A peptide model of insulin folding intermediate with one disulfide

HAN YAN,^{1,3} ZHAN-YUN GUO,^{2,3} XIAO-WEN GONG,² DAN XI,¹ AND YOU-MIN FENG²

¹Department of Bioengineering, Xi'an Jiaotong University, Xi'an 710049, China

²Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, The Chinese Academy of Sciences, Shanghai 200031, China

(RECEIVED October 18, 2002; FINAL REVISION January 2, 2003; ACCEPTED January 8, 2003)

Abstract

Insulin folds into a unique three-dimensional structure stabilized by three disulfide bonds. Our previous work suggested that during in vitro refolding of a recombinant single-chain insulin (PIP) there exists a critical folding intermediate containing the single disulfide A20–B19. However, the intermediate cannot be trapped during refolding because once this disulfide is formed, the remaining folding process is very quick. To circumvent this difficulty, a model peptide ([A20–B19]PIP) containing the single disulfide A20–B19 was prepared by protein engineering. The model peptide can be secreted from transformed yeast cells, but its secretion yield decreases 2–3 magnitudes compared with that of the wild-type PIP. The physicochemical property analysis suggested that the model peptide adopts a partially folded conformation. In vitro, the fully reduced model peptide can quickly and efficiently form the disulfide A20–B19, which suggested that formation of the disulfide A20–B19 is kinetically preferred. In redox buffer, the model peptide is reduced gradually as the reduction potential is increased, while the disulfides of the wild-type PIP are reduced in a cooperative manner. By analysis of the model peptide, it is possible to deduce the properties of the critical folding intermediate with the single disulfide A20–B19.

Keywords: Insulin; folding; disulfide; kinetics; thermodynamics

Since Anfinsen and coworkers first demonstrated that the three-dimensional structure of a globular protein is uniquely determined by its amino-acid sequence in the 1960s (Anfinsen 1973), a lot of advances have been made in the understanding of protein folding through experimental and

theoretical approaches. 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 (Weissman and Kim 1991; Creighton et al. 1996; Wu et al. 1998; Wedemeyer et al. 2000). Therefore, investigating the properties of kinetic intermediates is quite helpful and necessary to deeply understand the protein-folding mechanism.

Insulin is a structurally and functionally well-characterized, small globular protein containing A- and B-chains linked by three disulfides (one intrachain bond, A6–A11; two interchain bonds, A7–B7 and A20–B19). Its three-dimensional structure has been well defined by X-ray crystallography (The Peking Insulin Structure Research Group 1974; Baker et al. 1988) and NMR (Roy et al. 1990; Weiss et al. 1991). Although the separate A- and B-chains of insulin can be recombined successfully in vitro (Wang and Tsou 1991), a single-chain polypeptide (preproinsulin) was synthesized in vivo. When B29Lys and A1Gly were linked

Reprint requests to: You-Min Feng, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, The Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China; e-mail: fengym@sunm.shnc.ac.cn; fax: (86) 021-64338357.

³These authors contributed equally to this work.

Abbreviations: PIP, a recombinant single-chain insulin in which the C terminus of porcine insulin B-chain and the N terminus of porcine insulin A-chain were linked together by a dipeptide, Ala-Lys; IGF-1, insulin-like growth factor 1; BPTI, bovine pancreatic trypsin inhibitor; RNaseA, 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; UV, ultraviolet; CD, circular dichroism; NMR, nuclear magnetic resonance.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.0237203>.

together by a peptide bond directly, the mini-proinsulin still retained the three-dimensional structure identical to that of insulin (Derewenda et al. 1991; Hua et al. 1998). Our laboratory has constructed a single-chain insulin (PIP) that can fold correctly and can be secreted efficiently from transformed yeast cells (Zhang et al. 1996). It can be reasonably presumed that the three-dimensional structure of PIP is identical or very similar with that of insulin/mini-proinsulin.

Insulin folds into a stable three-dimensional structure mainly composed of three α -helical segments (A2–A8, A13–A19, and B9–B19) stabilized by its three disulfides. Deletion of the disulfide A6–A11 leads to the unfolding of the α -helix in the N terminus of A-chain (Hua et al. 1996a; Weiss et al. 2000). Removal of the disulfide A7–B7 causes more serious unfolding: Besides the α -helix in the N terminus of A-chain, a part of the α -helix in the C terminus of A-chain is also unfolded (Hua et al. 2001). When the disulfide A20–B19 was deleted, the PIP mutant cannot be secreted at all from the transformed yeast cells while the other two mutants with disulfide A6–A11 or A7–B7 deleted can be secreted but with decreased secretion yield (Guo and Feng 2001), which suggested that the disulfide A20–B19 is probably more important than the other two disulfides.

IGF-1 is an insulin-like, 70-residue, single-chain protein composed of B-, C-, A-, and D-domains (Humbel 1990). 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. IGF-1 adopts an insulin-like structure also stabilized by three disulfides (47–52, 6–48, and 18–61) corresponding to those of insulin (A6–A11, A7–B7, and A20–B19) (Cooke et al. 1991). Despite the sequence homology, insulin and IGF-1 have different folding properties: Insulin/PIP folds into a unique structure, while IGF-1 folds into two thermodynamically stable isomers characterized by different disulfide linkages (Hober et al. 1992, 1999; Miller et al. 1993). The unusual folding behavior of IGF-1 is probably a result of the high energetic state of the intra-A-domain disulfide (Hober et al. 1992, 1997, 1999) and can be overcome by its binding proteins (Hober et al. 1994). Just like insulin, the intact structure of IGF-1 also depends on its three disulfides. When disulfide 47–52 (corresponding to the disulfide A6–A11 of insulin) is reduced/deleted, the α -helix in the N terminus of A-domain is unfolded (Hua et al. 1996b). Deletion of the disulfide 6–48 (corresponding to the disulfide A7–B7 of insulin) causes a little more serious result on the structure (Narhi et al. 1993; Hober et al. 1997). When the disulfide 18–61 (corresponding to the disulfide A20–B19 of insulin) is deleted, the mutant IGF-1 cannot be expressed, but the IGF-1 mutant with the single disulfide 18–61 acquires a compact partially folded conformation (Narhi et al. 1993). Moreover, during *in vitro* refolding, the disulfide 18–61 is formed first (Rosenfeld et al. 1997; Milner et

al. 1999; Yang et al. 1999). These results suggest that first, the disulfide 18–61 is the most critical bond in maintaining the native structure of IGF-1; second, formation of the disulfide 18–61 is a kinetically preferred process that coupled with the formation of a partially folded conformation.

In vitro PIP can spontaneously fold into the native structure with correct pairing of its three disulfides (Qiao et al. 2001). During the refolding process, both the disulfides A6–A11 and A20–B19 form quickly. The trapped one-disulfide intermediate with disulfide A6–A11 lacks ordered structure (Qiao et al. 2001), moreover, the PIP analog containing the single disulfide A6–A11 cannot be secreted at all from transformed yeast cells (Guo and Feng 2001). These results suggested that the quick formation of the disulfide A6–A11 probably results from the proximity of the two cysteine residues in sequence, and formation of this disulfide doesn't couple with formation of a partially folded structure. However, the quick formation of the disulfide A20–B19 must result from the ordered structure formed in the early stage because the two cysteine residues are far in sequence. It is the ordered structure that takes the two cysteine residues into proximity in space and leads to its quick formation. In turn, this disulfide bond stabilizes the partially folded conformation and makes the later folding process possible. Therefore, the role of the disulfide A20–B19 during refolding is just like a template that ushers the later refolding process.

The previous result suggested that the disulfide A20–B19 is critical during the refolding of PIP, however, the folding intermediate with the single disulfide A20–B19 cannot be trapped because once this disulfide is formed, the remaining folding process is very quick. To circumvent this difficulty, a model peptide with the single disulfide A20–B19 was prepared by protein engineering. In the model peptide, the two cysteine residues of disulfide A6–A11 were replaced by Ala residues, because this disulfide bond is buried in the hydrophobic core of insulin; while the two cysteine residues of disulfide A7–B7 were substituted with Ser residues, because this disulfide bond is accessible to the solvent. Here, we purified the model peptide and investigated its physicochemical properties (including mobility rate on native PAGE, retention time on reversed-phase HPLC, and CD spectra), disulfide stability in redox buffer, *in vitro* refolding, and biological activity after conversion to the double-chain form. By analysis of the model peptide, we can reasonably deduce the properties of the critical folding intermediate with the single-disulfide A20–B19.

Results

DNA manipulation

The expression vector of the model peptide was constructed by site-directed mutagenesis on the basis of the expression vector pVT102-U/ α MFL-[A7Ser, B7Ser]PIP (Guo and

Feng 2001). The presence of the expected mutations was confirmed by DNA sequencing (data not shown). The expression vector of the model peptide was designated as pVT102-U/ α MFL-[A6Ala, A11Ala, A7Ser, B7Ser]PIP.

Expression and purification of the model peptide

The expression vector of the model peptide was transformed into yeast cells and then cultured in a 16-L fermenter. After purification, the model peptide is homogenous as judged by analytical C4 reversed-phase HPLC and native pH 8.3 PAGE (data not shown). Its molecular mass measured by MALDI-TOF mass spectrometry is 5868.7, consistent with the expected value that is 5866.7. We found that the secretion yield of the model peptide decreased significantly: Only about 400 μ g of purified model peptide was obtained from 8 L of fermentation supernatant, while usually >100 mg of purified wild-type PIP was obtained from the same volume of fermentation supernatant. So the secretion yield of the model peptide decreased 2–3 magnitudes compared with that of the wild-type PIP. The secretion yield of the model peptide was probably decreased for two reasons. The first relates to the folding thermodynamics of the model peptide. If its conformation were significantly different from that of its parent molecule, that is, more hydrophobic residues buried in the wild-type PIP were exposed in the model peptide, the model peptide would be trapped and finally degraded in the endoplasmic reticulum of the yeast cells because there is a quality control system in the secretory pathway (Ellgaard et al. 1999). The second relates to the folding kinetics of the model peptide. If the folding of the model peptide were impaired by deletion of the other two disulfides, then the model peptide could not fold efficiently and quickly and would be degraded in the secretory pathway, too.

Analysis of physicochemical properties of the model peptide

As deduced above, the decrease of the secretion yield of the model peptide was probably caused by its altered conformation. Therefore, its conformation was first analyzed comparing it with that of the wild-type PIP.

First, the conformational change of the model peptide was analyzed by native PAGE (Fig. 1). Here, the mobility rates of the model peptide, wild-type PIP, and two PIP mutants with disulfide A6–A11 or A7–B7 deleted were compared. The model peptide and the PIP mutant with disulfide A7–B7 deleted had almost identical mobility rates, while both of them ran much more slowly than wild-type PIP; the mobility rate of the PIP mutant with disulfide A6–A11 deleted was between that of wild-type PIP and the model peptide. Because the four molecules had the same number of charged residues and almost identical molecular

Figure 1. Native pH 8.3 PAGE analysis of the model peptide. 1–4 represents wild-type PIP, the model peptide, [A7Ser, B7Ser]PIP, and [A6Ala, A11Ala]PIP, respectively.

mass, their different mobility rates were mainly caused by their different conformation. As the conformation became looser, the mobility rate usually decreased. Because both of the two insulin mutants with disulfide A6–A11 or A7–B7 deleted adopted a partially folded conformation (Hua et al. 1996a, 2001; Weiss et al. 2000), we deduced that the model peptide also adopted a partially folded conformation probably similar to that of the insulin mutant with disulfide A7–B7 deleted.

Second, the model peptide was analyzed by C4 reversed-phase HPLC (Fig. 2). Here, we need to point out that in this analysis, acetonitrile and TFA were used in the elution buffer, and so the model peptide probably had been denatured in this analysis; while the wild-type PIP is likely still in native state because it is very robust and needs about 7 M guanidine chloride to be fully denatured (H. Yan, Z.-Y. Guo, and Y.-M. Feng, unpubl.). In this analysis, the model peptide shows a longer retention time on the reversed-phase column than the wild-type PIP. This was probably caused by two reasons. First, more hydrophobic residues buried in wild-type PIP were exposed in the nondenatured state of the model peptide; second, the model peptide had denatured in the elution buffer, which caused more hydrophobic residues exposed.

Third, the conformational change of the model peptide was analyzed by CD (Fig. 3). The far-UV spectrum of the model peptide is significantly different from that of the wild-type PIP. The α -helix content estimated from the CD spectrum of the model peptide is ~9%, while that of wild-type PIP is ~48%. The significant decrease of the α -helix content of the model peptide indicated that its conformation is partially unfolded. The α -helix content of the model peptide is similar to that of the PIP analog with disulfide A7–B7 deleted (Guo and Feng 2001), while somewhat lower than that of the PIP analog with disulfide A6–A11 removed (its α -helix content is 16%; H. Yan, Z.-Y. Guo, and Y.-M. Feng, unpubl.). The CD spectra in the aromatic region of the model peptide and the wild-type PIP are also significantly

Figure 2. C4 reversed-phase HPLC analysis of the model peptide. For each analysis, a 5- μ g sample was loaded onto the column and eluted by the gradient listed in Materials and Methods.

different. The microenvironment of the aromatic side-chains that is asymmetrical in the wild-type PIP has become almost completely symmetrical in the model peptide. The asymmetrical microenvironment of the aromatic side-chains depended upon the tertiary structure, so we deduced that the tertiary structure of the model peptide also changed significantly.

The above analysis suggested that the model peptide adopted a partially folded conformation probably similar to that of the insulin analog with disulfide A7–B7 deleted.

In vitro refolding of the model peptide

As deduced above, the decrease of the secretion yield of the model peptide is also probably caused by the impaired folding kinetics. If formation of the disulfide A20–B19 depended upon the formation of the other disulfides, the refolding rate of the model peptide would be impaired; the refolding yield would decrease and the refolding rate would become slow. Otherwise, the refolding would be quick and efficient. Here, the *in vitro* refolding of the reduced model peptide was carried out (Fig. 4). Under the redox potential (5 mM GSH and 1 mM GSSG) that favors for disulfide formation of the wild-type PIP (Guo et al. 2002), the reduced model peptide cannot quantitatively form its disulfide because there exists an equilibrium of the oxidized and the reduced model peptide. This implied that without the other two disulfides, the stability of disulfide A20–B19 decreased somewhat. Under a more oxidative redox potential (10 mM GSSG and 1 mM GSH), the reduced model peptide can refold quantitatively and quickly. The refolding process

lasted only ~ 3 min, and the refolding yield was $>85\%$ as calculated from the peak area. The present results suggested that the reduced model peptide can refold into its native structure quantitatively and quickly, that is, deletion of the other two disulfides didn't impair formation of the disulfide A20–B19. So the decrease of the secretion yield of the model peptide is most likely caused by the impaired folding thermodynamics but not by the impaired folding kinetics.

Disulfide stability of the model peptide in redox buffer

Under the redox potential of 5 mM GSH and 1 mM GSSG, the reduced model peptide cannot quantitatively form the disulfide A20–B19, while the disulfides of the wild-type PIP were stable under the same redox potential (Guo et al. 2002). Therefore, we deduced that the stability of the disulfide of the model peptide probably decreased somewhat compared with that of the wild-type PIP. To test this deduction, the disulfide stability of the model peptide was

Figure 3. CD analysis of the model peptide. The *top* panel shows the near-UV spectra; the *bottom* panel shows the far-UV spectra. The open circle represents the spectra of the model peptide; the filled circle represents the spectra of the wild-type PIP.

Figure 4. In vitro refolding of the model peptide analyzed by C4 reversed-phase HPLC. At the indicated time, 100 μ L refolding mixture was removed, acidified to pH 2.0 with TFA, and analyzed by C4 reversed-phase HPLC eluted with the gradient listed in Materials and Methods.

measured. As shown in Figure 5, the disulfide of the model peptide is apt to be reduced compared with that of the wild-type PIP. Moreover, the model peptide was reduced gradually as the reduction potential was increased, while the wild-type PIP was reduced in a cooperative manner.

Receptor-binding of (desB30)[A20–B19]insulin

By enzymatic cleavage, the model peptide was converted into the double-chain insulin analog, and its receptor-binding activity was measured (Fig. 6). The receptor-binding activity of (desB30)[A20–B19]insulin is too low to be quantified. Insulin binds to its receptor by the receptor-binding surface that depends upon the tertiary structure. When disulfide A6–A11 was deleted (substituted with Ala), the insulin analog retained 1.6% receptor-binding activity (H. Yan, Z.-Y. Guo, and Y.-M. Feng, unpubl.). When disulfide A7–B7 was removed, the insulin analog retained only 0.1% receptor-binding activity (Guo and Feng 2001). When both of the two disulfides were deleted, the activity of the analog further decreased.

Discussion

Based on the present analysis of the model peptide, we can reasonably deduce the property of the corresponding folding intermediate and reconstruct the folding events during the early refolding stage as sketched out in Figure 7. The fully reduced PIP is almost completely unfolded (Qiao et al. 2001); during refolding, the long-range interactions of some side-chains (mainly including the residues forming the hydrophobic core) take the polypeptide chain into a transient

Figure 5. Disulfide stability of the model peptide in redox buffer. Lanes 1–10 represent that in redox buffer the ratio of GSH to GSSG (mM/mM) was 0/0, 1/10, 2/5, 5/5, 5/1, 10/1, 20/1, 30/1, 40/1, and 50/1, respectively. When the disulfides were reduced in the redox buffer and then the free thiol groups were carboxymethylated, the molecule carried more negative charges and ran faster on the native PAGE, but the conformation also had an effect on their mobility rate. The gel was stained by Coomassie brilliant blue R250.

Figure 6. Receptor-binding analysis. The filled circle represents the plot of porcine insulin; the open circle represents the plot of (desB30)[A20-B19]insulin.

and unstable, partially folded conformation in which A20Cys and B19Cys are taken into proximity in space. The proximity results in the quick formation of the disulfide A20-B19; in turn, the disulfide A20-B19 stabilizes the partially folded conformation and makes the remaining refolding possible. We deduced that in the transient and the disulfide-stabilized, partially folded conformation, the central α -helix of B-chain probably has formed, while most parts of A-chain are probably still unfolded. Although the disulfide A6-A11 also forms quickly during PIP refolding, this is not a critical step; formation of this disulfide bond doesn't couple with formation of global, partially folded conformation (Qiao et al. 2001). When this disulfide is deleted, the reduced PIP mutant can quickly and efficiently form the remaining two disulfides (H. Yan, Z.-Y. Guo, and Y.-M. Feng, unpubl.). Therefore, during PIP refolding, the critical step in the early stage is the formation of disulfide A20-B19, but not the formation of disulfide A6-A11; it is the disulfide A20-B19 that stabilizes the partially unfolded intermediate and makes the later refolding events possible.

Together with the results of IGF-1 refolding, we can sketch out the common folding properties of insulin/PIP and IGF-1. These properties probably also exist in the refolding of other members of the insulin superfamily. Although there are different aspects of insulin/PIP and IGF-1 folding (insulin/PIP folds into a unique structure while IGF-1 folds into two disulfide isomers), the sequence homology determines that they share some common characters. First, their intact structure depends upon the three disulfide bonds, moreover, the role of each corresponding disulfide in insulin/PIP and IGF-1 is also similar; the disulfide A20-B19/18-61 is the most critical bond. Second, their refolding process shares common aspects, especially in the early stage where formation of the first disulfide 18-61/A20-B19 is a

critical step. Formation of this first disulfide and formation of a compact partially folded conformation are dependent upon each other. This early folding intermediate is the structural basis of the remaining folding processes. Therefore, formation of the disulfide 18-61/A20-B19 is a critical step during the folding of PIP and IGF-1 and even of other members of the insulin superfamily. This step converts the fully unfolded polypeptide into a partially folded intermediate.

The highly specific formation of the first disulfide during PIP and IGF-1 refolding suggested that formation of the transient folding intermediate is encoded only by the polypeptide chain itself, but not by the disulfide A20-B19. Additionally, we deduced that structure of the transient intermediate is like a molten globule state: At this folding stage, the intermediate still lacks specific side-chain packing that is characteristic of native structures. This phenomenon is similar to that observed in the α -lactalbumin: At low pH α -lactalbumin forms a molten globule state (Kuwajima 1996), but it has extensive native-like characters (Wu et al. 1995; Schulman et al. 1997); moreover, the α -lactalbumin mutant lacking all disulfides is nearly as compact as wild-type α -lactalbumin at low pH, and the stable core is formed by the segments of the polypeptide chain from both the N- and C-termini (Redfield et al. 1999). This result suggested that the overall architecture of the protein fold of α -lactalbumin is determined by the polypeptide sequence itself, and not by the result of cross-linking by disulfide bonds. During refolding of PIP and IGF-1, it is also the polypeptide sequence itself that determines the formation of the transient folding intermediate. However, the transient intermediate also has some characters different from that of α -lactalbumin: The folding intermediate of PIP and IGF-1 is a transient state that will quickly unfold; only when disulfide A20-B19/18-61 formed the intermediate was stabilized.

Materials and methods

Materials

The *Escherichia coli* strains used were DH12S and RZ1032 (*dur⁻, 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/ α MFL-[A7Ser, B7Ser]PIP was constructed previously (Guo and Feng 2001). The mutagenesis oligonucleotide primer was chemically synthesized. The chemical reagents used in the experiments were of analytical grade. The Pharmacia Biotech reversed-phase column (Sephasil Peptide C4 5 μ m ST 4.6/250), Gilson 306 HPLC system, and Gilson 115 UV detector were used. In HPLC analysis, a gradient elution was used. 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; 10 min, 60% solvent B; 36 min, 100% solvent B; 38 min, 100% solvent B; 40 min, 0% solvent B; 45 min, 0% solvent B. During analysis, the flow rate was 0.5 mL/min, and the detection wavelength was 230 nm.

Figure 7. Schematic representation of the early stage of PIP refolding.

DNA manipulation

The expression vector encoding the model peptide was constructed using a gapped duplex DNA approach for site-directed mutagenesis (Kramer et al. 1984). The plasmid pVT102-U/ α MFL-[A7Ser, B7Ser]PIP was used as a template for mutagenesis. The expected mutations were confirmed by DNA sequencing.

Expression and purification of the model peptide

The expression vector encoding the model peptide was transformed into *S. cerevisiae* XV700-6B (*Leu2, ura3, pep4*). The transformed yeast cells were cultured in a 16-L fermenter, and the model peptide was purified from the media supernatant according to previously described procedures (Zhang et al. 1996) with some modifications. First, the model 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 the ion-exchange column. Fourth, the eluted model peptide from the ion-exchange column was lyophilized then dissolved with 2–3 mL water, acidified to pH 2.0 with TFA, and centrifuged. Fifth, The pellet containing the model peptide was purified by C4 reversed-phase HPLC. The purity of the model peptide was analyzed by native pH 8.3 PAGE and analytical C4 reversed-phase HPLC.

Circular dichroism analysis

The model peptide and the wild-type PIP 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 CD spectropolarimeter. The spectra were recorded at room temperature and the protein concentration was adjusted to 0.2 mg/mL. The near-UV spectra were scanned from 320 nm to 245 nm using a cell with the path length of 1.0 cm; the far-UV spectra were scanned from 250 nm to 190 nm using a cell with the 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 stability of the model peptide in redox buffer

The model peptide was dissolved in the buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 8.7) containing different redox potential at the

final concentration of 0.15 mg/mL. In the redox buffer, the ratio (mM/mM) of GSH to GSSG was 1/10, 2/5, 5/5, 5/1, 10/1, 20/1, 30/1, 40/1, and 50/1, respectively. The total volume of each reaction was 30 μ L. At the same time, a negative control (the samples were dissolved in the buffer not containing redox potential) was carried out. The reaction was carried out at 4°C overnight. After incubation, some disulfides of proteins were reduced, then one-fifth volume of freshly prepared 0.5 M sodium iodoacetate solution was added to carboxymethylate, the free thiol groups of proteins. The carboxymethylation reaction was carried out at room temperature for 5 min. The modified mixture was then analyzed by native pH 8.3 PAGE. As a control, the disulfides stability of the wild-type PIP was analyzed at the same time.

In vitro refolding of the model peptide

The model peptide was dissolved in the buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 8.7) containing 2 M urea and 100 mM DTT. The final protein concentration was about 0.4 mg/mL. The reductive reaction was carried out at room temperature for 30 min. After reduction, the aliquot was removed, acidified to pH 2.0 with TFA, and then analyzed by C4 reverse-phase HPLC to determine whether the model peptide was fully reduced. The reduced model peptide was immediately exchanged to 10 mM HCl by gel filtration using a Sephadex G-25 column. The protein concentration of the reduced model peptide was determined by UV absorbance at 276 nm, and immediately stored at -80°C for later use. During refolding, the reduced-model peptide was added into the refolding buffer (0.1 M Tris-HCl, 1 mM EDTA, pH 9.2) containing a redox potential at the final concentration of 0.05 mg/mL. The refolding reaction was carried out at 16°C . At a different reaction time, 100 μ L refolding mixture was removed, acidified to pH 2.0 with TFA, and immediately analyzed by C4 reversed-phase HPLC.

Conversion of [A20–B19]PIP to (desB30)[A20–B19]insulin

The purified model peptide was dissolved in the reaction buffer (0.1 M NH_4HCO_3 , pH 8.5) at the final concentration of 3 mg/mL. Then Lys-C endoproteinase was added into the solution at a mass ratio of 1:500. The enzymatic cleavage was carried out at 25°C overnight. Then, (desB30)[A20–B19]insulin was purified by C4 reversed-phase HPLC.

Receptor-binding assay of (desB30)[A20–B19]insulin

The receptor-binding assay of the (desB30)[A20–B19]insulin with insulin receptor was performed using human placental membrane as previously described (Feng et al. 1982). The membrane insulin receptor (total protein is ~250 µg) was incubated with ¹²⁵I-insulin (cpm is ~10⁵) plus a selected amount of native insulin or sample in a total volume of 0.4 mL containing 50 mM Tris-HCl, 1% BSA, pH 7.5, at 4°C overnight. After incubation, the unbound ¹²⁵I-insulin was washed away with ice-cold 50 mM Tris-HCl, 0.1% BSA, pH 7.5 buffer three times and the radioactivity of precipitate was measured. The receptor-binding activity of the sample was calculated from the concentrations that caused 50% inhibition of ¹²⁵I-insulin binding to insulin receptor.

Acknowledgments

We thank the anonymous reviewers for their helpful suggestion and valuable discussion on this manuscript. This work was supported by the grants from the Chinese Academy of Sciences (KJ951-B1–606) and the National Foundation of Nature Science (No. 30170209).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Anfinsen, C.B. 1973. Principles that govern the folding of protein chains. *Science* **181**: 223–230.
- 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., et al. 1988. The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Phil. Trans. R. Soc. Lond.* **B319**: 369–456.
- Cooke, R.M., Harvey, T.S., and Campbell, I.D. 1991. Solution structure of human insulin-like growth factor 1: A nuclear magnetic resonance and restrained molecular dynamics study. *Biochemistry* **30**: 484–491.
- Creighton, T.E., Darby, N.J., and Kemmink, J. 1996. The roles of partly folded intermediates in protein folding. *FASEB J.* **10**: 110–118.
- 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.
- Ellgaard, L., Molinari, M., and Hellenius, A. 1999. Setting the standards: Quality control in the secretory pathway. *Science* **286**: 1882–1888.
- Feng, Y.-M., Zhu, J.-H., Zhang, X.-T., and Zhang, Y.-S. 1982. Studies on the mechanism of insulin action VIII. *Acta Biochimica et Biophysica Sinica*. **14**: 137–143.
- Guo, Z.-Y. and Feng, Y.-M. 2001. Effects of cysteine to serine substitutions in the two inter-chain disulfide bonds of insulin. *Biol. Chem.* **382**: 443–448.
- Guo, Z.-Y., Shen, L., and Feng, Y.-M. 2002. The different folding behavior of insulin and insulin-like growth factor I is mainly controlled by their B-chain/domain. *Biochemistry* **41**: 1556–1567.
- Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. 1992. Disulfide exchange folding of insulin-like growth factor I. *Biochemistry* **31**: 1749–1751.
- Hober, S., Hansson, A., Uhlen, M., and Nilsson, B. 1994. Folding of insulin-like growth factor is thermodynamically controlled by insulin-like growth factor binding proteins. *Biochemistry* **33**: 6758–6761.
- Hober, S., Uhlen, M., and Nilsson, B. 1997. Disulfides exchange folding of disulfide mutants of insulin-like growth factor I in vitro. *Biochemistry* **36**: 4616–4622.
- Hober, S., Ljung, J.L., Uhlen, M., and Nilsson, B. 1999. Insulin-like growth factor I and II are unable to form and maintain their native disulfides under in vivo redox conditions. *FEBS Lett.* **443**: 271–276.
- 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-I: Homologous protein sequences encoding non-homologous structures. *J. Mol. Biol.* **277**: 103–118.
- Hua, Q.-X., Nakagawa, S.H., Jia, W., Hu, S.-Q., Chu, Y.-C., Katsoyannis, P.G., and Weiss, M.A. 2001. Hierarchical protein folding: Asymmetric unfolding of an insulin analogue lacking the A7-B7 interchain disulfide bridge. *Biochemistry* **40**: 12299–12311.
- Hua, Q.-X., Narhi, L., Jia, W., Arakawa, T., Rosenfeld, R., Hawkins, N., Miller, J.A., and Weiss, M.A. 1996a. Mapping the functional surface of insulin by design: Structure and function of a novel A-chain analogue. *J. Mol. Biol.* **259**: 297–313.
- . 1996b. Native and non-native structure in a protein-folding intermediate: Spectroscopic studies of partially reduced IGF-I and an engineered alanine model. *J. Mol. Biol.* **259**: 297–313.
- Humbel, R.E. 1990. Insulin-like growth factor I and II. *Eur. J. Biochem.* **190**: 445–462.
- Kramer, W., Drusta, V., Jansen, H.W., Kramer, B., Pflugfelder, M., and Fritz, H.J. 1984. The gapped duplex DNA approach to oligonucleotide-directed mutation construction. *Nucleic Acid Res.* **12**: 9441–9454.
- Kuwajima, K. 1996. The molten globule state of α-lactalbumin. *FASEB J.* **10**: 102–109.
- Miller, J.A., Narhi, L., Hua, Q.-X., Rosenfeld, R., Arakawa, T., Rodhe, M., Prestrelski, S., Lauren, S., Stoney, K.S., Tsai, L., et al. 1993. Oxidative refolding of insulin-like growth factor 1 yields two products of similar thermodynamic stability: A bifurcating protein folding pathway. *Biochemistry* **32**: 5203–5213.
- Milner, S.J., Carver, J.A., Ballard, F.J., and Francis, G. 1999. Probing the disulfide folding pathway of insulin-like growth factor-I. *Biotech. Bioeng.* **62**: 693–703.
- 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. Role of native disulfide bonds in the structure and activity of insulin-like growth factor 1: Genetic models of protein-folding intermediates. *Biochemistry* **32**: 5214–5221.
- The Peking Insulin Structure Research Group. 1974. Studies on the insulin crystal structure: The molecule at 1.8 Å resolution. *Sci. Sin.* **17**: 752–778.
- 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.
- Redfield, C., Schulman, B.A., Milhollen, M.A., Kim, P.S., and Dobson, C.M. 1999. α-Lactalbumin forms a compact molten globule in the absence of disulfide bonds. *Nat. Struct. Biol.* **6**: 948–952.
- Rosenfeld, R.D., Miller, J.A., Narhi, L.O., Hawkins, N., Katta, V., Lauren, S.L., Weiss, M.A., and Arakawa, T. 1997. Putative folding pathway of insulin-like growth factor-I. *Arch. Biochem. Biophys.* **342**: 298–305.
- Roy, M., Lee, R.W., Brange, J., and Dunn, M.F. 1990. 1H NMR spectrum of the native human insulin monomer. Evidence for conformational differences between the monomer and aggregated forms. *J. Biol. Chem.* **265**: 5448–5453.
- Schulman, B.A., Kim, P.S., Dobson, C.M., and Redfield, C. 1997. A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat. Struct. Biol.* **4**: 630–634.
- Wang, C.C. and Tsou, C.L. 1991. The insulin A and B chains contain sufficient structural information to form the native molecule. *Trends Biochem. Sci.* **16**: 279–281.
- Wedemeyer, W.J., Welker, E., Narayan, M., and Scheraga, H.A. 2000. Disulfide bonds and protein folding. *Biochemistry* **39**: 4207–4216.
- Weiss, M.A., Hua, Q.-X., Frank, B.H., Lynch, C., and Shoelson, S.E. 1991. Heteronuclear 2D NMR studies of an engineered insulin monomer: Assignment and characterization of the receptor-binding surface by selective 2H and 13C labeling with application to protein design. *Biochemistry* **30**: 7373–7389.
- Weiss, M.A., Hua, Q.-X., Jia, W., Chu, Y.-C., Wang, R.-Y., and Katsoyannis, P.G. 2000. Hierarchical protein "un-design": Insulin's intrachain disulfide bridge tethers a recognition α-helix. *Biochemistry* **39**: 15429–15440.
- Weissman, J.S. and Kim, P.S. 1991. Reexamination of the folding of BPTI: Predominance of native intermediates. *Science* **253**: 1386–1393.
- Wu, J., Yang, Y., and Watson, J.T. 1998. Trapping of intermediates during the refolding of recombinant human epidermal growth factor (hEGF) by cyanation, and subsequent structural elucidation by mass spectrometry. *Protein Sci.* **7**: 1017–1028.
- Wu, L.C., Peng, Z.-Y., and Kim, P.S. 1995. Bipartite structure of the α-lactalbumin molten globule. *Nat. Struct. Biol.* **2**: 281–286.
- Yang, Y., Wu, J., and Watson, J.T. 1999. Probing the folding pathway of long R³ insulin-like growth factor-I (LR³IGF-I) and IGF-I via capture and identification of disulfide intermediates by cyanation methodology and mass spectrometry. *J. Biol. Chem.* **274**: 37598–37604.
- 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., et al. 1996. Secretary expression of a single-chain insulin precursor in yeast and its conversion into human insulin. *Sci. China (Series C)* **39**: 225–233.