Putative Disulfide-Forming Pathway of Porcine Insulin Precursor during Its Refolding in Vitro†

Zhi-Song Qiao, Zhan-Yun Guo, and You-Min Feng*

State Key Laboratory of Molecular Biology, Shanghai Institute of Biochemistry, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, People's Republic of China

Received July 12, 2000; Revised Manuscript Received January 2, 2001

ABSTRACT: Although the structure of insulin has been well studied, the formation pathway of the three disulfide bridges during the refolding of insulin precursor is ambiguous. Here, we reported the in vitro disulfide-forming pathway of a recombinant porcine insulin precursor (PIP). In redox buffer containing L-arginine, the yield of native PIP from fully reduced/denatured PIP can reach 85%. The refolding process was quenched at different time points, and three distinct intermediates, including one with one disulfide linkage and two with two disulfide bridges, have been captured and characterized. An intra-A disulfide bridge was found in the former but not in the latter. The two intermediates with two disulfide bridges contain the common A20−B19 disulfide linkage and another inter-AB one. Based on the time-dependent formation and distribution of disulfide pairs in the trapped intermediates, two different forming pathways of disulfide bonds in the refolding process of PIP in vitro have been proposed. The first one involves the rapid formation of the intra-A disulfide bond, followed by the slower formation of one of the inter-AB disulfide bonds and then the pairing of the remaining cysteines to complete the refolding of PIP. The second pathway begins first with the formation of the A20−B19 disulfide bridge, followed immediately by another inter-AB one, possibly nonnative. The nonnative two-disulfide intermediates may then slowly rearrange between CysA6, CysA7, CysA11, and CysB7, until the native disulfide bond A6−A11 or A7−B7 is formed to complete the refolding of PIP. The proposed refolding behavior of PIP is compared with that of IGF-I and discussed.

The studies on the pathway of protein folding and unfolding are attractive, not only for theoretical interests but also important for practical application (*1*). Proteins with disulfide linkages are frequently chosen for such studies because formation of disulfide bridges is always coupled with folding and assembly, and may serve as a unique probe for the study of both phenomena (*2*, *3*). Several proteins have been widely studied in this way, such as BPTI[1] (*4*−*9*), RNaseA (*10*−*12*), and EGF (*13*, *14*).

Members of the insulin superfamily, such as insulin, IGF-I, IGF-II, relaxin, bombyxin, etc., all contain three motif-specific disulfide linkages (*15*). Human IGF-I had its refolding process widely studied for being a singe-chain member of the family more readily to be studied (*16*−*20*). The bifurcating protein-refolding pathway of IGF-I has also attracted interest. The final refolding products always contain two isoforms, with 60% in the native form and 40% the disulfide-mismatched IGF-swap with disulfide linkages [6−47, 48−52, 18−61], corresponding to [B7−A6, A7−A11, B19−A20] in insulin (*17*). The disulfide linkage 18−61 (corresponding to B19−A20 in insulin) is found in all the intermediates trapped, suggesting that it is the first disulfide bridge formed during refolding. One native one-disulfide intermediate and at least three two-disulfide intermediates have been isolated and characterized. Of all three disulfide linkages, the one corresponding to the intra-A bridge of insulin is the last one formed. The putative refolding pathway of IGF-I has been proposed (*18*−*20*).

Although the structure of insulin has been well studied, the process forming the three disulfide bridges in the insulin precursor during refolding is still ambiguous. Previous work indicated that the intra-A disulfide linkage was formed first during the refolding of insulin as judged by comparing the speed of free cysteine disappearance in insulin and that in [A6, A11-Ser] insulin mutant (*21*). The results indicate that the intra-A disulfide bridge is the most stable one and its formation may help the A chain to form a stable helix structure which is necessary to complete the correct pairing with the B chain. But no intermediates containing the intra-A disulfide bridge have been isolated during refolding. So the conclusion remains a deduction without direct proof, and the order of formation of the remaining two interchain disulfide bridges during the refolding of insulin is still unknown.

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

* Correspondence should be addressed to this author at State Key Laboratory of Molecular Biology, Shanghai Institute of Biochemistry, SIBS, Chinese Academy of Sciences, 320 Yue-Yang Rd., Shanghai 200031, China. Tel: (86) 021-64374430. Fax: (86)021-64338357. E-mail: fengym@sunm.shcnc.ac.cn.

[1] Abbreviations: PIP, porcine insulin precursor; L-Arg, L-arginine; BPTI, bovine pancreatic trypsin inhibitor; RNaseA, ribonuclease A; EGF, epidermal growth factor; IGF-I, insulin-like growth factor-I; IGF-II, insulin-like growth factor-II; IAA, sodium salt of iodoacetic acid; GSH, reduced glutathione; GSSG, oxidized glutathione; DTT, reduced dithiothreitol; frdPIP, fully reduced/denatured PIP; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid; ESI-MS, electrospray ionization−mass spectrometry; UV, ultraviolet; PAGE, polyacrylamide gel electrophoresis; CD, circular dichroism; *n*SSPIP, intermediates of PIP containing *n* disulfide bonds.

10.1021/bi001613r CCC: $20.00
Published on Web 02/08/2001

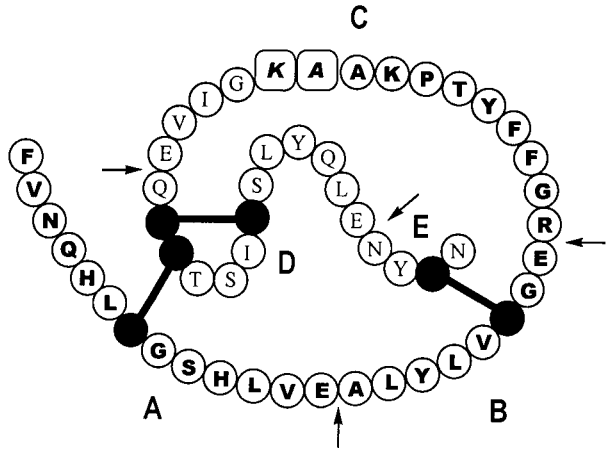

FIGURE 1: Amino acid sequence and disulfide linkage structure of single-chain porcine insulin precursor. Amino acids are shown in the one-letter code, and cysteine residues are shown as dark circles. The fragment containing boldface letters in a circle corresponds to the B chain of insulin. The two residues (K and A) in the rectangle represent the linkage peptide between the B chain and the A chain of insulin. The fragment following the linkage peptide corresponds to the A chain of insulin. The potential cleavage sites of protease V8 are indicated with arrows. Peptide fragments obtained after enzyme cleavage are designated with letters (A−E).

Porcine insulin precursor (PIP) is a recombinant single-chain insulin precursor whose B30Ala and A1Gly are linked by a dipeptide linkage, Ala−Lys, as shown in Figure 1. The recombinant PIP was expressed and secreted in a correctly folded and soluble form from *S. cerevasiae* (*22*). It can be converted into human insulin by means of transpeptidation. Because the crystal structures of the single-chain [B29Lys-A1Gly] insulin and porcine insulin are identical (*23*), it is reasonable to believe that the dipeptide cross-linkage in PIP does not affect its three-dimensional structure and PIP should have a similar tertiary structure as that of insulin. The three-dimensional structure of IGF-I was first modeled based on the X-ray-crystallographic structure of insulin (*24*). Later, the refined tertiary structure of IGF-I in solution was resolved by NMR, which showed that IGF-I and insulin have a very similar three-dimensional structure (*25*).

We have isolated and characterized three intermediates trapped during the in vitro refolding process of PIP. The present study brings some new insight into the forming pathway of disulfide linkage in PIP in vitro, and a refolding mechanism of insulin precursor by comparing PIP with IGF-I is suggested.

MATERIALS AND METHODS

Materials. Recombinant native PIP was produced and purified as described previously (*22*). Endoproteinase Lys-C and V8 were of sequencing grade and purchased from Boehringer Mannheim. The snakeskin dialysis tube with a molecular mass cutoff of 3500 daltons was the product of Pierce. Sodium salt of iodoacetic acid (IAA), reduced glutathione (GSH), and oxidized glutathione (GSSG) were ultrapure and purchased from Amerresco. Ultrapure dithiothreitol (DTT) and L-arginine (L-Arg) were purchased from BI. Ultrapure urea was the product of Promega. Acetonitrile and trifluoroacetic acid were HPLC grade. All the other reagents used in the experiment were analytical grade.

Refolding of Fully Reduced/Denatured PIP (frdPIP). Purified PIP was dissolved in 50 mM Tris-HCl buffer (pH 8.5) containing 8 M urea, 1 mM EDTA, and 0.1 M DTT. Reduction of the protein was carried out at room temperature for 30 min. Thereafter, the buffer was immediately exchanged with 10 mM HCl by gel filtration using Sephadex G-25 (Pharmacia, Sweden) and stored at −80 °C for later use. The refolding of reduced and denatured protein was initiated by adding the frdPIP into the prewarmed (16 °C) refolding buffer (0.5 M L-Arg, 50 mM Tris-HCl, 5 mM EDTA, 5 mM GSH, and 0.5 mM GSSG, pH 9.5) at a final protein concentration of 0.1 mg/mL. The refolding intermediates were trapped at different time points by the method described below.

Trapping Intermediates. Refolding intermediates were trapped in a time course manner by removing aliquots of protein solution which was quickly mixed with 1/5 volume of freshly prepared 0.5 M iodoacetic acid in 50 mM Tris-HCl (pH 8.5). The carboxymethylation of free thiol groups by IAA was carried out at room temperature for 5 min. To observe the temporal distribution of refolding intermediates, the solution was adjusted to pH 1.0 with TFA and immediately analyzed by HPLC under the conditions described below. For large preparation of the IAA-trapped intermediates, the solution containing IAA-trapped intermediates was thoroughly dialyzed against 20 mM NH_4HCO_3 at 4 °C and thereafter lyophilized. Then the refolding mixture was separated by Mono Q and HPLC under the conditions described below.

Separation and Purification of Refolding Intermediates. The separation of refolding intermediates by Mono Q was carried out on FPLC (Pharmacia, Sweden) with a linear NaCl gradient of 0−1.0 M. The UV detection was at 280 nm. Solvent A was 20 mM Tris-HCl (pH 7.6), 1 mM EDTA, and solvent B was 20 mM Tris-HCl (pH 7.6), 1 mM EDTA, and 1.0 M NaCl. The fractions were collected manually and then exchanged into 20 mM NH_4HCO_3 by gel filtration using Sephadex G-25; finally the sample was lyophilized for further use.

The separation of refolding intermediates by reverse-phase HPLC was carried out with a C8 column (Vydac, 0.4 cm × 18 cm). Solvent A was 0.15% aqueous trifluoroacetic acid. Solvent B was 60% acetonitrile containing 0.125% trifluoroacetic acid. The UV detection was at 280 nm, and the linear gradient was 40−80% solvent B in 30 min when separating the carboxymethylated intermediates. The UV detection was at 230 nm, and the linear gradient was 40−60% solvent B in 30 min when separating the fragment of intermediates digested by endoproteinase V8 or Lys-C. The flow rate is 0.8 mL/min. The HPLC fractions were collected manually and then lyophilized for further use.

Digesting Intermediates with Endoproteinase. The HPLC-purified intermediates were dissolved in 50 mM NH_4HCO_3 (pH 8.0) at a concentration of 0.5 mg/mL, and the endoproteinase Lys-C was added into the solution at a mass ratio of 1:1000 to the intermediate. When digested with V8, the intermediate was dissolved in 0.1 M phosphate buffer (pH 8.0), and V8 was added to the solution at a mass ratio of 1:20. All the reactions were carried out at 20 °C overnight. After digestion, the solution was adjusted with TFA to pH 1.0 and immediately separated by HPLC as described above.

Protein Analysis. The concentrations of frdPIP, PIP, and the intermediates were determined by measuring the absorption at 276 nm with a spectrophotometer and calculating the

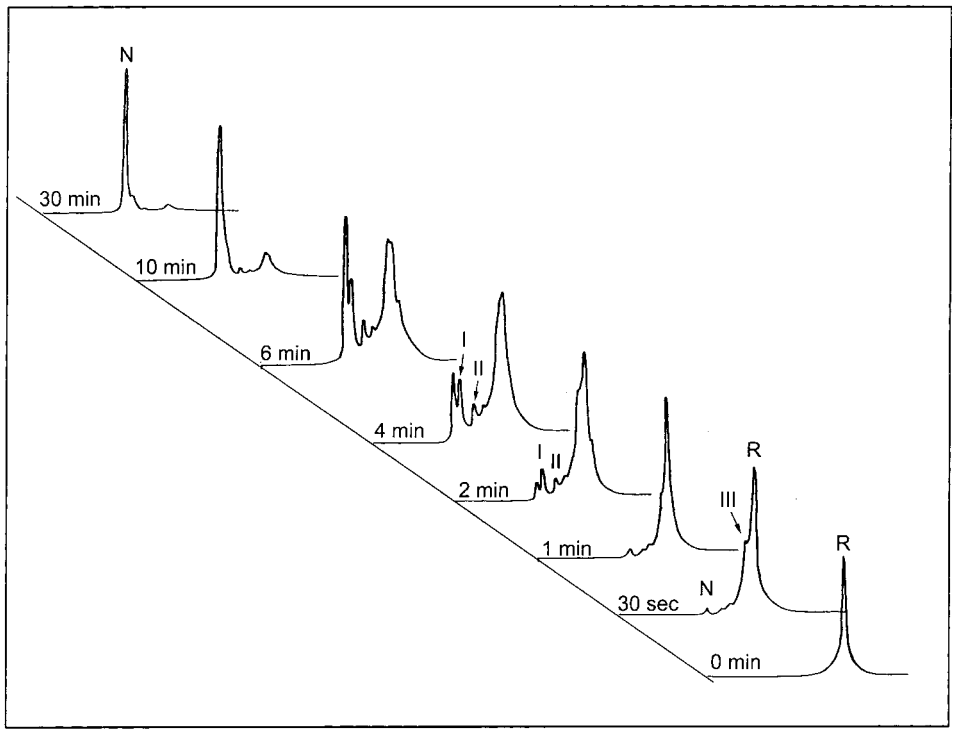

FIGURE 2: Temporal distribution of the intermediates trapped by IAA at different time points during the refolding of PIP in vitro. N, native PIP; R, IAA-modified fully reduced/denatured PIP. Peak I, 2SSPIPa; peak II, 2SSPIPb; peak III, 1SSPIP.

protein content with the absorption constant A_{276}(1 cm, 1.0 mg/mL) = 1.0. Molecular masses of samples were determined using ESI-MS.

Circular Dichroism Studies. The samples were dissolved in 10 mM HCl. The protein concentration was determined by the UV absorption at 276 nm and by the Lowry method. Circular dichroism measurements were performed on a Jasco-700 circular dichroism spectropolarimeter. The spectra were recorded at room temperature and scanned from 200 to 250 nm twice at a stepwise gradient of 0.1 nm. The cell path length was 0.1 cm, and the data were expressed as molar ellipticity.

RESULTS

Refolding of PIP in the Presence of L-*Arg.* The refolding reaction as described under Materials and Methods was allowed to carry on for 16 h; then a 100 μL sample was taken out, modified with iodoacetic acid (IAA), and run on a denatured pH 8.3 PAGE. After staining with Coomassie Brilliant Blue, the band was quantified by densitometry, which showed that about 85% frdPIP was refolded into the native PIP.

By changing the pH of the refolding solution to 8.0, 8.5, 9.0, 9.5, or 10.0, the recovery of native PIP became respectively 19%, 32%, 64%, 85%, and 84%, suggesting that higher pH favored refolding, possibly due to the more active cystine. However, in the absence of redox system, the refolding recovery could still reach 75%. The removal of L-Arg from the refolding solution lowered the refolding recovery drastically to only 15%. The results suggest the importance of L-Arg in the refolding of PIP and have been observed with many other proteins such as antibody Fab fragment (*26*), single-chain immunotoxins (*27*), lipase (*28*), and nicotinic acetylcholine receptor α-subunit (*29*).

Temporal Distribution of Modified Refolding Intermediates. After the refolding process was initiated by adding frdPIP, the reaction was stopped at different time points with IAA. Then the refolding mixture was adjusted to pH 1.0 by TFA and loaded onto a C8 rpHPLC column. The temporal distributions of modified refolding species are shown in Figure 2. In addition to the modified frdPIP (R) and the native PIP (N), three obvious intermediates, designated as peaks I, II, and III, can be observed. In theory, modified intermediates with two disulfide linkages should have a mass increase of 118 Da over PIP, and those modified with one disulfide should have a mass increase of 236 Da. Judged by ESI-MS, both peak II and peak I have a molecular mass of 6078 Da, and that of peak III is 6195 Da. In contrast to the 5959 Da mass of PIP, peaks I and II should be the intermediates with two disulfide linkages formed while peak III is the intermediate with one disulfide bond; accordingly, the intermediates of peaks I, II, and III are designated 2SSPIPa, 2SSPIPb, and 1SSPIP, respectively.

The HPLC profiles of the distribution of IAA-trapped intermediates during the refolding process (Figure 2) showed that in the presence of L-Arg and redox buffer, frdPIP was first converted into the intermediate with one disulfide bond in less than 30 s. The intermediates containing two disulfide bonds (2SSPIPa and 2SSPIPb) began to appear after 2 min of reaction, and native PIP accumulated to a visible fraction at 30 s eventually took up more than 90% of the total species at 10 min. Within 30 min, almost all the frdPIP was converted into native PIP.

Separation and Purification of the Intermediates. The 1SSPIP is difficult to separate from IAA-modified frdPIP (R) on HPLC, so the IAA-trapped refolding mixture at 30 s was first applied on a Mono-Q Sepharose column and separated by their charge difference. Then the 1SSPIP was purified by C8-rpHPLC. 2SSPIPa and 2SSPIPb were directly

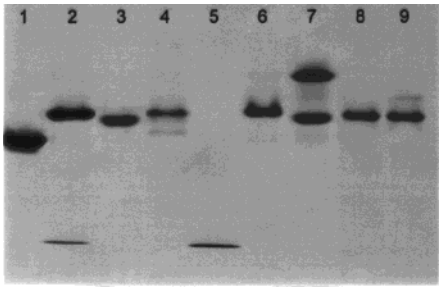

FIGURE 3: pH 8.3 native PAGE of the intermediates and the Lys-C-digested products. Lane 1: fully reduced/denatured PIP (frdPIP) modified with IAA. Lane 2: frdPIP digested with Lys-C. Lane 3: Purified 1SSPIP. Lane 4: 1SSPIP digested with Lys-C. Lane 5: S-Sulfonated insulin A chain. Lane 6: S-Sulfonated insulin B chain. Lane 7: PIP (upper) and insulin (lower) markers. Lane 8: Purified 2SSPIPa. Lane 9: Purified 2SSPIPb.

purified from the refolding mixture by HPLC. The purity of the three separated and purified intermediates was ascertained by HPLC (Figure 3) and native pH 8.3 electrophoresis (Figure 4). Each of the three intermediates gave a sharp peak on C8-HPLC at the retention time of 18.9, 20.3, and 22.7 min, respectively, as shown in Figure 3. All three intermediates have almost the same mobility rate as that of insulin (Figure 4) with 1SSPIP just a little faster than 2SSPIPa and 2SSPIPb.

Structure Elucidation of Captured Intermediate. PIP gave one single band at the position of insulin after the digestion with Lys-C (Figure 4), which cleaves at the C-terminal of the lysine residue. IAA-modified frdPIP showed two bands after digestion, with the slower moving one at the position of the S-sulfonated insulin B chain and the faster one barely visible by Coomassie Brilliant Blue staining, at the same position of the S-sulfonated insulin A chain. Purified 1SSPIP gave only one band at the same position of the S-sulfonated insulin B chain after digestion (Figure 4), suggesting that the disulfide bond formed was within the A chain. Because the disulfide bond exists between the A and B chains, the intermediate 1SSPIP digested with Lys-C should have only one product with one positive charge (lysine residue) removed from the molecule, and consequently moves faster than 1SSPIP on pH 8.3 PAGE. But in fact, 1SSPIP digested with Lys-C generates one band slower than itself, and this band is at the same position of the sulfonated insulin B chain. So it is sure that this disulfide bridge exists within the A chain. To further elucidate the position of the disulfide bond formed among the four cysteines within the A chain, 1SSPIP digested with V8 was analyzed on C8-rpHPLC. Five eluted peaks (Id, IId, IIId, IVd, and Vd) were obtained (Figure 5). The pooled fractions from peak Vd had a molecular mass of 1547 Da, indicating that disulfide linkage must exist in fragment D, possibly at A6−A11, A7−A11, or A6−A7.

To determine the disulfide linkage structure of 2SSPIPa and 2SSPIPb, we selected endoproteinase V8 to digest the two intermediates. There are four potential cleavage sites for V8 on the backbone of PIP. The frdPIP treated with V8 should generate five fragments, A, B, C, D, and E as shown in Figure 1, while V8-treated PIP should generate only three fragments, C, A+D, and B+E, because of the presence of inter-AB disulfide links. The results of digestion of PIP, 1SSPIP, 2SSPIPa, and 2SSPIPb by V8 were analyzed on C8-rpHPLC (Figure 5). The V8-digested PIP mixture gave three peaks, Ia, IIa, and IIIa, corresponding to fragments of

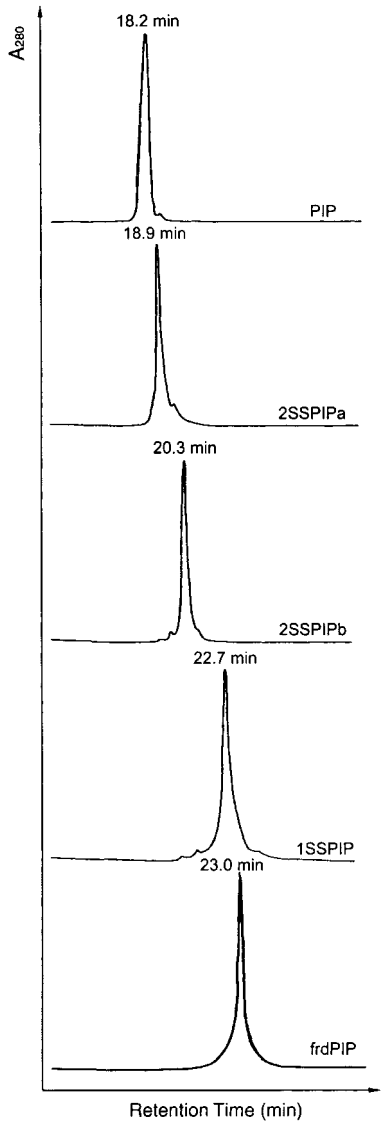

FIGURE 4: HPLC profiles of the three purified intermediates, frdPIP, and native PIP. Retention times of each species are shown on the tip of the peak. The linear elution gradient used was 40−80% solvent B during 30 min.

A+D, C, and B+E, respectively, as judged by the molecular mass of each peak (Table 1). The five peaks of V8-digested 1SSPIP, Id, IId, IIId, IVd, and Vd, corresponded to IAA-modified fragment E, A, B, C, and D, respectively, as judged by their molecular mass. The digestion of 2SSPIPa and 2SSPIPb showed similar results, with peaks I and II corresponding to peaks Ia and IIa, respectively, but peaks IIIb and IIIc are different from peak IIIa at a molecular mass increase of 118 Da (Table 1). The results of digestion suggested the presence of a disulfide bond between fragments A and D, with the other two cysteine residues carboxymethylated. Thus, for 2SSPIPa and 2SSPIPb, one disulfide bridge (A20−B19) must exist between fragments B and E, while the other disulfide bond exists between fragments A and D, possible A6−B7, A7−B7, or A11−B7, but not within the A chain.

CD Analysis of 2SSPIPa, 2SSPIPb, and 1SSPIP. Far-UV CD spectra in 10 mM HCl are shown in Figure 6. As compared with native PIP and frdPIP, the three intermediates of 2SSPIPa, 2SSPIPb, and 1SSPIP retain 69%, 50%, and 22% of the native helix content, respectively.

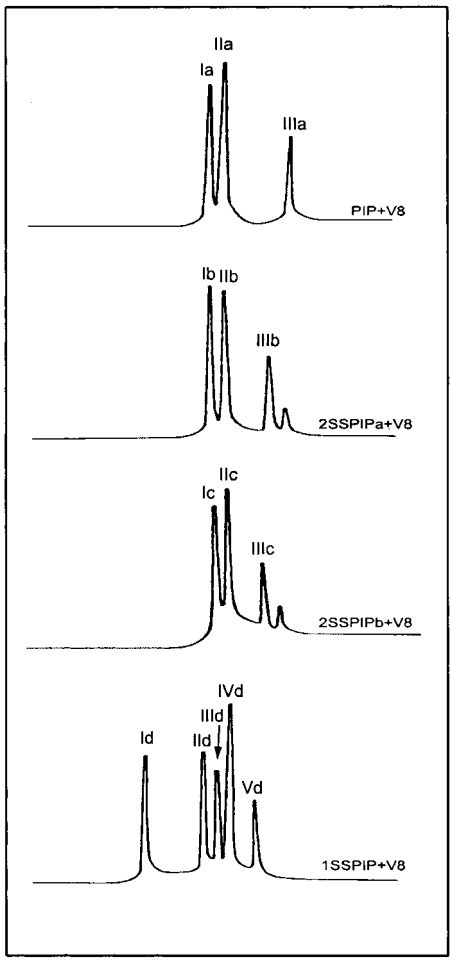

FIGURE 5: HPLC profiles of the V8-digested PIP and three intermediates.

Table 1: Molecular Mass of V8 Digestion Fragments of the Intermediates and PIP Determined by MS-ESI[a]

protein digested with V8	HPLC peaks in Figure 4	molecular mass	corresponding fragments in Figure 1
PIP	Ia	1377.0	B+E
	IIa	1684.0	C
	IIIa	2969.0	A+D
2SSPIPa	Ib	1377.0	B+E
	IIb	1684.0	C
	IIIb	3087.0	$A+D_s$
2SSPIPb	Ic	1377.0	B+E
	IIc	1684.0	C
	IIIc	3087.0	$A+D_s$
1SSPIP	Id	570.0	E_s
	IId	1540.0	A_s
	IIId	925.0	B_s
	IVd	1684.0	C
	Vd	1547.0	D_{intra}

[a] $A+D_s$ represents S-carboxymethylated fragment D linked with fragment A by an inter-AB disulfide bond. D_{intra} represents S-carboxymethylated fragment D with an intrachain disulfide bond. E_s, A_s, and B_s represent the corresponding S-carboxymethylated fragments.

DISCUSSION

Despite the report that alkylation of thiols with IAA occurred at the same time scale of thiol−disulfide rearrangement, so that the quenching method with IAA may not fully preclude the rearrangement of intermediates during the trapping procedure (*8*, *30*), we decided to use this irreversible quenching method in our study for the convenience in further structure characterization and elucidation after careful control trials. Because the alkylation of free thiols occurs on the same time scale as that of thiol−disulfide rearrangement, there should be a competition for free thiols between IAA and intramolecular disulfide bonds, and consequently the rearranged intermediates should increase with the decrease of IAA concentration. We found, however, there was no change in the fractions of the three captured intermediates when the concentration of IAA used in trapping free thiols varied from 10 mM to 200 mM, thus indicating the quenching method with IAA suitable for our work.

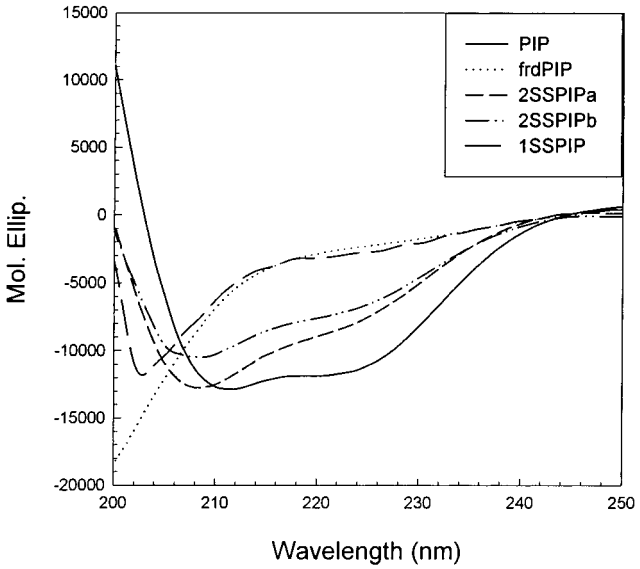

FIGURE 6: CD spectroscopic analysis of the secondary structures of frdPIP, PIP, 1SSPIP, 2SSPIPa, and 2SSPIPb.

Of the three intermediates captured, 1SSPIP contains the disulfide bond within the A chain of PIP. Based on the previous conclusion that the A6−A11 disulfide linkage formed first in the refolding of insulin precursor (*21*), we can deduce that this intra-A disulfide bond is most probably A6−A11. With the other two intermediates which contain two disulfide linkages, one must be A20−B19; the other one exists between CysB7 and one of the three cysteines in fragment D. Thus, the disulfide linkage may either be B7−A6, B7−A7, or B7−A11, among which only B7−A7 is the native one. It is certain that there is at least one non-native intermediate among 2SSPIPa and 2SSPIPb. We can observe that the behavior of 2SSPIPa on HPLC is closer to native PIP than that of 2SSPIPb (Figure 2), and 2SSPIPa retains up to 69% of the native helix structure of native PIP (Figure 6), while the amino-terminal helix of the insulin A chain takes up 30% of the native α-helix content (*31*, *32*). This led us to deduce that 2SSPIPa may be [A7−B7, A20−B19] 2SSPIP, while 2SSPIPb may be [A6−B7, A20−B19] 2SSPIP or [A11−B7, A20−B19] 2SSPIP, both containing a non-native disulfide bond.

Disulfide rearrangement is a general phenomenon in the disulfide-forming pathway of proteins such as BPTI (*33*), because it is the most rapid way of inserting a disulfide linkage into a protein when the result is a folded conformation in which that disulfide bond will become buried. Non-native disulfide bridges sometimes play an important role in the disulfide rearrangement. For example, the partly folded

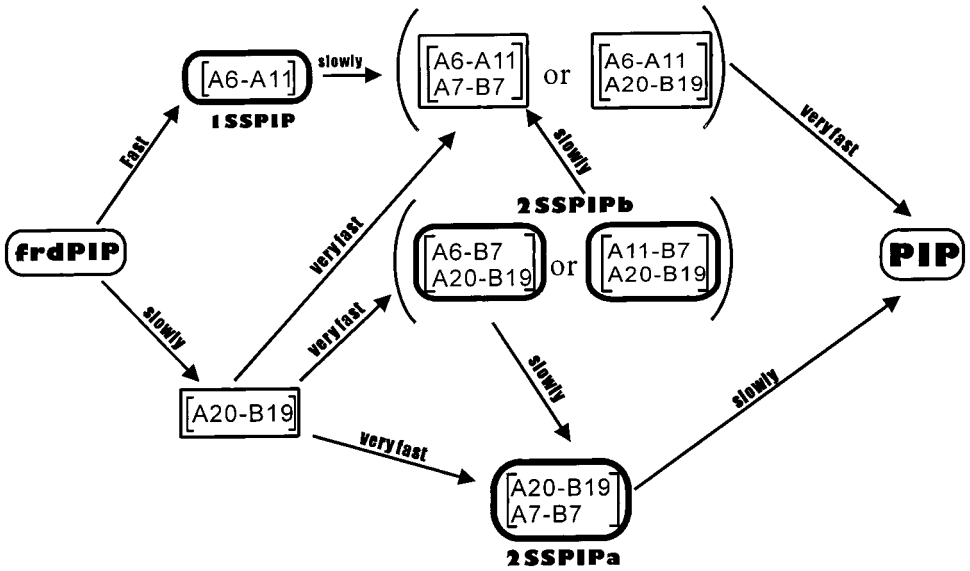

FIGURE 7: Schematic of the putative disulfide-forming pathway of PIP. Intermediates are represented by their disulfide linkage pattern. Intermediates with an overstriking rectangle represent the intermediates captured in this study.

non-native two-disulfide intermediates [30−51, 5−14] and [30−51, 5−38] of BPTI (*34*−*36*) have been shown to be essential to overcome the kinetic block encountered in forming the two native disulfide bridges 30−51 and 5−55 at the same time. The same phenomena have also been observed in the refolding of IGF-I (*20*).

It seems unlikely that both two-disulfide bonds species are off-pathway intermediates that will be cleared away from the subsequent refolding process, because the total fraction of 2SSPIPa and 2SSPIPb takes up more than 15% at 6 min in Figure 2, while the recovery of native PIP in the end was 85% under the same conditions. Then, what is the role of the two two-disulfide species in the refolding of PIP? As discussed above, 2SSPIPa and 2SSPIPb may be the consequence of disulfide rearrangement in the disulfide-forming pathway as that of BPTI or IGF-I. As the intra-A disulfide in 1SSPIP does not exist in the 2SSPIPa and 2SSPIPb, it is unlikely that all three intermediates are involved in the same folding pathway. So, we deduce that there are at least two folding pathways in the refolding of PIP in vitro (Figure 7). The first one is the direct folding pathway involving first formation of the intra-A disulfide bond. The second is the selective folding pathway involving disulfide rearrangement and begining with the disulfide bond A20−B19 formed first. In the former, the intra-A disulfide bond formed immediately after initiation of the refolding process; then the refolding intermediate 1SSPIP awaits for the formation of another inter-AB disulfide bond. Once one of the inter-AB disulfide bonds has formed, the rest of the inter-AB disulfide bond will immediately form, leading to the completed refolding of PIP. In the selective folding pathway, the disulfide bond A20−B19 formed first at a lower speed than that of the intra-A disulfide bond. Thereafter another inter-AB disulfide bond, possibly a non-native one, formed almost at the same time; then the second inter-AB disulfide bond slowly rearranged between CysA6, CysA7, CysA11, and CysB7 until the formation of a native two-disulfide intermediate, possibly [A20−B19, A7−B7] 2SSPIP or [A20−B19, A6−A11] 2SSPIP. Intermediate [A20−B19, A6−A11] 2SSPIP refolded into native PIP immediately following the formation of disulfide A7−B7, while the remaining CysA6 and CysA11 in [A20−B19, A7−B7] 2SSPIP pair slowly before the refolding of PIP is finished.

Interestingly, compared with the intermediates reported in the study of IGF-I and its analogues (*20*, *37*, *38*), it can be seen that 2SSPIPa corresponds to the intermediate [18−61, 6−48] of IGF-I and 2SSPIPb corresponds to the intermediate [18−61, 6−52] or [18−61, 6−47] of IGF-I. So, the above proposed selective folding pathway of PIP is very similar to the oxidative folding pathway of IGF-I. This is reasonable if we consider the high homology of amino acid sequence and tertiary structure between insulin and IGF-I (*25*). The only difference is that the non-native intermediate 2SSPIPb was refolded into native PIP, not to a PIP isomer as IGF-swap.

The present work showed that the A6−A11 disulfide bond formed first in the direct folding pathway of PIP, while the corresponding disulfide bond 47−52 in IGF-I is not formed first. This can be explained by comparison of the structural role of disulfide bonds 47−52 in IGF-I and A6−A11 in insulin. In the native state, disulfide bond 47−52 anchors an adjoining amphipathic α-helix (helix 2, residues 42−49, corresponding to the N-terminal helix of insulin A chain) against the hydrophobic core. Folding behavior and spectroscopic studies of partially reduced IGF-I [18−61, 6−48] and the corresponding engineered alanine model [Cys47A, Cys52A] IGF-I confirmed that disulfide bond 47−52 exists in the native molecule in a strained configuration and unstable (*39*−*41*), while the corresponding intra-A disulfide bonds in insulin have been demonstrated to be more difficult to reduce than the other two interchain disulfide bonds and play an important role in maintaining the tertiary structure of insulin (*42*). Therefore, it is not strange that disulfide bond 47−52 in IGF-I formed last while the corresponding intra-A disulfide bond of PIP can form first in the folding pathway.

The thermodynamically stable disulfide isomer of IGF-I contains the disulfide linkage pattern [6−47, 48−52, 18−61]. The corresponding isomer of insulin and proinsulin has not been identified during insulin-chain recombination or proinsulin folding, but has been generated from native insulin and proinsulin by disulfide reassortment under partially denaturing conditions (*43*). This metastable disulfide isomer

retains ordered secondary structure and a compact hydrophobic core. The existence of this insulin isomer suggests that insulin and IGF-I have similar folding information determined by the homology of the amino acid sequence, although the former is more specific than the latter. Therefore, the selective folding pathway of PIP, although not a dominant one, most probably exists because the similar folding behavior exists in IGF-I.

Except for the three intermediates captured during the refolding process, all the other intermediates appearing in the putative disulfide-forming pathway shown in Figure 7 did not accumulate to become visible fractions. The failure to capture such novel disulfide species may be due to one of the following three reasons. First, it is possible that such species are really present in the folding reactions but were aggregated or unresolved by HPLC for technical limitation. Second, it is possible that such species never formed during the folding for either kinetic or thermodynamic reasons. Finally, it is possible that such species did indeed form but were too short-lived to be detected. Intermediates [A20−B19], [A20−B19, A6−A11], and [A7−B7, A6−A11] are absent most possibly for such reasons. The escaping capture of [A20−B19, A6−A11] 2SSPIP can also be explained from its inferred structure. Deduced from the crystal structure of insulin and single-chain insulin (*31*, *32*), disulfide bond A6−A11 is totally buried and disulfide bond A20−B19 is substantially buried in the hydrophobic core, while A7−B7 is accessible at the surface of the molecule. So, with the remaining CysA7 and CysB7 exposed to the solvent and accessible to oxidized agents, [A20−B19, A6−A11] 2SSPIP was rapidly refolded into native PIP so as not to be captured. However, with the remaining free thiols CysA6 and CysA11 buried in the hydrophobic core, 2SSPIPa can only be slowly oxidized to native PIP and consequently accumulated to be captured.

ACKNOWLEDGMENT

We thank an unknown enthusiastic reviewer for helpful suggestions and valuable discussion on the manuscript.

REFERENCES

1. Creighton, T. E. (1985) *J. Phys. Chem. 89*, 2452−2459.
2. Kin, J. (1986) *Bio/Technology 4*, 297−303.
3. Creighton, T. E., and Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497−526.
4. Creighton, T. E., Darby, N. J., and Kemmink, J. (1996) *FASEB J. 10*, 110−118.
5. Bulaj, G., and Goldenberg, D. P. (1999) *Protein Sci. 8*, 1825−1842.
6. Darby, N. J., Morin, P. E., Talbo, G., and Creighton, T. E. (1995) *J. Mol. Biol. 249*, 463−477.
7. Weissman, J. S., and Kim, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9900−9904.
8. Weissman, J. S., and Kim, P. S. (1991) *Science 253*, 1386−1393.
9. Creighton, T. E., and Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497−526.
10. Wedemeyer, W. J., Welker, E., Narayan, M., and Scheraga, H. A. (2000) *Biochemistry 39*, 4207−4216.
11. Xu, X., Rothwarf, D. M., and Scheraga, H. A. (1996) *Biochemistry 35*, 6406−6417.
12. Rothwarf, D. M., and Scheraga, H. A. (1993) *Biochemistry 32*, 2671−2679.
13. Chang, J. Y., Schindler, P., Ramseier, U., and Lai, P. H. (1995) *J. Biol. Chem. 270*, 9207−9216.
14. Wu, J., Yang, Y., and Watson, J. T. (1998) *Protein Sci. 7*, 1017−1028.
15. Nagata, K., Hatanaka, H., Kohda, D., Kataoka, H., Nagasawa, H., Isogai, A., Ishizaki, H., Suzuki, A., and Inagaki, F. (1995) *J. Mol. Biol. 253*, 749−758.
16. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. (1992) *Biochemistry 31*, 1749−1756.
17. Miller, J. A., Narhi, L. O., Hua, Q. X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M. A. (1993) *Biochemistry 32*, 5203−5213.
18. Rosenfeld, R. D., Miller, J. A., Narhi, L. O., Hawkins, N., Katta, V., Lauren, S., Weiss, M. A., and Arakawa, T. (1997) *Arch. Biochem. Biophys. 342*, 298−305.
19. Milner, S. J., Carver, J. A., Ballard, F. J., and Francis, G. L. (1999) *Biotechnol. Bioeng. 62*, 693−703.
20. Yang, Y., Wu, J., and Watson, J. T. (1999) *J. Biol. Chem. 274*, 37598−37604.
21. Yuan, Y., Wang, Z. H., and Tang, J. G. (1999) *Biochem. J. 343, Pt. 1*, 139−144.
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) *Sci. China, Ser. C: Life Sci. 39*, 225−233.
23. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., and Markussen, J. (1991) *J. Mol. Biol. 220*, 425−433.
24. Blundell, T. L., Bedarkar, S., Rinderknecht, E., and Humbel, R. E. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 180−184.
25. Hua, Q. X., Shoelson, S. E., Kochoyan, M., and Weiss, M. A. (1991) *Nature 354*, 238−241.
26. Buchner, J., and Rudolph, R. (1991) *Bio/Technology 9*, 157−162.
27. Brinkmann, U., Buchner, J., and Pastan, I. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3075−3079.
28. Ahn, J. H., Lee, Y. P., and Rhee, J. S. (1997) *J. Biotechnol. 54*, 151−160.
29. Schrattenholz, A., Pfeiffer, S., Pejovic, V., Rudolph, R., Godovac-Zimmermann, J., and Maelicke, A. (1998) *J. Biol. Chem. 273*, 32393−32399.
30. Chang, J. Y. (1996) *Biochemistry 35*, 11702−11709.
31. The Peking Insulin Structure Research Group (1974) *Sci. Sin. 17*, 752−778.
32. 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) *Philos. Trans. R. Soc. London, Ser. B: Biol. Sci. 319*, 369−456.
33. Creighton, T. E. (1997) *J. Biol. Chem. 378*, 731−744.
34. Van Mierlo, C. P., Kemmink, J., Neuhaus, D., Darby, N. J., and Creighton, T. E. (1994) *J. Mol. Biol. 235*, 1044−1061.
35. Weissman, J. S., and Kim, P. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9900−9904.
36. Darby, N. J., Morin, P. E., Talbo, G., and Creighton, T. E. (1995) *J. Mol. Biol. 249*, 463−477.
37. Milner, S. J., Francis, G. L., Wallace, J. C., Magee, B. A., and Ballard, F. J. (1995) *Biochem. J. 308 (Pt. 3)*, 865−871.
38. Hober, S., Hansson, A., Uhlen, M., and Nilsson, B. (1994) *Biochemistry 33*, 6758−6761.
39. 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.
40. Narhi, L. O., Hua, Q. X., Arakawa, T., Fox, G. M., Tsai, L., Rosenfeld, R., Holst, P., Miller, J. A., and Weiss, M. A. (1993) *Biochemistry 32*, 5214−5221.
41. Hober, S., Uhlen, M., and Nilsson, B. (1997) *Biochemistry 36*, 4616−4622.
42. Cecil, R., and Weitzman, P. D. (1964) *Biochem. J. 93*, 1−11.
43. 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.

BI001613R