Ala} raise the question of why Cys^{A6} and Cys^{A11} are invariant among vertebrate insulins and insulin-like sequences (IGF-I, IGF-II, and relaxin; *23*). After all, an increasing number of native proteins are now known to be unfolded in part or in whole and to undergo folding transitions on binding to substrates or other macromolecules. One can readily imagine that an insulin-like sequence can evolve to exhibit high affinity (in absolute terms) without preformed structure in the A1−A8 segment. Such sequences may indeed be present in the recently completed genome of the nematode *C. elegans* (*49*,

[4] In the two-Zn T_6 crystal form (*27*), the A11 ϕ and ψ values are −150° and 150° and −144° and 163° in the two independent monomers, respectively. That this transfer term is likely to be negligible is indicated by (i) the distribution of all ϕ and ψ values in the ensemble of solution structures (*19*) already lies in the β region, (ii) tabulated β-sheet propensities are inapplicable as A11 does not participate in a β sheet, and (iii) in any case the latter β specific $\Delta\Delta G$ values between serine and alanine are small and context-dependent (*15* and references therein).

[5] Because insulin's active conformation is unknown, "native" structural elements may not correspond to the hormone's structure in a receptor complex.

[6] Pairwise substitution of the 47−52 cystine in IGF-I (see Figure 7B) with Ala results in similar attenuation of CD-derived helix content and 3% affinity for the type I IGF receptor relative to native IGF-I (*16*).

50; Figure 7C). Although the structure and receptor-binding properties of such putative insulin-like proteins have not been characterized, some of the sequences lack cystine A6−A11 but retain the A7−B7 and A20−B19 cystines and other conserved insulin side chains (arrows in Figure 7C). It is possible that the homologous A1−A8 segment in such proteins is nonetheless α-helical due to optimized side chain packing in respective hydrophibic cores. Alternatively, these proteins may provide naturally occurring examples of segmental induced fit. We speculate that the invariance of the A6−A11 cystine among vertebrate sequences reflects the combined effect of multiple selective constraints. These include not only receptor affinity but also the protein's thermodynamic stability to reduction in an organelle and in the bloodstream (*67*), dynamical stability in relation to proteolysis, and self-assembly in storage or binding to partner proteins (*24*). These constraints may be relaxed or inoperative in the nematode's simplified body plan.

ACKNOWLEDGMENT

We thank G. T. Burke (deceased) for assistance with the receptor-binding assay, R. Chance and B. H. Frank for the gift of human insulin and advice, K. Polonksy for support through the University of Chicago Diabetes Research & Training Center, K. Hallenga and D. Jones for assistance with NMR measurements, and G. G. Dodson, M. Karplus, S. Nakagawa, S. Rice, A. Rubenstein, S. E. Shoelson, D. F. Steiner, M. Zhao, and the late H. S. Tager for discussion. NMR spectra were obtained at the Biomolecular NMR Facility at The University of Chicago and the Cleveland Center for Structural Biology.

SUPPORTING INFORMATION AVAILABLE

Three figures showing sequential assignments and amide and aliphatic regions of the NOESY spectrum of DKP[A6−A11]^{Ala} at 600 MHz in relation to spectra of DKP[A6−A11]^{Ser} or DKP-insulin and four tables giving ^{1}H NMR resonance assignments of DKP[A6−A11]^{Ala}, secondary chemical shifts (relative to tabulated random-coil values), and differences in chemical shifts relative to DKP-insulin and DKP[A6−A11]^{Ser}. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Anfinsen, C. B. (1973) *Science 181*, 223−230.
2. Kauzmann, W. (1959) *Adv. Protein Chem. 14*, 1−63.
3. Tanford, C. (1972) *J. Am. Chem. Soc. 84*, 4240−4247.
4. Chothia, C. (1974) *Nature 248*, 338−339.
5. Eisenberg, D., and McLachlan, A. D. (1986) *Nature 319*, 199−203.
6. Hruby, V. J., and Sharma, S. D. (1991) *Curr. Opin. Biotechnol. 2*, 599−605.
7. Dobson, C. M., and Karplus, M. (1999) *Curr. Opin. Struct. Biol. 9*, 92−101.
8. Cooper, A. (1999) *Curr. Opin. Chem. Biol. 3*, 557−563.
9. Fauchere, J.-L., and Pliska, V. (1983) *Eur. J. Med. Chem. 18*, 369−375.
10. Karplus, P. A. (1997) *Protein Sci. 6*, 1302−1307.
11. Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M., and Baldwin, R. L. (1990) *Nature 344*, 268−270.
12. O'Neil, K. T., and DeGrado, W. F. (1990) *Science 250*, 646−651.
13. Chakrabartty, A., Kortemme, T., and Baldwin, R. L. (1994) *Protein Sci. 3*, 843−852.
14. Smith, C. K., Withka, J. M., and Regan, L. (1994) *Biochemistry 33*, 5510−5517.
15. Minor, D. L., Jr., and Kim, P. S. (1994) *Nature 367*, 660−663.
16. Narhi, L., Hua, Q.-X., Arakawa, T., Fox, G. M., Tsai, L., Rosenfeld, R., Holst, P., Miller, J. A., and Weiss, M. A. (1993) *Biochemistry 32*, 5214−5221.
17. Hua, Q. X., Narhi, L., Jia, W., Arakawa, T., Rosenfeld, R., Hawkins, N., Miller, J. A., and Weiss, M. A. (1996) *J. Mol. Biol. 259*, 297−313.
18. Dai, Y., and Tang, J. G. (1996) *Biochim. Biophys. Acta 1296*, 63−68.
19. 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) *J. Mol. Biol. 264*, 390−403.
20. Cosmatos, A., and Katsoyannis, P. G. (1975) *J. Biol. Chem. 250*, 5315−5321.
21. Katsoyannis, P. G., Okada, Y., and Zalut, C. (1973) *Biochemistry 12*, 2516−2525.
22. Sieber, P. S., Eisler, K., Kamber, B., Riniker, B., Rittel, W., Marki, F., and deGasparo, M. (1978) *Hoppe-Seyler's Z. Physiol. Chem. 359*, 113−123.
23. Steiner, D. F., and Chan, S. J. (1988) *Horm. Metab. Res. 20*, 443−444.
24. Dodson, G., and Steiner, D. (1998) *Curr. Opin. Struct. Biol. 8*, 189−194.
25. Hua, Q. X., Shoelson, S. E., Kochoyan, M., and Weiss, M. A. (1991) *Nature 354*, 238−241.
26. Olsen, H. B., Ludvigsen, S., and Kaar-sholm, N. C. (1996) *Biochemistry 35*, 8836−8845.
27. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, M. W., Reynolds, C. D., Sakabe, K., Sakabe, N., and Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London, Ser. B 319*, 369−456.
28. Bentley, G., Dodson, G., and Lewitova, A. (1978) *J. Mol. Biol. 126*, 871−875.
29. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C., Smith, G. D., Sparks, C., and Swenson, D. (1989) *Nature 338*, 594−596.
30. Chance, R. E., Hoffman, J. A., Kroeff, E. P., Johnson, M. G., Schirmer, W. E., and Bormer, W. W. (1981) in *Peptides: Proceedings of the Seventh American Peptide Symposium* (Rich, D. H., and Gross, E., Eds.) pp 721−728, Pierce Chemical Co., Rockford, IL.
31. Tang, J. G., and Tsou, C. L. (1990) *Biochem. J. 268*, 429−435.
32. Hua, Q.-X., Gozani, S. N., Chance, R. E., Hoffmann, J. A., Frank, B. H., and Weiss, M. A. (1995) *Nat. Struct. Biol. 2*, 129−138.
33. Okada, Y., Chanley, J. D., Burke, G. T., and Katsoyannis, P. G. (1981) *Hoppe Seyler's Z. Physiol. Chem. 362*, 629−638.
34. Kitagawa, K., Ogawa, H., Burke, G. T., Chanley, J. D., and Katsoyannis, P. G. (1984) *Biochemistry 23*, 1405−1413.
35. Kobayashi, M., Takata, Y., Ishibashi, O., Sasaoka, T., Iwasaki, M., Shigeta, Y., and Inouye, K. (1986) *Biochem. Biophys. Res. Commun. 137*, 250−257.
36. Mirmira, R., and Tager, H. S. (1989) *J. Biol. Chem. 246*, 6349−6354.
37. Schwartz, G. P., Burke, G. T., and Katsoyannis, P. G. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 458−461.
38. Wang, S. H., Hu, S. Q., Burke, G. T., and Katsoyannis, P. G. (1991) *J. Protein Chem. 10*, 313.
39. Mirmira, R. G., and Tager, H. S. (1991) *Biochemistry 30*, 8222−8229.
40. Nakagawa, S. H., and Tager, H. S. (1992) *Biochemistry 31*, 3204−3214.
41. Hu, S. Q., Burke, G. T., Schwartz, G. P., Ferderigos, N., Ross, J. B., and Katsoyannis, P. G. (1993) *Biochemistry 32*, 2631−2635.

42. Nanjo, K., Sanke, T., Miyano, M., Okai, K., Sowa, R., Kondo, M., Nishimura, S., Iwo, K., Miyamura, K., and Given, B. D. (1986) *J. Clin. Invest. 77*, 514−519.
43. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X. G., and Markussen, J. (1991) *J. Mol. Biol. 220*, 425−433.
44. Kaarsholm, N. C., Norris, K., Jorgensen, R. J., Mikkelsen, J., Ludvigsen, S., Olsen, O. H., Sorensen, A. R., and Havelund, S. (1993) *Biochemistry 32*, 10773−10778.
45. Olsen, H. B., Ludvigsen, S., and Kaarsholm, N. C. (1998) *J. Mol. Biol. 284*, 477−488.
46. van Mierlo, C. P., Darby, N. J., Neuhaus, D., and Creighton, T. E. (1991) *J. Mol. Biol. 222*, 373−390.
47. Staley, J. P., and Kim, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1519−1523.
48. Liu, Y., Breslauer, K., and Anderson, S. (1997) *Biochemistry 36*, 5323−5335.
49. Gregoire, F. M., Chomiki, N., Kachinskas, D., and Warden, C. H. (1998) *Biochem. Biophys. Res. Commun. 249*, 385−390.
50. Duret, L., Guex, N., Peitsch, M. C., and Bairoch, A. (1998) *Genome Res. 8*, 348−353.
51. Barany, G., and Merrifield, R. B. (1980) in *The Peptides* (Gross, E., and Meienhofer, J., Eds.) pp 3−284, Academic Press, New York.
52. Hua, Q. X., Zhao, M., Narayana, N., Nakagawa, S. H., Jia, W., and Weiss, M. A. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 1999−2004.
53. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., and Snel, L. (1988) *Nature 333*, 679−682.
54. 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* (Smith, J. A., and Rivier, J. E., Eds.) pp 26−28, ESCOM Science Publishers B. V., Leiden, The Netherlands.
55. Brems, D. N., Alter, L. A., Beckage, M. J., Chance, R. E., DiMarchi, R. D., Green, L. K., Long, H. B., Pekar, A. H., Shields, J. E., and Frank, B. H. (1992) *Protein Eng. 5*, 527−533.
56. Shoelson, S. E., Hua, Q.-X., Lynch, C. E., and Weiss, M. A. (1992) *Biochemistry 31*, 1757−1767.
57. Blaber, M., Zhang, X. J., and Matthews, B. W. (1993) *Science 260*, 1637−1640.
58. Kristensen, C., Kjeldsen, T., Wiberg, F. C., Schaffer, L., Hach, M., Havelund, S., Bass, J., Steiner, D. F., and Andersen, A. S. (1997) *J. Biol. Chem. 272*, 12978−12983.
59. Kobayashi, M., Ohgaku, S., Iwasaki, M., Maegawa, H., Shigeta, Y., and Inouye, K. (1982) *Biochem. Biophys. Res. Commun. 107*, 329−336.
60. Chothia, C., Lesk, A. M., Dodson, G. G., and Hodskin, D. C. (1983) *Nature 302*, 500−505.
61. Katsoyannis, P. G., Ginos, J., Schwartz, G. P., and Cosmatos, A. (1974) *J. Chem. Soc., Perkin Trans. 1*, 1311−1317.
62. Fischer, W. H., Saunders, D., Brandenburg, D., Wollmer, A., and Zahn, H. (1985) *Biol. Chem. Hoppe Seyler 366*, 521−525.
63. Brange, J., Dodson, G. G., Edwards, D. J., Holden, P. H., and Whittingham, J. L. (1997) *Proteins 27*, 507−516.
64. Oas, T. G., and Kim, P. S. (1988) *Nature 336*, 42−48.
65. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. (1992) *Biochemistry 31*, 1749−1756.
66. Miller, J. A., Narhi, L., Hua, Q.-X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M. A. (1993) *Biochemistry 32*, 5203−5213.
67. Hober, S., Hansson, A., Uhlen, M., and Nilsson, B. (1994) *Biochemistry 33*, 6758−6761.
68. States, D. J., Creighton, T. E., Dobson, C. M., and Karplus, M. (1987) *J. Mol. Biol. 195*, 731−739.
69. Eigenbrot, C., Randal, M., and Kossiakoff, A. A. (1990) *Protein Eng. 3*, 591−598.
70. Goldenberg, D. P. (1992) *Trends Biochem. Sci. 17*, 257−261.

BI001905S