# The Different Folding Behavior of Insulin and Insulin-like Growth Factor 1 Is Mainly Controlled by Their B-Chain/Domain †

**ARTICLE** *in* BIOCHEMISTRY · MARCH 2002

Impact Factor: 3.02 · DOI: 10.1021/bi011166v · Source: PubMed



**3 AUTHORS**, INCLUDING:

Zhan-Yun Guo
Tongji University
**61** PUBLICATIONS **589** CITATIONS
SEE PROFILE

You Min Feng
Chinese Academy of Sciences
**74** PUBLICATIONS **687** CITATIONS
SEE PROFILE

# The Different Folding Behavior of Insulin and Insulin-like Growth Factor 1 Is Mainly Controlled by Their B-Chain/Domain†

Zhan-Yun Guo, Lu Shen, and You-Min Feng*

*State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China*

*Received June 6, 2001; Revised Manuscript Received October 19, 2001*

ABSTRACT: Although insulin and insulin-like growth factor 1 (IGF-1) share homologous sequence, similar tertiary structure, weakly overlapped biological activity, and a common ancestor, the two highly homologous sequences encode different folding behavior: insulin folds into one unique stable tertiary structure while IGF-1 folds into two disulfide isomers with similar thermodynamic stability. To further elucidate the molecular mechanism of their different folding behavior, we prepared two single-chain hybrids of insulin and IGF-1, Ins(A)/IGF-1(B) and Ins(B)/IGF-1(A), as well as a mini-IGF-1 by means of protein engineering and studied their structure as well as folding behavior. Both mini-IGF-1 and Ins(A)/IGF-1(B) fold into two thermodynamically stable disulfide isomers in vivo and in vitro just like that of IGF-1, while Ins(B)/IGF-1(A) folds into one unique thermodynamically stable tertiary structure in vivo and in vitro just like that of insulin. So we deduce that the different folding behavior of insulin and IGF-1 is mainly controlled by their B-chain/domain. By V8 endoproteinase digestion and circular dichroism analysis, as well as insulin receptor binding assay, we deduce that Ins(B)/IGF-1(A), isomer 2 of mini-IGF-1, and isomer 2 of Ins(A)/IGF-1(B) adopt native IGF-1/insulin-like three-dimensional structure with native disulfides, while isomer 1 of mini-IGF-1 and isomer 1 of Ins(A)/IGF-1(B) adopt the swap IGF-1-like three-dimensional structure with swap disulfides.

Insulin is a structurally and functionally well-characterized small globular protein containing A- and B-chains linked by three disulfides (one intrachain bond, A6−A11; two interchain bonds, A7−B7 and A20−B19). Its three-dimensional structure has been well studied by X-ray crystallography (*1*, *2*) and NMR (*3*−*5*) since the 1970s (Figure 1A). Although the separate A- and B-chains of insulin can be recombined successfully in vitro (*6*), a single-chain polypeptide (pre-proinsulin) containing the signal peptide at the N-terminus of the B-chain as well as the C-peptide between the B- and A-chains was synthesized in vivo. When B29Lys and A1Gly were linked together by a peptide bond directly, the mini-proinsulin still retained the three-dimensional structure identical to that of insulin but lost biological activity completely (*7*, *8*) (Figure 1B). In the three-dimensional structure of insulin and mini-proinsulin, there are three common segments of α-helix (A2−A8, A13−A19, and B9−B19) which compose the frame and core of the insulin-like fold.

IGF-1[1] is a 70 residue single-chain globular protein composed of B-, C-, A-, and D-domains from the N-terminus to the C-terminus (*9*). Its B- and A-domains are homologous to the B- and A-chains of insulin, respectively; its 12 residue C-domain is analogous to the C-peptide of proinsulin, but they share no homology; its C-terminal 8 residue D-domain has no counterpart in the insulins. IGF-1 adopts an insulin-like structure (*10*) (Figure 1D) whose ordered structure also mainly includes three segments of α-helix (8−18, 42−49, and 54−61) in the A- and B-domains corresponding to those of insulin; the conformation of the C- and D-domains is highly flexible. So, the insulin-like structure of IGF-1 is mainly encoded by its A- and B-domains. When B28Pro and A1Gly were directly connected by a peptide bond, the mini-IGF-1 still retained the three helical segments presented in the native IGF-1, but the relative orientation of the three helixes had been changed (*11*) (Figure 1E) and its biological activity has been lost completely.

Protein folding is still challenging in biological sciences. Since Affinsen and co-workers first demonstrated that the three-dimensional structure of a globular protein is uniquely determined by its primary structure in 1960s (*12*), significant advances have been made in the understanding of protein folding through experimental and theoretical approaches. For small proteins with two-state folding, topology is a major determinant of the folding rate and greatly influences the

† This work was supported by the Chinese Academy of Sciences (KJ951-B1-606) and the National High Technology Program of China (863-103-13-01-01).

* Corresponding author. Tel: (86) 021-64374430. Fax: (86) 021-64338357. E-mail: fengym@sunm.shcnc.ac.cn.

[1] Abbreviations: IGF-1, insulin-like growth factor 1; IGF-2, insulin-like growth factor 2; PIP, recombinant porcine insulin precursor in which the C-terminus of the porcine insulin B-chain and the N-terminus of the porcine insulin A-chain were linked together by a dipeptide, Ala-Lys; BPTI, bovine pancreatic trypsin inhibitor; RNase A, ribonuclease A; EGF, epidermal growth factor; GSH, reduced glutathione; GSSG, oxidized glutathione; DTT, reduced dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid; TCA, trichloroacetic acid; PAGE, polyacrylamide gel electrophoresis; CD, circular dichroism; UV, ultraviolet; NMR, nuclear magnetic resonance; FPLC, fast protein liquid chromatography.

10.1021/bi011166v CCC: $22.00 
Published on Web 01/10/2002

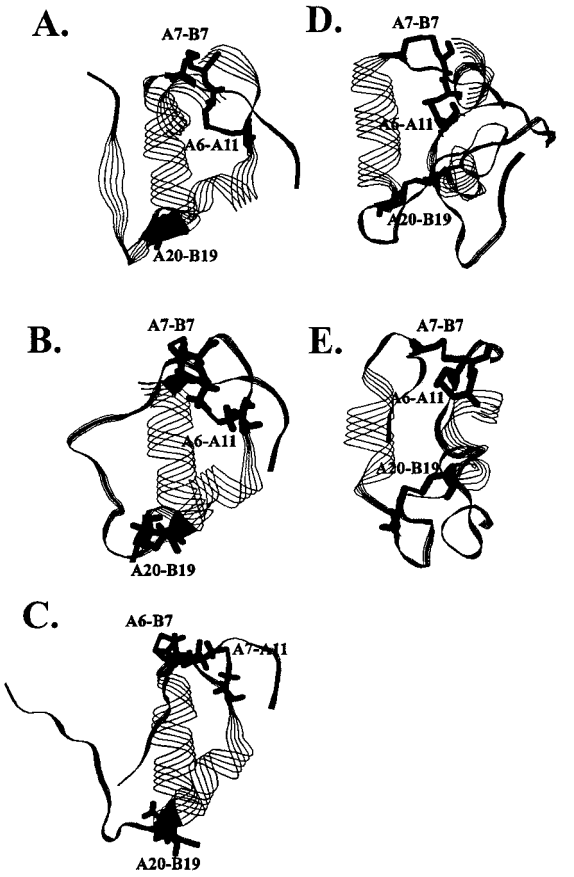



FIGURE 1: Strand models of (A) insulin [R state of crystal structure (*2*)], (B) mini-proinsulin [solution structure (*8*)], (C) swap insulin [solution structure (*28*)], (D) IGF-1 [solution structure (*10*)], and (E) mini-IGF-1 [solution structure (*11*)]. The disulfides were pointed in the molecules. For clarity the disulfides were all numbered as that of insulin. The solution structure of swap IGF-1 was similar to that of swap insulin (*38*), but we did not find its structure data in the Protein Data Bank (3-D).

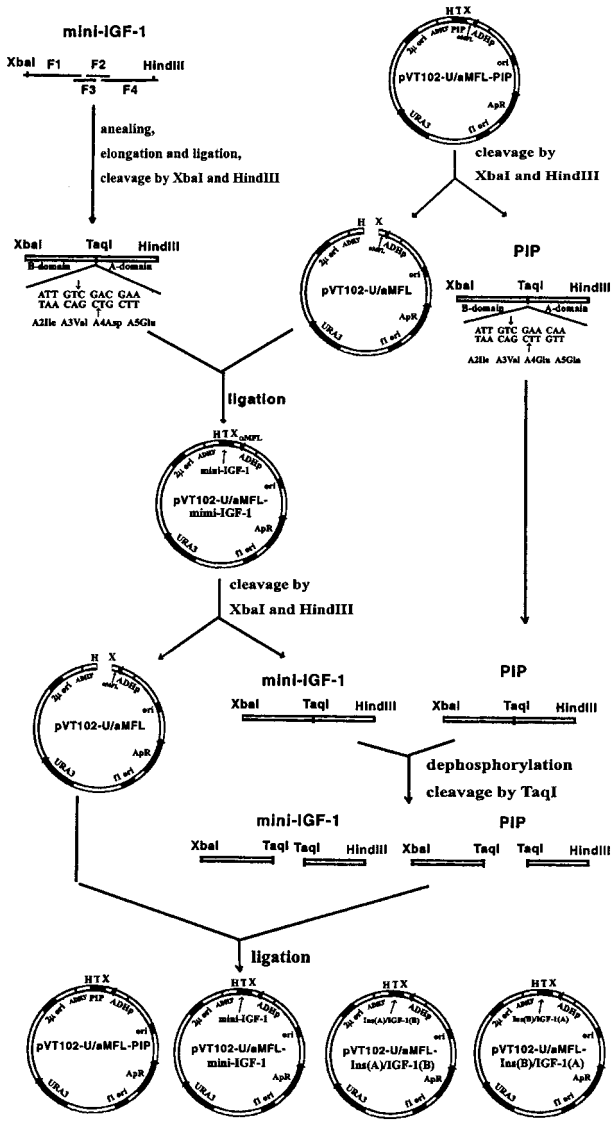
structure of the transition-state ensemble (*13*−*15*). Studies on the disulfide-coupled folding of some small globular proteins, such as BPTI, RNase A, and EGF, have revealed a sequence of preferred kinetic intermediates, which define a folding pathway (*16*−*23*). In vivo the protein folding is assisted by molecular chaperones, especially for large proteins (*24*−*26*); some chaperones even can provide the missing steric information for protein folding (*27*).

Although insulin and IGF-1 share highly homologous sequence, similar tertiary structure, and weakly overlapped biological activity (*9*), the folding behavior of the two homologous sequences is quite different: both insulin and recombinant single-chain insulin (PIP) fold into one unique thermodynamically stable tertiary structure with disulfides (A20−B19, A7−B7, A6−A11) (*28*, *29*), while IGF-1 folds into two disulfide isomers (native and swap) with different three-dimensional structure and different disulfide linkages but similar thermodynamic stability (*30*−*37*). The native IGF-1 adopts an insulin-like structure (*10*) with the disulfides 18−61, 6−48, and 47−52, corresponding to those of insulin; the swap IGF-1 adopts a different folding pattern (*38*) with disulfides 18−61, 6−47, and 48−52. The three-dimensional structure of swap IGF-1 is similar with that of swap insulin whose disulfide bridges are A20−B19, A6−B7, and A7−A11 (*28*) (Figure 1C). The α-helix II presented in the native form is unfolded, and the other two α-helical segments still exist in swap insulin and swap IGF-1. The swap IGF-1 is a thermodynamically controlled folding product with an energy state similar to that of native IGF-1, while the swap insulin is a kinetically controlled folding product that is thermodynamically unstable and only obtained as a kinetic trap (*28*). For most small globular proteins the final folding product is a unique thermodynamically stable three-dimensional structure (*39*), just like that of BPTI, RNase A, and insulin/PIP. Why can the folding of IGF-1 not produce a unique tertiary structure? Why does the highly homologous sequence of insulin and IGF-1 store different folding information? Which part of the sequence is the determinant of their different folding behavior? Although in previous work DiMarchi et al. have demonstrated that it is the A- and B-domains that lead to IGF-1 producing two disulfide isomers (*40*), further detailed information is still unknown. To find out the determinant sequence that controlled the different folding behavior of insulin and IGF-1, we prepared a mini-IGF-1 (the C-terminus of the IGF-1 B-domain and the N-terminus of the IGF-1 A-domain were linked together by a dipeptide, Ala-Lys, and B30Thr was also replaced by Lys), single-chain hybrid Ins(A)/IGF-1(B) (the C-terminus of the IGF-1 B-domain and the N-terminus of the insulin A-chain were linked together by a dipeptide, Ala-Lys, and B30Thr was also replaced by Lys), and single-chain hybrid Ins(B)/IGF-1(A) (the C-terminus of the insulin B-chain and N-terminus of the IGF-1 A-domain were linked together by a dipeptide, Ala-Lys). The amino acid sequence of these molecules is shown in Figure 2B. In the three molecules, the residues are numbered according to that of insulin, so the first residue of mini-IGF-1 and Ins(A)/IGF-1(B) is numbered B2. Here we report the preparation, structure analysis, and folding behavior study of mini-IGF-1 and the two single-chain hybrids of insulin and IGF-1 compared with that of PIP.

## MATERIALS AND METHODS

*Materials*. The *Escherichia coli* strain used was DH12S. *Saccharomyces cerevisiae* cells of strain XV700-6B (leu2, ura3, pep4) were kindly provided by Michael Smith (University of British Columbia, Vancouver, Canada). Plasmid pVT102-U/αMFL-PIP was constructed in our laboratory for secretary expression of PIP in yeast (*41*). PIP was a recombinant single-chain insulin in which the C-terminus of the porcine insulin B-chain and N-terminus of the porcine insulin A-chain were linked together by a dipeptide, Ala-Lys. The chemical reagents used in experiments were of analytical grade. The Pharmacia Biotech reverse-phase columns (Sephasil Peptide C8 5 μm ST 4.6/250 and Sephasil Peptide C4 5 μm ST 4.6/250), Gilson 306 HPLC system, and Gilson 115 UV detector were used. In all HPLC analysis in the experiment, a gradient elution was used with the flow rate of 0.8 mL/min. Solvent A was 0.15% aqueous TFA; solvent B was 60% acetonitrile containing 0.125% TFA. The elution gradient was as follows: 1 min, 0% solvent B; 1 min, 0% solvent B; 5 min, 40% solvent B; 35 min, 80% solvent B; 36 min, 100% solvent B; 38 min, 100% solvent B; 40 min, 0% solvent B; 45 min, 0% solvent B.

*DNA Manipulation*. The expression vectors of mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) were respectively constructed as shown in Figure 2A. To obtain the mini-IGF-1 gene, four partially complementary DNA fragments were




**A**

| | |
|---|---|
| PIP: | FVNQHLCGSHLVEALYLVCGERGFFYTPKA *AK* GIVEQCCTSICSLYQLENYCN |
| Mini-IGF-1: | GPETLCGAELVDALQFVCGDRGFYFNKPK *AK* GIVDECCFRSCDLRRLEMYCA |
| Ins(A)/IGF-1(B): | GPETLCGAELVDALQFVCGDRGFYFNKPK *AK* GIVEQCCTSICSLYQLENYCN |
| Ins(B)/IGF-1(A): | FVNQHLCGSHLVEALYLVCGERGFFYTPKA *AK* GIVDECCFRSCDLRRLEMYCA |

**B**

FIGURE 2: (A) The expression vector construction process of mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A). The expression vector of pVT102-U/αMFL-PIP was constructed in our laboratory for secretory expression of PIP in *S. cerevisiae* previously (*41*). There were cleavage sites of restriction enzymes *Xba*I and *Hin*dIII at the two ends of PIP gene. αMFL was the leading sequence (prepro) of yeast α-mating factor, which was linked to the N-terminus of PIP to help PIP secrete from yeast cells. In the endoplasmic reticulum and Golgi apparatus the αMFL was removed gradually, and then the mature and correctly folded PIP was secreted from yeast cells. H, T, and X on the expression vector represent restriction enzymes *Hin*dIII, *Taq*I, and *Xba*I, respectively. (B) The amino acid sequence of PIP, mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A). A linker dipeptide, Ala-Lys, was used to join the B-chain/domain and A-chain/domain together to form single-chain molecules.

designed and chemically synthesized. After annealing, elongation, and ligation, as well as cleavage by restriction enzymes *Xba*I and *Hin*dIII, the gene of mini-IGF-1 was constructed from the four DNA fragments and then cloned into the large fragment of the expression vector pVT102-U which was obtained by cleavage of the expression vector of PIP (*41*). The DNA sequence of mini-IGF-1 gene was confirmed by DNA sequencing, and the plasmid was designated as pVT102-U/αMFL-mini-IGF-1. To obtain the genes of Ins(A)/IGF-1(B) and Ins(B)/IGF-1(A), the mixture of the PIP gene and the mini-IGF-1 gene (respectively cleaved away from their expression vector by *Xba*I and *Hin*dIII, treated with calf intestine alkaline phosphatase, and mixed as the molar ratio of 1:1) was cleaved into two fragments by the restriction enzyme *Taq* I, whose cleavage site was on the common sequence encoding the N-terminal residues of the A-chain/domain of PIP and mini-IGF-1. Subsequently, the cleaved fragments of the PIP and the mini-IGF-1 gene were cloned into the large fragment of the expression vector pVT102-U (cleaved with *Xba*I and *Hin*dIII). The plasmids of pVT102-U/αMFL-Ins(A)/IGF-1(B) and pVT102-U/αMFL-Ins(B)/IGF-1(A) were selected from the mixture of the four possible vectors by DNA sequencing.

*Expression, Purification, and Identification of the Single-Chain Mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A).* The plasmids pVT102-U/αMFL-mini-IGF-1, pVT102-U/αMFL-Ins(A)/IGF-1(B), and pVT102-U/αMFL-Ins(B)/IGF-1(A) were transformed into *S. cerevisiae* cells of strain XV700-6B (leu2, ura3, pep4), respectively. The transformed yeast cells were cultured in a 16 L fermenter, and the secreted target protein was purified from the medium in four steps (*41*). First, the target protein was precipitated from the medium supernatant by TCA. Second, the precipitant was dissolved with 1 M acetic acid and applied to a Sephadex G-50 column. Third, the product was purified by ion-exchange chromatography on a DEAE-Sepharose CL-6B column. Fourth, the product was further purified by C8 reverse-phase HPLC using a gradient elution described in Materials and Methods and detected at 280 nm. Their purity was analyzed by analytical C8 reverse-phase HPLC and native pH 8.3 PAGE. Their molecular mass was measured by electrospray mass spectrometry.

*Circular Dichroism Measurement of the Single-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), and Ins(B)/IGF-1(A).* The protein concentration of the samples, isomers 1 and 2 of mini-IGF-1, isomers 1 and 2 of Ins(A)/IGF-1(B), Ins(B)/IGF-1(A), and PIP, was determined by Lowry method (*42*). The samples were dissolved in 5 mM HCl, and their final concentration was adjusted to 0.2 mg/mL, respectively. PIP was used as a single-chain standard with insulin-like structure. The measurements were performed on a Jasco-715 circular dichroism spectropolarimeter at room temperature. The cell path length for near-UV spectra (245−300 nm) and far-UV spectra (200−250 nm) measurements was 1.0 and 0.1 cm, respectively. The data were expressed as molar ellipticity. The software "J-700 for windows secondary structure estimation, Version 1.10.00" was used for secondary structural content estimation from far-UV circular dichroism (CD) spectra.

*Cleavage of the Single-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), and Ins(B)/IGF-1(A) with V8 Endoproteinase.* The five molecules, isomers 1 and 2 of mini-IGF-1, isomers 1 and 2 of Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A), were respectively digested with V8 endoproteinase to elucidate their disulfide linkages. Ins(B)/IGF-

1(A) and the two isomers of mini-IGF-1 were dissolved in 0.1 M phosphate buffer (pH 7.8), respectively; the two isomers of Ins(A)/IGF-1(B) were dissolved in 0.1 M $NH_4HCO_3$ buffer (pH 7.8), respectively. The protein concentration of the samples was about 0.5 mg/mL, and the V8 endoproteinase was added to the solution at about a mass ratio of 1:20. The reaction was carried out at 25 °C overnight. After digestion, the solution was adjusted with TFA to pH 2.0 and then separated by C8 reverse-phase HPLC eluted with the gradient described in Materials and Methods and detected at 230 nm. The fractions were collected manually and lyophilized. Isomer 2 of Ins(A)/IGF-1(B) was also digested with V8 endoproteinase in phosphate buffer (pH 7.8), and the digested fragments were separated by C4 reverse-phase HPLC and also eluted with the gradient described in Materials and Methods and detected at 230 nm. The molecular mass of the digested fragments was measured by electrospray mass spectrometry.

*Disulfide Thermodynamic Stability Measurement of the Single-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), Ins(B)/IGF-1(A), and PIP.* The samples, isomers 1 and 2 of mini-IGF-1, isomers 1 and 2 of Ins(A)/IGF-1(B), Ins(B)/IGF-1(A), and PIP, were respectively dissolved in a different redox buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 8.7) at a final concentration of 0.1 mg/mL. In the redox buffer, the ratio (mM/mM) of the reduced glutathione (GSH) and the oxidized glutathione (GSSG) was 1/10, 5/5, 10/1, 20/1, 30/1, and 50/1, respectively. At the same time a negative control (the samples were dissolved in the buffer not containing the redox potential) was carried out. The reaction was carried out at 4 °C overnight. After incubation, one-fifth volume of freshly prepared 0.5 M iodoacetic sodium solution was added to modify the free thiol groups. The carboxymethylation reaction was carried out at room temperature for 5 min. The modified mixture of the sample containing the intermediates with one or more disulfides was reduced, and free thiol groups were carboxymethylated and then analyzed by native pH 8.3 PAGE.

*Disulfide Rearrangement of the Single-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), and Ins(B)/IGF-1(A).* The purified isomers 1 and 2 of mini-IGF-1 and isomers 1 and 2 of Ins(A)/IGF-1(B) were respectively dissolved in 0.1 M Tris-HCl and 1 mM EDTA buffer (pH 8.7) containing 0.2 mM 2-mercaptoethanol at a final concentration of 0.1 mg/mL. The purified Ins(B)/IGF-1(A) was dissolved in 0.1 M Tris-HCl and 1 mM EDTA buffer (pH 10.3) or 0.1 M Tris-HCl and 1 mM EDTA buffer (pH 8.7) containing 0.2 mM 2-mercaptoethanol at a final concentration of 0.1 mg/mL. The disulfide rearrangement reaction was carried out at 10 or 25 °C, respectively. At different reaction times a 100 $\mu$L sample was removed and immediately adjusted to pH 2.0 with TFA to terminate the disulfide rearrangement and analyzed with analytical C8 reverse-phase HPLC. A gradient elution described in Materials and Methods was used and detected at 230 nm.

*In Vitro Refolding of the Single-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), and Ins(B)/IGF-1(A).* The purified isomers 1 and 2 of mini-IGF-1 and isomers 1 and 2 of Ins(A)/IGF-1(B) were respectively dissolved in 0.1 M Tris-HCl and 1 mM EDTA buffer (pH 8.7), while the purified Ins(B)/IGF-1(A) was dissolved in 0.1 M Tris-HCl and 1 mM EDTA buffer (pH 10.3) because of its lower solubility in pH 8.7 buffer, especially after being reduced. The final protein concentration was about 1 mg/mL. The reduction was started by addition of DTT to the final concentration of 50 mM. The reaction was carried out at 30 °C for 1 h. After reduction the aliquot was removed and carboxymethylated by a one-fifth volume of freshly prepared 0.5 M iodoacetic acid sodium salt solution and then analyzed by native pH 8.3 PAGE to determine whether the disulfides of the samples were fully reduced. The reduced samples were immediately exchanged to their refolding buffer (0.1 M Tris-HCl and 1 mM EDTA, pH 10.3, for Ins(B)/IGF-1(A); 0.1 M Tris-HCl and 1 mM EDTA, pH 8.7, for the other four molecules) by gel filtration using a Sephadex G-25 column and adjusted a final protein concentration to 0.1 mg/mL. The refolding reaction was carried out at 10 °C overnight by air oxidation. After incubation a 100 $\mu$L refolding solution was removed and acidified to pH 2.0 by TFA and then analyzed by C8 reverse-phase HPLC eluted by a gradient listed in Materials and Methods and detected at 230 nm.

*Self-Association Measurement of the Single-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), Ins(B)/IGF-1(A), and PIP in Refolding Buffer.* The self-association property of mini-IGF-1 (isomer 1 and 2), Ins(A)/IGF-1(B) (isomer 1 and 2), Ins(B)/IGF-1(A), and PIP in their refolding buffer was determined by gel filtration using FPLC: column, Superdex 75 (HR 10/30); flow rate, 0.4 mL/min; detection, 280 nm; elution buffer, 0.1 M Tris-HCl and 1 mM EDTA (pH 8.7) for mini-IGF-1 (isomer 1 and 2), Ins(A)/IGF-1(B) (isomer 1 and 2), PIP, and [B28Lys, B19Pro]insulin, and 0.1 M Tris-HCl and 1 mM EDTA (pH 10.3) for Ins(B)/IGF-1(A). The samples were dissolved with their corresponding elution buffer and adjusted to a final concentration of 1.0 mg/mL, respectively. In analysis a 100 $\mu$L sample was loaded onto the column. [B28Lys, B29Pro]-insulin was used as a non-self-association control.

*Conversion of the Single-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), and Ins(B)/IGF-1(A) to Double-Chain Forms.* The purified single-chain mini-IGF-1 (isomer 1 and 2) and single-chain Ins(A)/IGF-1(B) (isomer 1 and 2) were respectively dissolved in 0.1 M $NH_4HCO_3$ (pH 8.5) buffer at a final concentration of about 3 mg/mL; Ins(B)/IGF-1(A) was dissolved in 0.1 M $NH_4HCO_3$ (pH 8.5) containing 1 M urea at a final concentration of about 2 mg/mL because of its low solubility in pH 8.5 buffer not containing urea. The endoproteinase Lys-C was added to the solution at a mass ratio of 500:1. Digestion was carried out at 25 °C overnight. After cleavage, the double-chain molecules were respectively purified by C8 reverse-phase HPLC eluted by the gradient listed in Materials and Methods and detected at 280 nm. The purity of the five double-chain molecules was analyzed by analytical C8 reverse-phase HPLC and native pH 8.3 PAGE.

*Binding with Insulin Receptor of the Double-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), and Ins(B)/IGF-1(A).* The protein concentration of the five double-chain molecules was determined by the Lowry method (*42*). The receptor binding assay was performed using human placental membrane as previously described (*43*), using native porcine insulin as a standard. The membrane insulin receptor (total protein is about 250 $\mu$g) was incubated with $^{125}$I-labeled insulin (approximately $10^5$ cpm) plus a selected amount of native insulin or samples in

Table 1: Molecular Mass Measured by Electrospray Mass Spectrometry

| samples | measured values | theoretical values |
|---|---|---|
| mini-IGF-1 (isomer 1) | 5833.0 | 5833.8 |
| mini-IGF-1 (isomer 2) | 5833.0 | 5833.8 |
| Ins(A)/IGF-1(B) (isomer 1) | 5734.0 | 5734.6 |
| Ins(A)/IGF-1(B) (isomer 2) | 5733.0 | 5734.6 |
| Ins(B)/IGF-1(A) | 6058.0 | 6058.1 |

a total volume of 0.4 mL containing 50 mM Tris-HCl and 1.0% BSA, pH 7.5, buffer, at 4 °C overnight. After incubation the unbound $^{125}$I-labeled insulin was washed away with ice-cold 50 mM Tris-HCl and 0.1% BSA, pH 7.5, buffer three times by centrifugation, and the radioactivity of the precipitate was counted. The receptor binding activity of the samples was calculated from the dosages used for 50% inhibition of $^{125}$I-labeled insulin bound to the insulin receptor.

## RESULTS

*DNA Manipulation*. The construction process of the yeast expression vectors of mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) was shown in Figure 2A. Their DNA sequence was confirmed by DNA sequencing (data not shown). The amino acid sequence of PIP, mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) was shown in Figure 2B. In the four molecules the C-terminus of the B-chain/domain of insulin or IGF-1 and the N-terminus of the A-chain/domain of insulin or IGF-1 were respectively joined together by a dipeptide, Ala-Lys. The identical linker sequence of the four molecules made it convenient to compare their folding behaviors. The B30Thr of the B-domain of IGF-1 was mutated into Lys in mini-IGF-1 and Ins(A)/IGF-1(B) for conversion of the single-chain molecules into double-chain forms by endoproteinase Lys-C cleavage.

*Expression, Purification, and Identification of the Single-Chain Mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A).* The expression vectors pVT102-U/αMFL-mini-IGF-1, pVT102-U/αMFL-Ins(A)/IGF-1(B), and pVT102-U/αMFL-Ins(B)/IGF-1(A) were respectively transformed into *S. cerevisiae* cells. The transformants were cultured in a 16 L fermenter, and the expression products were purified in four steps as described in Materials and Methods, respectively. When purified using C8 reverse-phase HPLC, we found that both mini-IGF-1 and Ins(A)/IGF-1(B) contained two components. The component with the shorter retention time was named isomer 1; the component with the longer retention time was named isomer 2. Ins(B)/IGF-1(A) had only one component. The molecular mass of these molecules was measured by electrospray mass spectrometry, and the data are listed in Table 1. All of the measured values were consistent with the theoretical values, which indicated that the primary structure of these molecules was correct. The two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) had almost identical molecular mass, which implied that the two components probably were disulfide isomers just like that of IGF-1.

The five single-chain molecules, isomers 1 and 2 of mini-IGF-1, isomers 1 and 2 of Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A), were first analyzed by analytical C8 reverse-phase HPLC as shown in Figure 3. The sharpness of the peak reflected that the five molecules were homogeneous, respectively. The different retention time of the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) reflected that they had different secondary and/or tertiary structure although their amino acid sequence was identical.

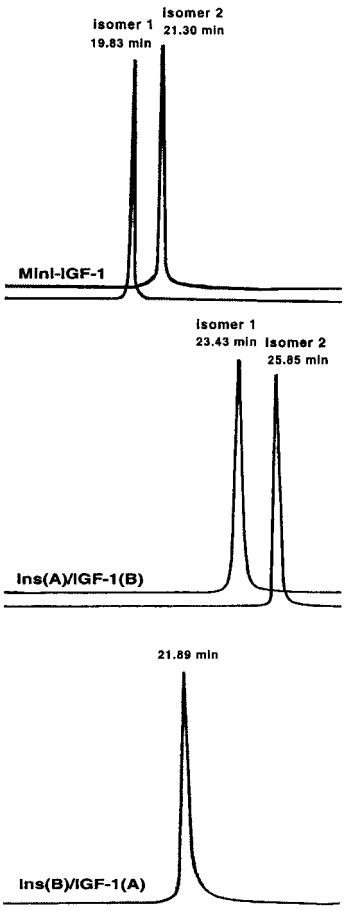



FIGURE 3: C8 reverse-phase HPLC analysis of the single-chain molecules. In the analysis a 100 μL sample (20 μg) was loaded onto a Pharmacia Biotech C8 column (Sephasil Peptide C8 5 μm ST 4.6/250) and eluted using the elution gradient listed in Materials and Methods and detected at 280 nm.

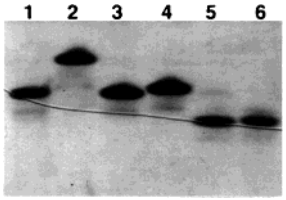



FIGURE 4: Native pH 8.3 PAGE analysis of the single-chain molecules. Lanes 1, 2, 3, 4, 5, and 6 were PIP, Ins(B)/IGF-1(A), isomer 2 of mini-IGF-1, isomer 1 of mini-IGF-1, isomer 2 of Ins(A)/IGF-1(B), and isomer 1 of Ins(A)/IGF-1(B), respectively. A Bio-Rad mini-protean II electrophoresis apparatus was used in the experiments. In each lane a 2−3 μg sample was loaded. The gel was stained by Comassie brilliant blue R250.

We also analyzed the five single-chain molecules by native pH 8.3 PAGE as shown in Figure 4. The two isomers of mini-IGF-1 had a slightly different mobility rate: isomer 2 ran a little faster than isomer 1, while the mobility rate of the two isomers of Ins(A)/IGF-1(B) was almost identical.

The secondary and tertiary structures of the five single-chain molecules were analyzed by far-UV and near-UV CD, respectively (Figure 5), using PIP as a single-chain standard with native insulin-like structure. The far-UV and near-UV

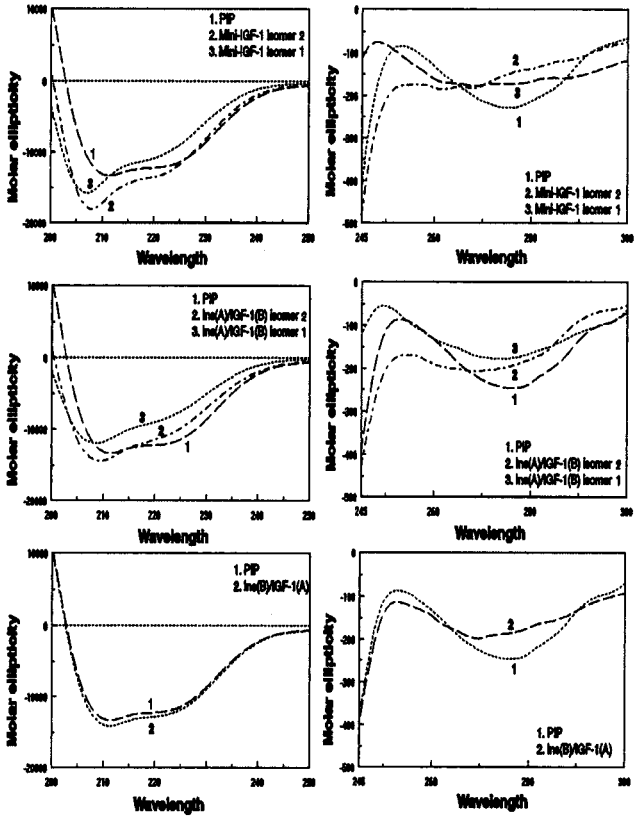



FIGURE 5: Circular dichroism analysis of mini-IGF-1 (isomer 1 and 2), Ins(A)/IGF-1(B) (isomer 1 and 2), Ins(B)/IGF-1(A), and PIP. The left column is the far-UV CD spectra; the right column is the near-UV CD spectra. For clarity the spectra of mini-IGF-1 (isomer 1 and 2), Ins(A)/IGF-1(B) (isomer 1 and 2), and Ins(B)/IGF-1(A) were compared with that of PIP, respectively.

spectra of Ins(B)/IGF-1(A) were very similar to that of PIP, especially the far-UV spectra. The α-helix content of PIP and Ins(B)/IGF-1(A) estimated from far-UV CD spectra was both about 50%, which was consistent with the α-helix content of insulin calculated from the crystal (R state) or NMR structure (49%). So we deduced that Ins(B)/IGF-1(A) adopted an insulin-like structure with native disulfide linkages. The far-UV CD spectra of the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) had some differences. The α-helix content of isomer 2 was close to that of PIP [47% for isomer 2 of mini-IGF-1 and 34% for isomer 2 of Ins(A)/IGF-1(B)], while the α-helix content of isomer 1 was lower than that of PIP [28% for isomer 1 of mini-IGF-1 and 21% for isomer 1 of Ins(A)/IGF-1(B)]. The α-helix content of isomer 2 of mini-IGF-1 estimated from CD was similar to the value calculated from the three helical segments of native IGF-1 (48%). The α-helix content of isomer 1 of mini-IGF-1 estimated from CD was similar to the value calculated from the two helical segments of swap IGF-1 (34%). The α-helix content of isomer 2 and isomer 1 of Ins(A)/IGF-1(B) estimated from CD was lower than the value calculated from native and swap IGF-1, respectively. But isomer 2 still contained more α-helix content than isomer 1. The near-UV spectra of the two isomers were also different, especially in the disulfide absorption region. The spectra of isomer 2 of mini-IGF-1 or Ins(A)/IGF-1(B) were more similar to that of PIP in the disulfide absorption region. So we deduced that isomer 2 adopted a native insulin/IGF-1-like structure with disulfide linkages identical to those of native insulin/IGF-1; isomer 1 adopted a swap insulin/IGF-1-like structure with disulfides identical to those of swap insulin/IGF-1.

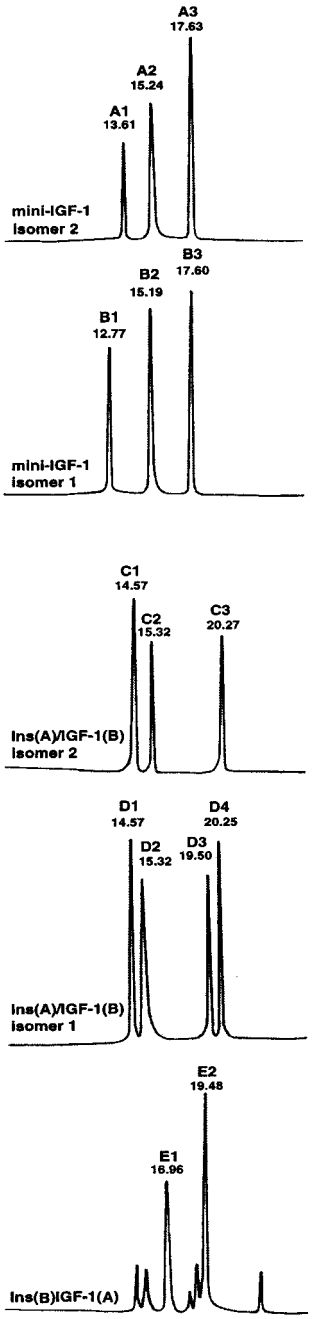



FIGURE 6: HPLC profile of V8-digested mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A). The V8-digested mixture was acidified by TFA and then loaded onto a C8 reverse-phase column and eluted by the gradient described in Materials and Methods and detected at 230 nm.

*Cleavage of the Single-Chain Mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) with V8 Endoproteinase.* The C8 reverse-phase HPLC chromatography of the V8 digested molecules is shown in Figure 6, and the measured molecular masses of the digested fragments as well as their corresponding sequences in the molecules are listed in Table 2. In phosphate buffer, the cleaved products of the two isomers of mini-IGF-1 both contained three peaks; peaks A1 and B1 had different retention times but identical molecular masses, corresponding to the sequence of B5Thr-B10Glu + A6Cys-A12Asp linked together by two disulfides. Their different

Table 2: Molecular Mass of V8 Digested Fragments of Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), and Ins(B)/IGF-1(A)

| samples | HPLC peaks in Figure 6 | retention time on HPLC (min) | measured molecular mass | corresponding sequence in molecule |
|---|---|---|---|---|
| mini-IGF-1 isomer 2 | A1 | 13.61 | 1421.0 | B5Thr-B10Glu + A6Cys-A12Asp |
| | A2 | 15.24 | 1869.0 | B22Arg-A5Glu |
| | A3 | 17.63 | 1336.0 | B14Ala-B21Asp + A18Met-A21Ala |
| mini-IGF-1 isomer 1 | B1 | 12.77 | 1421.0 | B5Thr-B10Glu + A6Cys-A12Asp |
| | B2 | 15.19 | 1869.0 | B22Arg-A5Glu |
| | B3 | 17.60 | 1336.0 | B14Ala-B21Asp + A18Met-A21Ala |
| Ins(A)IGF-1(B) isomer 2 | C1 | 14.57 | 1362.0 | B14Ala-B21Asp + A18Asp-A21Asn |
| | C2 | 15.32 | 1754.0 | B22Arg-A4Glu |
| | C3 | 20.27 | 1689.0 | B11Leu-B21Asp + A18Asp-A21Asn |
| Ins(A)/IGF-1(B) isomer 1 | D1 | 14.57 | 1362.0 | B14Ala-B21Asp + A18Asp-A21Asn |
| | D2 | 15.32 | 1754.0 | B22Arg-A4Glu |
| | D3 | 19.50 | 2078.0 | B5Thr-B10Glu + A5Gln-A17Glu |
| | D4 | 20.25 | 1689.0 | B11Leu-B21Asp + A18Asp-A21Asn |
| Ins(B)/IGF-1(A) | E1 | 16.96 | 1799.0 | B22Arg-A5Glu |
| | E2 | 19.48 | 1351.0 | B14Ala-B21Glu + A18Met-A21Ala |

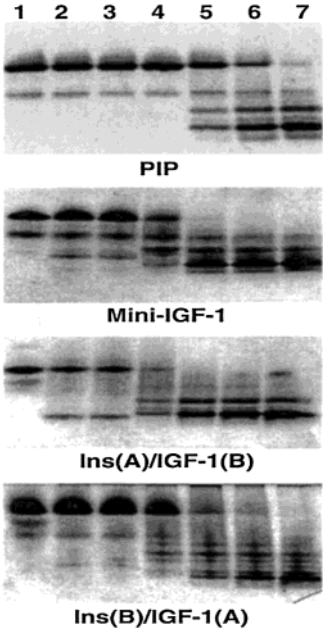



FIGURE 7: Disulfide thermodynamic stability in redox buffer. Lanes 1, 2, 3, 4, 5, and 6 represent that in redox buffer, and the ratio of GSH/GSSG (mM/mM) was 0/0, 1/10, 5/5, 10/1, 20/1, 30/1, and 50/1, respectively. When the disulfides were reduced in the redox buffer and then the free thiol groups were carboxymethylated, the molecule carried more negative charges and ran faster on the PAGE, but the conformation also had an effect on their mobility rate. The gel was stained by Comassie brilliant blue R250.

retention times must be caused by their different disulfide linkages. In $NH_4HCO_3$ buffer, the cleavage sites of V8 endoproteinase were not highly specific to Glu residue, so some sites of Asp were also cleaved. When separated by C8 reverse-phase HPLC, the fragment of isomer 2 of Ins(A)/IGF-1(B) corresponding to peak D3 of isomer 1 was not found. So we also digested isomer 2 of Ins(A)/IGF-1(B) in phosphate buffer and separated the fragments with C4 reverse-phase HPLC where a fragment with a molecular mass of 2078.0, corresponding to D3 of isomer 1, was found (data not shown). The peptide mapping demonstrated that the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) both contained the common disulfides A20−B19, so their different disulfide linkages must be among A6Cys, A7Cys, A11Cys, and B7Cys. Together with the results of the above circular dichroism analysis and their biological activity measured after conversion to double-chain forms, we deduced that isomer 1 was the swap form with disulfides identical to those of swap IGF-1 and isomer 2 was the native form with disulfides identical to those of insulin/native IGF-1. Digestion of Ins(B)/IGF-1(A) produced two major peaks, one corresponding to the sequence containing A20−B19 disulfide. The fragment containing the other two disulfides was not detected. However, its circular dichroism spectra and biological activity indicated that its disulfides were native form.

*Disulfide Thermodynamic Stability of Single-Chain Mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) in Redox Buffer.* Although PIP and native IGF-1 shared highly homologous sequence, their disulfide stability in redox buffer was different: the disulfides of PIP were more stable than those of IGF-1. In redox buffer which favors disulfide formation for most proteins, some disulfides of IGF-1 can be reduced, mainly disulfides 47−52 and 6−48 (*30*, *31*). Here we measured the disulfide thermodynamic stability of the single-chain mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) in redox buffer compared with that of PIP as shown in Figure 7. Since the two chains of insulin will separate in redox buffer, the disulfide stability comparison between insulin and the single-chain molecules is not suitable. The disulfide stability of the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) was similar, so only the PAGE pattern of isomer 2 was shown in this figure. When the disulfides were reduced in the redox buffer and then the free thiol groups were carboxymethylated, the molecules carried more negative charges and ran faster than those intact molecules on pH 8.3 PAGE. Generally, the more disulfides were reduced in the redox buffer, the more negative charges it carried, and the faster the molecules ran on the pH 8.3 PAGE, but the conformation of the intermediates also had influence on their mobility rate. The more stable the disulfides of the molecules were, the higher the reducing potential (the ratio of GSH/GSSG) needed to reduce the disulfides. From the PAGE pattern we can see that the disulfides of PIP were the most stable in redox buffer. The disulfides of mini-IGF-1 and Ins(A)/IGF-1(B) were the most unstable; the disulfide stability of Ins(B)/IGF-1(A) was between that of PIP and mini-IGF-1. So we deduced that the B-chain/domain was more important than the A-chain/domain in controlling the different disulfide stability of PIP and IGF-1.

*Disulfide Rearrangement of Single-Chain Mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A).* If both of the two isomers of mini-IGF-1 or Ins(A)IGF-1(B) were thermodynamically stable species, the two isomers should rearrange their disulfides in the alkaline buffer containing thiol catalyst and finally form an equilibrium. Here we measured the disulfide rearrangement of the two isomers of mini-IGF-1

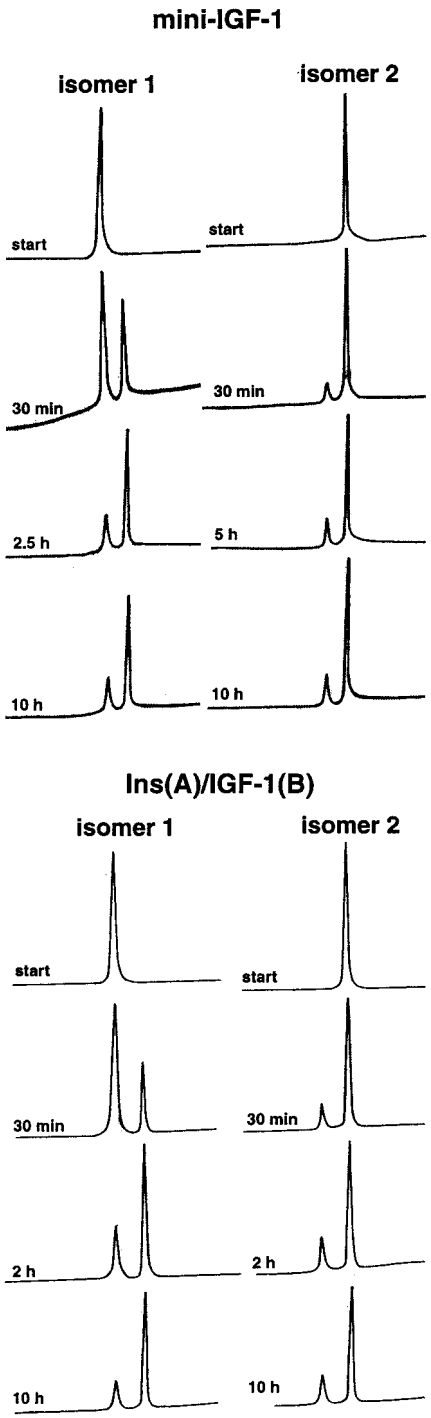



FIGURE 8: Disulfide rearrangement of the two isomers of mini-IGF-1 and Ins(A)/IGF-1(B). At different reaction times a 100 $\mu$L (10 $\mu$g) sample was removed and acidified by TFA and immediately loaded onto a C8 reverse-phase column. The elution gradient described in Materials and Methods was used and detected at 230 nm.

or Ins(A)/IGF-1(B) as shown in Figure 8. Without thiol catalyst the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) were stable in alkaline or acidic buffer; when 0.2 mM 2-mercaptoethanol was added into the alkaline buffer, the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) rearranged their disulfides gradually and finally formed an equilibrium mixture of the two isomers. The ratios of isomer 1 to isomer 2 (calculated from the peak area of HPLC) of mini-IGF-1 and Ins(A)/IGF-1(B) in equilibrium at 25 °C were both about 23:77. At 10 °C the ratios were both about 35:65. The folding

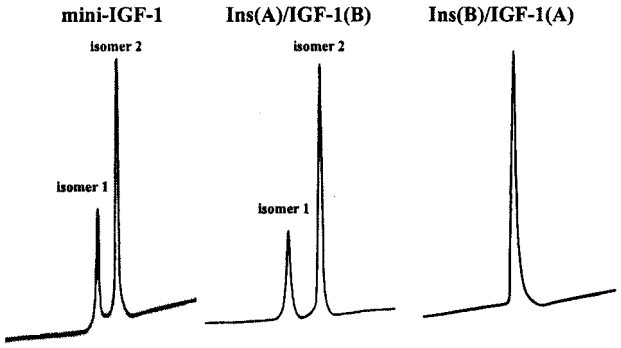



FIGURE 9: In vitro refolding of fully reduced mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A). After the fully reduced sample was incubated in the refolding buffer overnight, a 100 $\mu$L sample was removed and acidified by TFA and then analyzed by C8 reverse-phase HPLC. The elution gradient described in Materials and Methods was used and detected at 230 nm.

free energy difference ($\Delta\Delta G$) of isomer 2 to isomer 1 of mini-IGF-1 and Ins(A)/IGF-1(B) calculated from their equilibrium constant was both about −3.0 kJ/mol at 25 °C and −1.5 kJ/mol at 10 °C. The folding reaction was temperature dependent: at elevated temperatures, isomer 2 was the more favored folding product. These folding properties of mini-IGF-1 and Ins(A)/IGF-1(B) were similar with that of IGF-1. Here we demonstrated that the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) were thermodynamically controlled folding products with similar thermodynamic stability. The disulfide rearrangement of Ins(B)/IGF-1(A) was not detectable. So we deduced that it was the B-domain that mainly controlled IGF-1, forming two disulfide isomers with similar thermodynamic stability.

*In Vitro Refolding of the Single-Chain Mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A).* To further demonstrate that the sequence of mini-IGF-1 and Ins(A)/IGF-1(B) encoded two thermodynamically stable folding products while the sequence of Ins(B)/IGF-1(A) encoded a unique tertiary structure, in vitro refolding was carried out. After the samples were incubated in the alkaline buffer containing DTT, an aliquot was removed and treated with iodoacetic acid sodium salt solution and then analyzed with pH 8.3 PAGE. The PAGE pattern demonstrated that all of the disulfides of the incubated samples were reduced (data not shown). The refolding reaction was carried out at 10 °C overnight by air oxidation, and then the folding products were analyzed by C8 reverse-phase HPLC as shown in Figure 9. The HPLC profile of the refolding products of fully reduced isomer 1 and isomer 2 of mini-IGF-1 or Ins(A)/IGF-1(B) was similar, so only the refolding results of isomer 2 were shown here. Both mini-IGF-1 and Ins(A)/IGF-1(B) refolded into two products which had retention times identical to those of the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B), respectively. The refolding of fully reduced Ins(B)/IGF-1(A) produced only one product which had a retention time identical to that of the intact Ins(A)/IGF-1(B). These results further demonstrated that the amino acid sequence of mini-IGF-1 and Ins(A)/IGF-1(B) encoded two disulfide isomers with similar thermodynamic stability, while the sequence of Ins(B)/IGF-1(A) encoded a unique thermodynamically stable three-dimensional structure.

*Self-Association Analysis of the Single-Chain Mini-IGF-1, Ins(A)/IGF-1(B), Ins(B)/IGF-1(A), and PIP in Refolding*

*Buffer*. We analyzed the self-association property of these single-chain molecules in their refolding buffer by gel filtration (the chromatography results were not shown). The non-self-association control, [B28Lys, B29Pro]insulin, had a symmetric and sharp peak on the chromatograph. The PIP, isomer 1 of mini-IGF-1, and Ins(A)/IGF-1(B) (isomer 1 and 2) all appeared as symmetric peaks with retention times similar to that of [B28Lys, B29Pro]insulin on the chromatograph in their refolding buffer, which indicated that these samples did not self-associate even at a concentration of 1.0 mg/mL. So, in the refolding reaction with a final concentration of 0.1 mg/mL, these samples must not self-associate. The peaks of Ins(B)/IGF-1(A) and isomer 2 of mini-IGF-1 were not very symmetric at their refolding buffer, which indicated that they had somewhat self-associated at a concentration of 1.0 mg/mL in their refolding buffer. When PIP was analyzed in 0.1 M Tris-HCl and 1 mM EDTA (pH 7.4), its elution peak was very asymmetrical. So PIP self-associated at this condition but did not self-associate in the refolding buffer.

*Conversion of the Single-Chain Mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) to Double-Chain Forms.* When insulin bound with its receptor, the conformation of its C-terminus of the B-chain was changed to expose the conservative residues on the N-terminus of the A-chain (*44*). So mini-proinsulin lost biological activity completely although it still retained the three-dimensional structure identical to that of insulin (*7*, *8*). PIP retained only about 0.1% receptor binding activity (our unpublished data) although it contained the two linker residues as well as B30Ala. We deduced that the insulin receptor binding activities of the single-chain mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) were also too low to be accurately quantified. So, the five single-chain molecules were treated with endoproteinase Lys-C to convert them into double-chain forms, respectively. These double-chain molecules were respectively purified from the digestion mixture by C8 reverse-phase HPLC. All of them were homogeneous as analyzed by C8 analytical reverse-phase HPLC and pH 8.3 PAGE (data not shown).

*Binding with Insulin Receptor of the Double-Chain Mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A).* The insulin receptor binding activities of the five double-chain molecules, mini-IGF-1 (isomer 1 and 2), Ins(A)/IGF-1(B) (isomer 1 and 2), and Ins(B)/IGF-1(A), were measured as shown in Figure 10. Their relative insulin receptor binding activities calculated from the dosages used for 50% inhibition of $^{125}$I-labeled insulin bound with insulin receptor are listed in Table 3. The insulin receptor binding activity of isomer 2 of mini-IGF-1 or Ins(A)/IGF-1(B) was about 10-fold higher than that of the corresponding isomer 1, which indicated that isomer 2 had native disulfides while isomer 1 had swap disulfides. The insulin receptor binding activity of Ins(B)/IGF-1(A) was higher than that of native insulin (176% of insulin), which indicated that its disulfides were also the native form.

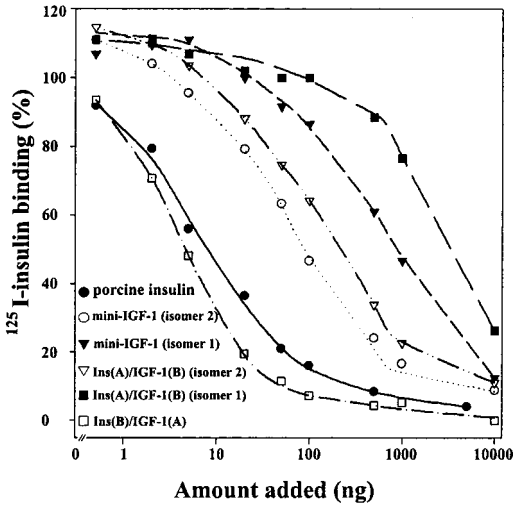



FIGURE 10: Insulin receptor binding activities of the double-chain mini-IGF-1 (isomer 1 and 2), Ins(A)/IGF-1(B) (isomer 1 and 2), and Ins(B)/IGF-1(A).

Table 3: Relative Insulin Receptor Binding Activity of the Double-Chain Mini-IGF-1 (Isomer 1 and 2), Ins(A)/IGF-1(B) (Isomer 1 and 2), and Ins(B)/IGF-1(A) (% of Native Porcine Insulin)

| | samples | | | | |
|---|---|---|---|---|---|
| | mini-IGF-1 | | Ins(A)/IGF-1(B) | | Ins(B)/ |
| | isomer 1 | isomer 2 | isomer 1 | isomer 2 | IGF-1(A) |
| insulin receptor binding activity | 0.9 | 10 | 0.2 | 3.8 | 176 |

## DISCUSSION

The insulin superfamily includes insulin, IGF-1, IGF-2, relaxin, bombyxin, and some other structural-related species. All of these members (native form) share a common insulin-like structure containing three motif-specific disulfides and three common helical segments (*45*−*47*). The three helical segments compose the frame and core of the insulin-like structure. The role of the three disulfides is to stabilize the three helical segments (*48*−*50*). All of these critical folding elements are located in the A- and B-chain/domain, so we deduced that the A- and B-chain/domain contained most of the folding information of the insulin-like structure. In this experiment we constructed two single-chain hybrids of insulin and IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A), as well as a single-chain mini-IGF-1 containing only the A- and B-domains of IGF-1. Together with our previously prepared PIP, we got a set of single-chain molecules containing the A- and B-chain/domain of insulin or IGF-1 with two linker residues. The four molecules are suitable for investigating the different folding behavior of insulin and IGF-1.

In CD analysis, the α-helix content of PIP, Ins(B)/IGF-1(A), and mini-IGF-1 (isomer 1 and 2) estimated from far-UV spectra is well consistent with the values calculated from the three (native) or two (swap) helical segments of the crystal or NMR structure of insulin or IGF-1. The α-helix content of Ins(A)/IGF-1(B) (isomer 1 and 2) estimated from CD spectra is somewhat lower than that calculated from NMR spectra of native and swap IGF-1, but isomer 2 still has higher α-helix content than isomer 1. Since the α-helix content estimated from CD is not very accurate, this result is also acceptable. This indicates that PIP, Ins(B)/IGF-1(A), isomer 2 of mini-IGF-1, and isomer 2 of Ins(A)/IGF-1(B) all contain three helical segments just like that of native insulin or native IGF-1, while isomer 1 contains two helical segments just like that of swap insulin or swap IGF-1. The

near-UV spectra indicate that Ins(B)/IGF-1(A), isomer 2 of mini-IGF-1, and isomer 2 of Ins(A)/IGF-1(B) adopt a native insulin-like three-dimensional structure, while isomer 1 adopts a swap IGF-1-like structure. When the sample concentration was diluted from 1.2 to 0.12 mg/mL, the CD spectra had little change, which implied that the concentration of these samples had little influence on the CD spectra in the 5 mM HCl buffer. By V8 endoproteinase cleavage, we demonstrated that mini-IGF-1 (isomer 1 and 2), Ins(A)/IGF-1(B) (isomer 1 and 2), and Ins(B)/IGF-1(A) all contained the common A20−B19 disulfide. After being converted to double-chain forms, the insulin receptor binding activities of isomer 2 of mini-IGF-1 or Ins(A)/IGF-1(B) were 10-fold higher than that of the corresponding isomer 1; the Ins(B)/IGF-1(A) was more potent that native insulin. These results all indicated that Ins(B)/IGF-1(A), isomer 2 of mini-IGF-1, and isomer 2 of Ins(A)/IGF-1(B) were native form while isomer 1 was swap form.

Mini-proinsulin (B29Lys and A1Gly were connected directly) adopts a three-dimensional structure identical to that of insulin (*7*, *8*). We thought that the three-dimensional structure of PIP was also identical to that of insulin. The previous reported mini-IGF-1 (B28Pro and A1Gly were connected directly and contained the D-domain) adopts a three-dimensional structure different from that of native IGF-1: although the three helical segments still existed, their relative orientation had changed slightly (*11*). However, the present mini-IGF-1 contained a dipeptide linker between the A- and B-domains as well as a B29Thr residue and lacked the D-domain. The structural change of the previously reported mini-IGF-1 perhaps was caused by the direct connection of A- and B-domains (1−28 residue). Containing the dipeptide linker and B29Thr residue, we thought the present mini-IGF-1 (native form) perhaps adopts a conformation identical to that of native IGF-1. The growth-promoting activity of the present single-chain mini-IGF-1 (native) was a little higher than that of its double-chain form (our unpublished data), while the previous mini-IGF-1 lost biological activity completely.

The major differences of the folding behavior of PIP, mini-IGF-1, Ins(A)/IGF-1(B), and Ins(B)/IGF-1(A) are as follows: PIP and Ins(B)/IGF-1(A) secreted from yeast as a single form, while mini-IGF-1 and Ins(A)/IGF-1(B) secreted from yeast as two isomers. In disulfide rearrangement analysis, the two isomers of mini-IGF-1 or Ins(A)/IGF-1(B) can rearrange their disulfides gradually and form an equilibrium finally, while the disulfide rearrangement of Ins(B)/IGF-1(A) and PIP was not detectable. In refolding analysis, fully reduced PIP and Ins(B)/IGF-1(A) produced one final product, while fully reduced mini-IGF-1 and Ins(A)/IGF-1(B) produced two disulfide isomers. In disulfide thermodynamic stability analysis, the disulfide stability of the molecule containing the B-chain of insulin was higher than that of the molecule containing the B-domain of IGF-1. These results all demonstrated that the B-domain played a critical role in controlling the different folding behavior of insulin and IGF-1.

We analyzed the self-association property of these single-chain molecules in their refolding buffer at a concentration of 1.0 mg/mL. PIP, isomer 1 of mini-IGF-1, and Ins(A)/IGF-1(B) (isomer 1 and 2) do not self-associate on the chromatograph, so they must not self-associate in the refolding condition with 10-fold lower concentration. Isomer 2 of mini-IGF-1 and Ins(B)/IGF-1(A) somewhat self-associate on the chromatograph, but whether they self-associate in the refolding condition (0.1 mg/mL) is not clear. According to the present and some other results we deduced that the folding behavior had no evident relationship with the self-association property. PIP did not self-associate in the refolding condition, but it folded into one product. When B12Val was replaced by Ser or Ala, PIP lost the self-association ability even at neutral pH, but they still refolded into one product (our unpublished data). The self-association of native mini-IGF-1 was somewhat unusual considering that IGF-1 was a monomer; we thought the present result of native mini-IGF-1 needed other methods to confirm. Anyway, the recent results demonstrated that the folding behavior and self-association property had no evident relationship.

IGF-1 and PIP not only had different folding thermodynamics but also had different folding kinetics (*29*). In the in vitro refolding of PIP, there are two folding pathways: in one pathway the disulfide A6−A11 forms first; in the other pathway the disulfide A20−B19 forms first. But its final folding product is only the native PIP. In the in vitro refolding of IGF-1, the first formed disulfide is 18−61 (corresponding to disulfide A20−B19 in PIP), and the intra-A-domain disulfide forms slowly. Its final folding product is two thermodynamically stable isomers. The fast formation of the A20−B19 disulfide in PIP and 18−61 disulfide in IGF-1 is the result that the ordered structure formed in refolding takes the two Cys residues into proximity in space. The disulfide deletion analysis has proved that this disulfide is the most important for maintaining the order structure of insulin and IGF-1 (*36*, *50*). The fast formation of the A6−A11 disulfide in PIP is due to this disulfide having the lowest free energy (*51*) as well as the two Cys residues near in sequence; since disulfide 47−52 in IGF-1 is a strained bond with high energy, it forms slowly in refolding although the two Cys residues are near in sequence. From the above analysis, we see that the different folding kinetics of PIP and IGF-1 are influenced by the different energy state of the disulfides. Here we deduce that the B-chain/domain is more important than the A-chain/domain in controlling the different in vitro refolding kinetics of PIP and IGF-1.

The free energy difference of the two isomers of IGF-1, mini-IGF-1, and Ins(A)/IGF-1(B) is too small to preclude the swap isomer in the folding controlled by thermodynamics. This phenomenon is somewhat similar to the energy landscape of the funnel model (*52*). The swap form is analogous to the folding intermediate as well as the globule state, but the energy state of this intermediate/globule state is similar to that of the native from. The structure of swap IGF-1 is analogous to the folding intermediate/globule state because the α-helix II present in the native form is unfolded in the swap form. In the structure of folding intermediates (or intermediate models) of insulin and IGF-1 with the deletion of the intra-A-chain/domain disulfide the α-helix II is also unfolded (*49*, *53*, *54*). The unfolding of the α-helix II must elevate the energy state of the molecule, which resulted in an unstable structure. In swap IGF-1 the free energy elevation is somewhat counteracted by the swap disulfide linkages because disulfide 47−52 in native IGF-1 is a high-energy bond (*30*). In swap insulin/PIP, the free energy elevation cannot be counteracted by the swap disulfide linkages since

disulfide A6−A11 is still a stable bond. So insulin/PIP folds into one ground state while IGF-1 folds into two ground states. The energy state of the intra-A-chain/domain disulfide probably directly resulted in the different folding behavior of insulin/PIP and IGF-1. We found that the intra-A-domain disulfide of Ins(A)/IGF-1(B) was also a high-energy bond like that of IGF-1, while the intra-A-domain disulfide of Ins(B)/IGF-1(A) was stable (our unpublished data). So the B-domain probably plays a critical role in controlling the energy state of the intra-A-chain/domain disulfide. But how B-domain affects the energy state of the intra-A-chain/domain disulfide is still not clear. A detailed dissection of the B-chain/domain must cast more information on this issue.

In the insulin superfamily, insulin and IGF-1 share high sequence homology, similar tertiary structure, weakly overlapped biological activity, and a common ancestor (*55*), but the reason they evolved different folding behavior is still unknown. In vivo IGF-1 is bound with IGF binding proteins which do not bind to insulins. The pioneering work of the Nilsson laboratory demonstrated that the folding of IGF-1 was thermodynamically controlled by its binding proteins (*37*). Considering the fact that the B-domain of IGF-1 is critical for interaction with IGF binding proteins (*56*−*58*), we deduce that the unusual folding property of IGF-1 is coevolved with its binding proteins. The residues critical for binding with IGF binding proteins on the B-domain of IGF-1 probably are important to the unusual folding behavior of IGF-1. Three residues, 48Phe, 49Arg, and 50Ser, in the A-domain of IGF-1, are also important for interaction with IGF binding proteins, but when the corresponding residues of PIP were replaced by the three residues of IGF-1, the folding property of PIP had little change (our unpublished data). So the three residues in the A-domain were not responsible for the different folding behavior of insulin and IGF-1.

## ACKNOWLEDGMENT

We thank the anonymous reviewers for the helpful suggestions and valuable discussions on the manuscript.

## REFERENCES

1. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N., and Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London B319*, 369−456.
2. The Peking Insulin Structure Research Group (1974) *Sci. Sin. 17*, 752−778.
3. Weiss, M. A., Hua, Q.-X., Frank, B. H., Lynch, C., and Shoelson, S. E. (1991) *Biochemistry 30*, 7373−7389.
4. Roy, M., Lee, R. W.-K., Brange, J., and Dunn, M. F. (1990) *J. Biol. Chem. 265*, 5448−5453.
5. Olsen, H. B., Ludvigsen, S, and Kaarsholm, N. C. (1996) *Biochemistry 35*, 8836−8845.
6. Wang, C. C., and Tsou, C. L. (1991) *Trends Biochem. Sci. 16*, 279−281.
7. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X. G., and Markussen, J. (1991) *J. Mol. Biol. 220*, 425−433.
8. Hua, Q.-X., Hu, S.-Q., Jia, W., Chu, Y.-C., Burke, G. T., Wang, S.-H., Wang, R.-Y., Katsoyannis, P. G., and Weiss, M. A. (1998) *J. Mol. Biol. 277*, 103−118.
9. Humbel, R. E. (1990) *Eur. J. Biochem. 190*, 445−462.
10. Cook, R. M., Harvey, T. S., and Campbell, I. D. (1991) *Biochemistry 30*, 5484−5491.
11. Wolf, E. D., Gill, R., Geddes, S., Pitts, J., Wollmer, A., and Grotzinger, J. (1996) *Protein Sci. 5*, 2193−2202.
12. Anfinsen, C. B. (1973) *Science 181*, 223−230.
13. Plaxco, K. W., Simons, K. T., and Baker, D. (1998) *J. Mol. Biol. 277*, 985−994.
14. Viguera, A. R., Serrano, L., and Wilmanns, M. (1996) *Nat. Struct. Biol. 3*, 874−880.
15. Baker, D. (2000) *Nature 405*, 39−42.
16. Creighton, T. E. (1974) *J. Mol. Biol. 87*, 603−642.
17. Creighton, T. E., Darby, N. J., and Kemmink, J. (1996) *FASEB J. 10*, 110−118.
18. Darby, N. J., Morin, P. E., Talbo, G., and Creighton, T. E. (1995) *J. Mol. Biol. 249*, 463−447.
19. Weissman, J. S., and Kim, P. S. (1991) *Science 253*, 1386-1393.
20. Weissman, J. S., and Kim, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 497−526.
21. Wedemeyer, W. J., Welker, E., Narayan, M, and Scheraga, H. A. (2000) *Biochemistry 39*, 4207−4216.
22. Rothwarf, D. M., and Scheraga, H. A. (1993) *Biochemistry 32*, 2671−9216.
23. Wu, J., Yang, Y., and Watson, J. T. (1998) *Protein Sci. 7*, 1017−1028.
24. Grantcharova, V., Alm, E. J., Baker, D., and Horwich, A. (2001) *Curr. Opin. Struct. Biol. 11*, 70−81.
25. Ellgaard, L., Molinari, M., and Helenius, A. (1999) *Science 286*, 1882−1888.
26. Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) *Science 286*, 1888−1893.
27. Barnhart, M. M., Pinkner, J. S., Soto, G. E., Sauer, F. G., Langermann, S., Waksman, G., Frieden, C., and Hultgren, S. J. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 7709−7714.
28. 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.
29. Qiao, Z.-S., Guo, Z.-Y., and Feng, Y.-M. (2001) *Biochemistry 40*, 2662−2668.
30. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. (1992) *Biochemistry 31*, 1749−1751.
31. Hober, S., Ljung, J. L., Uhlen, M., and Nilsson, B. (1999) *FEBS Lett. 443*, 271−276.
32. Miller, J. A., Narhi, L., Hua, Q.-X., Rosenfeld, R., Arakawa, T., Rodhe, M., Prestrelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M. A. (1993) *Biochemistry 32*, 5203−5213.
33. Hober, S., Ljung, J. L., Uhlen, M., and Nilsson, B. (1997) *Biochemistry 36*, 4616−4622.
34. Rosenfeld, R. D., Miller, J. A., Narhi, L. O., Hawkins, N., Katta, V., Lauren, S. L., Weiss, M. A., and Arakawa, T. (1997) *Arch. Biochem. Biophys. 342*, 298−305.
35. Milner, S. J., Carver, J. A., Ballard, F. J., and Francis, G. (1999) *Biotechnol. Bioeng. 62*, 693−703.
36. Narhi, L. O., Hua, Q. X., Arakawa, T., Fox, G. M., Tsai, L., Rosenfeld, R, Holst, P., Miller, J. M., and Weiss, M. A. (1993) *Biochemistry 32*, 5214−5221.
37. Hober, S., Hansson, A., Uhlen, M., and Nilsson, B. (1994) *Biochemistry 33*, 6758−6761.
38. Sato, A., Koyama, S., Yamada, H., Suzuki, S., Tamura, K., Kobayashi, M., Niwa, M., Yasuda, T., Kyogoku, Y., and Kobayashi, Y. (2000) *J. Pept. Res. 56*, 218−230.
39. Kim, P. S., and Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
40. DiMarchi, R. D., Mayer, J. P., Fan, L., Brems, D. N., Frank, B. H., Green, L. K., Hoffmann, 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 12th American Peptide Symposium* (Smith, J. A., and River, J. E., Eds.) pp 26−28, ESCOM Science Publisher B.V., Leiden, The Netherlands.
41. Zhang, Y.-S., Hu, H.-M., Cai, R.-R., Feng, Y.-M., Zhu, S.-Q., He, Q.-B., Tang, Y.-H., Xu, M.-H., Xu, Y.-G., Liu, B., and Liang, Z.-H. (1996) *Sci. China* (*Ser. C*) *39*, 225−233.
42. Lowrey, O. H., Rosebrough, N. O., Farr, A. L., and Randall, R. (1951) *J. Biol. Chem. 193*, 265−275.

43. Feng, Y. M., Zhu, J. H., Zhang, X. T., and Zhang, Y. S. (1982) *Acta Biochim. Biophys. Sin. 14*, 137−143.
44. Hua, Q.-X., Shoelson, S. E., Kochoyan, M., and Weiss, M. A. (1992) *Nature 354*, 238−241.
45. Terasawa, H., Kohda, D., Hatanaka, H., Nagata, K., Higashihashi, N., Fujiwara, H., Sakano, K., and Inagaki, F. (1994) *EMBO J. 13*, 5590−5597.
46. Eigenbrot, C., Randal, M., Quan, C., Burnier, J., O'Connell, L., Rinderknecht, E., and Kossiakoff, A. A. (1991) *J. Mol. Biol. 221*, 15−21.
47. Nagata, K., Hatanaka, H., Kohda, D., Kataoka, H., Nagasawa, H., Isogai, A., Ishizaki, H., Suzuki, A., and Inagaki, F. (1995) *J. Mol. Biol. 253*, 749−758.
48. 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.
49. 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.
50. Guo, Z.-Y., and Feng, Y.-M. (2001) *Biol. Chem. 382*, 443−448.
51. Cecil, R., and Weitzman, J. (1964) *Biochem. J. 93*, 1−11.
52. Onuchic, J. N., Nymeyer, H., Garcia, A. E., Chahine, J., and Socci, N. D. (2000) *Adv. Protein Chem. 53*, 87−152.
53. 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.
54. Weiss, M. A., Hua, Q.-X., Jia, W., Chu, Y.-C., Wang, R.-Y., and Katsoyannis, P. G. (2000) *Biochemistry 39*, 15429−15440.
55. Chan, S. J., Cao, Q. P., and Steiner, D. F. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9319−9323.
56. De Vroede, M. A., Rechler, M. M., Nissley, S. P., Joshi, S., Burke, G. T., and Katsoyannis, P. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3010−3014.
57. Chen, Z. Z., Schwartz, G. P., Zong, L., Burke, G. T., Chanley, J. D., and Katsoyannis, P. G. (1988) *Biochemistry 27*, 6105−6111.
58. Bayne, M. L., Applebaum, J., Chicchi, G. G., Hayes, N. S., Green, B. G., and Cascieri, M. A. (1988) *J. Biol. Chem. 263*, 6233−6239.

BI011166V