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†

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

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

Received February 26, 2002; Revised Manuscript Received June 6, 2002

ABSTRACT: Insulin and insulin-like growth factor 1 (IGF-1) share homologous sequence, similar three-dimensional structure, and weakly overlapping biological activity, but different folding information is stored in their homologous sequences: the sequence of insulin encodes one unique thermodynamically stable three-dimensional structure while that of IGF-1 encodes two disulfide isomers with different three-dimensional structure but similar thermodynamic stability. Their different folding behavior probably resulted from the different energetic state of the intra A-chain/domain disulfide: the intra A-chain disulfide of insulin is a stable bond while that of IGF-1 is a strained bond with high energy. To find out the sequence determinant of the different energetic state of their intra A-chain/domain disulfide, the following experiments were carried out. First, a local chimeric single-chain insulin (PIP) with the A8−A10 residues replaced by the corresponding residues of IGF-1 was prepared. Second, the disulfide stability of two global hybrids of insulin and IGF-1, Ins(A)/IGF-1(B) and Ins(B)/IGF-1(A), was investigated. The local segment swap had no effect on the fidelity of disulfide pairing and the disulfide stability of PIP molecule although the swapped segment is close to the intra A-chain/domain disulfide. In redox buffer which favors the disulfide formation for most proteins, Ins(A)/IGF-1(B) cannot form and maintain its native disulfides just like that of IGF-1, while the disulfides of Ins(B)/IGF-1(A) are stable in the same condition. One major equilibrium intermediate with two disulfides of Ins(A)/IGF-1(B) was purified and characterized. V8 endoproteinase cleavage and circular dichroism analysis suggested that the intra A-chain/domain disulfide was reduced in the intermediate. Our present results suggested that the energetic state of the intra A-chain/domain disulfide of insulin and IGF-1 was not controlled by the A-chain/domain sequence close to this disulfide but was mainly controlled by the sequence of the B-chain/domain.

Insulin is an extensively studied small globular protein with 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-defined by X-ray crystallography (*1*, *2*) and NMR (*3*−*5*) since the 1970s. IGF-1[1] is an insulin-like 70-residue single-chain protein composed of B-, C-, A-, and D-domains (*6*). The B- and A-domains of IGF-1 are homologous to the B- and A-chains of insulin, respectively; the C-domain is analogous to the C-peptide of proinsulin, but they share no sequence homology; the D-domain has no counterparts in insulins. The three-dimensional structure of IGF-1 is very similar with that of insulin (*7*). The architecture of insulin and IGF-1 mainly consists 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-chain/domains (Figure 1A,B). The three α-helical segments are stabilized by the three disulfides (A6−A11, A7−B7, A20−B19 in insulin; 47−52, 6−48, 18−61 in IGF-1) (*8*−*10*). The conformation of the C- and D-domains of IGF-1 is highly flexible. When the B-chain/domain (1−29 residue for insulin and 1−28 residue for IGF-1) and the A-chain/domain were connected by a peptide bond directly, the mini-proinsulin still retained the insulin-like three-dimensional structure (*11*, *12*) while the conformation of the mini-IGF-1 changed slightly: although the three α-helical segments were still retained, their relative orientation had changed slightly (*13*).

Protein folding is still challenging in biological sciences. Since Anfinsen and co-workers first demonstrated that the three-dimensional structure of a globular protein is uniquely determined by its amino acid sequence (*14*), 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), the National High Technology Program of China (863-103-13-01-01), and the National Foundation of Nature Science (No. 30170209).

* 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 porcine insulin B-chain and the N-terminus of 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; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid; PAGE, polyacrylamide gel electrophoresis; CD, circular dichroism; NMR, nuclear magnetic resonance.

10.1021/bi020165f CCC: $22.00
Published on Web 07/31/2002

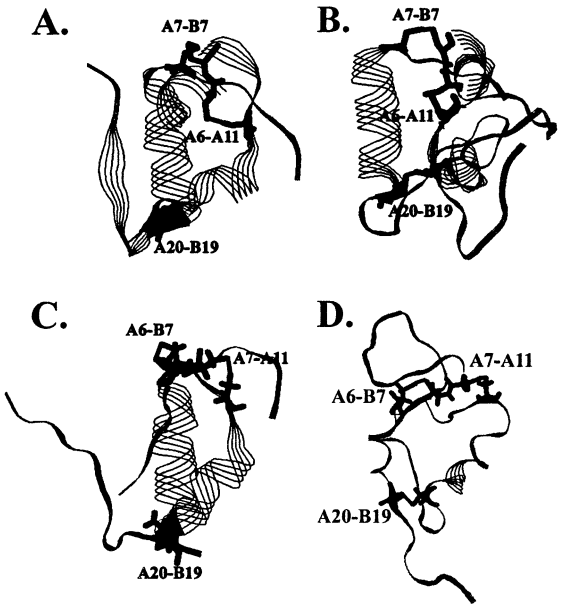

FIGURE 1: Strand models of (A) insulin [R state of crystal structure (*2*)], (B) IGF-1 [solution structure (*7*)], (C) swap insulin [solution structure (*31*)], and (D) swap IGF-1 [solution structure (*41*)]. For clarity the disulfides were all numbered as those of insulin.

structure of the transition-state ensemble (*15*−*17*). 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 (*18*−*25*). In vivo the protein folding is assisted by the molecular chaperons, especially for large proteins (*26*−*28*), some chaperons even can provide the missing steric information for protein folding (*29*).

Although insulin and IGF-1 share homologous sequence, similar three-dimensional structure, weakly overlapping biological activity (*6*), and a common ancestor (*30*), the folding information stored in their sequences is different: insulin and its recombinant single-chain precursor (PIP) fold into one unique thermodynamically stable tertiary structure (*31*, *32*), while IGF-1 folds into two disulfide isomers (native and swap) with different disulfide linkages and different three-dimensional structure but similar thermodynamic stability. The native IGF-1 with three disulfides identical to those of insulin adopts an insulin-like structure; the swap IGF-1 and swap insulin with identical disulfides (18−61, 48−52, 6−47 for swap IGF-1; A20−B19, A7−A11, A6−B7 for swap insulin) also adopt similar three-dimensional structure (*31*, *41*). The α-helix in the N-terminus of the A-chain/domain presented in the native form is unfolded, and the other two α-helical segments still exist in swap insulin and swap IGF-1 (Figure 1C,D). Although swap IGF-1 and swap insulin adopt similar three-dimensional structure with identical disulfides, they have a different energetic state: the swap IGF-1 is a thermodynamically controlled folding product with an energetic state similar to that of native IGF-1, while the swap insulin is a kinetic trap and is thermodynamically unstable (*31*). For most small globular proteins the final folding product is a unique thermodynamically stable three-dimensional structure (*42*), just like that of BPTI, RNase A, and insulin/PIP. Our previous work suggested that the different folding behavior of insulin/PIP and IGF-1 was mainly controlled by their B-chain/domain (*43*), but the detailed mechanism of B-chain/domain exerting its role is still unknown. In IGF-1 the intra A-domain disulfide, 47−52 (corresponding to the disulfide A6−A11 in insulin), is a "strained" bond with high energetic state (*33*). In the redox buffer that favors the disulfide formation for most proteins this disulfide can be reduced. However, the intra A-chain disulfide, A6−A11, of insulin was a stable bond (*44*). The different energetic state of the intra A-chain/domain disulfide probably led to the different folding behavior of insulin and IGF-1. However, the sequence determinant of the different energetic state of the intra A-chain/domain disulfide is still unknown. To answer this question, the following experiments were carried out. First, we constructed a local hybrid PIP molecule with A8Thr-A9Ser-A10Ile of insulin replaced by the corresponding residues, 49Phe-50Arg-51Ser, of IGF-1 (the local hybrid was designated as [FRS]PIP), since the three residues are close to the intra A-chain/domain disulfide. The three residues, A8Thr-A9Ser-A10Ile, were not conservative in insulins coming from different species, but in natural insulins they had never been replaced by the corresponding residues, 49Phe-50Arg-51Ser, of IGF-1. Second, we investigated the intra A-chain/domain disulfide stability of two single-chain global hybrids of insulin and IGF-1, Ins(A)/IGF-1(B) and Ins(B)/IGF-1(A). Here we report the effect of the local and global sequence swap on the stability of the intra A-chain/domain disulfide.

MATERIALS AND METHODS

Materials. The *Escherichia coli* strains used were DH12S and RZ1032 (dut^-, ung^-). *Saccharomyces cerevisiae* XV700-6B (Leu2, ura3, pep4) and helper phage R408 were kindly provided by Michael Smith (University of British Columbia, Vancouver, Canada). Plasmid pVT102-U/αMFL-PIP was constructed in our laboratory for secretory expression of PIP in yeast (*45*). The mutagenesis oligonucleotide primer for constructing the local hybrid was chemically synthesized. The two single-chain global hybrids, Ins(A)/IGF-1(B) (the C-terminus of human IGF-1 B-domain and the N-terminus of human insulin A-chain were linked together by a dipeptide, Ala-Lys, and its B29Thr was replaced by Lys) and Ins(B)/IGF-1(A) (the C-terminus of porcine insulin B-chain and N-terminus of human IGF-1 A-domain were linked together by a dipeptide, Ala-Lys), were prepared in our laboratory (*43*). The sequence of human insulin and porcine insulin was identical except for the B30 residue; in porcine insulin it is Ala and in human insulin it is Thr. The chemical reagents used in the 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 HPLC analysis a gradient elution was used and the detection was at 230 nm. Solvent A was 0.15% aqueous TFA; 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, 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.

Construction of the Local Hybrid. The local hybrid, [FRS]-PIP, was constructed by site-directed mutagenesis on the basis of pVT102-U/αMFL-PIP using a gapped duplex

approach (*46*). The presence of the expected mutation was confirmed by DNA sequencing. The expression plasmid of the local hybrid was designated as pVT102-U/αMFL-[RFS]-PIP.

Expression and Purification of the Local Hybrid. The plasmid pVT102-U/αMFL-[FRS]PIP was transformed into *S. cerevisiae* XV700-6B (Leu2, ura3, pep4). The transformed yeast cells were cultured in a 16 L fermenter, and the [FRS]-PIP was purified from the media supernatant in four steps (*45*). First, the [FRS]PIP was precipitated from the media supernatant by trichloroacetic acid. Second, the precipitate was dissolved with 1 M acetic acid and applied to a Sephadex G-50 column. Third, the [FRS]PIP was purified by chromatography on a DEAE-Sepharose CL-6B column. Fourth, the [FRS]PIP was further purified by C8 reverse-phase HPLC. The purity of the [FRS]PIP was analyzed by native pH 8.3 polyacrylamide gel electrophoresis (pH 8.3 PAGE) and analytical C8 reverse-phase HPLC. Its molecular mass was determined by electrospray mass spectrometry.

In Vitro Refolding of the Local Hybrid. The purified [FRS]-PIP was dissolved in 0.1 M Tris-HCl and 1 mM EDTA, pH 9.7, buffer containing 4 M urea (the final concentration of [FRS]PIP is about 0.5 mg/mL); then DTT was added to the final concentration of 0.1 M. The reduction was carried out at room temperature for 1 h. After incubation an aliquot was removed, carboxymethylated, and then analyzed by native pH 8.3 PAGE to detect if the disulfides were fully reduced or not. The fully reduced [FRS]PIP was exchanged into the refolding buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 9.7) by gel filtration using a Sephadex G-25 column. The refolding reaction was carried out in redox buffer containing 5 mM GSH and 5 mM GSSG or by air oxidation. The refolding solution was incubated at 4 °C overnight. After incubation 100 μL of refolding product was removed, acidified to pH 2.0 with TFA, and then analyzed by C4 reverse-phase HPLC eluted with the gradient described in Materials and Methods with the flow rate of 0.5 mL/min.

Disulfide Stability of the Local and the Global Hybrids in Redox Buffer. The single-chain local hybrid, [FRS]PIP, and the single-chain global hybrids, Ins(A)/IGF-1(B) (native form) and Ins(B)/IGF-1(A), were respectively dissolved in the redox buffer [0.1 M Tris-HCl and 1 mM EDTA, pH 8.7, buffer for Ins(A)/IGF-1(B); 0.1 M Tris-HCl and 1 mM EDTA, pH 9.7, buffer for [FRS]PIP and Ins(B)/IGF-1(A) because of their low solubility in pH 8.7 buffer]. The final protein concentration was adjusted to 0.1 mg/mL, respectively. In the redox buffer the ratio (mM/mM) of GSH to GSSG was 1:10 and 5:5, respectively. At the same time a negative control in which the samples were incubated in the buffer not containing redox potential was carried out. The reaction was carried out at 4 °C overnight. After incubation, a 100 μL sample was removed, acidified to pH 2.0 with TFA, and then analyzed by C8 or C4 reverse-phase HPLC eluted with the gradient listed in Materials and Methods.

Separation of the Major Equilibrium Intermediate of Ins(A)/IGF-1(B). For separation of the major equilibrium intermediate, native Ins(A)/IGF-1(B) was dissolved in the redox buffer (0.1 M Tris-HCl and 1 mM EDTA, pH 8.7, containing 1 mM GSH and 10 mM GSSG) at the final concentration of 0.3 mg/mL. The reaction was carried out at 4 °C overnight. After incubation the reaction mixture was adjusted to pH 2.0 with TFA and then separated by C8 reverse-phase HPLC eluted with the gradient described in Materials and Methods with the flow rate of 0.8 mL/min. The elution fractions were collected manually and lyophilized. The molecular mass was measured by electrospray mass spectrometry.

Circular Dichroism Analysis. The samples, PIP, [FRS]-PIP, native Ins(A)/IGF-1(B), swap Ins(A)/IGF-1(B), and the major equilibrium intermediate of Ins(A)/IGF-1(B), were dissolved in 5 mM HCl, respectively. The protein concentration was determined by UV absorbance at 276 nm. CD measurements were performed on a Jasco-715 circular dichroism spectropolarimeter at room temperature. The spectra were scanned from 250 to 190 nm in the far-UV region and from 320 to 245 nm in the near-UV region. Cell path length was 0.1 and 1.0 cm for far-UV and near-UV spectra, respectively. The protein concentration for the local and global hybrids was 0.2 and 0.12 mg/mL, 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 CD spectra.

Peptide Mapping of the Major Equilibrium Intermediate of Ins(A)/IGF-1(B). The major equilibrium intermediate was dissolved in 0.1 M phosphate buffer (pH 7.8), and V8 endoproteinase was added to the solution at about a mass ratio of 1:20 to the sample. The reaction was carried out at 25 °C overnight. After digestion, the solution was adjusted with TFA to pH 2.0 and separated by C4 reverse-phase HPLC eluted with the gradient listed in Materials and Methods with the flow rate of 0.5 mL/min. The fractions were collected manually and lyophilized. Their molecular mass was measured by electrospray mass spectrometry.

Refolding of the Major Equilibrium Intermediate of Ins(A)/IGF-1(B). The major equilibrium intermediate of Ins(A)/IGF-1(B) was dissolved in the refolding buffer (0.1 M Tris-HCl and 1 mM EDTA, pH 8.7) at the final concentration of 0.1 mg/mL. 2-Mercaptoethenol was added to the final concentration of 0.2 mM to initiate the refolding. The refolding reaction was carried out at 10 °C. At different time points 100 μL of the reaction mixture was removed, acidified to pH 2.0 with TFA, and then immediately analyzed by C8 reverse-phase HPLC eluted with the gradient listed in Materials and Methods with the flow rate of 0.8 mL/min.

RESULTS

Construction, Expression, and Purification of the Local Hybrid. The presence of the expected mutation on the PIP gene was confirmed by DNA sequencing (data not shown). The [FRS]PIP was purified from the fermentation media in four steps (*45*). Its molecular mass determined by electrospray mass spectrometry is 6047.0, which is consistent with the theoretical value of 6048.0. The purified [FRS]PIP is homogeneous as judged by the native pH 8.3 PAGE and C8 reverse-phase HPLC (data not shown). During purification we found that [FRS]PIP had only one component, which suggested that the local hybrid probably folds into a unique tertiary structure.

Circular Dichroism Analysis of the Local Hybrid. Did the local swap affect the structure of the local hybrid compared with that of its parent molecule? To answer this question, the far-UV and near-UV CD spectra of [FRS]PIP were

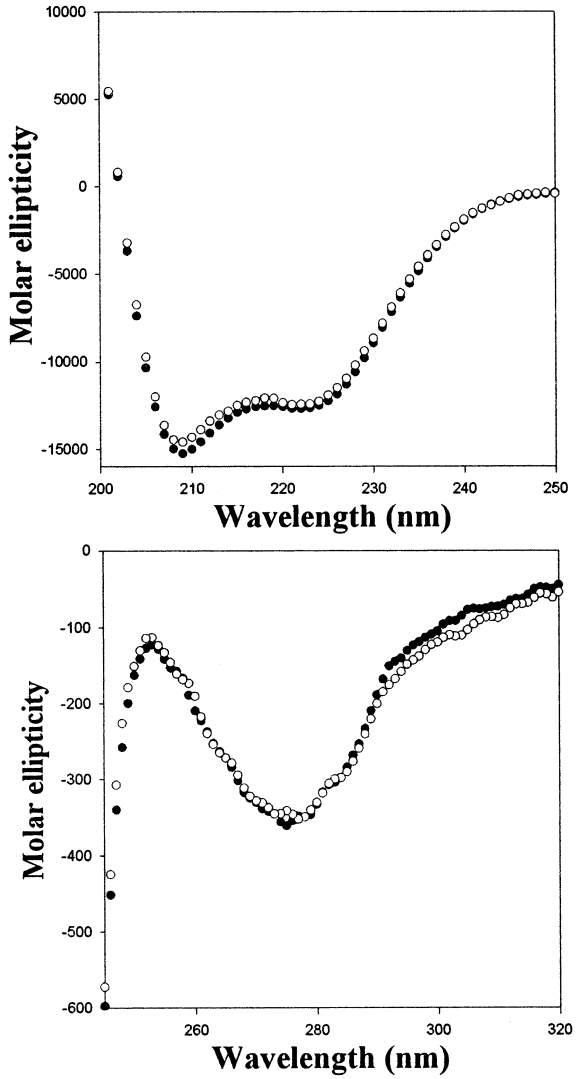

FIGURE 2: Circular dichroism analysis of the local hybrid. The spectrum of [FRS]PIP was compared with that of the wild-type PIP. The above is the far-UV spectra while the below is the near-UV spectra. The open circle represents the spectra of the local hybrid, [FRS]PIP; the filled circle represents the spectra of the wild-type PIP.

analyzed as shown in Figure 2. The far-UV spectra of [FRS]-PIP and the wild-type PIP were very similar, which suggested that they share similar or identical secondary structure. The α-helix content estimated from CD spectra of [FRS]PIP and the wild-type PIP both was 49%, which was consistent with the value calculated from the insulin crystal structure according to the three α-helical segments. The local hybrid and the wild-type PIP also share similar near-UV spectra, which indicated that the microenvironment of the aromatic residues as well as the disulfide bonds was similar in the local hybrid and its parent molecule. The similarity of the CD spectra between the local hybrid and its parent molecule suggested that the local hybrid adopts a fold similar/identical to that of the wild-type PIP; that is, the local segment swap had little effect on the conformation of the PIP molecule.

In Vitro Refolding of the Local Hybrid. To further determine if the local hybrid refolded into a unique tertiary structure or not, its in vitro refolding was carried out. As shown in Figure 3, the final folding product had only one peak with the retention time identical to that of the standard

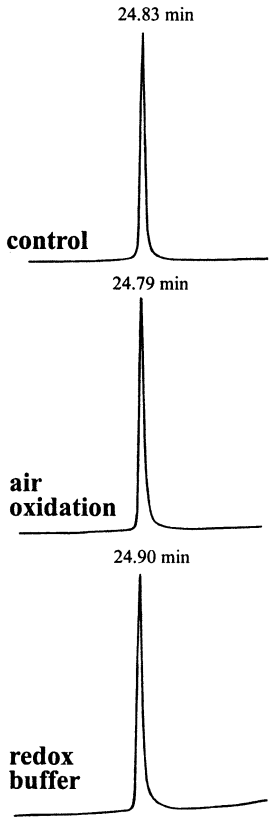

FIGURE 3: In vitro refolding of the local hybrid by air oxidation or in redox buffer. The redox potential of the redox buffer was 5 mM GSH and 5 mM GSSG. After incubation 100 μL of refolding solution was removed, acidified to pH 2.0 with TFA, and then analyzed by C4 reverse-phase HPLC eluted with a gradient described in Materials and Methods with the flow rate of 0.5 mL/min. The control is the purified [FRS]PIP.

[FRS]PIP when analyzed by C4 reverse-phase HPLC no matter whether the refolding was carried out in redox buffer or by air oxidation. This result suggested that the local hybrid refolded into a unique tertiary structure in vitro. Together with the fact that the local hybrid secreted from the yeast cells as a single form with the CD spectra very similar to that of the wild-type PIP, we deduced that the local swap had no effect on the fidelity of disulfide pairing of PIP; that is, the local hybrid folded into one unique thermodynamically stable three-dimensional structure with the disulfide linkages identical to those of insulin.

Disulfide Stability of the Local Hybrid in Redox Buffer. The local swap had no effect on the fidelity of disulfide pairing; did it have an effect on the disulfide stability of PIP? Here we analyzed the disulfide stability of the local hybrid in redox buffer compared with that of its parent molecule (data not shown). As analyzed by C4 reverse-phase HPLC, the local hybrid incubated in the redox buffer (the ratio of GSH to GSSG is 1:10 and 5:5, respectively) appeared as a single peak on the HPLC profile with the retention time identical to that of the control. This result is identical to that of the wild-type PIP but different from that of IGF-1 significantly since IGF-1 cannot maintain its intact disulfides in the above redox buffer (*34*). The present result suggested that the local swap had little effect on the disulfide stability of PIP.

Disulfide Stability of the Global Hybrids, Ins(A)/IGF-1(B) and Ins(B)/IGF-1(A), in Redox Buffer. The local swap had

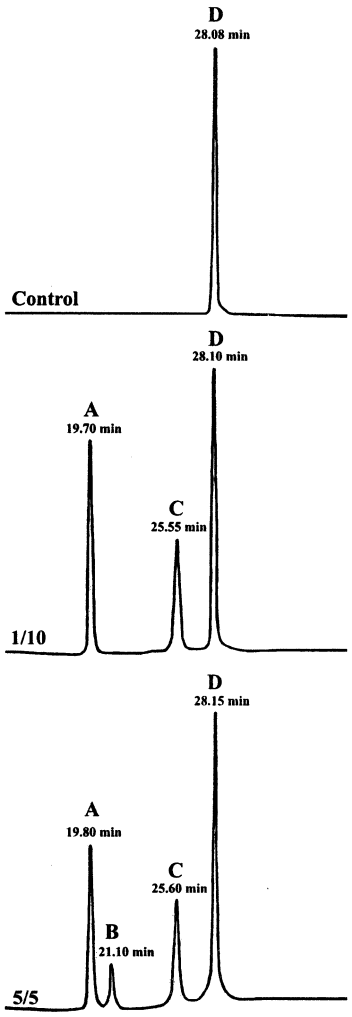

FIGURE 4: Disulfide stability of Ins(A)/IGF-1(B) in redox buffer. After the sample was incubated in the redox buffer overnight, 100 μL of reaction mixture was removed, acidified to pH 2.0 with TFA, and then analyzed by C8 reverse-phase HPLC eluted with the gradient described in Materials and Methods using the flow rate of 0.8 mL/min. 1/10 and 5/5 represented the ratio of GSH to GSSG (mM/mM) in the redox buffer was 1/10 and 5/5, respectively. In the control the sample was incubated in the buffer not containing redox potential.

no effect on the fidelity of disulfide pairing as well as disulfide stability of the PIP molecule. Which is the determinant of the different energetic state of the intra A-chain/domain disulfide of insulin and IGF-1? Here the disulfide stability of two global hybrids, Ins(A)/IGF-1(B) and Ins(B)/IGF-1(A), was analyzed. As shown in Figure 4, native Ins(A)/IGF-1(B) cannot form and maintain its native disulfides in redox buffer which favors the disulfide formation for most proteins, just like that of IGF-1 [swap Ins(A)/IGF-1(B) gave an identical result after incubation in redox buffer]. In alkaline buffer not containing thiol reagent the native Ins(A)/IGF-1(B) was stable as shown by the control. In alkaline buffer containing 1 mM GSH and 10 mM GSSG, three peaks appeared on the chromatograph. Peak D is the native form whose disulfides are identical to those of insulin/IGF-1. The retention time of peak C is identical to that of the swap Ins(A)IGF-1(B) (the control of the swap form was not shown here); together with other analysis (such as the native PAGE) and previous study (*43*), we deduce that peak C is the swap form of Ins(A)/IGF-1(B). The swap and the

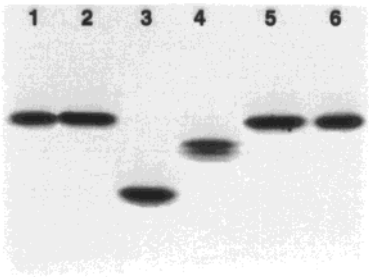

FIGURE 5: Native pH 8.3 PAGE analysis of the four peaks of Figure 4. In each lane about 2 μg of sample was loaded onto the native gel where the proteins were separated according to their net charges, molecular mass, and the conformation. A Bio-Rad mini-protean II electrophoresis apparatus was used in the experiments. The gel was stained with Comassie brilliant blue R250. Lanes 1 and 2 were native and swap Ins(A)/IGF-1(B), respectively; lanes 3−6 represented peaks A, B, C, and D in Figure 4, respectively.

native Ins(A)/IGF-1(B) have identical sequence but different disulfide linkages and different three-dimensional structure (*43*). The disulfides of the swap Ins(A)/IGF-1(B) were identical to those of swap insulin/swap IGF-1. Peak A is a new peak with the molecular mass of 6346.0, and we named it the major equilibrium intermediate. In contrast to the 5734.6 mass of the intact Ins(A)/IGF-1(B), peak A has a mass increase of 611.4, which is equal to the mass of two GSH groups, so we deduced that peak A was an equilibrium intermediate with one disulfide reduced and the two free thiol groups respectively modified with a GSH group; that is, the thiol group of GSH forms a disulfide bond with the free thiol group of the protein. In buffer containing 5 mM GSH and 5 mM GSSG, a small peak B appeared, and we named it the minor equilibrium intermediate, but we did not characterized this peak because of its low content. In the buffer with the same redox potential as above, the chromatography of Ins(B)/IGF-1(A) appeared as only one peak with identical retention time and peak area as the control (data not shown), so the disulfides of Ins(B)/IGF-1(A) were stable in such condition.

The four peaks in Figure 4 were analyzed by native pH 8.3 PAGE as shown in Figure 5. Peaks C and D had a mobility rate identical to those of the swap and native Ins(A)/IGF-1(B), respectively. Together with previous study (*43*) we deduce that peaks C and D were swap and native Ins(A)-IGF-1(B), respectively. The high mobility rate of peak A indicated that it carried more negative charges than the intact molecule, which was consistent with the modification by GSH groups. The mobility rate of peak B was between that of the intact molecule and peak A. Peak B was not homogeneous as judged by native PAGE.

Circular Dichroism Analysis of the Major Equilibrium Intermediate of Ins(A)/IGF-1(B). The secondary structure of the major equilibrium intermediate of Ins(A)/IGF-1(B) was analyzed by far-UV circular dichroism compared with that of the native and swap Ins(A)/IGF-1(B) as shown in Figure 6. The relative α-helix content estimated from CD spectra of the native Ins(A)/IGF-1(B), swap Ins(A)/IGF-1(B), and the major equilibrium intermediate was 30%, 21%, and 21%, respectively. The α-helix content decrease of the swap form and the major equilibrium intermediate was both about 30% compared with that of the native form. In the solution structure of swap insulin (*31*) and swap IGF-1 (*41*) the α-helix in the N-terminus of the A-chain/domain was

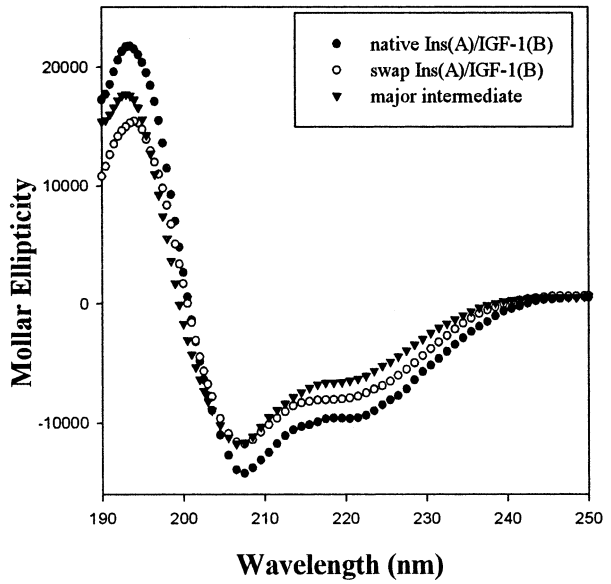

FIGURE 6: The far-UV circular dichroism spectra of the native Ins(A)/IGF-1(B), swap Ins(A)/IGF-1(B), and the major equilibrium intermediate of Ins(A)/IGF-1(B).

unfolded while the other two α-helical segments were still retained. In insulin and IGF-1 with the intra A-chain/domain disulfide deleted/reduced (*47*−*49*) the α-helix in the N-terminus of the A-chain/domain was also unfolded while the other two α-helixes were retained. The α-helix in the N-terminus of the A-chain/domain takes up about 30% of the total α-helix content of insulin and IGF-1. So we deduced that the α-helix in the N-terminus of the A-chain/domain of the swap Ins(A)/IGF-1(B) and the major equilibrium intermediate was unfolded; the reduced disulfide of the major equilibrium intermediate was most likely the intra A-chain one.

Cleavage of the Major Equilibrium Intermediate of Ins(A)/IGF-1(B) by V8 Endoproteinase. We analyzed the disulfide linkages of the major equilibrium intermediate of Ins(A)/IGF-1(B) by V8 endoproteinase digestion as shown in Figure 7. After cleavage three major peaks appeared on C4 reverse-phase HPLC chromatography, and their molecular masses as well as their corresponding sequences were listed in Table 1. Peak C was the fragment between the A- and B-chain/domain. Peak A was the fragment containing the disulfide A20−B19, and this disulfide was intact in the intermediate. Peak B was the fragment containing B7Cys, A6Cys, A7Cys, and A11Cys. Among the four Cys residues, only one disulfide bond formed and it must be between A- and B-chain/domains, and the other two Cys residues were modified by the GSH group, respectively. Together with circular dichroism measurement we deduced that the reduced disulfide was A6−A11 in the intermediate, so the intra A-chain/domain disulfide of Ins(A)/IGF-1(B) was the most unstable disulfide bond just like that of IGF-1.

Refolding of the Major Equilibrium Intermediate of Ins(A)/IGF-1(B). If the deduced structure was correct, the major equilibrium intermediate of Ins(A)/IGF-1(B) can refold into the intact forms, the native and the swap Ins(A)/IGF-1(B). When the major equilibrium intermediate was dissolved in the alkaline buffer containing 0.2 mM 2-mercaptoethanol, it indeed refolded into two forms with the retention time

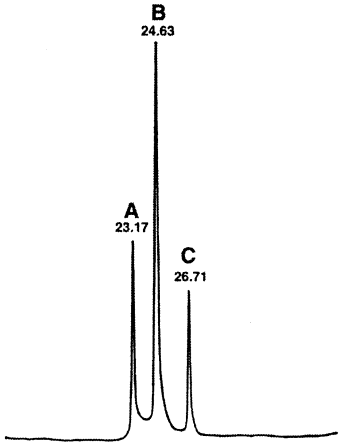

FIGURE 7: C4 reverse-phase HPLC profile of the V8 endoproteinase digestion mixture of the major equilibrium intermediate of Ins(A)/IGF-1(B). 100 μL of the digested mixture was removed, acidified to pH 2.0 with TFA, and then analyzed by C4 reverse-phase HPLC eluted with the gradient described in Materials and Methods using the flow rate of 0.5 mL/min.

Table 1: Molecular Mass of V8 Digested Fragments of the Major Equilibrium Intermediate of Ins(A)/IGF-1(B)

peak no. in Figure 4	retention time (min)	measured value	expected value	corresponding sequence
A	23.17	1362.0	1362.5	B13Ala-B20Asp + A18Asn-A21Asn
B	24.63	2691.0	2692.0	B4Thr-B9Glu + A5Gln-A17Glu
C	26.71	1754.0	1754.1	B21Arg-A4Glu

identical to that of the native and swap Ins(A)/IGF-1(B), respectively (Figure 8). This result further demonstrated that the sequence of Ins(A)/IGF-1(B) encodes two three-dimensional structures with similar thermodynamic stability.

DISCUSSION

Considering the high sequence homology of insulin and IGF-1, their different folding behavior is puzzling. Our previous work suggested that the B-chain/domain mainly manipulated the different folding behavior of insulin and IGF-1 (*43*). But the detailed mechanism of the B-chain/domain exerting its effect is still unknown. The intra A-domain disulfide of IGF-1 is a strained bond with high energetic state (*33*) while the intra A-chain disulfide of insulin is a stable bond (*44*). The different energetic state of the intra A-chain/domain disulfide is probably relevant to their different folding property: insulin/PIP folds into one unique structure while IGF-1 folds into two disulfide isomers. In swap IGF-1 and swap insulin the α-helix in the N-terminus of the A-chain/domain presented the native form as unfolded. The unfolding of the α-helix must elevate the energetic state of the molecule, which led to one unstable structure. In swap IGF-1 the free energy increase is somewhat counteracted by the swap disulfide linkages because the 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 the disulfide A6−A11 is still a stable bond. But which part of the sequence is the determinant of the different energetic state of the intra A-chain/domain disulfide of insulin and IGF-1? Our present

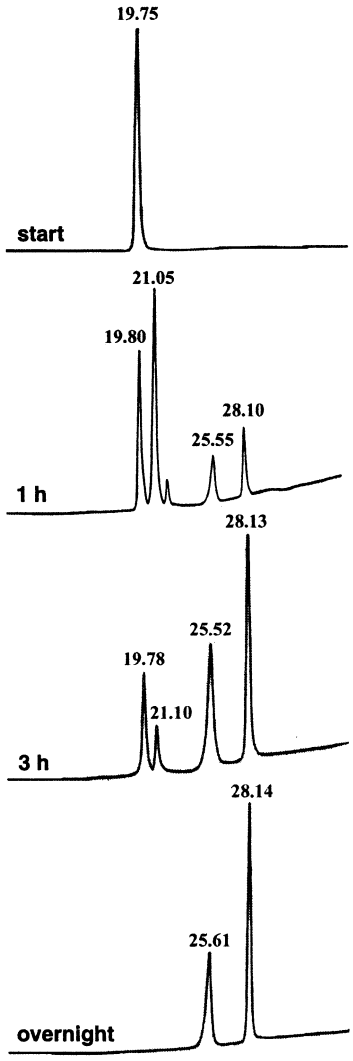

FIGURE 8: In vitro refolding of the major equilibrium intermediate of Ins(A)/IGF-1(B). At the appointed reaction time 100 μL of reaction mixture was removed, acidified to pH 2.0 with TFA, and immediately analyzed by C8 reverse-phase HPLC eluted with the gradient described in Materials and Methods using the flow rate of 0.8 mL/min. The retention time (min) of each peak was indicated.

results suggested that it is not the A-chain/domain sequence close to the intra A-chain/domain disulfide but the sequence of the B-chain/domain that mainly controlled the different energetic state of the intra A-chain/domain disulfide.

Why did insulin and IGF-1 evolve the different folding property from a common ancestor, amphioxus insulin-like peptide (ILP)? Our previous results showed that a recombinant single-chain ILP (*50*) containing the deduced B- and A-domains folds into a unique three-dimensional structure in vivo and in vitro (our unpublished data). So the appearance of the unusual folding property of IGF-1 probably coupled with some other properties. In vivo IGF-1 was bound with IGF binding proteins which did not bind with insulin. The existence of IGF binding protein just overcame the folding problem of IGF-1 (*40*), so we deduced that the unusual folding property of IGF-1 was coevolved with the IGF binding proteins. Our deduction was consistent with the fact that the B-domain of IGF-1 was critical for interaction with IGF binding proteins (*51*−*53*).

ACKNOWLEDGMENT

We thank the anonymous reviewers for helpful suggestions and valuable discussion 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. Humbel, R. E. (1990) *Eur. J. Biochem. 190*, 445−462.
7. Cook, R. M., Harvey, T. S., and Campbell, I. D. (1991) *Biochemistry 30*, 5484−5491.
8. 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.
9. 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.
10. Guo, Z.-Y., and Feng, Y.-M. (2001) *Biol. Chem. 382*, 443−448.
11. 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.
12. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X. G., and Markussen, J. (1991) *J. Mol. Biol. 220*, 425−433.
13. Wolf, E. D., Gill, R., Geddes, S., Pitts, J., Wollmer, A., and Grotzinger, J. (1996) *Protein Sci. 5*, 2193−2202.
14. Anfinsen, C. B. (1973) *Science 181*, 223−230.
15. Plaxco, K. W., Simons, K. T., and Baker, D. (1998) *J. Mol. Biol. 277*, 985−994.
16. Viguera, A. R., Serrano, L., and Wilmanns, M. (1996) *Nat. Struct. Biol. 3*, 874−880.
17. Baker, D. (2000) *Nature 405*, 39−42.
18. Creighton, T. E. (1974) *J. Mol. Biol. 87*, 603−642.
19. Creighton, T. E., Darby, N. J., and Kemmink, J. (1996) *FASEB J. 10*, 110−118.
20. Darby, N. J., Morin, P. E., Talbo, G., and Creighton, T. E. (1995) *J. Mol. Biol. 249*, 463−447.
21. Weissman, J. S., and Kim, P. S. (1991) *Science 253*, 1386−1393.
22. Weissman, J. S., and Kim, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 497−526.
23. Wedemeyer, W. J., Welker, E., Narayan, M, and Scheraga, H. A. (2000) *Biochemistry 39*, 4207−4216.
24. Rothwarf, D. M., and Scheraga, H. A. (1993) *Biochemistry 32*, 2671−9216.
25. Wu, J., Yang, Y., and Watson, J. T. (1998) *Protein Sci. 7*, 1017−1028.
26. Grantcharova, V., Alm, E. J., Baker, D., and Horwich, A. (2001) *Curr. Opin. Struct. Biol. 11*, 70−81.
27. Ellgaard, L., Molinari, M., and Helenius, A. (1999) *Science 286*, 1882−1888.
28. Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) *Science 286*, 1888−1893.
29. 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.
30. Chan, S. J., Cao, Q. P., and Steiner, D. F. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9319−9323.
31. 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.
32. Qiao, Z. S., Guo, Z. Y., and Feng, Y. M. (2001) *Biochemistry 40*, 2662−2668.
33. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. (1992) *Biochemistry 31*, 1749−1751.

34. Hober, S., Ljung, J. L., Uhlen, M., and Nilsson, B. (1999) *FEBS Lett. 443*, 271−276.
35. 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.
36. Hober, S., Ljung, J. L., Uhlen, M., Nilsson, B. (1997) *Biochemistry 36*, 4616−4622.
37. 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.
38. Milner, S. J., Carver, J. A., Ballard, F. J., and Francis, G. (1999) *Biotechnol. Bioeng. 62*, 693−703.
39. 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.
40. Hober, S., Hansson, A., Uhlen, M., and Nilsson, B. (1994) *Biochemistry 33*, 6758−6761.
41. 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.
42. Kim, P. S., and Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
43. Guo, Z.-Y., Shen, L., and Feng, Y.-M. (2002) *Biochemistry 41*, 1556−1567.
44. Cecil, R., and Weitzman, J. (1964) *Biochem. J. 93*, 1−11.
45. 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.
46. Kramer, W., Drusta, V., Jansen, H. W., Kramer, B., Pflugfelder, M., and Fritz, H. J. (1984) *Nucleic Acids Res. 12*, 9441−9454.
47. 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.
48. Hua, Q. X., Narhi, L., Jia, W., Arafawa, T., Rosenfeld, R., Hawkins, N., Miller, J. A., and Weiss, M. A. (1996) *J. Mol. Biol. 259*, 297−313.
49. Weiss, M. A., Hua, Q. X., Jia, W., Chu, Y. C., Wang, R. Y., and Katsoyannis, P. G. (2000) *Biochemistry 39*, 15429−15440.
50. Shen, L., Guo, Z.-Y., Chen, Y., Liu, L.-Y., and Feng, Y.-M. (2001) *Acta Biochim. Biophys. Sin. 33*, 629−633.
51. 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.
52. Chen, Z. Z., Schwartz, G. P., Zong, L., Burke, G. T., Chanley, J. D., and Katsoyannis, P. G. (1988) *Biochemistry 27*, 6105−6111.
53. 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.

BI020165F