Journal of Cellular Biochemistry 31:19–26 (1986)
Protein Structure, Folding, and Design 163–170

Refolding of Serine Proteinases

Bovine trypsinogen and chymotrypsinogen were successfully refolded as the mixed disulfide of glutathione using cysteine as the disulfide interchange catalyst. The native structures were regenerated with yields of 40%–50% at pH 8.6 and 4°C, and the half-time for the refolding was approximately 60–75 min. We then refolded threonine-neochymotrypsinogen, which is a two-chain structure held together by disulfide bonds and produced on cleavage of Tyr 146-Thr 147 in native chymotrypsinogen [Duda CT, Light A, J Biol Chem 257 9866–9871, 1982]. Neochymotrypsinogen was denatured and fully reduced, and the thiols were converted to the mixed disulfide of glutathione. The two polypeptide fragments, representing the amino- and carboxyl-terminal domains, were separated on Sephadex G-75. Mixtures of the polypeptide fragments varying in the ratio of their concentration from 1:5 to 5:1 were refolded with yields of 21–28%. The lack of dependence on the concentration of either fragment and the relatively high yields suggest independent folding of the amino- and carboxyl-terminal domains. When the globular structures of the domains formed, they then interacted with one another and produced the native intermolecular disulfide bridge and the proper geometry of the active site.

Key words: protein folding, serine proteinases, folding pathway, neochymotrypsinogen, protein

The pancreatic proteolytic enzymes trypsin, chymotrypsin, and elastase are the best studied members of the serine proteinase family of proteins. The amino acid sequences are homologous [1] and the three-dimensional structures are almost the same [2]. The proteins have two domains of approximately equal size, linked by a short stretch of polypeptide chain and by disulfide bonds [3,4]. The active site serine-195 and histidine-57 are located in the carboxyl- and amino-terminal domains, respectively, and the close proximity of the two in space is a consequence of the folding of the polypeptide chain and the interaction of the domains. In an attempt to explain the

Received April 17, 1985; revised and accepted November 13, 1985.

relationship of structure and function, we have been studying the refolding of intact bovine trypsinogen [5,6] and chymotrypsinogen [6,7] and mixtures of their domains. We developed suitable experimental conditions for refolding these proteins, and we now report our results and a proposed folding pathway.

MATERIALS AND METHODS

Materials

Bovine trypsinogen, chymotrypsinogen, trypsin and α-chymotrypsin were purchased from Worthington. Iodoacetate, cysteine, glutathione, guanidine hydrochloride, N-tosyl arginine methyl ester, N-acetyl tyrosine ethyl ester, and Tris were from Sigma. Dithioerythritol was from Aldrich. Glutathione disulfide was prepared from glutathione by oxidation with hydrogen peroxide [6].

Preparation of Protein Derivatives

Threonine-neochymotrypsinogen was prepared by autolysis with added chymotrypsin as described previously [7], and the conditions used restricted the cleavage to Tyr 146-Thr 147. Protein samples were fully reduced with 10 mM dithioerythritol in 6 M guanidine HCl, at pH 9.1, 37°C, for 2.5 hr. The fully reduced protein in the solution of 6 M guanidine hydrochloride was adjusted to pH 8.6 and treated with 0.11 M glutathione disulfide [5,7]. After 3 hr, the mixed disulfide derivative was diluted directly into the refolding buffer solution, or, alternatively, the protein was recovered by gel filtration, dialysis, and lyophilization.

The two polypeptide fragments of the fully reduced threonine-neochymotrypsinogen, as the mixed disulfide derivative, were separated on a column of Sephadex G-75SF (1.5 × 85 cm) in a solution of 0.1 M acetic acid containing 3 M guanidine hydrochloride. The appropriate fractions were dialyzed, and the samples were recovered after lyophilization. The purity of the fragments was established with polyacrylamide gel electrophoresis and on amino acid analysis.

Refolding

The mixed disulfide derivative was refolded at pH 8.6, 4°C, in the presence of 4 mM cysteine, with a protein concentration of 20 to 25 μg/ml, and under a nitrogen barrier [5,7]. At selected timed intervals, samples were removed and quenched with a 25-fold molar excess of iodoacetate, which immediately inhibited further disulfide interchange [2]. Trypsinogen and the regenerated protein were activated with bovine enterokinase as described by Odorzynski and Light [5]. Chymotrypsinogen, neochymotrypsinogen, and the regenerated proteins were activated with added trypsin [6,7].

Analytical

Trypsin activity was measured with N-tosyl-arginine methyl ester as substrate [8], and chymotrypsin with N-acetyl-tyrosine ethyl ester [9]. Polyacrylamide gel electrophoresis was performed as described by Reisfeld et al [10] and SDS-polyacrylamide gel electrophoresis by the procedure of Laemmli [11]. Amino acid analyses were performed on a Durrum D-550 analyzer on samples hydrolyzed in 6 N HCl at 110°C for 25 hr *in vacuo*. The mixed disulfide derivatives were oxidized with performic acid prior to acid hydrolysis [5,6].

RESULTS

Attempts to refold fully reduced trypsinogen and chymotrypsinogen by the procedures used successfully [12,13] with disulfide-containing proteins were unsuccessful because the unfolded polypeptide chain rapidly aggregated and became insoluble [14]. Lowering the protein concentration; changing the pH, the temperature, and the disulfide interchange catalyst; and using a slow removal of the denaturant were all tested but none of these conditions altered the pathway from the production of high-molecular aggregates to that of the native structure [5,6,14]. However, the proteins were refolded successfully when the mixed disulfide of glutathione and the fully reduced protein (ie, Protein-S-SG) was used (Fig. 1). Refolding was initiated on the addition of 4 mM cysteine, which catalyzed disulfide interchange, and refolding was inhibited when aliquots of the refolding mixture were treated with a high concentration of iodoacetate. The rate of protein disulfide formation, (Protein-(S-S),) from the mixed disulfide was rapid ($t_{1/2}$ of 2 min), and the polypeptide chain initially formed a loose globular structure linked by non-native disulfide bonds [15]. Further disulfide interchange and chain folding continued until the native structure was formed. Both trypsinogen and chymotrypsinogen refolded with the same half-time of 60 to 75 min and with yields of 40% to 50% (Fig. 1).

The progress of the refolding of trypsinogen was also followed by examining samples by gel filtration [15] and more recently by HPLC [unpublished studies of Higaki and Light] on columns of Toya Soda G-2000 (data not shown). With time, globular structures with a large Stokes radius were replaced by more compact structures, and finally by the native molecule. The refolding pathway was a series of

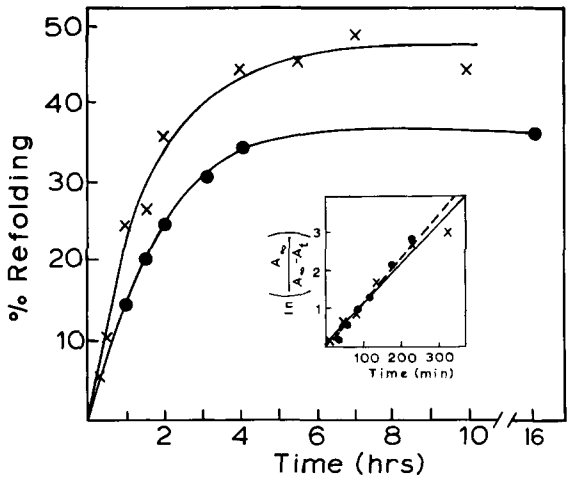

Fig. 1. Refolding the mixed disulfide derivatives of trypsinogen (●) and chymotrypsinogen (×). The mixed disulfide derivative (approximately 20 μg/ml) was maintained at pH 8.6 and 4°C in the presence of 4 mM cysteine under strictly anaerobic conditions. Samples were removed as a function of time, quenched with a large excess of iodoacetate (25-fold molar excess) and activated to the corresponding enzyme, and the yield was calculated based on the enzyme units produced and the amount of the mixed disulfide of the protein used in refolding. The ordinate is given as percent refolding, which is equivalent to the fraction of active enzyme recovered. The inset is a first-order plot of the refolding of the proteins.

sequential events suggesting that obligatory early intermediates behaved as precursors to the native structure.

It seemed reasonable to presume that an important intermediate on the refolding pathways would be the globular structure of the amino- and carboxyl-terminal domains. This possibility was examined by refolding mixtures of the unfolded domains prepared from neo-chymotrypsinogen [7]. The mixed disulfide derivative of the polypeptide fragments were purified on columns of Sephadex G-75 (Fig. 2). The amino acid compositions of the corresponding dialyzed and lyophilized fragments are given in Table I. The absence of methionine in the amino-terminal domain and the lack of phenylalanine and histidine in the carboxyl-terminal domain served as an additional test of their purity.

We recently demonstrated that the mixed disulfide of neochymotrypsinogen refolded and formed the native globular structure of the protein [7]. In these studies, the two fragments (domains) were used as mixtures without further purification. We now report on the refolding of mixtures of the purified fragments (Table II). The 1:1 mixture of the fragments gave similar results to those reported previously [7]. The purification procedure used for the isolation of the polypeptide fragments had not altered the covalent structure and the mixed disulfide remained intact.

The possibility that one or the other polypeptide fragment acted as a critical nucleation site in the refolding pathway was explored by mixing various amounts of each of the two polypeptide fragments. Essentially identical yields were obtained when the amino-terminal domain was present in five times the amount of the carboxyl-terminal domain and with the reverse ratio (Table II). Apparently, the refolding pathway did not appear to require the folding of either one of the domains as a prerequisite for the folding of the other.

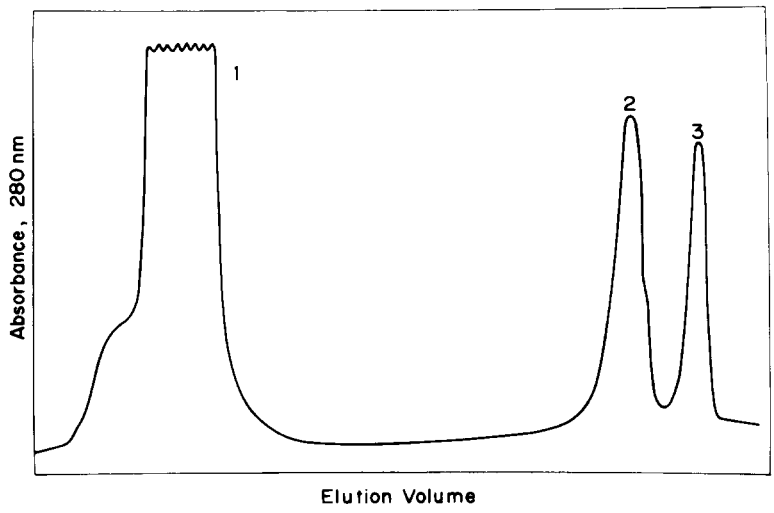

Fig. 2. Purification of the polypeptide fragments from fully reduced threonine-neochymotrypsinogen as the mixed disulfide derivatives on a column of Sephadex G-75SF (1.5 × 85 cm). The eluting solvent was 0.1 M acetic acid containing 3 M guanidine hydrochloride. Flow rate, 6 ml/hr, full-scale was 0.1 OD. Peak 1 was intact chymotrypsinogen; peak 2, the carboxyl-terminal fragment residues, 147–245; and peak 3, the amino-terminal fragment residues, 1–146.

TABLE I. Amino Acid Composition of the Mixed Disulfide of the Purified Fragments of Threonine-Neochymotrypsinogen*

	Polypeptide fragment			
	(1–146)		(147–245)	
	Found	Theory	Found	Theory
Asp	12.7	13	9.6	10
Thr	12.0	13	9.2	10
Ser	15.3	17	11.3	11
Glu[a]	15.8	16	8.9	9
Pro	6.1	5	5.1	4
Gly[a]	18.0	18	15.2	15
Ala	12.1	12	10.1	10
½CyS	—	5	—	5
Val	13.9	15	7.5	8
Met	0	0	2.0	2
Ile	6.6	7	2.8	3
Leu	10.9	11	7.9	8
Tyr	1.8	2	1.8	2
Phe	5.5	6	0	0
His	2.0	2	0.4	0
Lys	8.3	8	6.1	6
Arg	2.0	2	1.9	2
Trp	—	0	—	4
Total	—	152	—	105
Mr		14,087		10,055

*Tryptophan and ½-cystine were not determined.
[a]The content of glu and gly was increased by five residues because each fragment contained five glutathione residues as the mixed disulfide derivative.

TABLE II. The Effect of the Concentration of the Polypeptide Chains on the Regeneration of Threonine-Neochymotrypsinogen*

Polypeptide fragments		
(1–146)	(147–245)	Yield
(nmoles/ml)		(%)
1.0	1.0	21
1.0	2.5	28
1.0	5.0	26
2.5	1.0	27
5.0	1.0	26

*Amino-terminal (1–146) and carboxyl-terminal (147–245) fragments were each dissolved in 6 M guanidine hydrochloride, pH 3.0. Refolding was initiated with a 60-fold dilution into refolding buffer (0.05 M Tris-HCl, pH 8.6, containing 4 mM cysteine), and the mixture was maintained for 3 hr, at 4°C, at pH 8.6, under a nitrogen atmosphere. Refolding yields were calculated as the fraction of regenerated native protein recovered based on the amount of the mixed disulfide, the concentration of which was established by amino acid analysis. The amount of regenerated native protein was determined from the enzymatic activity produced after activation with bovine trypsinogen.

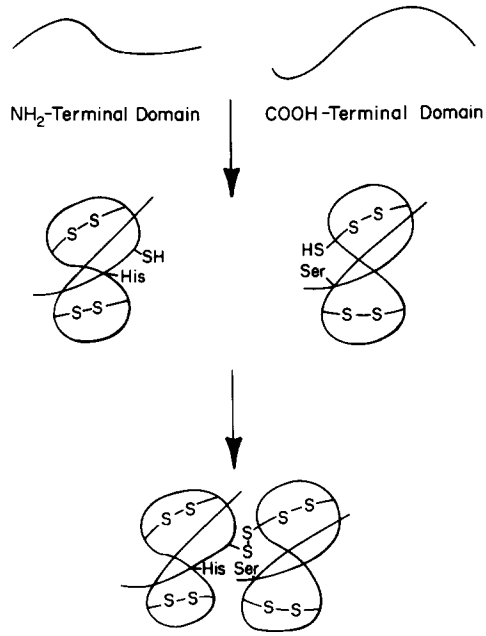

Fig. 3. Proposed folding pathway for threonine-neochymotrypsinogen from the fully reduced mixed disulfide derivative. The location of the histidine (HIS) and serine (SER) residues of the active site and the representation of the three-dimensional structure are given pictorially.

DISCUSSION

The pancreatic serine proteinase family of enzymes is characterized by a three-dimensional structure composed of two domains, each barrel-shaped with six anti-parallel strands in the pleated sheet conformation [16,17,18]. The active site residues of serine 195 and histidine 57 are in the carboxyl-, and amino-terminal domains, respectively. They are close in space, because the domains are held together by a short stretch of polypeptide chain and by disulfide bonds. Because the domains have similar three-dimensional structures, we hypothesize that the refolding of the serine proteinases proceeds at the domain level.

We initially examined experimental conditions that could be used to refold the fully reduced proteins [14]. We eventually selected the mixed disulfide of the protein and glutathione to successfully refold trypsinogen and chymotrypsinogen [5,6,7]. The proteins remained soluble and refolded with yields of 40 to 50%. The mixed disulfide derivative maintained the solubility of the unfolded protein and avoided aggregation in the critical early stages of refolding [5,6,19]. Furthermore, the kinetics of refolding showed that trypsinogen and chymotrypsinogen had half-times of 60 to 75 min, suggesting that both zymogens had the same rate-limiting step. Since disulfide interchange is known to be one of the rate-limiting steps in the refolding of disulfide-containing proteins, therefore the same disulfide-linked intermediates apparently were present in both proteins. It should be recalled that trypsinogen had six disulfide bonds and chymotrypsinogen, five [1]. Although many non-native disulfide bonds can form, only a selected few common to both proteins may be critical precursors leading to the

native protein. Alternatively, the formation of one of the native disulfide bonds later in the refolding process could be the rate-limiting step. Studies now in progress are examining these possibilities, and we hope to obtain evidence that will permit a general description of the folding pathway.

In an attempt to simplify the experimental approach required to identify the disulfide bonds in intermediate species [2], the present report shows that the immediate precursors of the native protein were most likely the two folded domains. We showed previously that the two-chain structure of neochymotrypsinogen could be refolded with a yield that was as high as 50% of that found with the intact chymotrypsinogen [7]. We now show that the refolding of neochymotrypsinogen proceeded equally well when either the amino- or carboxyl-terminal domains was present as the limiting component (Table II).

We propose a mechanism of folding shown in Figure 3 where each domain refolds independently of the other. Only after the individual domains refold to globular structures do they then recognize one another and interact as two complementing protein molecules to form the required intermolecular disulfide bonds and the active site geometry of the native protein. Furthermore, we have to assume that the independent refolding of the amino- and carboxyl-terminal domains proceeded at the same rate, since we could not find a concentration dependency for either one. This mechanism of refolding accounts for the observed high yield since the bimolecular interaction of the two domains could only be successful if the domain structures recognized one another. Such recognition sites are possible only after each domain formed a globular, native-like structure.

ACKNOWLEDGMENTS

These studies were supported by NIH grant GM22261 and by the Department of Chemistry, Purdue University. The authors appreciate the helpful contributions of Jeffrey N. Higaki. We thank Chris Gobble for her assistance in the preparation of the manuscript.

REFERENCES

1. Dayhoff MD: "Atlas of Protein Sequence and Structure," Vol. 5. Silver Springs, Maryland: National Biomedical Research Foundation, 1972.
2. Creighton TE: "Proteins. Structures and Molecular Principles." New York: W. H. Freeman and Co., 1983.
3. Bode W, Fehlhammer H, Huber R: J Mol Biol 106:325, 1976.
4. Fehlhammer H, Bode W, Huber R: J Mol Biol 111:415, 1977.
5. Odorzynski TW, Light A: J Biol Chem 254:4291, 1979.
6. Moore WGI: PhD Thesis, Purdue University, 1981.
7. Duda CT, Light A: J Biol Chem 257:9866, 1982.
8. Walsh KA, Wilcox PE: Methods Enzymol 19:31–41, 1970.
9. Wilcox PE: Methods Enzymol 19:64–112, 1970.
10. Reisfeld RA, Lewis VJ, Williams DE: Nature 195:281, 1962.
11. Laemmli UK: Nature 227:680, 1970.
12. Anfinsen CB, Haber E: J Biol Chem 236:1361, 1961.

13. Anfinsen CB: Science 181:223, 1973.
14. Sinha NK, Light A: J Biol Chem 250:8624, 1975.
15. Odorzynski TW: PhD Thesis, Purdue University, 1978.
16. McLachlan AD: J Mol Biol 128:48, 1979.
17. Schulz GE, Schirmer RH: "Principles of Protein Structure." New York: Springer-Verlag, 1979.
18. Richardson J: Advances in Protein Chemistry 34:167–339, 1981.
19. Light A: BioTechniques 3:298, 1985.