The Sequence Determinant Causing Different Folding Behaviors of Insulin and Insulin-like Growth Factor-1†

Qi-Long Huang,‡ Jie Zhao,‡ Yue-Hua Tang,‡,§ Su-Qing Shao,‡ Gen-Jun Xu,‡,§ and You-Min Feng*,‡,§

College of Life Science, Zhejiang Sci-Tech University, Hangzhou 310018, China, and Institute of Biochemistry and Cell Biology, SIBS, Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China

Received August 18, 2006; Revised Manuscript Received October 25, 2006

ABSTRACT: Although insulin and insulin-like growth factor-1 (IGF-1) belong to the insulin superfamily and share highly homologous sequences, similar tertiary structure, and a common ancestor molecule, amphioxus insulin-like peptide, they have different folding behaviors: IGF-1 folds into two thermodynamically stable tertiary structures (native and swap forms), while insulin folds into one unique stable structure. To further understand which part of the sequence determines their different folding behavior, based on previous reports from the laboratory, two peptide models, [B9A][1−4]porcine insulin precursor (PIP) and [B10E][1−4]PIP, were constructed. The plasmids encoding the peptides were transformed into yeast cells for expression of the peptides; the results showed that the former peptide was expressed as single component, while the latter was expressed as a mixture of two components (isomer 1 and isomer 2). The expression results together with studies of circular dichoism, disulfide rearrangement, and refolding lead us to deduce that isomer 1 corresponds to the swap form and the isomer 2 corresponds to the native form. We further demonstrate that the sequence 1−4 plus B9 of IGF-1 B-domain can make PIP fold into two structures, while sequence 1−5 of insulin B-chain can make IGF-1 fold into one unique structure. In other words, it is the IGF-1 B-domain sequence that 1−4 allows IGF-1 folding into two thermodynamically stable tertiary structures; this sequence plus its residue B9E can change PIP folding behavior from folding into one unique structure to two thermodynamically stable structures, like that of IGF-1.

One of the most interesting but puzzling phenomena in protein folding is that insulin-like growth factor-1 (IGF-1)[1] can fold into two thermodynamically stable structures both in vivo and in vitro (*1*−*3*); that is to say, one protein sequence encodes two sets of folding information.

IGF-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 (*4*). Although the mature insulin consists of two chains, A-chain (21 residues) and B-chain (30 residues) linked by two interchain disulfides, it is synthesized as a single-chain polypeptide (named proinsulin) in vivo and folded well before processing into the mature insulin (*5*). The B- and A-domains of IGF-1 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 eight residue D-domain has no counterpart in insulin. IGF-1 adopts an insulin-like structure (*6*) whose ordered structure also mainly includes three segments of α-helix (8−18, α-helix I; 42−49, α-helix II; and 54−61, α-helix III) 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.

Insulin and IGF-1 belong to the insulin superfamily and share highly homologous sequences, similar tertiary structure, and a common ancestor molecule, amphioxus insulin-like peptide. However, insulin folds into one unique thermodynamically stable structure with disulfides (A20−B19, A7−B7, A6−A11) (*7*, *8*) (Figure 1B), while IGF-1 folds into two disulfide isomers (native and swap) with different three-dimensional (3D) structures and different disulfide linkages but similar thermodynamic stability (*2*, *3*, *9*). The native form adopts an insulin-like structure (*6*) with the disulfides 18−61, 6−48, and 47−52 (Figure 1A), corresponding to those of insulin; the swap form adopts a different folding pattern (*10*) with disulfides 18−61, 6−47, and 48−52 (corresponding A20−B19, A6−B7, A7−A11 in insulin, respectively). Disulfide A20−B19/18−61 exists in both native and swap forms of insulin/IGF-1 for its crucial role in their structure and folding (*11*, *12*). The 3D structure of swap IGF-1 is similar to that of swap insulin whose disulfide bridges are A20−B19, A6−B7, and A7−A11 (*7*). The α-helix II present in the native form is unfolded, and the other two α-helical segments still exist in the swap insulin and swap IGF-1 forms

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

* Address correspondence to You-Min Feng, Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China. Tel: (86) 21-54921133. Fax: (86) 21-54921011. E-mail: fengym@sunm.shcnc.ac.cn.

‡ Zhejiang Sci-Tech University.

§ Chinese Academy of Sciences.

[1] Abbreviations: IGF-1, insulin-like growth factor-1; PIP, recombinant porcine insulin precursor in which the C-terminus of porcine insulin B chain and the N-terminus of porcine insulin A chain are linked together by a dipeptide, Ala-Lys; DTT, reduced dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid; PAGE, polyacrylamide gel electrophoresis; CD, circular dichroism; UV, ultraviolet.

10.1021/bi0616798 CCC: $37.00
Published on Web 12/16/2006

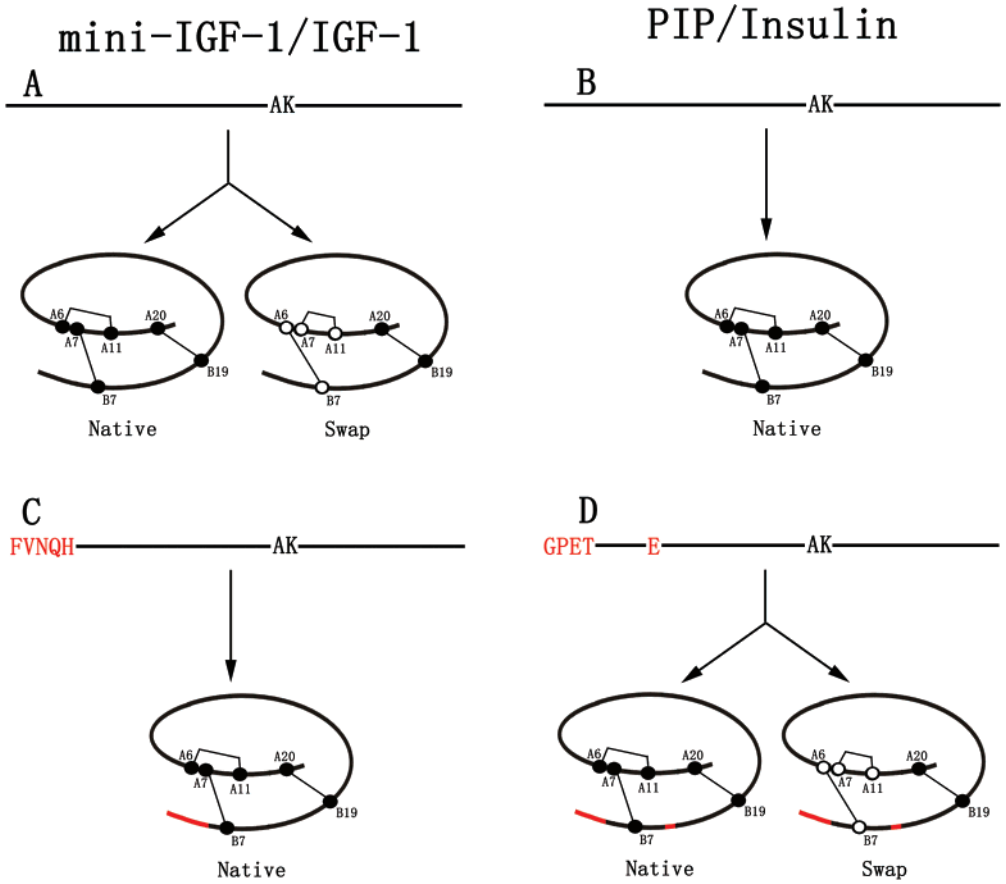

FIGURE 1: Primary structure and disulfide pairing of isomers of mini-IGF-1, PIP, and their mutants. Mini-IGF-1 and PIP are constructed by linking their B-domain/-chain (represent with long straight line) and A-domain/-chain (represented with a short straight line) with a dipeptide, AK. Despite their similarity in primary and tertiary structure, mini-IGF-1 and PIP exhibit different folding behaviors: mini-IGF-1 can fold into two thermodynamically stable structures (A), while PIP folds into only one (B). The two isomers of mini-IGF-1 are different in their disulfide pairings: in native form, the disulfide pattern is A6−A11, A7−B7 (solid circles) and A20−B19, while in swap form, the pattern is A6−B7, A7−A11 (hollow circles) and A20−B19. When the sequence 1−4 of mini-IGF-1 is replaced by its counterpart in PIP, FVNQH (highlighted in red), the consequent peptide ([*1*−*5*]mini-IGF-1, C) abolishes the bifurcate folding behavior and adopts a PIP-like one. When the sequence 1−5 plus 10 of PIP is replaced by its counterpart in mini-IGF-1, GEPT and E (highlighted in red), the consequent peptide ([1−4][B10E]PIP, D) adopts a bifurcate folding behavior, like mini-IGF-1.

(*10*). On the basis of the flexibility of the C-peptide/domain of insulin/IGF-1, PIP and mini-IGF-I were designed. Porcine insulin precursor (PIP) is a recombinant porcine insulin precursor, in which the C-peptide was replaced by a dipeptide, Ala-Lys (*13*, *14*). The folding of PIP is completely correct since the human insulin converted from PIP has been successfully used in the clinic. Further when B29Lys and A1Gly were directly linked together by peptide bond, the mini-proinsulin still retained the 3D structure identical to that of insulin but lost its biological activity completely (*15*, *16*). In mini-IGF-1, the C-domain is replaced by the same dipeptide, Ala-Lys, and the D-domain is deleted, which also folded into two thermodynamically stable structures (*17*). Other mini-IGF-1, like mini-proinsulin, in which B28Pro and A1Gly were directly connected by a peptide bond, still retained the three helical segments presented in the native IGF-1, although the orientation of these three helixes has been changed and its biological activity has been lost completely (*18*).

To understand why one sequence (IGF-1) folded into two structures and which part of the sequence controlled the different folding behaviors of insulin and IGF-1, DiMarchi et al. (*19*) had reported that it was the A- and B-domains that cause IGF-1 to have two different disulfide isomers, which indicated that the C- and D-domains did not influence the folding. Guo et al. (*17*, *20*) demonstrated that the different folding behaviors and the different energetic states of the intra A-chain/-domain disulfide of insulin and IGF-1 were mainly controlled by their B-chain/B-domain, which hinted that the A-chain/-domain did not influence the folding. Chen et al. (*21*) have further prepared four model peptides: [1−10]-mini-IGF-1 and [1−5]mini-IGF-1, in which the B-domain sequence 1−9 or 1−4 was replaced by the PIP B-chain sequence 1−10 or 1−5, respectively, [1−9]PIP and [1−4]-PIP, in which the B-chain sequence 1−10 or 1−5 was replaced by the IGF-1 B-domain sequence 1−9 or 1−4. Among the four peptides, the [1−10]mini-IGF-1 folded into one thermodynamically stable structure, while [1−9]PIP folded into two thermodynamically stable structures, which indicated that the sequence of B-chain/domain 1−10/1−9 of insulin and IGF-1 determined their different folding behaviors. They also observed that [1−5]mini-IGF-1 and [1−4]PIP both folded into one thermodynamically stable structure, which demonstrated that insulin B-chain sequence 1−5 has abolished the generation of swap IGF-1, while IGF-1 B-domain sequence 1−4 could not make PIP generate swap PIP. To further understand the sequence determinant determining the different folding behavior of insulin and IGF-

1, two PIP mutants, [B9A][1−4]PIP and [B10E][1−4]PIP, were prepared by means of protein engineering and analyzed. Here we report the preparation, structure analysis, and folding behavior studies of the mutants compared with that of PIP and mini-IGF-1.

EXPERIMENTAL PROCEDURES

Materials. The *Escherichia coli* strain used was DH12S. *Saccharomyces cerevisiae* XV700-6B (Leu2, ura3, pep4) was kindly provided by Michael Smith (University of British Columbia, Vancouver, Canada). Plasmid pVT102-U/α-MFL-[1−4]PIP was constructed in our laboratory. The mutagenesis oligonucleotide primers were chemically synthesized. The chemical reagents used in the experiments were of analytical grade. The DIKMA (GL Science Inc.) reverse-phase columns (Inertsil, C8−3 5 μm 4.6 × 250 mm 100 Å), Gilson 306 HPLC system, and Gilson 115 UV detector were used. In HPLC analysis, a gradient elution was used. Solvent A was 10% acetonitrile containing 1.5% TFA; solvent B was 60% acetonitrile containing 0.125% TFA. The elution gradient was 25−75% B solvent in 30 min. During analysis, the flow rate is 1.0 mL/min, and the detection wavelength was set at 230 nm.

DNA Manipulation. The plasmids encoding the peptides were constructed following the product manual of Stratagene's Quick Change II Site-Directed Mutagenesis kit. The expected mutations were confirmed by DNA sequencing. The plasmids were designated as pVT102-U/α-MFL-[1−4][B9A]-PIP and pVT102-U/α-MFL-[1−4][B10E]PIP, respectively.

Expression and Purification of the [1−4][B9A]PIP and [1−4][B10E]PIP (Isomers 1 and 2). The plasmids pVT102-U/α-MFL-[1−4][B9A]PIP and pVT102-U/α-MFL-[1−4]-[B10E]PIP were transformed into *S. cerevisiae* XV700-6B (Leu2, ura3, pep4) cells, respectively. The transformed yeast cells were cultured in a 16-L fermenter, and the secreted peptides were purified from the media supernatant according to the procedure of Guo et al. (*14*, *17*). Briefly, first, the peptide was precipitated from the media supernatant by trichloroacetic acid. Second, the precipitate was dissolved with 1 M acetic acid and applied to a Sephadex-G50 column. Third, the product was purified by ion-exchange chromatography on a DEAE-Sepharose F.F. column. Fourth, the eluted peptide was further purified by C8 reverse-phase HPLC with gradient elution described in Materials. The purity of the peptide was analyzed by native pH 8.3 PAGE and analytical C8 reverse-phase HPLC. Their molecular mass was measured by electrospray mass spectrometry.

Circular Dichroism Studies of the [1−4][B9A]PIP and [1−4][B10E]PIP (Isomers 1 and 2). The peptides were dissolved in 1 mM HCl. The protein concentration was determined by UV absorbance at 276 nm, and the final concentration was adjusted to 0.2 mg/mL. Circular dichroism (CD) measurements were performed on a Jasco-715 spectropolarimeter at room temperature. The far-UV spectrum from 250 to 190 nm was scanned 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 CD spectra.

Disulfide Rearrangements of the [1−4][B9A]PIP and [1−4][B10E]PIP (Isomers 1 and 2). The purified peptides were dissolved in 0.1 M Tris-HCl buffer (pH 8.7) containing 1 mM EDTA and 0.2 mM 2-mercaptoethanol at 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 of sample was removed, immediately adjusted to pH 2.0 with 50% TFA to terminate the disulfide rearrangement reaction, and analyzed with analytical C8 reverse-phase HPLC. Gradient elution as described under Materials was used with detection at 230 nm.

In Vitro Refolding of [1−4][B9A]PIP and the Isomer 2 of [1−4][B10E]PIP. To prepare the fully reduced unfolding peptides, the peptides were dissolved, respectively, in the unfolding buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 8.7) at the final protein concentration of 1 mg/mL. The reduction reaction was started by addition of DTT to a concentration of 10 mM. The reduction reaction was carried out at 30 °C for 30 min. After reduction, an aliquot was removed and carboxymethylated by a one-fifth volume of fresh prepared 0.5 M iodoacetic acid sodium salt solution and then analyzed by native pH 8.3 PAGE to determine whether the peptides were fully reduced. The fully reduced peptides were immediately stored at −80 °C for later use.

In the refolding reaction, the fully reduced peptide was diluted in refolding buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 9.2) at final concentration of 0.1 mg/mL. The oxidative refolding reaction was initiated by adding oxidized glutathione to the refolding buffer making the final concentration of oxidized glutathione at 1 mM. The refolding reaction was carried out at 16 °C for 12 h. Then the refolding mixture was acidified to pH 2.0 with 50% TFA and immediately analyzed by C8 reverse-phase HPLC.

RESULTS

Expression and Purification of [B9A][1−4]PIP and [B10E][1−4]PIP. The plasmids of pVT102U/α-MFL-[B9A]-[1−4]PIP and pVT102U/α-MFL-[B10E][1−4]PIP were confirmed to be correct by DNA sequencing and were transformed into yeast cells for expression in a 16-L fermenter. A four-step procedure described in Experimental Procedures was applied to separate and purify the peptides. Two isomers of [B10E][1−4]PIP were separated from the media supernatant (Figure 2B), while only a single component for [B9A]-[1−4]PIP (Figure 2A). The isomers of [B10E][1−4]PIP were further separated and purified (Figure 2C for isomer 1 and Figure 2D for isomer 2). All of the purified peptides were analyzed by native pH 8.3 PAGE for checking of the purity (Figure 2E). The two isomers of [B10E][1−4]PIP showed different mobility rates despite their identity in amino acids: isomer 2 ran a little faster than isomer 1, indicating that isomer 2 has a tighter secondary and/or tertiary structure. The relative molecular mass of [B9A][1−4]PIP, isomer 1 and isomer 2 of [B10E][1−4]PIP was 5702, 5710, 5710, which was identical to the theoretical value of 5702, 5710, 5710, respectively. From the data, we assume that the two components of [B10E][1−4]PIP were the two disulfide isomers just like those of IGF-1.

Circular Dichroism Analysis of [B9A][1−4]PIP, [B10E]-[1−4]PIP (Isomers 1 and 2) and PIP. The secondary structures of the four peptides were analyzed by far-UV circular dichroism spectra, using PIP as a standard. Isomer 2 of [B10E][1−4]PIP and [B9A][1−4]PIP showed a profile

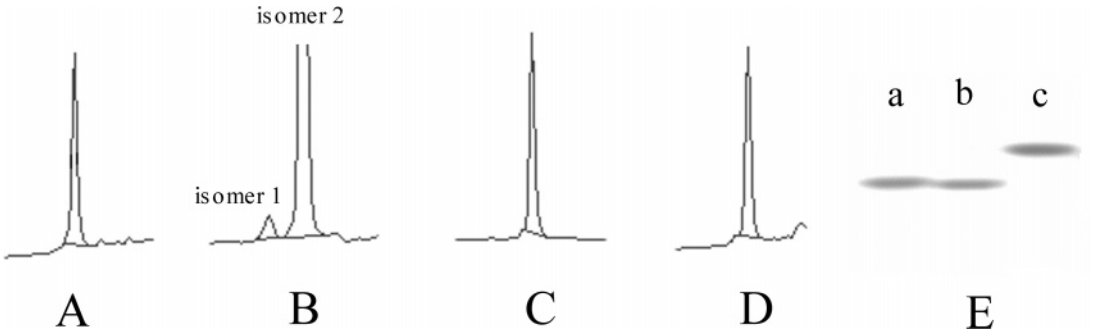

FIGURE 2: [B9A][1−4]PIP (A) was expressed in yeast cells as a single component, while [B10E][1−4]PIP (B) is a mixture of two components. Isomer 1 and isomer 2 were separated and purified for latter experiments (C as isomer 1 and D as isomer 2). E shows the electrophoresis behavior of isomer 2, isomer 1 of [B10E][1−4]PIP and [B9A][1−4]PIP (from lane a to c). Isomer 2 ran a little faster than isomer 1 indicating that a tighter secondary/tertiary structure been adopted. [B9A][1−4]PIP ran the slowest for it carried less negative charge than did [B10E][1−4]PIP.

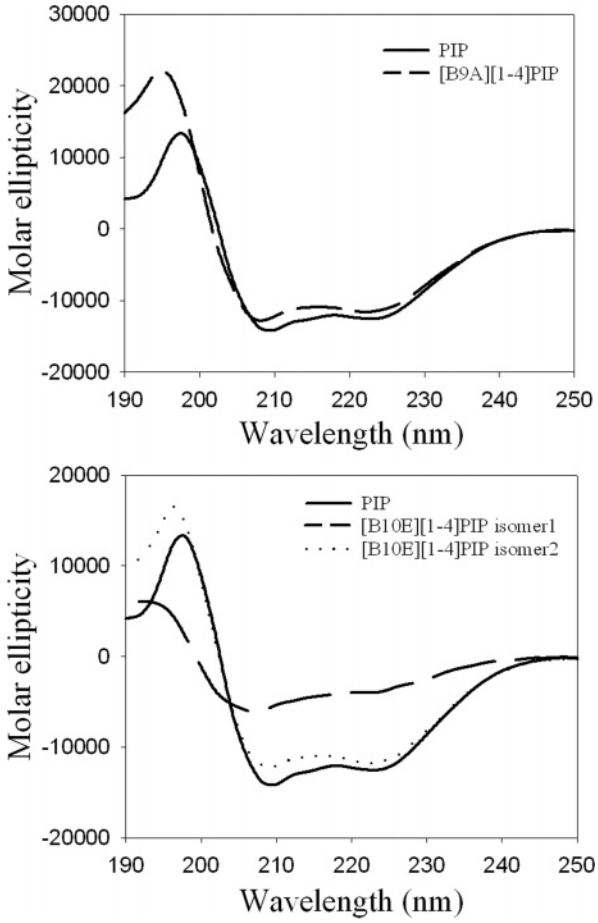

FIGURE 3: Far UV circular dichroism spectra of PIP, [B9A][1−4]PIP, and [B10E][1−4]PIP (isomers 1 and 2). For clarity, [B9A]-[1−4]PIP and [B10E][1−4]PIP were compared with PIP, respectively.

similar to PIP, while the isomer 1 of [B10E][1−4]PIP showed a large difference (Figure 3). The helix content of these four peptides, [B9A][1−4]PIP, isomer 1, and isomer 2 of [B10E][1−4]PIP as well as PIP were 41.3, 9.1, 44.6, and 41.3%, respectively. On the basis of their behavior in RP-HPLC and PAGE, and the CD data, together with the works of Guo et al. and Chen et al., we deduced that [B9A]-[1−4]PIP adopted an insulin-like structure; isomer 1 of [B10E][1−4]PIP adopted the swap IGF-1-like structure; isomer 2 of [B10E][1−4]PIP adopted a native IGF-1-like structure (also an insulin-like structure); for the proteins with

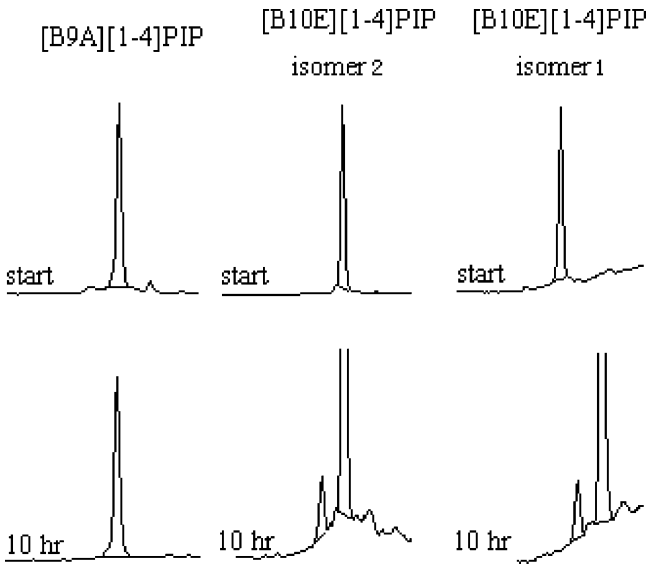

FIGURE 4: Disulfide rearrangement of [B9A][1−4]PIP and the two isomers of [B10E][1−4]PIP. After 10 h of incubation in alkaline buffer containing 0.2 mM 2-mercaptoethanol, a 100 μL (10 μg) sample was removed and acidified by 50% TFA and immediately loaded onto a C8 reverse phase column and analyzed by HPLC according to Experimental Procedures. The two isomers of [B10E]-[1−4]PIP were interconvertable and could reach the same quilibrium.

a bifurcating folding pathway (such as IGF-1, [1−4]PIP), the native form has a tighter structure and a longer retention time than the swap form. And for the protein that can only fold into one stable product (such as insulin/PIP, [1−5]IGF-1), the native form can be adopted.

Disulfide Rearrangement of [B9A][1−4]PIP and [B10E]-[1−4]PIP (*Isomers 1 and 2*). When protein with disulfide bonds is dissolved in solution containing a thiol reagent, its disulfide bond can be rearranged. The disulfide bonds in isomers of [B10E][1−4]PIP can transfer in such solution through disulfide rearrangement; when isomer 1 was incubated in the condition described in Experimental Procedures, it could convert into isomer 2 spontaneously through disulfide rearrangement; finally an equilibrium was reached, in which the ratio of isomer 1 to isomer 2 was about 1:10 (calculated from peak area); the same equilibrium could be reached when isomer 2 was incubated in such condition. But when the same operation was applied to [B9A][1−4]PIP, no component other than itself was detected (Figure 4).

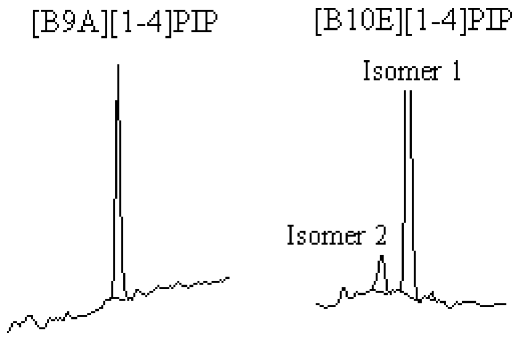

FIGURE 5: In vitro refolding of fully reduced [B9A][1−4]PIP and [B10E][1−4]PIP. After incubation in the refolding buffer at 16 °C for 12 h, the fully reduced isomer 2 of [B10E][1−4]PIP folded into two isomers, while [B9A][1−4]PIP folded into a unique one.

In Vitro Refolding of Isomer II of [1−4][B10E]PIP and [B9A][1−4]PIP. To further demonstrate the sequence of the [B10E][1−4]PIP encoding two thermodynamically stable folding products and [1−4][B9A]PIP encoding an unique one, an in vitro refolding experiment was carried out. Since the fully reduced isomer 2 and isomer 1 are identical, only the refolding of isomer 2 has been carried out. After the samples were incubated in alkaline buffer containing DTT, an aliquot was removed and treated with iodoacetic sodium salt solution, and then applied to native pH 8.3 PAGE to check whether the samples were completely reduced (data not shown). After incubation of the sample in the refolding buffer at 16 °C for 12 h, the fully reduced isomer 2 of [B10E]-[1−4]PIP folded into two isomers, while the [B9A][1−4]-PIP folded into a unique one (Figure 5).

DISCUSSION

The Sequence Determinant of the Different Folding Behaviors of Insulin and IGF-1. The present results show that when the B-chain sequence 1−5, FVNQH, and B10H of PIP were substituted by the corresponding IGF-1 B-domain sequence 1−4, GPET, and B9E, the mutant PIP, [B10E][1−4]PIP, folded into two structures: native and swap forms (Figure 1D), while [B9A][1−4]PIP folded into only one structure. A previous report from our laboratory indicated that when the mini-IGF-1 B-domain sequence 1−4 was replaced by the PIP B-chain sequence 1−5, the mutant mini-IGF-1, [1−5]mini-IGF-1, folded into only one structure, native form (*21*). Both results demonstrate that it is the IGF-1 B-domain sequence 1−4 that determines its bifurcated folding behavior. The present results demonstrate that the sequence 1−4 plus B9E (corresponding to B10H of PIP) of IGF-1 B-domain make the mutant PIP, [B10E][1−4]PIP, fold into two structures, native and swap PIP. However, the sequence 1−4 or B9E alone does not cause the PIP mutants, [1−4]PIP (*21*) and [B10E]insulin/IGF-1(C) (*22*), to fold into two structures. The relationship between sequence and folding behavior is summarized in Figure 6.

Deducing How the Sequence Determinant Determines the Folding Behavior. Yun et al. (*23*) compared the crystal structure of mini-IGF-1 (native form) with that of insulin and observed that the insulin N-terminal of the B-chain ran backward close to the A-chain, and several hydrogen bonds were formed between A6−A11 and B1−B7. In the same region of IGF-1 and mini-IGF-1, such hydrogen bonds do not exist. They deduced that the B1F and B5H play an important role in the formation of such region and those hydrogen bonds, while there are no such Phe and His residues in the corresponding positions in IGF-1 B-domain. From the structures of insulin and IGF-1/mini-IGF-1 (native form), we may elucidate that the replacing of the B-domain sequence of mini-IGF-1 by the corresponding B-chain sequence 1−5 of PIP can lead the B-domain sequence 1−5 of [1−5]mini-IGF-1 backward close to the A-domain (just like that of insulin) and the formation of several hydrogen bonds between the A-domain 47−52 (corresponding to insulin A6-A11) and the B-domain B1−B6 (corresponding to insulin B1−B7). Therefore, [1−5]mini-IGF-1 folds into one unique structure. The IGF-1 B-domain 1−4 sequence alone does not make [1−4]PIP fold into two structures (*20*); however, when the B10H of [1−4]PIP was replaced by Glu (corresponding to mini-IGF-1B9), the mutant PIP, [B10E]-[1−4]PIP, does fold into two structures. Inspection of the insulin crystal structure (*24*) suggests that the B5His and B10His residues locate closely at both sides of the disulfide region formed by the disulfide A6−A11 and A7−B7 (Figure 7). Therefore, the residue B10His may, like B5His, also play an important role in stabilizing the disulfide A6−A11 and A7−B7. This may explain why PIP B-chain sequence 1−5

PIP/Insulin:

B chain		A chain	Number of Folded Structure	References
FVNQHLCGSHLVEALYLVCGERGFFYTPKA	AK	GIVEQCCTSICSLYQLENYCN	1	(11)
GPETLCGAELVDALQFVCGDRGFYFNKPT	AK	GIVEQCCTSICSLYQLENYCN	2	(15)
*GPETLCGAE*LVEALYLVCGERGFFYTPKA	AK	GIVEQCCTSICSLYQLENYCN	2	(17)
*GPET*LCGSHLVEALYLVCGERGFFYTPKA	AK	GIVEQCCTSICSLYQLENYCN	1	(17)
*GPET*LCG*A*HLVEALYLVCGERGFFYTPKA	AK	GIVEQCCTSICSLYQLENYCN	1	Present Work
FVNQHLCGS*E*LVEALYLVCGERGFFYTPKA	*GYGSSSRRAPQG*	GIVEQCCTSICSLYQLENYCN	1	(20)
*GPET*LCGS*E*LVEALYLVCGERGFFYTPKA	AK	GIVEQCCTSICSLYQLENYCN	2	Present Work

Mini-IGF-1/IGF-1:

B domain		A domain		
GPETLCGAELVDALQFVCGDRGFYFNKPT	AK	GIVDECCFRSCDLRRLEMYCA	2	(13)
FVNQHLCGSHLVEALYLVCGERGFFYTPKA	AK	GIVDECCFRSCDLRRLEMYCA	1	(15)
*FVNQHLCGSE*LVDALQFVCGDRGFYFNKPT	AK	GIVDECCFRSCDLRRLEMYCA	1	(17)
*FVNQH*LCGAELVDALQFVCGDRGFYFNKPT	AK	GIVDECCFRSCDLRRLEMYCA	1	(17)

FIGURE 6: Relationship between sequence and folding behavior. The upper panel shows sequence and numbers of thermodynamically stable folded structures of PIP/insulin mutants. The lower panel shows that of mini-IGF-1/IGF-1 mutants. The colored underlined italic characters represent the mutated parts. These parts in insulin are represented by their counterparts in IGF-1, vice versa. From the figure, we can see that B1−B4 is the sequence determinant of IGF-1's folding behavior and B1−B5 plus B10 is the sequence determinant of insulin's folding behavior.

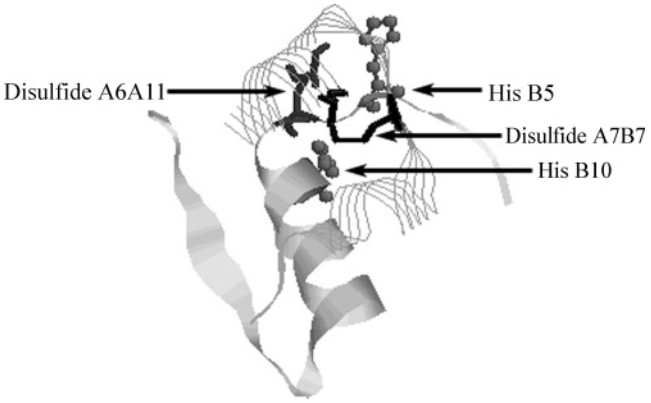

FIGURE 7: The crystal structure of insulin. It shows that residues B5H and B10H locate closely by the disulfide A6−A11 and A7−B7, so the side chain of B5 and B10 may stabilize the disulfide A6−A11 and/or A7−B7 through hydrogen bonds formed between imidazole groups and relative main chain N or O atoms of residues.

(including B5His) or B10His alone was replaced by IGF-1 B-domain sequence 1−4 or B9E formed PIP mutants, [1−4]-PIP and [B10E]insulin/IGF-1(C), still folded into one unique structure, since both mutants retain one His residue at position B5 or B10. In contrast, when both His residues were replaced by the corresponding residues of IGF-1, [1−4]-[B10E]PIP folded into two structures. Therefore, to understand the structure in detail, the crystal structure or NMR analysis of [1−4][B10E]PIP and [1−5]mini-IGF-1 as compared with the structures of insulin and IGF-1 will be necessary.

Amphioxus insulin-like peptide (aILP) not only belongs to the insulin superfamily but also is considered to be the common ancestor molecule of insulin and IGF-1 (*25*, *26*). The folding behavior of aILP shows the characteristics of both insulin and IGF-1 (*27*, *28*), which on one hand suggests that the different folding behaviors of insulin and IGF-1 are acquired through a bifurcating evolution, but on the other hand, provides some useful clues for elucidation of the sequence determinant that leads to different folding behaviors of insulin and IGF-1. For example, when the residue at position B10/B9 of insulin, aILP, and IGF-1 is His, Thr, and Glu, respectively, the electron charge of the side chains changes from positive to neutral to negative. Concerning the ratio of native form to swap form, insulin is 100/0, aILP is 98/2 (*28*), IGF-1 is 71/29 (*29*), and [B10E][1−4]PIP is 90/10 (present results), so we may suggest that the residue His or Glu at B10/B9 plays an important role in determining the folding behavior of insulin or IGF-1.

ACKNOWLEDGMENT

We thank the unknown enthusiastic reviewer for helpful suggestions on the manuscript. We thank Hong-Yu Hu and Xiao-jing Lin for great help in CD analysis, and Fei Chen for introducing the method of site directed mutagensis.

REFERENCES

1. Elliott, S., Fagin, K. D., Narhi, L. O., Miller, J. A., Jones, M., Koski, R., Peters, M., Hsieh, P., Sachdev, R., Rosenfeld, R. D. (1990) Yeast-derived recombinant human insulin-like growth factor I: production, purification, and structural characterization, *Protein Chem. 9*, 95−104.
2. 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.
3. 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.
4. Humbel, R. E. (1990) Insulin-like growth factors I and II, *Eur. J. Biochem. 190*, 445−462.
5. Steiner, D. F., Kemmler, W., Clark, J. L., Oyer, P. E., and Rubenstein, A. H. (1972) in *Handbook of Physiology* (Steiner, D. F., and Freinkel, N., Eds.) pp 175-198, Williams and Wilkins, Baltimore, MD.
6. Cook, 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.
7. 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.
8. 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.
9. Narhi, L. O., Hua, Q. X., Arakawa, T., Fox, G. M., TsaiL, Rosenfeld, R., Holst, P., Miller, J. A., Weiss, M. A. (1993) Role of native disulfide bonds in the structure and activity of insulin-like growth factor 1: genetic models of protein-folding intermediates, *Biochemistry 18*, 5214−5221.
10. 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.
11. Yan, H., Guo, Z. Y., Gong, X. W., Xi, D., and Feng, Y. M. (2003) A peptide model of insulin folding intermediate with one disulfide, *Protein Sci. 4*, 768−775.
12. Hua, Q. X., Mayer, J. P., Jia, W., Zhang, J., Weiss, M. A. (2006) The folding nucleus of the insulin superfamily: a flexible peptide model foreshadows the native state, *J. Biol. Chem. 2*, 28131−28142.
13. Markussen, J., Damgaard, U., Diers, I., Fiil, N., Hansen, M. T., Larsen, P., Norris, F., Norris, K., Schou, O., Snel, L., Thim, L., and Voigt, H. O. (1987) Biosynthesis of human insulin in yeast via single-chain precursors, in (Theodoropoulos, D., Ed.) *Peptides 1986*, pp 189−194, de Gruyter, Berlin.
14. 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) Secretory expression of a single-chain insulin precursor in yeast and its conversion into human insulin, *Sci. China* (*Ser. C*) *39*, 225−233.
15. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X. G., and Markussen, J. (1991) X-ray analysis of the single chain B29-A1 peptide-linked insulin molecule. A completely inactive analogue, *J. Mol. Biol. 220*, 425−433.
16. 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) Mini-proinsulin and mini-IGF-1: homologous protein sequences encoding non-homologous structures, *J. Mol. Biol. 277*, 103−118.
17. 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.
18. Wolf, E. D., Gill, R., Geddes, S., Pitts, J., Wollmer, A., and Grotzinger, J. (1996) Solution structure of a mini IGF-1, *Protein Sci. 5*, 2193−2202.
19. 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.
20. 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.
21. Chen, Y., You, Y., Jin, R., Guo, Z. Y., and Feng, Y. M. (2004) Sequences of B-chain/domain 1-10/1-9 of insulin and insulin-like growth factor 1 determine their different folding behavior, *Biochemistry 43*, 9225−9233.

22. Guo, Z. Y., and Feng, Y. M. (2002) The thermodynamic stability of insulin disulfides is not affected by the C-domain of insulin-like growth factor 1, *Sci. China (Series C) 45*, 245−250.
23. Yun, C. H., Tang, Y. H., Feng, Y. M., An, X. M., Chang, W. R., Liang, D. C. (2005) 1.42 Å crystal structure of mini-IGF-1(2): an analysis of the disulfide isomerization property and receptor binding property of IGF-1 based on the three-dimensional structure, *Biochem. Biophys. Res. Commun. 326*, 52−59.
24. 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 Zn pig insulin crystals at 1.5 Å resolution. *Philos. Trans. R. Soc. London B Biol. Sci. 319*, 369−456.
25. 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.
26. 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.
27. Chen, Y., Jin, R., Dong, H. Y., and Feng, Y. M. (2004) In vitro refolding/unfolding pathways of amphioxus insulin-like peptide: implications for folding behavior of insulin family proteins, *J. Biol. Chem. 279*, 55224−55233.
28. 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.
29. Iwai, M., Yokoyama, H., Yamada, H., Niwa, M., Kobayashi, M. (2000) Direct identification of a novel disulfide bond linkage system of new isolated isomer (isomer V) in recombinantly produced h-IGF-1, *Chem. Pham. Bull. (Tokyo) 48*, 1304−1309.

BI0616798