The amino acid sequence of ERp57, which functions in the endoplasmic reticulum together with the lectins calreticulin and calnexin to achieve folding of newly synthesized glycoproteins, is highly similar to that of protein disulfide isomerase (PDI), but they have their own distinct roles in protein folding. We have characterized the domain structure of ERp57 by limited proteolysis and N-terminal sequencing and have found it to be similar but not identical to that of PDI. ERp57 had three major protease-sensitive regions, the first of which was located between residues 120 and 150, the second between 201 and 215, and the third between 313 and 341, the data thus being consistent with a four-domain structure abb′a′. Recombinant expression in Escherichia coli was used to verify the domain boundaries. Each single domain and a b′a′ double domain could be produced in the form of soluble, folded polypeptides, as verified by circular dichroism spectra and urea gradient gel electrophoresis. When the ability of ERp57 and its a and a′ domains to fold denatured RNase A was studied by electrospray mass analyses, ERp57 markedly enhanced the folding rate at early time points, although less effectively than PDI, but was an ineffective catalyst of the overall process. The a and a′ domains produced only minor, if any, increases in the folding rate at the early stages and no increase at the late stages. Interaction of the soluble ERp57 domains with the P domain of calreticulin was studied by chemical cross-linking in vitro. None of the single ERp57 domains nor the b′a′ double domain could be cross-linked to the P domain, whereas cross-linking was obtained with a hybrid ERpaab′PDia′c peptide but not with ERpaabPDiab′a′c, indicating that multiple domains are involved in this protein-protein interaction and that the b′ domain of ERp57 cannot be replaced by that of PDI.

ERp57 is a protein resident in the endoplasmic reticulum (ER) that functions as a co-chaperone in glycoprotein folding with the lectins calnexin (CNX) and its soluble homologue calreticulin (CRT) (for reviews, see Refs. 1 and 2). ERp57 is a member of the protein disulfide isomerase (PDI) oxidoreductase family (3, 4) and has two thioredoxin-like domains with -Cys-Gly-His-Cys- catalytic site motifs in positions corresponding to the thioredoxin-like domains a and a′ of PDI, which consists altogether of domains a, b, b′, and a′ plus an acidic C-terminal extension c (5). ERp57 also shows a significant sequence similarity to PDI in regions corresponding to domains b and b′ (6). No similarity is found in the organization of the human genes or gene loci for ERp57 and PDI, however (7). Despite the 56% overall amino acid sequence similarity between PDI and ERp57, the latter does not substitute for PDI as the β subunit in the assembly of the collagen prolyl 4-hydroxylase αβtetramer (6), an enzyme involved in the synthesis of collagens (8–10), but the N-terminal domains c and b′ of ERp57 can in part substitute for those of PDI in this assembly (11).

ERp57 possesses disulfide oxidoreductase activity in vitro, but the level is lower than that of PDI (12–14). Interestingly, this activity is markedly increased in the presence of CNX or CRT (15). ERp57 present in transient complexes with its substrates, and CNX and CRT catalyzes the formation and isomerization of disulfide bonds in the synthesis of glycoproteins in vitro (16). It is also involved in the folding of a major histocompatibility complex class I heavy chain in a system of semipermeabilized cells, leading to disulfide bond formation in this protein when in interaction with ERp57 (17). ERp57 even interacts with glycoproteins having no cysteine residues, and thus its role in glycoprotein folding may be more extensive than that of a basic oxidoreductase (18, 19). PDI can likewise become bound to glycoproteins, but it is not co-precipitated from cells with CNX and CRT, indicating that PDI and the lectins are not associated partners in glycoprotein folding in vivo (16, 20).

The exact mode of ERp57 action in glycoprotein folding is still partly unknown. A freshly attached core oligosaccharide of an unfolded glycoprotein is first trimmed by glucosidases I and II, after which the glycoprotein becomes bound to CNX or CRT, this being followed by binding of the complex to ERp57 (18, 21). The oligosaccharide motif does not act as the binding site between ERp57 and the lectins, but instead, ERp57 becomes directly bound to CNX and CRT (20). The oligosaccharide binding function of CRT and CNX has been mapped to a distinct lectin domain of the polypeptides (22). These P domains have nexitin; CRT, calreticulin; ESIMS, electrospray ionization mass spectrometry; HIS-ERp57, ERp57 containing an N-terminal histidine tag; PDI, protein disulfide isomerase; H5, High Five.

Laura Silvennoinen‡, Johanna Myllyharju‡, Margherita Ruoppolo§§, Stefania Orru‡, Marianna Caterino‡, Kari I. Kivirikko‡, and Peppi Koivunen‡‡

From the ‡Collagen Research Unit, Biocenter Oulu and Department of Medical Biochemistry and Molecular Biology, University of Oulu, P. O. Box 5000, FIN-90014 Oulu, Finland, the §Department of Biochemistry and Medical Biotechnologies, University of Naples “Federico II,” 80138 Naples, Italy and the ¶Centro di Ingegneria Genetica, Biotecnologie Avanzate, scrl, Naples 80138, Italy

Published, JBC Papers in Press, January 19, 2004, DOI 10.1074/jbc.M313054200

Received for publication, December 1, 2003, and in revised form, January 9, 2004

Printed in U.S.A.

© 2004 by The American Society for Biochemistry and Molecular Biology, Inc.
ERp57 Domains

been found to form a hairpin motif (23), and the tip of the hairpin is believed to be involved in the protein-protein interactions and to bind ERp57 (2, 23).

All four PDI domains have the thioredoxin fold, although domains b and d show no amino acid sequence similarity to thioredoxin (24–26). No three-dimensional structure is available for ERp57, but nuclear magnetic resonance (NMR) characterization of the b’ domain has shown that it is structurally related to the corresponding PDI domain and also possesses the thioredoxin fold (27).

The present work sets out to study the ERp57 domain architecture by limited proteolysis and recombinant protein expression methods. The structures of the soluble individual domains were studied in more detail by circular dichroism (CD) spectrum analysis and urea gradient gel electrophoresis. The abilities of ERp57 and the a and c’ domains to fold denatured RNase A were also examined, as was the ability of the individual domains to interact with the P domain of CRT.

EXPERIMENTAL PROCEDURES

Construction of a Baculovirus Transfer Vector and Generation of a Recombinant ERp57 Baculovirus—An ERp57 expression vector containing its N terminus was synthesized by PCR using human ERp57 cDNA (6) as a template. Two fragments, named Eag3His and 3HisAvr, were produced. Eag3His contained an EagI site at its 5’ end, nucleotides coding for the ERp57 signal peptide, and 9 additional nucleotides coding for 3 histidines at its 3′ end, whereas 3HisAvr contained nucleotides coding for 3 histidines at its 5′ end followed by nucleotides coding for the N terminus of the mature ERp57 polypeptide, extending to an internal AvrII site at 420 bp. The PCR products were blunt-ended and digested with the corresponding restriction enzymes. These fragments were then ligated together with an AvrII-EcoRI-digested fragment of human ERp57 cDNA (6) and cloned into the Eagl-EcoRI sites in pVL1392 (Invitrogen). The resulting plasmid encoding the ERp57 signal peptide followed by 6 histidines and the mature human ERp57 was co-transfected into Spodoptera frugiperda (SF9) insect cells with a modified Autographa californica nuclear polyhedrosis virus DNA (BaculoGold, Pharmingen) by calcium phosphate transfection (28). The resultant viral pool was collected 4 days later, amplified twice, and used for recombinant protein production.

Production of ERp57 in Insect Cells and Purification of the Protein—SF9 cells were cultured in suspension in TBM-FH medium (Sigma) supplemented with 10% fetal bovine serum (Bioclear) at 27 °C. The cells were infected with a baculovirus coding for the untagged ERp57 at a multiplicity of 5, whereas the histidine-tagged ERp57 (His-ERp57) was produced in High Five (H5) insect cells in SF-900 II SFM medium (Invitrogen) without any serum supplement. The cells expressing the untagged ERp57 were homogenized in a solution of 0.15 M NaCl, 0.02 M Tris-HCl, pH 7.8, and the fractions containing the protein were pooled and concentrated with a 1-mm quartz cuvette at 20 °C, resulting in an ATG sequence. This plasmid was digested with EcoRI and HindIII, and the fragment obtained was subsequently ligated into the corresponding sites in PET21 (Novagen). Expression constructs for ERp1–111, ERp1–113, ERp107–216, ERp107–326, ERp211–348, ERp211–468, ERp222–326, and ERp222–348 were generated by PCR. The forward primers contained a BamHI site at their 5′ end, and the reverse primers contained a HindIII site at their 3′ end. The PCR products were first ligated into PQE30 (Qiagen) and then excised from the vector with EcoRI and HindIII, generating fragments that contained a ribosome binding site and nucleotides encoding the amino acids Met-Arg-Gly-Ser-(His)8-Gly-Ser before the first ERp57 amino acid indicated for the construct. These plasmids were transformed into ET348 (Novagen). Expression constructs for ERp1–123 and ERp1–119, containing NcoI and HindIII sites at their 5′ and 3′ ends, were generated by PCR and cloned into the corresponding sites in PET32a (Novagen). ERp349–468 was generated as described earlier (27).

Production of ERp57 and Its Domains in E. coli and Purification of the Polypeptides—The expression constructs generated in PET21 (+) and PET32a (Novagen) were transformed into E. coli BL21 (DE3) Star or Origami (DE3) (Novagen) and induced by isopropyl-1-thio-D-galactopyranoside. The cells were harvested 3 h after induction, suspended in 0.5 mM imidazole, 0.5 mM NaCl, 0.1% Triton X-100, and 0.02 mM Tris, pH 8, supplemented with Complete EDTA-free protease inhibitor mixture (Roche Applied Science) and sonicated on ice. After centrifugation at 60,000 × g for 45 min at 4 °C, the supernatants were applied to Ni2+ affinity columns, washed, and eluted as described above for HIS-ERp57 produced in insect cells, and fractions containing the desired proteins were pooled. The remaining pellets were solubilized in 1% SDS and analyzed by SDS-PAGE. Fractions of ERp57, ERp1–113, and ERp107–216 were then applied to a Sephacryl S-100 high-resolution column (Amersham Biosciences) and fractions of ERp211–348 and ERp211–468 were applied to a Superdex 75 HR column (Amersham Biosciences), equilibrated and eluted with 0.2 mM NaCl, 0.02 mM Tris-HCl, pH 7.5. Fractions containing the protein were pooled and concentrated, and the protein concentrations were determined as described above. ERp349–468 was precipitated earlier (27). The cell pellets were lyophilized and stored at –80 °C.

Transverse Urea Gradient Gel Electrophoresis—Electrophoresis of HIS-ERp57 (data not shown), ERp1–113, ERp107–216, ERp211–348, ERp211–468 (data not shown), and ERp349–468 was carried out in a urea gradient increasing from 0 on the left to 8 on the right, with a compensating 11–7% acrylamide gradient, in 0.5% Tris acetate, pH 4.0. Light-induced polymerization of the gels was initiated with 2 mM methylene, and after overnight incubation, they were rotated 90°, and samples of 50 μl were loaded. The gel assembly was cooled to 4 °C, pre-electrophoresed for 30 min at 100 V, and electrophoresed for 4 h at 150 V for the domains and at 200 V for HIS-ERp57, after which the gel was analyzed by Coomassie Blue staining.

Circular Dichroism Spectra—Circular dichroism spectra of HIS-ERp57, ERp1–113 (domain a), ERp107–216 (domain b), ERp211–348 (domain b’), ERp211–468 (domains b’a’), and ERp349–468 (domain a’) were measured with a Jasco J715 spectropolarimeter in a 1-mm quartz cuvette at 20 °C. The spectra were averages of five scans. All the proteins analyzed were dissolved in H2O except for ERp107–216 (dissolved in 10 mM sodium phosphate buffer, pH 6). RNAs were precipitated from the protein solution with ethanol and then re-suspended in water. The RNase A Folding Analysis—RNase A was reduced and denatured as described earlier (29, 30). The folding reactions were carried out at 25 °C in a nitrogen atmosphere in the presence of 1 mg/ml RNase A, 1.5 mM reduced glutathione (GSH), 0.3 mM oxidized glutathione (GSSG), and 10 or 50 μM HIS-ERp57, ERp1–113 (domain a), or ERp349–468 in 0.05 M sodium phosphate buffer (pH 7.0) at 25 °C. The samples were pre-incubated in the presence of 1.5 mM GSH/0.3 mM GSSG redox buffer for 10 min at 25 °C. Aliquots of the folding mixtures were taken at different time intervals, and the protein samples were alkylated by the addition of iodoacetamide (29, 30). Electrospray ionization mass spec-
Migration of the fragments generated with chymotrypsin (lane 1, Triton X-100-insoluble, SDS-soluble proteins from H5 cells infected with a baculovirus encoding human ERp57) were pooled and concentrated (Fig. 1, lane 3). The 17-kDa fraction was then applied to an Ni²⁺ charged chelating Sepharose column eluted by an imidazole gradient. The fractions were analyzed by 8% SDS-PAGE under reducing conditions, and those containing the ERp57 protein were pooled and concentrated (Fig. 1, lane 3).

Digestion of HIS-ERp57 with chymotrypsin gave several major bands, of which those of about 31, 20, 19, 17, and 14.5 kDa (Fig. 2, lane 1) were identified by N-terminal sequencing. The 31-kDa fragment was found to begin at Ile⁴¹², the 20-kDa fragment was found to begin at Val¹¹¹, and the 19-kDa fragment was found to begin at Ser¹²⁹ (Fig. 2, lane 1). The 17-kDa band gave two N-terminal sequences, one of which began at Leu³²⁹ and the other of which began 12 residues later, at Leu³⁴¹ (Fig. 2, lane 1). The sizes of these fragments suggested that they all extended to about residues 470–481, i.e. close to the C terminus of the polypeptide. The 14.5-kDa fragment

trometry (ESIMS) analyses of the samples were carried out using a Quattro Micro triple quadrupole mass spectrometer equipped with an electrospray ion source (Micromass, Manchester, UK, 30).

Chemical Cross-linking—A cDNA for the P domain of CRT was amplified by PCR from a rat fetal liver cDNA pool (BD Biosciences), cloned into pET32(+) (Novagen), expressed as a fusion protein with thioredoxin in E. coli, and purified by exploiting the histidine tag at the C terminus of thioredoxin. The fusion partner was baculovirus encoded enterokinase (Sigma). Histidine tags were introduced between the ERp57 signal peptide and the N terminus of the two baculovirus expression constructs ERp⁴⁸⁸PDIc’ and ERp⁴⁶⁶PDid’a’c’ generated earlier, coding for amino acids 1–352 of ERp57 and 351–491 of PDI and for 1–218 of ERp57 and 219–491 of PDI, respectively (11). The hybrid polypeptides were expressed in H5 insect cells and purified by exploiting the histidine tags at their N termini, as described earlier for HIS-ERp57.

Cross-linking studies with 1 μg of the purified ERp57 domains and the hybrid ERp⁴⁸⁸PDIc’ and ERp⁴⁶⁶PDid’a’c’ polypeptides with 1 μg of the CRT P domain and disuccinimidyl glutarate were carried out on ice for 2 h (31). The reactions were analyzed by reducing 10 and 15% SDS-PAGE followed by silver staining.

RESULTS

Gel Filtration Reveals That ERp57 Expressed in Insect Cells Is a Monomer.—To study whether ERp57, like PDI (32), is a homodimer, a recombinant baculovirus encoding the human ERp57 polypeptide without a histidine tag was used to infect Sf9 insect cells. The cells were harvested 72 h after infection and homogenized in a buffer containing Triton X-100, and the supernatant was applied to a Sephacryl S-200 gel filtration column calibrated with polypeptide standards. The elution position of ERp57 was determined by SDS-PAGE and non-denaturing PAGE analyses of the fractions and was found to correspond to that of a molecular mass of about 60 kDa (data not shown), thus indicating that the polypeptide was eluted as a monomer.

Limited Proteolysis of ERp57 Indicates a Four-domain Structure Similar to That of PDI.—To purify the recombinant ERp57 polypeptide for limited proteolysis studies, a baculovirus encoding human ERp57 with a histidine tag in its N terminus was amplified by PCR from a rat fetal liver cDNA pool (BD Biosciences), cloned into pET32(+) (Novagen), expressed as a fusion protein with thioredoxin and homodimer, a recombinant baculovirus encoding the human ERp57 polypeptide without a histidine tag was used to infect H5 insect cells. The cells were harvested 72 h after infection, homogenized in a buffer containing Triton X-100, and centrifuged. The pellet was solubilized in 1% SDS, and the Triton X-100-soluble and SDS-soluble proteins were analyzed by 8% SDS-PAGE under reducing conditions followed by Coomassie Blue staining. Most of the recombinant protein was found in the Triton X-100-soluble fraction (Fig. 1, lanes 1 and 2). This fraction was then applied to an Ni²⁺ charged chelating Sepharose column eluted by an imidazole gradient. The fractions were analyzed by 8% SDS-PAGE under reducing conditions, and those containing the ERp57 protein were pooled and concentrated (Fig. 1, lane 3).

Trypsin, thermolysin, proteinase K, chymotrypsin, and V8 protease were used to digest the purified HIS-ERp57 polypeptide to determine the conditions favoring its limited degradation to protease-resistant fragments. Chymotrypsin and V8 protease were found to be the most suitable enzymes for this purpose. The molecular masses of the protease-resistant fragments were determined by SDS-PAGE under reducing conditions, and the major bands were blotted onto polyvinylidene difluoride membranes for N-terminal sequencing.

Digestion of HIS-ERp57 with chymotrypsin gave several major bands, of which those of about 31, 20, 19, 17, and 14.5 kDa (Fig. 2, lane 1) were identified by N-terminal sequencing. The 31-kDa fragment was found to begin at Ile⁴¹², the 20-kDa fragment was found to begin at Val¹¹¹, and the 19-kDa fragment was found to begin at Ser¹²⁹ (Fig. 2, lane 1). The 17-kDa band gave two N-terminal sequences, one of which began at Leu³²⁹ and the other of which began 12 residues later, at Leu³⁴¹ (Fig. 2, lane 1). The sizes of these fragments suggested that they all extended to about residues 470–481, i.e. close to the C terminus of the polypeptide. The 14.5-kDa fragment

![Fig. 1. Analysis of the insect cell expression and purification of HIS-ERp57 by SDS-PAGE under reducing conditions. Lane 1, Triton X-100-soluble proteins from H5 cells infected with a baculovirus encoding HIS-ERp57; lane 2, Triton X-100-insoluble, SDS-soluble proteins from the same cells; lane 3, HIS-ERp57 eluted from a Ni²⁺ affinity column. The migration of molecular mass markers is shown on the left, and that of HIS-ERp57 is shown by the arrow. The samples were analyzed by 8% SDS-PAGE and Coomassie Blue staining.](http://www.jbc.org/)

![Fig. 2. SDS-PAGE analysis under reducing conditions of HIS-ERp57 fragments produced by limited proteolysis with chymotrypsin and V8 protease. Migration of the fragments generated with chymotrypsin (lane 1) and V8 protease (lane 2) is shown on 15% SDS-PAGE analyzed by Coomassie Blue staining. The first 10 amino acids of the fragments obtained by N-terminal sequencing and the migration of molecular mass markers are also shown.](http://www.jbc.org/)
began at the histidine tag followed by the first amino acids of the ERp57 polypeptide (Fig. 2, lane 1), and its C terminus was probably around residue 120, possibly after Phe121, as deduced from the cleavage specificity of chymotrypsin. The digestion patterns obtained with V8 protease consisted of numerous bands, of which the N termini of nine were sequenced (Fig. 2, lane 2). The highest molecular mass polypeptide, of about 55 kDa, began with the histidine tag followed by the N-terminal amino acids of ERp57, thus corresponding to an undigested polypeptide (Fig. 2, lane 2). The 41-kDa fragment began at Phe121, and the 38-kDa fragment began at Phe147 (Fig. 2, lane 2), both probably extending to about the C terminus of the polypeptide. The N termini of the 32- and 31-kDa fragments corresponded to Gln201 and Asn215, and these fragments probably also extended approximately to the C terminus of the polypeptide. The 24-kDa fragment was identical to that of the 41-kDa fragment, i.e. Phe121 (Fig. 2, lane 2), and it probably ended after Glu326. The 18-kDa fragment began at Arg327, whereas the 16.5-kDa fragment began with the histidine tag (Fig. 2, lane 2) and probably extended to about residue 120. The N termini of fragments generated by thermolysin were not strictly in agreement with its substrate specificity. The digestions with proteinase K resulted in the largest population of bands. The N termini of the 16.5- and 14.5-kDa fragments began at the histidine tag followed by the first amino acids of the ERp57 polypeptide (Fig. 2, lane 1) and probably extended to about the C terminus of the polypeptide. The N termini of fragments generated by thermolysin were not strictly in agreement with its substrate specificity. The digestions with proteinase K resulted in the largest population of bands. The N termini of the 16.5- and 14.5-kDa fragments began at the histidine tag followed by the first amino acids of the ERp57 polypeptide (Fig. 2, lane 1) and probably extended to about the C terminus of the polypeptide.
fragments were sequenced, and both were found to begin with the histidine tag and were likely to extend to about residues 140 and 120, respectively (Fig. 3).

The above experiments identified three major protease-sensitive regions in the HIS-ERp57 polypeptide that probably correspond to four protease-resistant domains. The first region, identified by all the proteases studied except for trypsin, begins at the histidine tag located at the N terminus of the polypeptide. The molecular weights of six fragments obtained with three proteases suggested that four of them ended at about residue 120, one ended at about 130, and one ended at about 140. In agreement with this, the N termini of two fragments were at Phe123 and one at Phe147. The first protease-sensitive region must thus begin around residue 120, which seems to be the main site, and extend until residue 140–150.

Seven fragments obtained by four proteases had their N termini between residues 201 and 215, suggesting that the second protease-sensitive region is relatively short and well defined. Interestingly, we did not find the second protease-resistant domain in the form of a single fragment in any of the digestions but only together with either the third or the third and fourth protease-resistant domains. Eight fragments obtained with three proteases had their N termini between residues 313 and 341, six of them between 327 and 341. The molecular weights of three fragments obtained with the V8 protease suggested that their C termini were likewise between residues 326 and 334. The third protease-resistant region is thus located between residues 313 and 341, and more particularly, between 326 and 341. The proteolysis data thus suggest that ERp57, like PDI (24), consists of four domains, termed a, b, b', and a', that are separated by three protease-sensitive regions, the locations of which appear to be slightly different in the two polypeptides, however. The first region has a slightly more C-terminal location in the ERp57 polypeptide than in PDI, whereas the second and third regions have slightly more N-terminal locations (Fig. 3). It may be noted that these domain boundaries are not in strict agreement with the predicted secondary structure elements of the ERp57 polypeptide (Fig. 3).

Expression of ERp57 Domains a, b, b' and a'—To verify that the protease-resistant fragments obtained from the full-length ERp57 represent individual protein domains, recombinant polypeptides coding for residues 1–123, 1–119, 1–115, 1–113, 1–111, 107–216, 107–326, 211–348, 211–468, 222–326, and 222–348 were generated and expressed in E. coli, an additional polypeptide ERp349–468 being produced and purified as described earlier (27). Most of the domain constructs contained codons for a few additional amino acids, extending the polypeptide at its N and C termini as compared with the domain boundaries suggested by the protease digestions, which was assumed to help folding of the polypeptides in E. coli. All the recombinant polypeptides contained a histidine tag at their N termini, and ERp1–123 and ERp1–119 also contained a thioredoxin fusion partner there. ERp1–123, coding for the a domain, was found in the soluble cell fraction (data not shown). When the thioredoxin fusion partner was cleaved off, two bands with intact N termini verified by sequencing appeared in SDS-PAGE (data not shown). Mass spectrometry indicated that the minor band corresponded to a polypeptide 1–123, whereas the major band corresponded to 1–113, thus indicating degradation of ERp1–123 (data not shown). Four shorter constructs, ERp1–119, ERp1–115, ERp1–113, and ERp1–111, were therefore prepared. ERp1–119 and ERp1–115 failed to produce any significant amounts of recombinant protein (data not shown), whereas ERp1–113 and ERp1–111 both expressed a soluble a domain, although the expression level of ERp1–111 was lower than that of ERp1–113, the Triton X-100-soluble fraction of ERp1–113 being shown in lane 1 in Fig. 4A. Expression constructs ERp107–216, ERp211–468, and 349–468, coding for the b, b’ a’ and a’ domains, all expressed soluble polypeptides, as shown in SDS-PAGE analysis in Fig. 4A, lanes 2–4. All these polypeptides were purified to essential homogeneity by Ni2+-affinity chromatography (Fig. 4C, lanes 1, 2, 4, and 5). Most of the polypeptides 107–326, 311–348, 222–326, and 222–348, corresponding to the double domain bb’ and three length variants of the bb’ domain, were found in the Triton X-100-insoluble fraction solubilized by SDS (Fig. 4B, lanes 2, 4, 6, and 8), although ERp211–348, coding for the bb’ domain, could be purified to essential homogeneity from the Triton X-100-soluble fraction by Ni2+-affinity chromatography (Fig. 4C, lane 3).

ERp57 containing a histidine tag at its C terminus was also expressed in E. coli and purified, and the purified protein was analyzed by SDS-PAGE (Fig. 4C, lane 6). In addition to the full-length ERp57 polypeptide, fragments of 34 and 24 kDa were obtained (Fig. 4C, lane 6). These also co-eluted with the full-length ERp57 in Sephacryl S-200 and S-100 gel filtration columns (data not shown). N-terminal sequencing indicated that they began at Ala307 and Met314 (Fig. 4A). No such proteolysis was found to occur in insect cells (Fig. 1).

Analysis of ERp57 Domains by Transverse Urea Gradient Gel Electrophoresis and CD Spectroscopy—Transverse urea gradient gel electrophoresis performed on the purified ERp57 a, b, b’, and a’ domains (Fig. 5, A–D) showed that increasing urea concentration markedly retarded the migration of all these polypeptides, thus indicating denaturation of their folds. CD spectra of the four domain polypeptides along with HIS-ERp57 and the b’a’ double domain were recorded in the far UV region (Fig. 6). These results support the suggestion that the
independent domains and the double domain b’a’ correspond to individual folded polypeptides, as each spectrum is typical of a folded protein (Fig. 6). The CD spectra of domains b and b’ were slightly different from those of domains a and a’, the former having their most negative molar ellipticity values between 205 and 210 nm and the latter between 225 and 235 nm (Fig. 6), which suggests minor differences in their secondary structure components.

ERp57 but Not Its Domains a and a’ Catalyze the Folding of RNase A—Reduced, denatured RNase A was incubated in the presence of 10 μM HIS-ERp57 or domain a or a’, and the folding intermediates present at different time intervals were withdrawn, trapped by alkylation of the free thiol groups, and identified and analyzed by ESI-MS (30, 33, 34). The relative amounts of the intermediates formed during the folding catalyzed by HIS-ERp57 or its domains a and a’ are shown in Fig. 7, A–C.

The reduced species 8H (Fig. 7A) rapidly disappeared within the first 4 min in the presence of HIS-ERp57, as compared with 20 min in spontaneous RNase A folding (33). The 3S2H intermediates predominated until about 5–10 min, after which the relative concentration of the 4S species, i.e., the fully oxidized protein, became the most prominent. The fully oxidized protein rapidly accumulated to an extent of about 55% at 40 min, but its rate of formation then declined, and its abundance reached only 90% even at 24 h (1440 min, Fig. 7A). The relative amount of the fully oxidized species at 40 min in the spontaneous process is only about 5%, but it likewise reaches a value of about 90% at 1440 min (33). In the folding catalyzed by PDI, the 2S4H intermediate is the most prominent species until 20 min, when the 4S species becomes the most abundant one and reaches 100% at 40 min (30). The predominance of the 2S4H species over 3S2H observed in the PDI-catalyzed folding was also seen in the spontaneous process (30). ERp57 thus markedly increased the rate of folding of RNase A at early time points, although it was not as effective as PDI (Table I), but it showed much less efficient catalysis of the overall process than PDI (compare Ref. 30 and Fig. 7A). ERp57 also altered the distribution of the folding intermediates relative to the PDI-catalyzed and spontaneous reactions (compare Ref. 30 and Fig. 7A).

The isolated ERp57 domains a and a’ catalyzed the disappearance of the fully reduced RNase A to a relative amount of about 5–10% during the first 4 min, but this species did not fully vanish even during 420 min (Fig. 7, B and C). The 2S4H species were the most prominent ones in the folding reactions catalyzed by both domains, followed by the 1S6H, 3S2H, and 4S species until 150 min in the presence of domain a or 80 min in the case of domain a’, when the 4S species became the most prominent one (Fig. 7, B and C). The appearance of the fully oxidized species showed a lag phase of about 40 min, during which it reached a level of about 10% in the presence of domain a or 15% in the presence of domain a’ (Fig. 7, B and C), as compared with about 5% in the spontaneous process (33). The overall folding pathway observed in the presence of the single domains a and a’ is very similar to that of spontaneous folding (33), suggesting that these domains alone produce only minor, if any, increases in the rate of folding at early time points by comparison with the spontaneous process and no increase at later time points. The corresponding domains of PDI clearly catalyzed the disappearance of the reduced protein and formation of species containing intramolecular disulfides at the early stages in the process, whereas they showed negligible catalysis of the appearance of the fully oxidized protein (30), being even less efficient than the spontaneous process or the ERp57 domains a and a’ (Table I).

Folding experiments were also performed using a 50 μM concentration of HIS-ERp57 and the domains a and a’ to increase the level of activity (data not shown). The folding pathways were very similar to those observed with a 10 μM concentration of the polypeptides, indicating that the concentration of the catalysts is not a limiting factor but that they have a markedly different behavior from that exhibited by PDI and its domains (30).

Association with the Calreticulin P Domain in Vitro Requires Multiple ERp57 Domains—Chemical cross-linking of ERp57 and its domains a, b, b’, b’a’, and a’ with the P domain of rat CRT in the presence of disuccinimidyl glutarate resulted in the appearance of an additional band corresponding to a polypeptide with a molecular mass about 11 kDa larger than that of the original polypeptide only in the presence of ERp57 (Fig. 8, lane 14), suggesting that no association had taken place between the P domain of CRT and the individual ERp57 domains nor the b’a’ double domain (Fig. 8, lanes 3, 5, 7, 9, and 11). Further experiments performed with two soluble full-length hybrid polypeptides, ERpabb’PDIa’c and ERpabPDIb’a’c, showed formation of an additional band with retarded migration in the case of ERpabb’PDIa’c but not ERpabPDIb’a’c (Fig. 8, lanes 16 and 18, respectively). Our data thus demonstrate that none of

---

**Fig. 6.** CD spectrum analysis of purified ERp57 and its domains. Far-UV CD spectra of the polypeptides are shown.
the isolated domains nor the domain pair \( b'a' \) is sufficient for interaction with the P domain of CRT, whereas the ERp57 domain construct \( abb' \) fulfills the minimum requirement.

**DISCUSSION**

Cloning of the human ERp57 polypeptide indicated that its amino acid sequence is 29% identical and more than 50% similar to that of PDI, which suggested that ERp57 may also possess a four-domain structure similar to that of PDI (6). The data presented here provide biochemical evidence for this suggestion, although some minor differences were found between these two polypeptides in the locations of the protease-sensitive regions and the presumed lengths of some of the domains.

The first protease-sensitive region was found to begin around residue 120, which appeared to be the main site, and to extend until residue 140–150, although recombinant expression in *E. coli* suggested that domain \( a \) may extend only to residue 113. Our preliminary NMR data in \( ^{15}N \) and \( ^{15}N/^{13}C \)-labeled ERp1–113 polypeptides indicate that the last \( \alpha \) helix of the folded domain extends to Lys\(^{105} \) (Fig. 3). It is currently unknown whether there is a short linker region between domains \( a \) and \( b \) or whether domain \( b \) begins around residue 113, although the protease-sensitive region begins only around residue 120. Predictions of secondary structures using two programs suggest that an \( \alpha \) helix possibly located in domain \( b \) may begin around Glu\(^{118} \) (Fig. 3).

2 L. Silvennoinen, P. Koivunen, J. Myllyharju, K. I. Kivirikko, I. Kilpelainen, and P. Permi, manuscript in preparation.
TABLE I
Folding rates at early stages of the folding and times required for the appearance of 50% of the 4S species in the folding of RNase A

| Catalyst | 4S, a,b | Appearance of 50% of 4S b |
|----------|--------|--------------------------|
|          | %/h    | min                      |
| Spontaneous folding c | 10     | 400                      |
| ERp57    | 105    | 36                       |
| ERp57 domain a | 9      | 420                      |
| ERp57 domain a’ | 15     | 420                      |
| PDI      | 170    | 27                       |
| PDI domain a’ d | 6      | >420                     |
| PDI domain a’ d | 6      | >420                     |

The formation rate of the 4S species (%/h) was calculated from a linear reaction curve when 35% of the 4S species had been formed, and the levels of 1S, 2S, and 3S intermediates were relatively stable, indicating that the forward and reverse reactions were balanced and that the species were at a temporary kinetic equilibrium. The rate was found to be linear in all reactions except that catalyzed by PDI (30), most probably due to it being limited in other reaction steps. The reactions were carried out in the presence of 10 μM of the catalyst.

No catalyst was added, see Ref. 33. See Ref. 30.

The second protease-resistant fragment, domain b’, but seven protease-resistant fragments had been obtained only along with the rest of the polypeptide or together with domain b’, but seven protease-resistant fragments had their N termini between residues 201 and 215. The location of a domain boundary within this region is supported by data indicating that the expression of the constructs ERp107–216 and ERp211–348 gave folded polypeptides, evidently corresponding to domains b and b’. Secondary structure predictions suggest that the last a helix of domain b may extend to Glu314 or Asn315, i.e. to the C-terminal end of the protease-sensitive region (Fig. 3). Domain b may thus be slightly shorter than the corresponding PDI domain (26), due to shortening of its C-terminal end (Fig. 3).

The third protease-sensitive region was found to be located between residues 313 and 341, especially in the interval 326–348, giving folded polypeptides, evidently corresponding to domains b’ and a’. The actual length of domain b’ is likely to be very similar to that of PDI.

NMR data indicate that the last a helix of domain a’ ends at Arg258 (27), ERp57 thus having 23 additional residues in the C terminus of this domain. This stretch is distinctly shorter than the 37-residue highly acidic C-terminal extension c present in the C terminus of PDI (5). This PDI extension contains the ER retention signal -Lys-Asp-Glu-Leu, but otherwise the presence or absence of this extension has no effect on any of the main functions of the polypeptide (37). The extension present in the C terminus of ERp57 also contains the ER retention signal in the form of -Glu-Glu-Asp-Leu, but this extension differs distinctly from that of PDI in containing many basic rather than acidic residues.

Domains a and a’ of ERp57 have been shown2 (27) to have the thioredoxin fold. The CD spectra recorded here for domains b and b’ were slightly different from those of domains a and a’, suggesting minor differences in their secondary structure components (Fig. 6). Domains b and b’ of PDI also have the thioredoxin fold, and secondary structure predictions (Fig. 3) suggest that these two ERp57 domains may well also have this fold. It is nevertheless very likely that there are specific differences in the structures of the corresponding domains between ERp57 and PDI, as ERp57 domains b and a’ cannot substitute for the corresponding PDI domains in the assembly of the collagen prolyl 4-hydroxylase αβ tetramer and as ERp57 domains a and b can only in part substitute for the corresponding PDI domains in this assembly (11).

ERp57 was found to markedly enhance the folding of denatured RNase A, a protein with four intramolecular disulfide bridges, during the early stages of the process (Fig. 7A). Nevertheless, ERp57 was not as efficient as PDI during the early stages, and it was a much less efficient catalyst of the overall process than PDI. ERp57 differs distinctly from PDI in being a co-chaperone in glycoprotein folding (1, 2) and interacting with its glycoprotein substrates only after being bound to CNX or CRT, which enhances its activity (15, 20). It is not surprising therefore that ERp57 was an inefficient catalyst in the folding of denatured RNase A, as this enzyme contains no post-translationally added carbohydrate moieties and as the experiments were consequently performed in the absence of any lectin.

The recombinant catalytic domains a and a’ of ERp57 were found to produce only minor, if any, increases in the folding rate at early time points, being clearly less effective than the corresponding domains of PDI in catalyzing the disappearance of the reduced protein and the formation of species containing intramolecular disulfides at the early stages in the process (30), whereas these domains showed negligible catalysis of the appearance of the fully oxidized protein in both polypeptides (Table I). The differences in the catalytic behavior between the a and a’ domains of ERp57 and PDI must be due to intrinsic properties of the domains, as no additional domains were present.

The primary peptide substrate binding site in PDI is located in domain b’, and this domain alone is sufficient for the binding of peptides of 10–15 amino acids (38). The binding of longer peptides such as scrambled RNaseA nevertheless requires the presence of either PDI bbb’ or b’a’ (11, 38). In the binding of long peptides to PDI, the a’ domain can be replaced by the a domain, whereas this substitution cannot be made in the assembly of the collagen prolyl 4-hydroxylase αβ tetramer, the PDI domains b’ and a’ being absolute requirements (11). We found here no association of the P domain of CRT with any single ERp57 domain, nor with the double domain b’a’, indicating that this association requires several domains and that, unlike the case of the binding of RNase A to PDI (11), the ERp57 double domain b’a’ is not sufficient. The finding that the P domain could be cross-linked to the hybrid ERp57-PDIc’ but not to ERp57-PDIb’a’c indicates that the b’ domain of ERp57 is an absolute requirement, but our data do not indicate whether the double domain b’b’ would have been sufficient or...
whether this binding, like that of long peptides to PDI, requires all three domains α, β, and β'.

Acknowledgments—The expert technical assistance of Anne Kokko, Eeva Lehtimäki, Outi Manty, Merja Nissilä, and Jaana Träskelin is acknowledged with thanks.

REFERENCES
1. High, S., Lecomte, F. J. L., Russell, S. J., Abell, B. M., and Oliver, J. D. (2000) FEBS Lett. 476, 38–41
2. Ellgaard, L., and Helenius, A. (2003) Nat. Rev. Mol. Cell. Biol. 3, 181–191
3. Freedman, R. B., Hirst, T. R., and Tuite, M. F. (1994) Trends Biochem. Sci. 19, 331–336
4. Ferrari, D. M., and Soling, H.-D. (1999) Biochem. J. 339, 1–10
5. Kemmink, J., Darby, N. J., Dijkstra, K., Nilges, M., and Creighton, T. E. (1997) Curr. Biol. 7, 235–245
6. Koivunen, P., Helaakoski, T., Annunen, P., Veijola, J., Räisänen, S., Pihlajaniemi, T., and Kivirikko, K. I. (1996) Biochem. J. 316, 599–605
7. Koivunen, P., Horelli-Kuitunen, N., Helaakoski, T., Karvonen, P., Jaakkola, M., Palöst, A., and Kivirikko, K. I. (1997) Genomics 42, 397–404
8. Kivirikko, K. I., and Myllyharju, J. (1998) Matrix Biol. 17, 357–368
9. Kivirikko, K. I., and Pihlajaniemi, T. (1998) Adv. Enzymol. Relat. Areas Mol. Biol. 72, 325–398
10. Myllyharju, J. (2003) Matrix Biol. 22, 15–24
11. Pirneskoski, A., Ruddock, L. W., Klapa, P., Freedman, B. B., Kivirikko, K. I., and Koivunen, P. (2001) J. Biol. Chem. 276, 11287–11293
12. Srivastava, S. P., Fuchs, J. A., and Holtzman, J. L. (1995) Biochem. Biophys. Res. Commun. 193, 971–976
13. Bourdi, M., Demady, D., Martin, J. L., Jabbour, S. K., Martin, B. M., George, J. W., and Pohl, L. R. (1995) Arch. Biochem. Biophys. 323, 397–403
14. Hirano, N., Shibasaka, F., Sakai, R., Tanaka, T., Nishida, J., Yazaki, Y., Takenawa, T., and Hirai, H. (1995) Eur. J. Biochem. 234, 336–342
15. Zupan, A., Darby, N. J., Tessier, D. C., Michalak, M., Bergeron, J. M. J., and Thomas, D. Y. (1998) J. Biol. Chem. 273, 6009–6012
16. Molinari, M., and Helenius, A. (1999) Nature 402, 90–93
17. Farmery, M. R., Allen, S., Allen, A. J., and Bulleid, N. J. (2000) J. Biol. Chem. 275, 14933–14938
18. Oliver, J. D., van der Wal, F. J., Bulleid, N. J., and High, S. (1997) Science 275, 86–88
19. Elliott, J. G., Oliver, J. D., and High, S. (1997) J. Biol. Chem. 272, 13849–13855
20. Oliver, J. D., Roderick, H. L., Llewellyn, D. H., and High, S. (1999) Mol. Biol. Cell 10, 2575–2582
21. Hammond, C., Braakman, I., and Helenius, A. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 913–917
22. Peterson, J. R., and Helenius, A. (1999) J. Cell Sci. 112, 2775–2784
23. Ellgaard, L., Riek, R., Herrmann, T., Güntert, P., Helenius, A., and Wuthrich, K. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 3133–3138
24. Kemmink, J., Darby, N. J., Dijkstra, K., Nilges, M., and Creighton, T. E. (1996) Biochemistry 35, 7684–7691
25. Dijkstra, K., Karvonen, P., Pirneskoski, A., Koivunen, P., Kivirikko, K. I., Darby, N. J., van Straaten, M., Scheek, R. M., and Kemmink, J. (1999) J. Biol. NMR 14, 195–196
26. Kemmink, J., Dijkstra, K., Mariani, M., Scheek, R. M., Penka, E., Nilges, M., and Darby, N. J. (1999) J. Biol. NMR 13, 357–368
27. Silvennoinen, L., Karvonen, P., Koivunen, P., Myllyharju, J., Kivirikko, K. I., and Kilpeläinen, I. (2001) J. Biol. NMR 20, 385–386
28. Gruenwald, S., and Heitz, J. (1994) Baculovirus Expression Vector System, Procedures and Methods Manual, Pharmingen, San Diego, CA
29. Terrella, C., Ruoppolo, M., Marino, G., and Pucci, P. (1994) FEBS Lett. 352, 301–306
30. Ruoppolo, M., Orrù, S., Talamo, F., Ljung, J., Pirneskoski, A., Kivirikko, K. I., Marino, G., and Koivunen, P. (2003) Protein Sci. 12, 939–952
31. Frickel E.-M., Riek, R., Jelesarov, I., Wuthrich, K., and Ellgaard, L. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 19584–19589
32. Lambert, N., and Freedman, R. B. (1983) Biochem. J. 213, 225–234
33. Ruoppolo, M., Lundström-Ljung, J., Talamo, F., Pucci, P., and Marino, G. (1997) Biochemistry 35, 11259–11267
34. Ruoppolo, M., Vinci, F., Klink, T. A., Raines, R. T., and Marino, G. (2000) Biochemistry 39, 12033–12042
35. Freedman, R. B., Gane, P. J., Hawkins, H. C., Hidman, R., McLaughlin, S. H., and Parry, J. W. (1998) Biol. Chem. 379, 321–329
36. Alanen, H. I., Salo, K. E., Pekkala, M., Siekkinen, H. M., Pirneskoski, A., and Ruddock, L. W. (2003) Antioxid. Redox Signal. 5, 367–374
37. Koivunen, P., Pirneskoski, A., Karvonen, P., Ljung, J., Helaakoski, T., Noth, H., and Kivirikko, K. I. (1999) EMBO J. 18, 65–74
38. Klappa, P., Ruddock, L. W., Darby, N. J., and Freedman, R. B. (1998) EMBO J. 17, 927–935
