Sequences of B-Chain/Domain 1−10/1−9 of Insulin and Insulin-like Growth Factor 1 Determine Their Different Folding Behavior †

ARTICLE *in* BIOCHEMISTRY · AUGUST 2004

Impact Factor: 3.02 · DOI: 10.1021/bi049710y · Source: PubMed

Sequences of B-Chain/Domain 1−10/1−9 of Insulin and Insulin-like Growth Factor 1 Determine Their Different Folding Behavior†

Yan Chen, You You, Rui Jin, Zhan-Yun Guo, and You-Min Feng*

Key Laboratory of Proteomics, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences Graduate School, The Chinese Academy of Sciences, Shanghai 200031, China

Received February 9, 2004; Revised Manuscript Received May 14, 2004

ABSTRACT: Although insulin and insulin-like growth factor-1 (IGF-1) belong to one family, insulin folds into one thermodynamically stable structure, while IGF-1-folds into two thermodynamically stable structures (native and swap forms). We have demonstrated previously that the bifurcating folding behavior of IGF-1 is mainly controlled by its B-domain. To further elucidate which parts of the sequences determine their different folding behavior, by exchanging the N-terminal sequences of mini-IGF-1 and recombinant porcine insulin precursor (PIP), we prepared four peptide models: [1−9]PIP, [1−10]mini-IGF-1, [1−4]PIP, and [1−5]mini-IGF-1 by means of protein engineering, and their disulfide rearrangement, V8 digestion, circular dichroic spectra, disulfide stability, and in vitro refolding were investigated. Among them only [1−9]-PIP, like mini-IGF-1/IGF-1, was expressed in yeast as two isomers: isomer 1 (corresponding to swap IGF-1) and isomer 2 (corresponding to native IGF-1), which are supported by the experimental results of disulfide rearrangements, peptide mapping of V8 endoprotenase digests, circular dichroic analysis, in vitro refolding, and disulfide stability analysis. The other peptide models, [1−10]mini-IGF-1, [1−4]PIP, and [1−5]mini-IGF-1, fold into one stable structure as PIP does, which indicates that sequence 1−4 of mini-IGF-1 is important for the folding behavior of mini-IGF-1/IGF-1 but not sufficient to lead to a bifurcating folding. The results demonstrated that the folding information, by which mini-IGF-1/IGF-1-folds into two thermodynamically structures, is encoded/written in its sequence 1−9, while sequences 1−10 of B chain in insulin/PIP play an important role in the guide of its unique disulfide pairing during the folding process.

Insulin is a structurally and functionally well-characterized small globular protein with A- and B-chains linked by three disulfides. Its three-dimensional structure has been well defined by X-ray crystallography (*1*) and NMR (*2*−*4*) since the 1970s. Insulin-like growth factor-1 (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 (*5*). 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 insulin. The three-dimensional structure of IGF-1 has also been well-defined by X-ray crystallography (*6*, *7*) and NMR (*8*, *9*). Insulin and IGF-1 belong to the same superfamily. They have a common structural motif (*10*, *11*) and share homologous sequence, similar three-dimensional structure (*1*, *6*, *7*), and a common ancestor (*12*, *13*). Both mainly consist of three α-helical segments (A2−A8, A13−A19, and B9−B19 in insulin; 8−18, 42−49, and 54−61 in IGF-1) in the A- and B-chains/domains; the three α-helical segments are stabilized by three identical disulfides (A6−A11, A7−B7, and A20−B19 in insulin; 47−52, 6−48, and 18−61 in IGF-1); and when IGF-1 and the insulin T-form are superimposed on the Cα positions of their respective helical segments, the RMSD is only 0.47 Å (*6*).

However, the folding behaviors of insulin and IGF-1 are quite different. Insulin can fold into one unique three-dimensional structure, while IGF-1 is a protein that encodes two thermodynamically stable folding structures. In 1992, Hober et al. (*14*) observed that in vitro refolding of IGF-1 resulted in two major products with different disulfide pairing. One product was native IGF-1, the other product was named swap IGF-1 because of its nonnative disulfide bonds (6−47 and 48−52). Native IGF-1 and swap IGF-1 have different three-dimensional structures. Native IGF-1 has three α-helical segments (*6*, *7*), while swap IGF-1 has only two α-helical segments (*15*). The bifurcating folding behavior of IGF-1 is also present under in vivo-like conditions (*16*). Different from IGF-1, the swap form of insulin (with the disulfide pairings A7−B6, A6−A11, and A20−B19) was only obtained as a kinetic trap with the existence of guanidine hydrochloride and DTT (*17*).

† This work was supported by the National Foundation of Natural Sciences (30170209) and the Chinese Academy of Sciences (KJ951-B1-606).

* Corresponding author: Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China. Tel (86) 021-54921133; fax (86) 021-54921011; e-mail fengym@sunm.shcnc.ac.cn.

[1] Abbreviations: IGF-1, insulin-like growth factor 1; PIP, recombinant single-chain insulin in which the C-terminus of porcine insulin B-chain and the N-terminus of porcine insulin A-chain were linked together by a dipeptide, Ala-Lys; aILP, amphioxus insulin-like peptide; ESI MS, electrospray ionization mass spectrometry; 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; RMSD, root means square deviation.

10.1021/bi049710y CCC: $27.50
Published on Web 06/25/2004

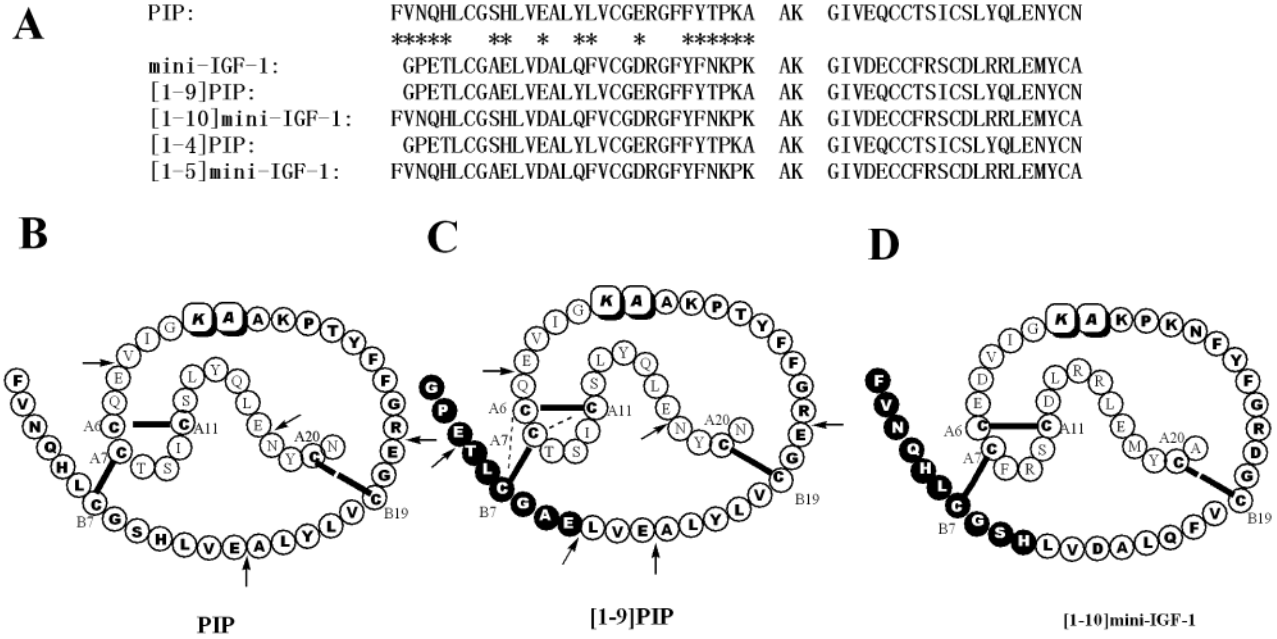

FIGURE 1: Design of the four peptide models. (A) Amino acid sequences of PIP, mini-IGF-1, [1−9]PIP, [1−10]mini-IGF-1, [1−4]PIP, and [1−5]mini-IGF-1. A linker dipeptide, Ala-Lys, was used to join the B-chain/domain and A-chain/domain together to form single-chain molecules. Asterisks indicate nonidentical amino acids between the B-chain/domain of insulin and mini-IGF-1. (B) Primary structure of PIP. (C) Primary structure of [1−9]PIP, in which the B-chain sequence 1−10 of PIP was replaced by the B-domain sequence 1−9 of mini-IGF-1 (residues in solid circles). (D) Primary structure of [1−10]mini-IGF-1, in which the B-domain sequence 1−9 of mini-IGF-1 was replaced by the B-chain sequence 1−10 of PIP (residues in solid circles). In panels B−D, for convenient description, the sequence number refers to insulin A- and B-chains. The native disulfides (A6−A11, A7−B7, and A20−B19) are indicated by solid lines; nonnative disulfides (A7−A11 and A6−B7) are indicated by dotted lines. The arrows indicate the potential cleavage sites of V8 endoproteinase in panels B and C.

Genetic information can be transferred from DNA to polypeptide through a simple triplet coding mechanism. However, the mechanism that governs the transfer of information in protein from primary sequence to three-dimensional structure is still unclear. As early as the 1960s, Anfinsen and co-workers first demonstrated that the three-dimensional structure of a globular protein is uniquely determined by its amino acid sequence (*18*). Since then, significant advances have been made in the understanding of protein folding through experimental and theoretical approaches. However, such a "one sequence, one structure" theory is being challenged by discoveries that one sequence may fold into more than one three-dimensional structure; for example, prion protein (*19*), formin binding protein domain (*20*), and phosphoglycerate kinase (*21*) can adopt totally different folded conformations. Yet it is still ambiguous why one protein has different folding structures. Revealing the structural basis by which IGF-1 (one sequence) folds into two thermodynamically stable folding forms (two structures) while insulin folds into one unique structure would not only have implications for understanding the different folding behavior of insulin and IGF-1 but also help to understand how protein folding information is encoded/written in its sequence.

Porcine insulin precursor (PIP) is a recombinant single-chain insulin precursor, in which the C-peptide of porcine proinsulin was replaced by a dipeptide, Ala-Lys (Figure 1B) (*22*). PIP not only can be converted into recombinant human insulin, which had been used clinically, but also can be used as a peptide model to study the folding behavior of insulin (*23*). Mini-IGF-1 (the C-terminus of IGF-1 B-domain and the N-terminus of IGF-1 A-domain were linked together by a dipeptide, Ala-Lys, and B30Thr was also replaced by Lys) has bifurcating folding behavior similar to that of IGF-1 (*24*). Previously, on the basis of studies of folding behavior of mini-IGF-1, PIP, and hybrids Ins(A)/IGF-1(B) and Ins(B)/IGF-1(A), we have demonstrated that the different folding behavior of insulin/PIP and IGF-1/mini-IGF-1 is mainly controlled by their B-chain/domains (*24*). The different energetic states of the intra-A-chain/domain disulfide (*14*, *25*), which had been suggested to correspond to the different folding behavior of insulin and IGF-1, are also controlled by their B-chain/domains (*26*). Further mapping in the B-chain/domain of insulin/IGF-1 is needed to reveal the structural basis of the different folding behavior of insulin and IGF-1. We noted that the N-terminal sequences 1−9 of mini-IGF-1 and 1−10 of PIP are quite different in sequence and contain one Cys, which participates in the swap and native disulfide pairing (Figure 1A). It is possible that the sequences within 1−9/1−10 of IGF-1 and insulin may encode the information for their different folding behavior. So we use mini-IGF-1 (*24*) and PIP (*22*) as peptide models to test this hypothesis. We prepared the following four peptide models by means of protein engineering (*22*, *24*)—[1−9]PIP, in which the sequence 1−10 of PIP was replaced by the sequence 1−9 of IGF-1; [1−10]mini-IGF-1, in which 1−9 of IGF-1 was substituted with 1−10 of PIP; as well as [1−4]PIP and [1−5]mini-IGF-1 (Figure 1)—and compared their structure and folding behavior with that of mini-IGF-1 and PIP. The results demonstrated that [1−9]PIP, like mini-IGF-1/IGF-1, folds into two thermodynamically stable structures, while [1−10]mini-IGF-1, [1−4]PIP, and [1−5]-mini-IGF-1 fold into one stable structure as PIP does. Thus we deduced that the folding information, by which mini-

IGF-1/ IGF-1 folded into two thermodynamically structures, is encoded/written in its sequence 1−9, while sequence 1−10 of B-chain in insulin/PIP plays an important role in the guide of its unique disulfide pairing during the folding process.

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, BC, Canada). Plasmids pVT102-U/αMFL-PIP and pVT102-U/αMFL-mini-IGF-1 were previously constructed in our laboratory for secretory expression of PIP and mini-IGF-1 in yeast (*16*, *19*). The chemical reagents used in experiments were analytical-grade. The Pharmacia Biotech reverse-phase column (Sephasil Peptide C8 5μm ST 4.6/250), Gilson 306 HPLC system, and Gilson 115 UV detector were used. In HPLC analysis, gradient elution was used. Solvent A was 0.15% aqueous TFA, and solvent B was 60% acetonitrile containing 0.125% TFA. The elution gradient was as follows: 0 min, 0% solvent B; 1 min, 0% solvent B; 5 min, 45% solvent B; 32 min, 90% solvent B; 33 min, 100% solvent B; 35 min, 100% solvent B; 38 min, 0% solvent B; 42 min, 0% solvent B. During analysis, the flow rate is 0.5 mL/min, and the detection wavelength is 230 nm.

DNA Manipulation. The expression plasmids encoding [1−9]PIP and [1−4]PIP were constructed by a gapped duplex DNA approach for site-directed mutagenesis on the basis of pVT102-U/αMFL-PIP (*19*). The expression plasmids encoding [1−10]mini-IGF-1 and [1−5]mini-IGF-1 were constructed in the same way on the basis of pVT102-U/αMFL-mini-IGF-1. The expected mutations were confirmed by DNA sequencing. The four plasmids constructed were designated as pVT102-U/αMFL-[1−9]PIP, pVT102-U/αMFL-[1−4]PIP, pVT102-U/αMFL-[1−10]mini-IGF-1, and pVT102-U/αMFL-[1−5]mini-IGF-1.

Expression, Purification, and Identification of [1−9]PIP, [1−4]PIP, [1−10]Mini-IGF-1, and [1−5]Mini-IGF-1. The plasmids pVT102-U/αMFL-[1−9]PIP, pVT102-U/αMFL-[1−4]PIP, pVT102-U/αMFL-[1−10]mini-IGF-1, and pVT102-U/αMFL-[1−5]mini-IGF-1 were transformed into *S. cerevisiae* cells of strain XV700-6B (leu2 ura3 pep4). The transformed yeast cells were cultured in a 16-L fermenter and the secreted target protein was purified from the medium in four steps (*16*, *19*). Briefly, 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 with gradient elution as described under Materials 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.

Disulfide Rearrangements of [1−9]PIP (Isomers 1 and 2), [1−4]PIP, [1−10]Mini-IGF-1, and [1−5]Mini-IGF-1. The purified proteins were 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 disulfide rearrangement reaction was carried out at 10 °C. After 10 h of incubation, a 100 μL sample was removed, immediately adjusted to pH 2.0 with TFA to terminate the disulfide rearrangement, and analyzed with analytical C8 reverse-phase HPLC. Gradient elution as described under Materials was used, with detection at 230 nm.

Peptide Mapping of V8 Endoproteinase Digests. PIP and [1−9]PIP isomers 1 and 2 were dissolved in 0.1 M phosphate buffer (pH 7.8), and then V8 endoproteinase was added to the solution at about a mass ratio of 1:10 (enzyme to protein). Digestion was performed at 37 °C for 16 h. After digestion, the solution was adjusted to pH 2.0 with TFA and analyzed by C8 reverse-phase HPLC, with elution by the gradient described under Materials.

Circular Dichroic Analysis. The samples were dissolved in 1 mM HCl. The protein concentration was determined by UV absorbance at 280 nm, and their final concentration was adjusted to 0.2 mg/mL. Circular dichroic measurements were performed on a Jasco-715 spectropolarimeter at room temperature. The near-UV spectra were scanned from 300 to 245 nm in a cell with a path length of 1.0 cm; the far-UV spectra were scanned from 250 to 190 nm in a cell with a path length of 0.1 cm. 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 spectra.

In Vitro Refolding of [1−9]PIP (Isomers 1 and 2) and [1−10]Mini-IGF-1. The purified isomers 1 and 2 of [1−9]-PIP, as well as [1−10]mini-IGF-1, were dissolved in 0.1 M Tris-HCl and 1 mM EDTA buffer (pH 8.7) with 8 M urea. The final protein concentration was 1 mg/mL. Adding DTT to a final concentration of 100 mM initiated the reduction. The reaction was carried out at 37 °C for 1 h. The reduced samples were immediately exchanged to 10 mM HCl by gel filtration on a Sephadex G-25 column and stored at −80 °C for later use. The refolding reaction was initiated by diluting the reduced samples with the refolding buffer (0.5M L-Arg, pH 9.6) at a final protein concentration of 0.1 mg/mL. Refolding was carried out at 25 °C overnight by air oxidization. After incubation, 100 μL of refolding solution was removed, acidified to pH 2.0 by TFA, and then analyzed by C8 reverse-phase HPLC, with elution by the gradient described under Materials and detection at 230 nm.

Disulfide Thermodynamic Stability Comparison of Isomer 2 of [1−9]PIP, [1−10]Mini-IGF-1, Native Form of Mini-IGF-1, and PIP. The sample, [1−9]PIP (isomer 2), [1−10]-mini-IGF-1, native form of mini-IGF-1, and PI, were dissolved in the different redox buffer (0.1 M Tris-HCl and 1 mM EDTA, pH 8.7) at a final concentration of 0.1 mg/mL. In the redox buffer, the ratio of GSH and GSSG was 5 mM/5 mM. At the same time a negative control (the samples were dissolved in buffer not containing the redox components) was carried out. The reaction was carried out at 4 °C for 16 h. After incubation, samples were adjusted to pH 2.0 with TFA to terminate the reaction and analyzed by analytical C8 reverse-phase HPLC. Gradient elution as described under Materials was used, with detection at 230 nm.

RESULTS

[1−9]PIP Was Expressed in Yeast as Two Isomers while [1−10]Mini-IGF-1, [1−4]PIP, and [1−5]Mini-IGF-1 Were Expressed as Single Component. Plasmids pVT102U/

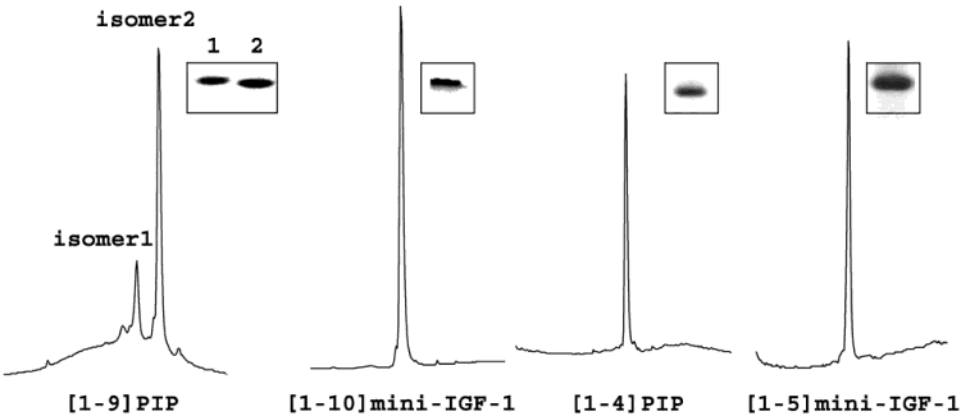

FIGURE 2: [1−9]PIP was expressed in yeast as two isomers while [1−10]mini-IGF-1, [1−4]PIP, and [1−5]mini-IGF-1 were expressed as single components. The C8 reverse-phase HPLC analysis and PAGE patterns of [1−9]PIP, [1−10]mini-IGF-1, [1−4]PIP, and [1−5]mini-IGF-1 were performed according to Materials and Methods. The four peptide models are homogeneous in PAGE. Among the four peptide models, only [1−9]PIP has two isomers like mini-IGF-1. The [1−9]PIP isomer with shorter retention time was named isomer 1; the isomer with longer retention time was named isomer 2. On native PAGE, isomer 1 move a little slower than isomer 2, which implies that the conformation of isomer 1 is less compact than that of isomer 2.

α-MFL-[1−9]PIP, pVT102U/α-MFL-[1−10]mini-IGF-1, pVT102U/α-MFL-[1−4]PIP, and pVT102U/α-MFL-[1−5]-mini-IGF-1 were constructed with correct DNA sequences (data not shown), transformed into yeast cells, and expressed in a 16-L fermenter, and the expression products [1−9]PIP, [1−10]mini-IGF-1, [1−4]PIP, and [1−5]mini-IGF-1 were purified in four steps, as described under Materials and Methods. Their molecular masses were measured by electrospray mass spectrometry: the measured values for [1−9]-PIP, [1−10]mini-IGF-1, [1−4]PIP, and [1−5]mini-IGF-1 are 5694, 6099, 5717, and 6075, which are virtually identical with their theoretical values 5694, 6099, 5718, and 6075, respectively. The purity of the peptides was analyzed by native pH 8.3 PAGE and C8 reverse-phase HPLC (Figure 2). Our previous work has shown that when mini-IGF-1 and PIP were analyzed by C8 reverse-phase HPLC, mini-IGF-1 has two peaks like IGF-1: the peak with shorter retention time was named as isomer 1 (swap form with two nonnative disulfides, 48−52 and 6−47, and one native disulfide, 18−61) and the other peak was named as isomer 2 (native form with three native disulfides, 47−52, 6−48, and 18−61); while PIP has only one peak with disulfides A6−A11, A7−B7, and A20−B19, which are identical with native disulfides 47−52, 6−48, and 18−61 of mini-IGF-1. Using the same procedures, we analyzed the four peptide models, in which only [1−9]PIP, like mini-IGF-1, has two peaks; these were also named isomers 1 and 2 (Figure 2). Therefore, we suggest that replacement of the N-terminal sequence 1−10 of PIP with the N-terminal sequence 1−9 of mini-IGF-1 results in the exchange of folding behavior between PIP and mini-IGF-1. In other words, [1−9]PIP, like mini-IGF-1, folds into two thermodynamically stable structures (Figure 1C), while the other three peptides, [1−10]mini-IGF-1 (Figure 1D), [1−4]PIP, and [1−5]mini-IGF-1, exhibit folding behavior like PIP; that is, they fold into only one thermodynamically stable structure.

Disulfide Rearrangement of the Disulfide Isomers of [1−9]-PIP, [1−10]Mini-IGF-1, [1−4]PIP, and [1−5]Mini-IGF-1. If both of the isomers of [1−9]PIP 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 incubate the purified isomer 1 or 2 in an alkaline buffer containing 0.2 mM 2-mercaptoethanol, and this equilibrium process was monitored by HPLC as shown in Figure 3. Both isomers will reshuffle their disulfide bonds to form a mixture of isomers 1 and 2 during this process, and the final equilibration ratio of isomer 1 to 2 is 1:7, no matter whether the reaction is initiated from isomer 1 or isomer 2 of [1−9]PIP. As a control, both isomers are stable in similar buffer without free thiol catalyst (data not shown). This disulfide rearrangement experiment further confirmed that these two isomers of [1−9]-PIP are thermodynamically controlled folding products with different disulfide linkages. If we compare the disulfide reshuffling process of isomers 1 and 2 at different time points, we can see that the conversion from isomer 1 into isomer 2 is faster than the reverse direction (data not shown). This implies that isomer 2 of [1−9]PIP is more kinetically preferred than isomer 1 during the folding process. Disulfide rearrangement of [1−10]mini-IGF-1, [1−4]PIP, and [1−5]-mini-IGF-1 was also carried out under the same conditions and the results show that they remained stable as a single fraction and no other isomers can be observed during this process, which indicated that their folding behavior is similar to that of PIP (Figure 3).

From the results mentioned above we demonstrated that the folding information, by which mini-IGF-1/IGF-1-folds into two thermodynamically structures, is encoded/written in its sequence 1−9. The folding behavior of [1−4]PIP and [1−5]mini-IGF-1 is similar to that of PIP, which indicates that sequence 1−4 of mini-IGF-1 cannot change the folding behavior of PIP, but the substitution of 1−4 of mini-IGF-1 with 1−5 of PIP makes [1−5]mini-IGF-1 abolish the folding behavior of mini-IGF-1, which suggests that sequence 1−4 of mini-IGF-1 is important for the folding behavior of mini-IGF-1/IGF-1 but not sufficient to lead to a bifurcated folding.

Peptide Mapping of V8 Endoproteinase-Digested [1−9]-PIP Isomers 1 and 2. V8 endoproteinase can cleave at the carboxyl terminus of Glu: the potential V8 cleavage sites in [1−9]PIP are indicated by arrows in Figure 1C. Previously, we have identified V8 digestion fragments of PIP (Figure 1B) (*23*). Because the sequences of [1−9]PIP and PIP are identical except for their N-terminal sequences, V8 digestion of [1−9]PIP and PIP should result in some identical peptide

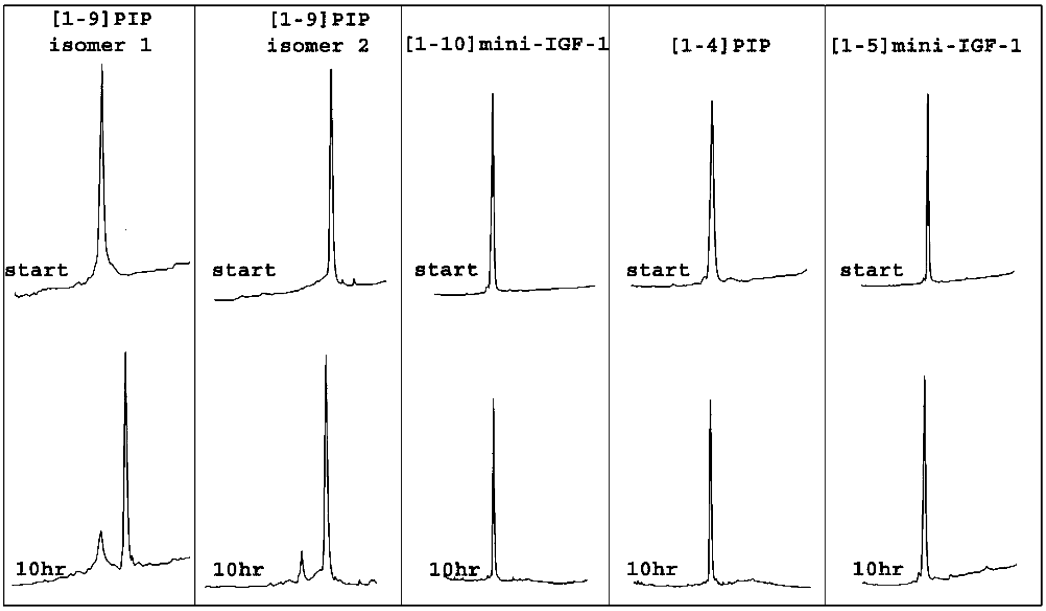

FIGURE 3: Disulfide rearrangement of the two isomers of [1−9]PIP, [1−10]mini-IGF-1, [1−4]PIP, and [1−5]mini-IGF-1. After 10 h of incubation in an alkaline buffer containing 0.2 mM 2-mercaptoethanol, a 100 μL (10 μg) sample was removed and acidified by TFA and immediately loaded onto a C8 reverse-phase column and analyzed by HPLC according to Materials and Methods. The two isomers of [1−9]PIP were interconvertible and can reach equilibrium.

fragments. To understand the disulfide linkages of the isomers of [1−9]PIP, we compare the HPLC profiles of V8-digested [1−9]PIP isomers with that of V8-digested PIP. As shown in Figure 4, peak B1 in [1−9]PIP isomer 1 and peak C1 in [1−9]PIP isomer 2 have almost identical retention times with peak A1 of PIP (it was identified to be the fragment B14A−B21E + A18N−A21N, which contains disulfide A20−B19) (*23*). Hence we deduce that both isomers 1 and 2 of [1−9]PIP contain disulfide A20−B19. The difference in disulfide pairing of isomers 1 and 2 must result from the other two disulfides formed between B7Cys, A6Cys, A7Cys, and A11Cys.

Circular Dichroic Analysis of [1−9]PIP Isomers and [1−10]Mini-IGF-1. To obtain more structural information, the CD spectra of [1−9]PIP two isomers compared with PIP and [1−10]mini-IGF-1 compared with the two isomers of mini-IGF-1 were analyzed as shown in Figure 5. The far-UV CD spectrum of [1−9]PIP isomer 2 is very similar to that of PIP, and their estimated α-helix contents are very close: 38.2% for PIP and 38.9% for [1−9]PIP isomer 2. The estimated α-helix content of [1−9]PIP isomer 1 is only 20%, which is close to that of the swap form of mini-IGF-1 (28.2%) The correct disulfide pairing is crucial for the maintenance of the three α-helical segments in insulin/PIP. Thus we inferred that [1−9]PIP isomer 2 retains three α-helical segments of insulin/PIP and adopts native disulfide pairing of insulin (A20−B19, A7−B7, and A6−A11), while [1−9]PIP isomer 1 adopts the swap form of mini-IGF-1 with nonnative disulfides A6−B7 and A7−A11. The α-helix content of [1−10]mini-IGF-1 (37.5%) is close to that of the native form of mini-IGF-1 (38.8%), so we inferred that [1−10]mini-IGF-1 should have native disulfide linkages (corresponding to insulin disulfides A20−B19, A7−B7, and A6−A11) and a nativelike structure. The near-UV spectra also show that [1−9]PIP isomer 2 is similar to PIP, isomer 1 is similar to mini-IGF-1 swap form, and [1−10]mini-IGF-1 is similar to mini-IGF-1 native form.

In Vitro Refolding of [1−9]PIP and [1−10]Mini-IGF-1. To further demonstrate that the sequence of [1−9]PIP encoded two thermodynamically stable folding products while the sequence of [1−10]mini-IGF-1 encoded a unique 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 by 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, which showed that the reduced [1−9]PIP can also refold into two products, which had retention times identical to those of the two isomers of [1−9]PIP (Figure 6). The refolding of fully reduced [1−10]mini-IGF-1 produced only one product, which had a retention time identical to that of the intact [1−10]mini-IGF-1. These results further demonstrated that the amino acid sequence of [1−9]PIP encoded two disulfide isomers with similar thermodynamic stabilities, while the sequence of [1−10]mini-IGF-1 encoded one unique thermodynamically stable structure.

Disulfide Stability of [1−9]PIP Native Form and [1−10]-Mini-IGF-1. Although PIP and native IGF-1 share highly homologous sequences, their disulfide stability in redox buffer is different: the disulfides of PIP are more stable than those of mini-IGF-1 (*24*). In vivo, the redox state in the secretary pathway exhibits as oxidative with a ratio of GSH/GSSG between 1/1 and 1/3 (*27*). IGF-1 is unable to maintain its native disulfides under such in vivo redox conditions (*16*). Here we compared the disulfide stability of [1−9]PIP native form, PIP, [1−10]mini-IGF-1, and native mini-IGF-1 in vitro under an in vivo-like redox state (where the ratio of GSH/GSSG is 1/1). The results are shown in Figure 7: PIP is very stable (Figure 7B), while native mini-IGF-1 is very unstable (Figure 7D); of the peptide models, [1−9]PIP native form, like native mini-IGF-1, is unstable (Figure 7A) and [1−10]mini-IGF-1, like PIP, is much more stable

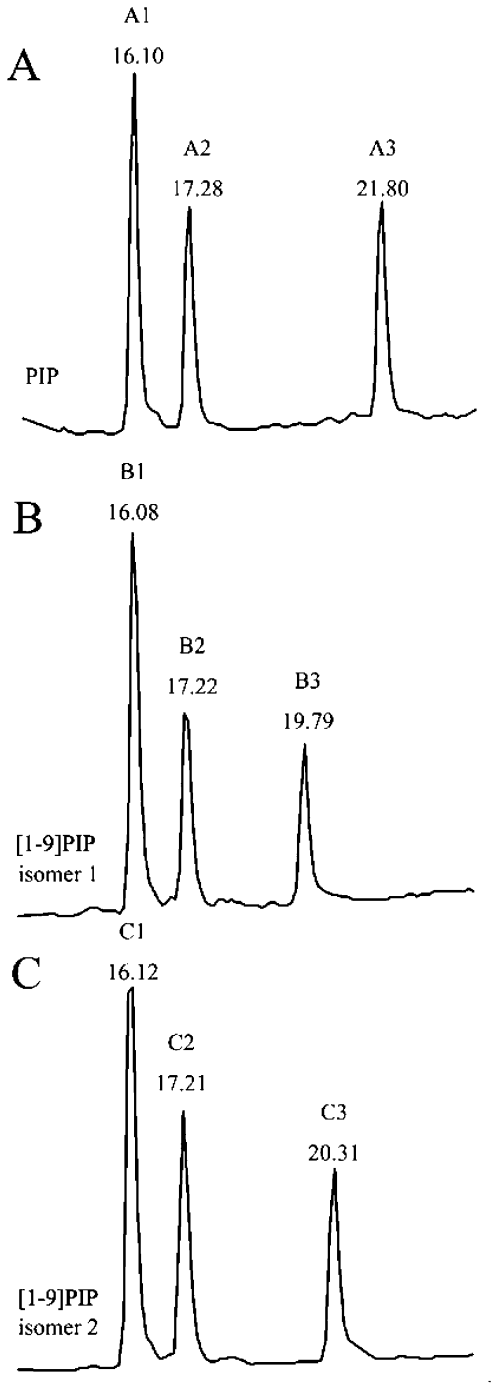

FIGURE 4: HPLC profiles of V8 digests of PIP (A), [1−9]PIP isomer 1 (B), and [1−9]PIP isomer 2 (C). After digestion, samples were adjusted to pH 2.0 with TFA and analyzed by C8 reverse-phase HPLC. The retention time of each peak is indicated. Peak A1 of PIP was previously identified to contain disulfide A20−B19. Because peaks B1 and C1 have the identical retention time as peak A1, we deduced that both isomers of [1−9]PIP contain disulfide A20−B19.

(Figure 7C) under such redox conditions. These results demonstrated that by swapping the N-terminal 10 amino acids between PIP and mini-IGF-1, the disulfide stability was also exchanged accordingly. So we propose that the N-terminal amino acids 1−10 of PIP/insulin are important for its high disulfide stability and correct disulfide pairing, while the N terminal amino acids 1−9 of IGF-1 are responsible for its low disulfide stability and bifurcating folding behavior.

DISCUSSION

What is the molecular basis by which insulin folds into one ground-state structure while IGF-1 folds into two thermodynamically stable structures? Our previous results demonstrated that the different folding behavior of insulin and IGF-1 is mainly controlled by their B-chain/domains (*24*, *26*). Here we have extended this approach to investigate the role of their N-terminal sequences in the different folding behavior of insulin and IGF-1. By exchanging the N-terminal 1−9/1−10 sequences of B domain/chain in mini-IGF-1 and PIP, we have exchanged the different folding behavior of mini-IGF-1 and PIP: [1−10]mini-IGF-1 was secreted from yeast as a single form as PIP was, while [1−9]PIP was secreted from yeast as two isomers as mini-IGF-1 was. The disulfide rearrangement experiment shown that the two isomers of [1−9]PIP can rearrange their disulfides gradually between the forms of native and swap and finally reach equilibrium, while the disulfide rearrangement of [1−10]-mini-IGF-1 was not detectable. In refolding analysis, fully reduced [1−10]mini-IGF-1 refolded into one final product, while fully reduced [1−9]PIP produced two disulfide isomers. Thus we demonstrate that the N-terminal 1−10 sequence of insulin B-chain encodes the information for correct disulfide pairing, while the corresponding N-terminal 1−9 sequence in IGF-1 determines a bifurcating disulfide pairing. Studies on [1−5]mini-IGF-1 demonstrated that the substitution of 1−4 of mini-IGF-1 with 1−5 of PIP makes [1−5]mini-IGF-1 abolish the folding behavior of mini-IGF-1, which suggests that sequence 1−4 of mini-IGF-1 is important for the folding behavior of mini-IGF-1/IGF-1. The importance of sequence 1−4 of mini-IGF-1 is consistent with Milner's finding that N-terminal extension increased the yield of the native isomer, the analogue Long-[Arg^3]IGF-1 with maximal yield of native isomer (*28*). They proposed that the N-terminal extension imparts a steric constraint at a crucial point in folding, thus allowing native disulfide bonds to form efficiently. As in the case of [1−5]mini-IGF-1, it is possible that steric constraint was introduced into mini-IGF-1, because amino acid sequence GPET was replaced by FDNQH. Yet sequence 1−4 of mini-IGF-1 is not sufficient to lead to a bifurcating folding behavior, because when it was introduced into PIP, [1−4]PIP still folds into a unique product as PIP/insulin does. Taking all these together, we propose that all amino acids 1−9 in IGF-1 contributes to its bifurcating folding behavior, while amino acids 1−10 of PIP/insulin play an important role in determining the disulfide pairing of thermodynamically stable forms.

Why does the different folding behavior of insulin and IGF-1 depend on the N-terminal sequences of their B-chain/domains? Insulin and IGF-1 have similar three-dimensional structures and identical disulfide pairing, and their folding pathways are also similar in the early stage. The formation of the A20−B19 (in insulin)/18−61(in IGF-1) disulfide bond is crucial to the initiation of folding (*23*, *29*, *30*). The B-chain of insulin was proposed to provide a template to guide folding of the A-chain (*31*). So different N-terminal sequences the B-chain/domain guide different folding of A-chain/domain in the latter stage of folding in insulin/IGF-1. Both insulin and IGF-1 contain three α-helical segments (*1*, *6*, *7*). The swap IGF-1 and kinetically trapped swap insulin also adopt similar structures. In both swap form

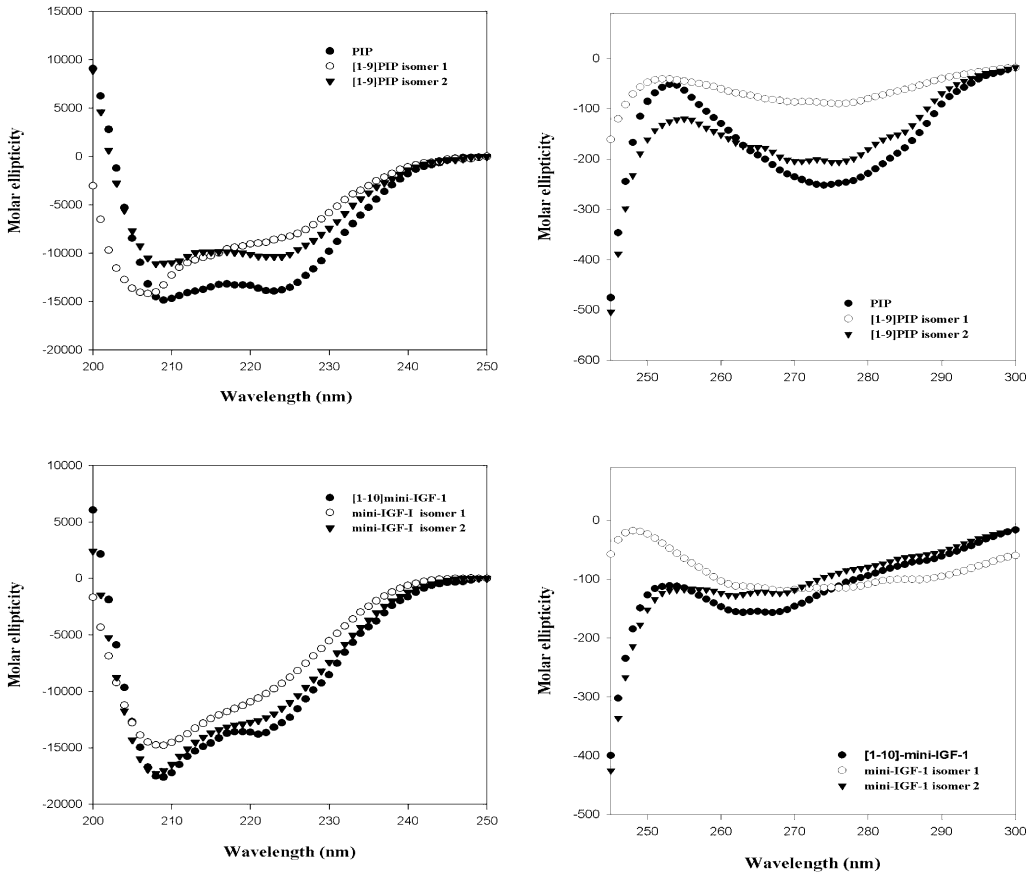

FIGURE 5: Circular dichroic analysis of [1−9]PIP isomers and [1−10]mini-IGF-1. The left column is the far-UV CD spectra; the right column is the near-UV CD spectra. For clarity the spectra of [1−9]PIP (isomer 1 and 2) were compared with that of PIP; the spectra of [1−10]mini-IGF-1 were compared with that of mini-IGF-1 isomer 1 (swap form) and isomer 2 (native form).

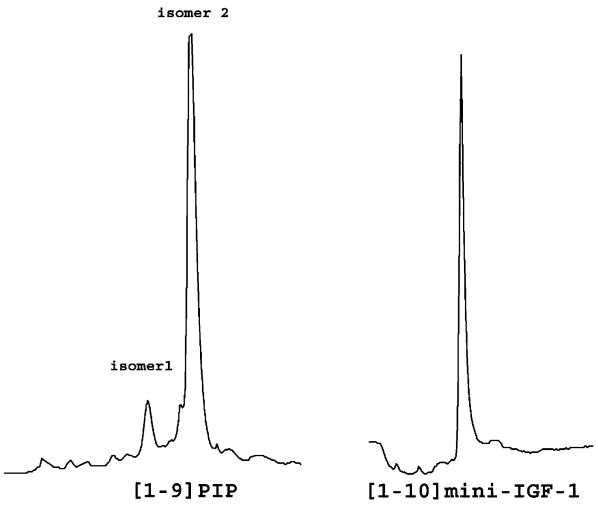

FIGURE 6: In vitro refolding of fully reduced [1−9]PIP and [1−10]-mini-IGF-1. After the fully reduced sample was incubated in the refolding buffer overnight, a 100 μL sample was removed and acidified by TFA and then analyzed by C8 reverse-phase HPLC. The reduced [1−9]PIP refolded into two products that had retention times identical to those of the two isomers of [1−9]PIP. The refolding of fully reduced [1−10]mini-IGF-1 produced only one product, which had a retention time identical to that of the intact [1−10]mini-IGF-1.

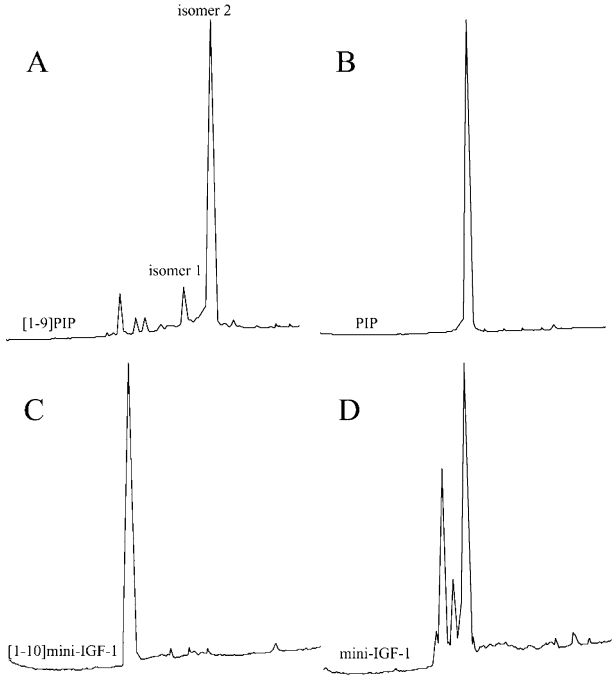

FIGURE 7: Disulfide stability comparison of [1−9]PIP isomer 2 (A), PIP (B), [1−10]mini-IGF-1 (C), and mini-IGF-1 isomer 2 (D) in redox buffer (5 mM GSH/5 mM GSSG). After the sample was incubated in the redox buffer at 4 °C for 16 h, 100 μL of reaction mixture was removed, acidified to pH 2.0 with TFA, and then analyzed by C8 reverse-phase HPLC. PIP and [1−10]mini-IGF-1 were stable in the redox buffer used here, while [1−9]PIP isomer 2 (native form) and mini-IGF-1 isomer 2 (native form) were unstable.

structures, helix 2 no longer exists (*15*, *17*). This is the most striking three-dimensional structure difference between swap form and native form. The stabilization of helix 2 needs the correct disulfide pairing of A7−B7 (6−48 in IGF-1) and A6−A11 (47−52 in IGF-1) disulfide. Breaking either

disulfide results in unfolding of helix 2 in both insulin (*31*−*33*) and IGF-1 (*34*). So, the disulfide stability of A7−B7 and A6−A11 determine the stability of helix 2. IGF-1 was unable to form and maintain the native disulfides under in vivo redox conditions (*16*). In such conditions, IGF-1 is apt to undergo partial unfolding at helix 2 as well as disulfide reshuffling to swap form. This makes swap form an energetic state as favorable as native form. As the disulfide stability experiment shows, the N-terminal 9 amino acids of IGF-1 decrease [1−9]PIP's disulfide stability (when compared with PIP). As a consequence, [1−9]PIP folded into two thermodynamically stable isoforms. On the other hand, the N-terminal 10 amino acids of PIP increase [1−10]mini-IGF-1's disulfide stability (when compared with mini-IGF-1), thus making [1−10]mini-IGF-1 fold into one ground state both in vivo and in vitro.

The native and swap forms of IGF-1 are different structures. The swap form has nonnative disulfides 6−47 and 48−52 and does not have α-helix 2. This indicates that one sequence (IGF-1) folds into two thermodynamically stable structures. Although prion protein (*19*), formin binding protein WW domain (*20*), phosphoglycerate kinase (*21*), and so on can adopt totally different folded conformations, they depend on the environment such as pH, temperature, and solvents. The coexistence of native and swap forms of IGF-1 is under the same environment. So the bifurcating folding ability is an intrinsic property of the IGF-1 molecule. Recently, we revealed that the bifurcating folding behavior of IGF-1 was evolved from aILP (*35*), which has been thought to be the common ancestor of insulin and IGF-1 (*13*, *14*). Our current investigation further suggests that the bifurcating folding behavior of IGF-1 has evolved from mutations in the N-terminal residues 1−9 of B-domain.

REFERENCES

1. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N., and Vijayan, N. M. (1988) The structure of 2Zn pig insulin crystals at 1.5 A resolution. *Philos. Trans. R. Soc. London B Biol. Sci. 319*, 369−456.
2. Weiss, M. A., Hua, Q. X., Lynch, C. S., Frank, B. H., and Shoelson, S. E. (1991) Heteronuclear 2D NMR studies of an engineered insulin monomer: assignment and characterization of the receptor-binding surface by selective ^{2}H and ^{13}C labeling with application to protein design, *Biochemistry 30*, 7373−7389.
3. Roy, M., Lee, R. W., Brange, J., and Dunn, M. F. (1990) ^{1}H NMR spectrum of the native human insulin monomer. Evidence for conformational differences between the monomer and aggregated forms, *J. Biol. Chem. 265*, 5448−5452.
4. Olsen, H. B., Ludvigsen, S., and Kaarsholm, N. C. (1996) Solution structure of an engineered insulin monomer at neutral pH, *Biochemistry 35*, 8836−8845.
5. Humbel, R. E. (1990) Insulin-like growth factors I and II, *Eur. J. Biochem. 190*, 445−462.
6. Vajdos, F. F., Ultsch, M., Schaffer, M. L., Deshayes, K. D., Liu, J., Skelton, N. J., and de Vos, A. M. (2001) Crystal structure of human insulin-like growth factor-1: detergent binding inhibits binding protein interactions, *Biochemistry 40*, 11022−11029.
7. Brzozowski, A. M., Dodson, E. J., Dodson, G. G., Murshudov, G. N., Verma, C., Turkenburg, J. P., de Bree, F. M., and Dauter, Z. (2002) Structural origins of the functional divergence of human insulin-like growth factor-I and insulin, *Biochemistry 41*, 9389−9397.
8. Cooke, R. M., Harvey, T. S., and Campbell, I. D. (1991) Solution structure of human insulin-like growth factor 1: a nuclear magnetic resonance and restrained molecular dynamics study, *Biochemistry 30*, 5484−5491.
9. Sato, A., Nishimura, S., Ohkubo, T., Kyogoku, Y., Koyama, S., Kobayashi, M., Yasuda, T., and Kobayashi, Y. (1993) Three-dimensional structure of human insulin-like growth factor-I (IGF-I) determined by ^{1}H NMR and distance geometry, *Int. J. Pept. Protein Res. 41*, 433−440.
10. Murray-Rust, J., Mcleod, A. N., Blundell, T. L., and Wood, S. P. (1992) Structure and evolution of insulins: implications for receptor binding, *BioEssays 14*, 325−331.
11. Chen, H., and Feng, Y. M. (2001) Contribution of the residue Glu9, Glu46, and Phe49 to the biological activity of insulin-like growth factor-1, *IUBMB Life 51*, 33−37.
12. Chan, S. J., Cao, Q. P., and Steiner, F. D. (1990) Evolution of the insulin superfamily: cloning of a hybrid insulin/insulin-like growth factor cDNA from amphioxus, *Proc. Natl. Acad. Sci. U.S.A. 87*, 9319−9323.
13. Guo, Z. Y., Shen, L., Gu, W., Wu, A. Z., Ma, J. G., and Feng, Y. M. (2002) In vitro evolution of amphioxus insulin-like peptide to mammalian insulin, *Biochemistry 41*, 10603−10607.
14. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. (1992) Disulfide exchange folding of insulin-like growth factor I, *Biochemistry 31*, 1749−1756.
15. Sato, A., Koyama, S., Yamada, H., Suzuki, S., Tamura, K., Kobayashi, M., Niwa, M., Yasuda, T., Kyogoku, Y., and Kobayashi, Y. (2000) Three-dimensional solution structure of a disulfide bond isomer of the human insulin-like growth factor-I, *J. Pept. Res. 56*, 218−230.
16. Hober, S., Lundstrom Ljung, J., Uhlen, M., and Nilsson, B. (1999) Insulin-like growth factors I and II are unable to form and maintain their native disulfides under in vivo redox conditions, *FEBS Lett. 443*, 271−276.
17. Hua, Q. X., Gozani, S. N., Chance, R. E., Hoffmann, J. A., Frank, B. H., and Weiss, M. A. (1995) Structure of a protein in a kinetic trap, *Nat. Struct. Biol. 2*, 129−138.
18. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science 181*, 223−230.
19. Hornemann, S., and Glockshuber, R. A (1998) A scrapie-like unfolding intermediate of the prion protein domain PrP(121−231) induced by acidic pH, *Proc. Natl. Acad. Sci. U.S.A. 95*, 6010−6014.
20. Nguyen, H., Jager, M., Moretto, A., Gruebele, M., and Kelly, J. W. (2003) Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation, *Proc. Natl. Acad. Sci. U.S.A. 100*, 3948−3953.
21. Damaschun, G., Damaschun, H., Gast, K., and Zirwer, D. (1999) Proteins can adopt totally different folded conformations, *J. Mol. Biol. 291*, 715−725.
22. 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., Zhang, X. T., Liu, B., and Liang, Z. H. (1996) Secretory expression of a single-chain insulin precursor in yeast and its conversion into human insulin, *Sci. China* (*Ser. C*). *39*, 225−233.
23. Qiao, Z. S., Guo, Z. Y., and Feng, Y. M. (2001) Putative disulfide-forming pathway of porcine insulin precursor during its refolding in vitro, *Biochemistry 40*, 2662−2668.
24. Guo, Z. Y., Shen, L., and Feng, Y. M. (2002) The different folding behavior of insulin and insulin-like growth factor 1 is mainly controlled by their B-chain/domain, *Biochemistry 41*, 1556−1567.
25. Cecil, R., and Weitzman, P. D. (1964) The electroreduction of the disulphide bonds of insulin and other proteins, *Biochem. J. 93*, 1−11.
26. Guo, Z. Y., Shen, L., and Feng, Y. M. (2002) The different energetic state of the intra A-chain/domain disulfide of insulin and insulin-like growth factor 1 is mainly controlled by their B-chain/domain, *Biochemistry 41*, 10585−10592.
27. Hwang, C., Sinskey, A. J., and Lodish, H. F. (1992) Oxidized redox state of glutathione in the endoplasmic reticulum, *Science 257*, 1496−1502.
28. Milner, S. J., Francis, G. L., Wallace, J. C., Magee, B. A., and Ballard, F. J. (1995) Mutations in the B-domain of insulin-like growth factor-I influence the oxidative folding to yield products with modified biological properties, *Biochem J. 308* (Pt 3), 865−871.
29. Qiao, Z. S., Min, C. Y., Hua, Q. X., Weiss, M. A., and Feng, Y. M. (2003) In vitro refolding of human proinsulin. Kinetic intermediates, putative disulfide-forming pathway folding initiation site, and potential role of C-peptide in folding process, *J. Biol. Chem. 278*, 17800−17809.

30. Yang, Y., Wu, J., and Watson, J. T. (1999) Probing the folding pathways of long R(3) insulin-like growth factor-I (LR(3)IGF−I) and IGF−I via capture and identification of disulfide intermediates by cyanylation methodology and mass spectrometry, *J. Biol. Chem. 274*, 37598−37604.
31. Hua, Q. X., Nakagawa, S. H., Jia, W., Hu, S. Q., Chu, Y. C., Katsoyannis, P. G., and Weiss, M. A. (2001) Hierarchical protein folding: asymmetric unfolding of an insulin analogue lacking the A7-B7 interchain disulfide bridge, *Biochemistry 40*, 12299−12311.
32. Weiss, M. A., Hua, Q. X., Jia, W., Chu, Y. C., Wang, R. Y., and Katsoyannis, P. G. (2000) Hierarchical protein "un-design": insulin's intrachain disulfide bridge tethers a recognition alpha-helix, *Biochemistry 39*, 15429−15440.
33. Guo, Z. Y., and Feng, Y. M. (2001) Effects of cysteine to serine substitutions in the two interchain disulfide bonds of insulin, *Biol. Chem. 382*, 443−448.
34. Hua, Q. X., Narhi, L., Jia, W., Arakawa, T., Rosenfeld, R., Hawkins, N., Miller, J. A., and Weiss, M. A. (1996) Native and non-native structure in a protein-folding intermediate: spectroscopic studies of partially reduced IGF-I and an engineered alanine model, *J. Mol. Biol. 259*, 297−313.
35. Wang, S., Guo, Z. Y., Shen, L., Zhang, Y. J., and Feng, Y. M. (2003) Refolding of amphioxus insulin-like peptide: implications of a bifurcating evolution of the different folding behavior of insulin and insulin-like growth factor 1, *Biochemistry 42*, 9687−9693.

BI049710Y