# Two Distinct Forms of Peptidylprolyl-*cis-trans*-isomerase Are Expressed Separately in Periplasmic and Cytoplasmic Compartments of *Escherichia coli* Cells†

Toshiya Hayano, Nobuhiro Takahashi,* Setsuko Kato, Noboru Maki, and Masanori Suzuki
*Corporate Research and Development Laboratory, Tonen K. K., 1-3-1 Nishi-tsurugaoka, Ohi-machi, Iruma-gun, Saitama 354, Japan*

*Received August 10, 1990; Revised Manuscript Received December 6, 1990*

ABSTRACT: Peptidylprolyl-*cis-trans*-isomerase (PPIase) is thought to be essential for protein folding in the cell. Two forms, *a* and *b*, of PPIase and their corresponding genes were isolated from *Escherichia coli* cells. Despite their insensitivity to cyclosporin A (CsA), both amino acid sequences were homologous and related to that of pig cyclophilin, a protein that has PPIase activity sensitive to CsA (Takahashi et al., 1989). PPIase *a* is found to be identical with the *E. coli* ORF 190 gene product that was sequenced by Kawamukai et al. (1989) and overexpressed by Liu and Walsh (1990). It is translocated into *E. coli* periplasmic space with the signal sequence. PPIase *b* lacks a hydrophobic amino acid stretch which could serve as a signal sequence or a transmembrane domain, and it is detected mainly in the bacterial cytoplasm. These findings indicate that proteins with the ability to assist folding of various polypeptides are located on both sides of the inner membrane. Thus, we propose that the folding of some exported proteins may be catalyzed by the periplasmic proline isomerase and, in turn, that some proteins which have isomerized may not be translocated efficiently.

Peptidylprolyl-*cis-trans*-isomerase (PPIase) catalyzes the cis-trans isomerization of Xaa–Pro peptide bonds in oligopeptides and accelerates the slow, rate-limiting steps in the refolding of several proteins in vitro (Fischer et al., 1984; Lang et al., 1987; Lin et al., 1988); it is thought to be essential for protein folding during protein synthesis in the cell (Fischer et al., 1984). The action of PPIase, however, appears to be restricted to specific proteins in vitro (Lang et al., 1987; Lin et al., 1988). Protein–disulfide isomerase (PDI), another enzyme known to catalyze protein folding, was proposed to accelerate folding, not by reshuffling incorrect disulfide bonds, but in the same way as PPIase by catalyzing proline isomerization (Pain, 1987). However, acceleration of the reshuffling of disulfide bridges during reoxidation of some proteins is unquestionably catalyzed by PDI. Furthermore, it has been confirmed that PDI and PPIase act differently and independently as catalysts of protein folding (Lang & Schmid, 1988).

PPIase is found to be identical with cyclophilin, a protein known to bind the immunosuppressive drug cyclosporin A (CsA) (Takahashi et al., 1989), and its activity is inhibited by CsA (Takahashi et al., 1989; Fischer et al., 1989). These findings led to the hypothesis that the effect of CsA is mediated through inhibition of the peptidylprolyl-*cis-trans* isomerizing activity of PPIase in T-cell activation (Takahashi et al., 1989). After these findings, a cellular binding protein for FK506, an immunosuppressant remarkably similar to CsA, was also found to have the same enzymatic activity as cyclophilin (Siekierka et al., 1989; Harding et al., 1989). However, cyclophilin and FK506 binding protein (FKBP) are quite distinct in terms of ligand specificity: cyclophilin binds to and is inhibited by CsA but shows no recognition of FK506, whereas the converse holds for FKBP (Siekierka et al., 1989; Harding et al., 1989). These two PPIases are unrelated to each other in their entire amino acid sequences, indicating that PPIases are classified into at least two different superfamilies, cyclophilin and FKBP superfamilies (Maki et al., 1990). On the basis of earlier data, which indicated that CsA and FK506 act on T-lymphocytes in an essentially equivalent fashion (Thomson, 1989), these findings indicate that the two drugs act through distinct pathways but that their mode of action converges on PPIase activities (Freedman, 1989), and suggest that PPIase activity is essential for T-cell activation.

One action of CsA thought to be central to many of its immunosuppressive effects is its ability to suppress the events occurring early in T-cell activation, such as lymphokine gene transcription in response to signals initiated at the antigen receptor (Elliott et al., 1984; Kronke et al., 1984). Recently CsA was found to specifically inhibit the appearance of DNA binding activity of nuclear proteins involved in T-cell activation (Emmel et al., 1989). Since CsA inhibits the activity of PPIase, the PPIase activity may be required for DNA binding and/or transcriptional activation of nuclear proteins (Emmel et al., 1989).

The CsA binding and PPIase activities are present in nearly all organs as well as in species as varied as mammals and fungi (Koletsky et al., 1986; Fischer et al., 1984). In *Neurospora crassa* and *Saccharomyces cerevisiae*, cyclophilin mediates the cytotoxic effect of CsA (Tropschug et al., 1989). This finding, coupled with the result that the PPIase activity of cyclophilin is inhibited by CsA, suggests that cytotoxic effects of CsA on many organisms of widely diverging phylogenic origin are also mediated through inhibition of the PPIase activity. However, it is not yet known whether or not the mechanism of CsA action in these organisms is similar to that proposed for T-cell activation, e.g., that CsA inhibits the appearance of DNA binding and/or transcriptional activities of nuclear proteins in T-cells. In any case, the presence of cyclophilin-like PPIase activity across many species and the widespread effects of CsA, in conjunction with its inhibitory effect on PPIase, suggest that the control of the proline isomerization may be a ubiquitous regulatory mechanism in some cellular events, such as those in cell growth or differentiation.

On the basis of the probable participation of a cyclophilin-like protein in the formation of rhodopsin in the pho-

† The nucleic acid sequences in this paper have been submitted to GenBank under Accession Numbers M55429 (isomerase *a*) and M55430 (isomerase *b*).

* To whom correspondence should be addressed.

0006-2960/91/0430-3041$02.50/0 

toreceptor cell of *Drosophila*, it was proposed that the PPIase activity is necessary for the posttranslational folding of rhodopsin (Schnuewly et al., 1989; Shieh et al., 1989). In this case, however, the existence of a cell-specific isoform was postulated to explain the specific action of cyclophilin-like PPIase on the formation of rhodopsin in the photoreceptor cells. In fact, a number of cyclophilin-related DNA sequences have also been detected in mammalian genomes (Haendler et al., 1987; Danielson et al., 1988), suggesting the presence of a large family of cyclophilin-like PPIases encoded by multiple genes, and the functional diversity of cyclophilin-like PPIase within a single species. In this report, we show the presence of two distinct forms of cyclophilin-like PPIase encoded by two separate genes and their possible functional roles in different cellular compartments of the *Escherichia coli* cell.

## Experimental Procedures

*Purification of Escherichia coli PPIases a and b.* Two hundred grams of *E. coli* ST249 cells (Kajie & Anraku, 1986) was treated with 2 g of lysozyme in 500 mL of 0.1 M Tris-HCl buffer (pH 7.8) containing 5 mM 2-mercaptoethanol for 30 min. The soluble fraction was concentrated with ammonium sulfate fractionation and dialyzed against 10 mM Tris-HCl buffer (pH 7.8) containing 0.05% $NaN_3$. The dialyzate was applied on a DEAE-Sepharose CL-6B column (2.5 cm × 40 cm) and eluted by stepwise increasing NaCl concentrations: 0, 0.05, 0.1, 0.2, 0.3, and 0.5 M. The PPIase activity was detected in the eluents with 0 and 0.1 M NaCl. The two fractions with PPIase activity were collected separately, and each of the collections was applied on a Sephadex G-75 column (2.5 cm × 90 cm) equilibrated with 0.15 M NaCl/10 mM Tris-HCl buffer (pH 8.0) containing 0.05% $NaN_3$. Form *b* was purified as a single band by SDS/PAGE at this step. Form *a* was further purified on a CM-Sepharose CL-6B column (1.5 cm × 20 cm) with linear gradient elution (from 0 to 0.25 M NaCl/10 mM sodium acetate buffer, pH 6.0). Electrofocusing was done with Ampholine, PAGPLATE, pH 3.5–9.5, by a Multiphor II electrophoresis apparatus (Pharmacia LKB Biotechnology, Tokyo, Japan). SDS/PAGE was done by a method modified from that of Laemmli with a gradient gel from 12% to 30% (Laemmli, 1970; Takahashi et al., 1987).

*Assay of PPIase Activity and Effects of CsA and FK506 on the Activity.* The cis-trans isomerization of the Ala–Pro peptide bond in the synthetic peptide *N*-succinyl-Ala-Ala-Pro-Phe-4-methylcoumaryl-7-amide (MCA) was measured in a coupled assay with chymotrypsin, based on the ability of this protease to cleave the synthetic peptide only when the Ala-Pro is in trans configuration. The synthetic peptide (50 μL of a 1.68 mM solution) was preincubated with the appropriate concentrations of PPIase and CsA in 2 mL of 0.035 M HEPES buffer containing 5 mM 2-mercaptoethanol, pH 7.8, and the assay was started by mechanical mixing with 20 μL of 0.76 mM chymotrypsin (Sigma) in a spectrophotometer cell (Hitachi Spectrophotometer U-3210 with temperature controller SPR-7 and stirrer) at 25 °C. The trans peptide was cleaved by the deadline; hydrolysis of the MCA in the cis peptide is limited in rate by cis-trans isomerization at Ala-Pro and was monitored by the increase in absorbance at 360 nm. The inhibitory effects of CsA and FK506 on PPIase activity were measured by determining the residual activity of PPIase preincubated in the presence of various concentrations of CsA and FK506, respectively.

*Protein Sequencing Analysis.* The amino acid sequence analysis was carried out by automated Edman degradation using an Applied Biosystems 477A protein sequencer equipped with a Model 120A on-line PTH analyzer. The native protein (50–200 μg) was subjected to separate digestions with L-1-(tosylamido)-2-phenylethyl chloromethyl ketone treated trypsin and cyanogen bromide. The digests were separated on an Aquapore RP-300 column or on a Spheri-5RP-18 column using a 130-Å separation system (Applied Biosystems). The peptides were eluted from the columns at a flow rate of 200 μL/min with a linear gradient of acetonitrile (0–100%) containing 0.1% trifluoroacetic acid or 0.1% heptafluorobutyric acid for 45 min. The amino acid analysis was carried out with a JEOL JLC-300 amino acid analyzer (Nihon Denshi, Tokyo, Japan) after hydrolysis with 6 M HCl for 24 h.

*Construction and Screening of the Genomic Library.* To detect genomic DNA fragments carrying the *E. coli* PPIase *b* gene, the high molecular weight DNA isolated from *E. coli* HB101 (Marmur, 1961) was digested with various restriction endonucleases and was subjected to Southern blot analysis (Southern, 1975) with the radiolabeled mixed oligonucleotide probes 5′ATGAAA/GCAA/GAAA/GGCNACCAAA-GAACC3′, corresponding to the amino acid sequence Met-Lys-Gln-Lys-Ala-Thr-Lys-Glu-Pro which was obtained from protein sequencing of PPIase *b*. A genomic DNA fragment of about 1 kb generated by *Hin*dIII/*Bgl*II digestion was hybridized to the probes. The *Hin*dIII/*Bgl*II fragments, whose sizes were roughly 1 kb, were fractionated by 0.8% agarose gel electrophoresis subcloned into the *Hin*dIII/*Bam*HI site of pUC118 (TAKARA SHUZO, Kyoto, Japan). *E. coli* JM109 cells were transformed with the plasmid library, and the transformants were screened for the *E. coli* PPIase *b* gene with the probe described above. The hybridization was performed overnight at 37 °C in 50 mM Tris-HCl buffer, pH 7.5, containing 1 M NaCl, 10 mM EDTA, 0.1% sodium dodecyl sarcosinate, 0.1% poly(vinylpyrrolidone), 0.1% Ficoll, 0.1% bovine serum albumin, and 20 μg/mL sheared salmon sperm DNA. The filters were washed with 3 × SSC (1 × SSC: 0.15 M NaCl/0.015 M sodium citrate, pH 7.0) and 0.1% sodium dodecyl sarcosinate at 37 °C for 30 min. Autoradiography was carried out at −70 °C with an intensifying screen. One out of approximately 1500 recombinant clones, designated as pEPPIb, gave a positive hybridization signal. To isolate a genomic DNA fragment coding for the PPIase *a* gene, a genomic library for *E. coli* strain HB101 was constructed as follows. The high molecular weight DNA isolated from *E. coli* HB101 cells (Marmur, 1961) was partially digested with *Sau*3AI, followed by fractionation through a sucrose density gradient (Maniatis et al., 1982). The 10–20-kb genomic DNA fragments were recovered and ligated to EMBL4 *Bam*HI arms (STRATAGENE). After in vitro packaging, *E. coli* strain LE392 was transfected with the EMBL4 library, and the transfectants were screened for PPIase *a* gene inserts with radiolabeled mixed oligonucleotide probes 5′GAACAA/GATGCAA/GCAA/GAAA/GAAACC3′, corresponding to the amino acid sequence Glu-Gln-Met-Gln-Gln-Lys-Lys-Pro, obtained from protein sequencing of PPIase *a*. The hybridization was performed as in the case of the screening for the PPIase *b* gene described above. Of approximately $10^4$ phages screened, 5 clones gave positive signals after subtraction using the PPIase *b* gene as a probe. The restriction endonuclease analysis showed that these five clones overlapped with each other; one of them, designated as λEPPIa, was further analyzed. A fragment of approximately 2.5 kb size was generated by *Ava*I digestion of λEPPIa DNA which hybridized to the oligonucleotide probe used for the screening described above. It was subcloned into the *Sma*I site of pUC119 after filling its protruding ends with Klenow fragments (pEPPIa).

*DNA Sequencing Analysis.* The entire insert DNA for pEPPIb and the restricted DNA insert containing the PPIase *a* gene for pEPPIa were sequenced after subcloning into M13mp18, M13mp19, pUC118, or pUC119(TAKARA SHUZO). Single-strand templates were prepared according to standard procedures, and dideoxy sequencing analyses (Sanger et al., 1977) were performed with a 7-deaza sequencing kit (TAKARA SHUZO or TOYOBO), using the universal primer except for the oligonucleotide (5′ATAAAGATACGGATCGCGCCCTTCGCCGAA3′) used for the sequencing of a part of the PPIase *a* gene, according to the procedures recommended by the manufacturer.

*Computer Analysis.* The computer-assisted homology analysis compared the *E. coli* PPIase genes and proteins to the data in the NBRF, GenBank, and EMBL databases accessed through DDBJ (1989) [DNA Data Bank of Japan (National Institute of Genetics Information Analysis, Mishima, Shizuoka 411, Japan)]. SEQFP and SEQHP programs for the homology search (Goad & Kanehisa, 1982) and the HPLOT program for hydropathy calculation (Kyte & Doolittle, 1982) were provided by DDBJ.

*Southern Blot Analysis for E. coli Total DNA.* High molecular weight DNA (2 μg) isolated from *E. coli* W3110 cells (Marmur, 1961) was digested with *Eco*RV, *Pst*I, or *Pvu*II (TAKARA SHUZO or NIPPON GENE) according to the procedures recommended by the vendors. The digests were subjected to 0.8% agarose gel electrophoresis and transferred to nylon membranes (Hybond N, Amersham) as described (Southern, 1975). The filters were probed with the *Pvu*I/*Eco*RI DNA fragment of pEPPIb or the *Ava*II DNA fragment radiolabeled with random primer DNA labeling kit (TAKARA SHUZO). The hybridization was carried out overnight at 68 °C in 6 × SSC containing 0.5% SDS, 0.2% poly(vinylpyrrolidone), 0.2% Ficoll 400, 0.2% bovine serum albumin, and 20 μg of sheared salmon sperm DNA. Following hybridization, the filters were washed with 0.1 × SSC containing 0.1% SDS at 68 °C for 30 min. Autoradiography was performed at −70 °C using an intensifying screen for the appropriate times.

*Separation of Periplasmic and Cytoplasmic Proteins.* Ten grams of *E. coli* cells (wet weight) was suspended in 80 mL of 20% sucrose/10 mM Tris-HCl buffer, pH 7.5, and incubated with the addition of 5.4 mL of 0.25 M EDTA solution (adjusted pH to 7.5) at 4 °C for 15 min. *E. coli* cells were collected by centrifugation, and the periplasmic fraction was obtained by the osmotic shock method, in which the cells were resuspended in 120 mL of cold water by vortexing. After centrifugation, the precipitate was suspended in 50 mL of 0.1 M Tris-HCl (pH 7.5) containing 5 mM 2-mercaptoethanol and treated with 11.8 mg of lysozyme at 30 °C for 1 h. The supernatant contained the cytoplasmic fraction. Protein determination was done by using a Bio-Rad protein assay kit (Richmond, CA). β-Galactosidase activity was measured by the method described by Maniatis et al. (1982).

The periplasmic and cytoplasmic fractions were separated on a TSK-gel DEAE-5PW column (7.5 mm i.d. × 7.5 cm, Toso, Tokyo, Japan) equilibrated with 10 mM Tris-HCl buffer, pH 8.0, containing $NaN_3$. Proteins were eluted at a flow rate of 1 mL/min with a linear gradient of NaCl (0–0.25 M) during 60 min after the elution of the equilibration buffer for 10 min, and the eluent was collected at an interval of 1 min. Some of the major peaks obtained were purified further by reversed-phase chromatography on an Aquapore RP-300 column (2 mm i.d. × 3 cm, Applied Biosystems) to identify the proteins eluted and to compare their levels in the cytoplasmic and periplasmic fractions. The column was eluted at a flow rate of 200 μL/min with a linear gradient of acetonitrile (0–100%) containing 0.1% trifluoroacetic acid for 45 min. Identification of the purified proteins was done by computer search on the databases described above after determination of their amino-terminal sequences. The protein content was compared by calculating the total absorbance of the corresponding protein peaks originating from the cytoplasmic and periplasmic fractions.

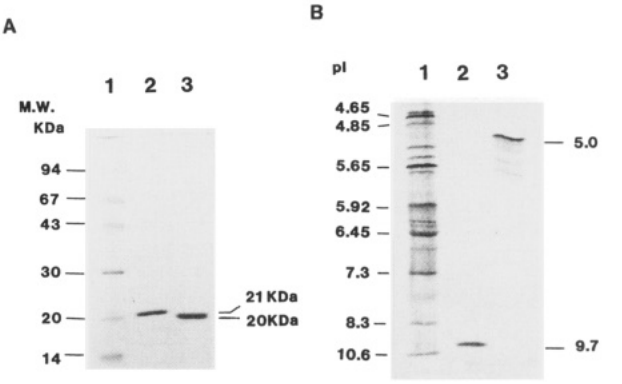


FIGURE 1: Two forms of *E. coli* PPIase. (A) SDS/PAGE and (B) electrofocusing of the purified *E. coli* PPIases *a* and *b*. Lane 1, standard; lane 2, *E. coli* PPIase *a*; and lane 3, *b*.

## Results

*Two Distinct Forms of PPIase Are Present in E. coli Cells.* Two peaks of PPIase activity were detected in the chromatogram of the extract of *E. coli* ST249 cells on a DEAE-Sepharose CL-6B column; one passed through the column, and the other was eluted with 0.1 M NaCl. They were separately purified further as described under Experimental Procedures. Each of the two peaks contained a single component of PPIase. The two PPIases obtained were designated as *a* and *b* which have apparently different isoelectric points of 9.7 and 5.0, and molecular weights of 21 000 and 20 000, respectively (Figure 1A,B). *a* and *b* were present in a ratio of about 1:9, and their specific PPIase activities were almost equal when the synthetic peptide described (Takahashi et al., 1989) was used for the substrate in the enzyme assay (*a*, 56 600 units/mg; *b*, 51 500 units/mg). These values were very close to that of pig PPIase (57 700 units/mg). A comparison of their tryptic peptide maps indicated that they are apparently different molecular species (data not shown). This is in contrast to the case of the isoforms of pig PPIase reported previously; the structural difference between the pig isoforms is probably the result of incomplete posttranslational modification at the amino terminus of the PPIase molecule (Takahashi et al., 1989). The amino-terminal sequence analyses of the two PPIases gave entirely different sequences for the first 18 amino acid residues, i.e., Ala-Lys-Gly-Asp-Pro-His-Val-Leu-Leu-Thr-Thr-Ser-Ala-Gly-Asn-Ile-Glu-Leu for *a* and Met-Val-Thr-Phe-His-Thr-Asn-His-Gly-Asp-Ile-Val-Ile-Lys-Thr-Phe-Asp-Asp- for *b*. These results, in combination with those of the peptide mapping, indicate that at least two distinct PPIases are present in *E. coli* cells.

*E. coli PPIases Are Insensitive to both CsA and FK506.* Since the two forms of PPIase were found to be present in *E. coli* cells, their sensitivity to immunosuppressants, CsA and FK506, was examined to show the relationships to the known type of PPIases, CsA-sensitive and FK506-sensitive PPIases. However, the activity of neither of the PPIases was inhibited by CsA in the same concentration range as required for the inhibition of pig PPIase (Takahashi et al., 1989a; Fischer et

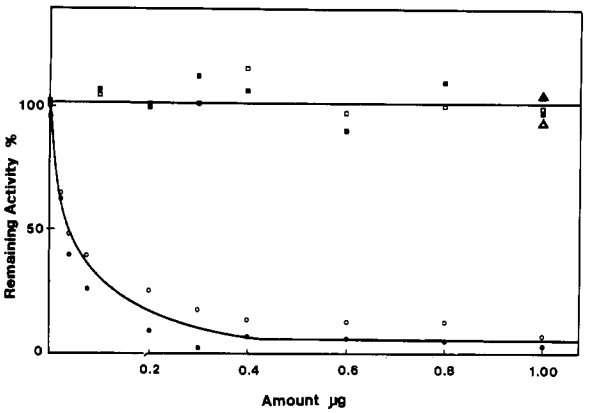


FIGURE 2: Effects of CsA and FK506 on the activities of *E. coli* PPIases. The activity of PPIases was determined by measuring the catalysis of the cis-trans interconversion of *cis-N*-succinyl-Ala-Ala-Pro-Phe-MCA (Takahashi et al., 1989). Residual PPIase activity is plotted against increasing amounts of CsA and FK506. PPIase (27 ng/2 mL) was preincubated in the presence of varying concentrations of CsA or FK506, respectively, and the remaining PPIase activity was determined. For CsA: (□) *E. coli* PPIase *a*; (■) *E. coli* PPIase *b*; (●) yeast PPIase; (○) pig PPIase. Only the remaining activity at the highest concentration is plotted for FK506: (▲) *E. coli* PPIase *a*; (△) *E. coli* PPIase *b*.

al., 1989), nor was either sensitive to FK506 (Figure 2). The result for CsA is consistent with our observation that the growth of *E. coli* cells was not affected by the presence of CsA in the culture medium, and with the fact that no CsA binding activity was detected in *E. coli* extracts despite its ubiquitous presence in all eukaryotes so far analyzed (Koletsky et al., 1986; Fischer et al., 1984).

*Cloning of the Two PPIase Genes.* Genomic libraries were constructed from the high molecular weight DNA isolated from *E. coli* HB101 cells. The libraries were screened with the synthetic oligonucleotide probes designed on the basis of the amino acid sequences obtained by direct protein sequencing of the purified PPIases as described under Experimental Procedures. Five positive clones for the PPIase *a* gene and one positive clone, designated pEPPIb, for the *b* gene were isolated. The five clones for the *a* gene overlapped each other, and one of them, designated λEPPIa, was further analyzed. Southern blot and restriction mapping analyses of the λEPPIa resulted in the finding of an approximately 2.5-kb *Ava*I DNA fragment containing the PPIase *a* gene. This fragment was then subcloned into pUC19 (pEPPIa) and sequenced. The insert of pEPPIa contained an entire open reading frame encoding a polypeptide of 190 amino acids with a molecular weight of 20430 (Figure 3a). On the other hand, sequencing of the insert of pEPPIb revealed that it contained an open reading frame encoding a polypeptide of 164 amino acids with a molecular weight of 18184 (Figure 3b).

The amino acid sequences deduced by translation of the gene sequences indicated the presence of a single Met, an amino terminal to the amino-terminal sequence determined directly for purified PPIase *b*, and of a sequence of 24-residue amino terminal to that determined for purified PPIase *a*. The amino acid sequences of the peptides determined by direct protein sequencing of the cyanogen bromide and tryptic peptides obtained from the purified PPIases gave no discrepancies in the amino acid sequences deduced from the cloned DNA sequences for both PPIases except two amino acid residues in the sequence of form *a*, in which Ala and Asp were identified by direct peptide sequencing instead of Thr and Ser at positions 124 and 125, respectively (Figure 3a). Furthermore, the amino acid compositions and the molecular weight values calculated from the deduced amino acid sequences were in good agreement with those determined by amino acid analysis and SDS/PAGE for both PPIases isolated. These results indicate that the isolated DNA sequences represent the genes coding for *E. coli* PPIases *a* and *b*, respectively. Thus, the two *E. coli* PPIases are encoded separately by two different genes in the *E. coli* genome.

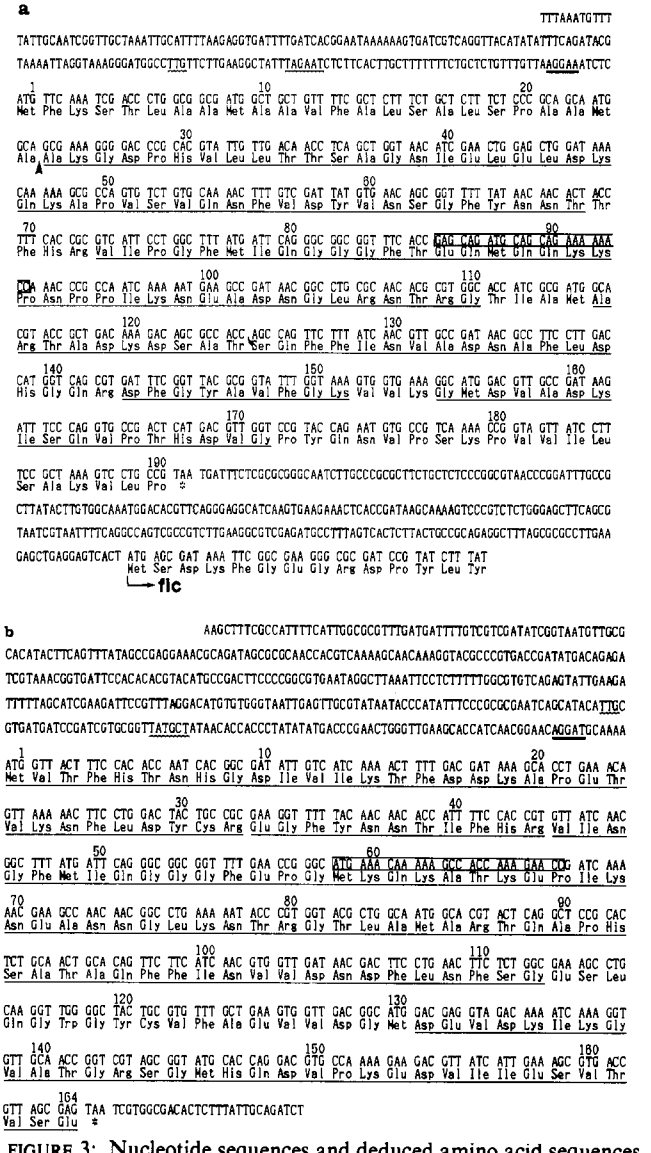


FIGURE 3: Nucleotide sequences and deduced amino acid sequences of *E. coli* PPIases *a* and *b*. (a) The nucleotide sequence of a portion of the subcloned *Ava*I fragment is shown for PPIase *a*. The deduced amino acid sequence is shown under the corresponding nucleotide sequence, and found to be identical with that of the hypothetical protein encoded by the ORF 190 sequence adjacent to the *fic* gene which is involved in the cell filamentation induced by cyclic AMP in *E. coli* (Kawamukai et al., 1989). The initiation codon of the *fic* gene starts 262 bp downstream of the termination codon of the PPIase *a* gene. The first 14 amino acid sequence of *fic* is indicated under the corresponding nucleotide sequence. The signal peptide processing site is indicated by an arrow. (b) the nucleotide sequence of the *Hin*dIII/*Bgl*II region of the isolated genomic fragment is shown for PPIase *b*. Amino acid sequences determined by the protein sequencing of tryptic and cyanogen bromide peptides of the purified PPIases *b* and *a*, respectively, are underlined. The oligonucleotides used as probes are indicated by a box. The termination codon is indicated by an asterisk. The predicted ribosome binding site is shown by a double line. Putative −35 and −10 (Pribnow box) regions are indicated by a wavy line.

To investigate the genomic complexity of *E. coli* PPIase genes, Southern blot analysis was performed. *E. coli* genomic DNA was separately digested with three different restriction enzymes and hybridized with the fragments of the *a* or *b* gene

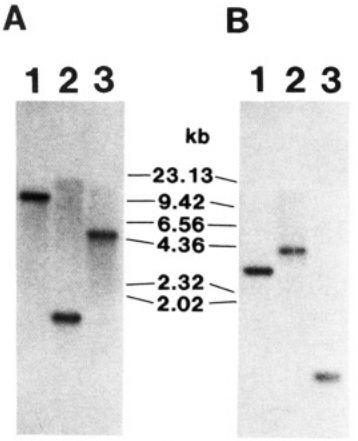


FIGURE 4: Genomic Southern blot analyses. *E. coli* W3110 DNA (2 μg) was digested with *Eco*RV (lane 1), *Pst*I (lane 2), and *Pvu*II (lane 3), respectively, subjected to 0.8% agarose gel electrophoresis, transferred to a nylon membrane, and probed with $^{32}$P-labeled *Ava*II DNA fragment for PPIase *a* (A) and *Pvu*I/*Eco*RI fragment for PPIase *b* (B). Marker sizes of the electrophoresis are indicated between each autoradiogram.

insert as probes under highly stringent conditions (Figure 4). Only one band for any of the restriction fragments was found to be hybridized with each of the two PPIase inserts, suggesting that only one copy of each of the two PPIase genes is present in *E. coli* genome.

Northern blot analysis showed that both *a* and *b* genes are transcribed in approximately 600-base RNA species (data not shown). Judging from the sizes of their coding regions, it can be concluded that both of the *E. coli* PPIase genes are transcribed in a monocistronic manner.

*Two E. coli PPIases Are Related to Cyclophilin in Their Amino Acid Sequences.* Despite the insensitivity of *E. coli* PPIases to CsA, the amino acid sequences of the known cyclophilins and *E. coli* PPIases are similar to one another. To search for the possible occurrences of other protein sequences homologous to the *E. coli* PPIases, computer-assisted homology analysis was done by comparing the *E. coli* PPIase genes and proteins to several databases. The amino acid sequences of *E. coli* PPIases were confirmed to have homology with that of bovine cyclophilin as expected, and were shown to have no sequence similarity to any other proteins or peptide segments in these databases used for the homology search. Despite the fact that the amino acid and nucleotide sequences of cyclophilin from many species other than bovine have been reported, these data had not yet been included in the databases updated as of the end of 1989. During a search of the literature, *E. coli* PPIase *a* was found to be identical with the ORF 190 gene that was sequenced by Kawamukai et al. (1989). A comparison of the amino acid sequences of *E. coli* PPIases and those of the other species indicated that the former, which are about 50% identical with each other, have about 25% homology with those of mammalian (Takahashi et al., 1989; Haendler et al., 1987; Danielson et al., 1988; Harding et al., 1986), fly (Schnuewly et al., 1989), fungi (Tropschug et al., 1988), and yeast PPIases (Haendler et al., 1989; Dietmeier & Tropschug, 1990), suggesting some conservation in the structures of PPIases from prokaryote to eukaryote (Figure 5). Although the homology among all molecules is much lower than that observed among eukaryotic cyclophilin/PPIases, two highly homologous regions were found in the middle of the sequences across several species including *E. coli*; one is located around residues 40–70 and the other around residues 90–120 of the PPIase sequences (i.e., the *E. coli* PPIases). On the other hand, the amino-terminal and carboxy-terminal regions of *E. coli* PPIase differ entirely from those of other species. On the basis of these structural features, the two regions with high homology seem to be of great significance for the PPIase activity. In contrast to the clear resemblance of *E. coli* PPIases with known cyclophilins, no significant homology to FKBP was found.

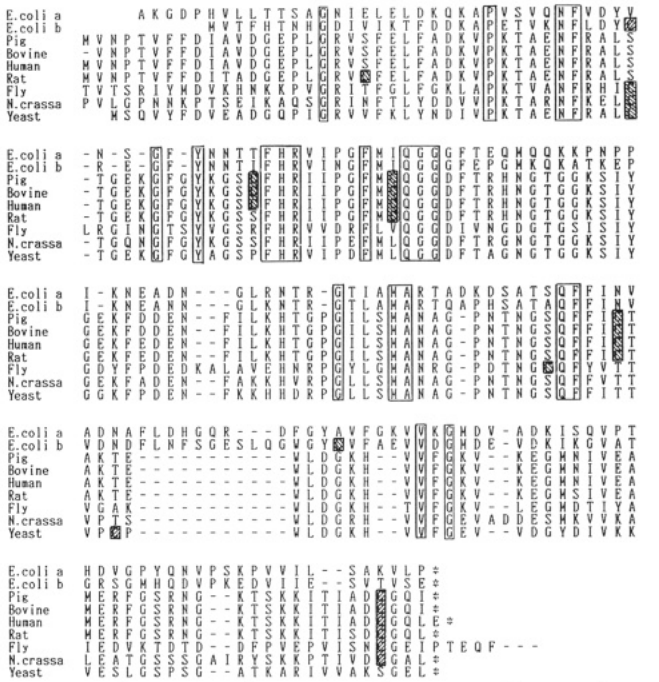


FIGURE 5: Comparison of the amino acid sequences of PPIases from mammal, fly, fungi, yeast, and *E. coli*. The amino acid residues identical throughout all species are boxed. Gaps are shown by dashes and are inserted in the sequences to maximize the sequence homology. The amino-terminal sequences of *N. crassa* and fly and the carboxy-terminal sequence of fly are not shown. Cysteine residues are shaded.

Although the involvement of a single sulfhydryl group in the activity of pig PPIase was suggested (Fischer et al., 1989), no conserved cysteine residues were found in the aligned PPIase sequences (Figure 5). In addition, *E. coli* PPIase *a* does not contain any cysteine residue in its amino acid sequence, indicating that the sulfhydryl group is not always involved in the active site of PPIase. In other words, the sulfhydryl group is not necessarily essential for PPIase to exhibit the enzymatic activity (Liu & Walsh, 1990). In fact, a sulfhydryl-modifying reagent, *p*-(hydroxymercuri)benzoate, which effectively inhibited the activity of pig PPIase, did not affect that of *E. coli* PPIase *a* and only slightly affected that of PPIase *b* (data not shown).

*Form a Has a Hydrophobic Signal Sequence.* To investigate the structural features of the *E. coli* PPIases, their hydropathic patterns were calculated by the HPLOT program (Kyte & Doolittle, 1982) (Figure 6). The hydropathy plots along the amino acid sequences corresponding to those of the isolated *E. coli* PPIases indicated that their overall patterns are very similar, as expected from their sequence homology. However, the first 24 amino acid residues in the amino acid sequence deduced from the nucleotide sequence of the *a* gene, which is not present in the isolated PPIase *a*, showed a striking hydrophobic character. This 24-residue sequence has the following characteristics of a signal peptide: (1) there is a basic amino acid residue very close to the amino terminus (lysine at the third position from the initiator methionine); (2) there is a hydrophobic amino acid stretch in the middle part; and (3) an amino acid residue having a small side chain (alanine) occupies the carboxyl-terminal site. On the basis of these

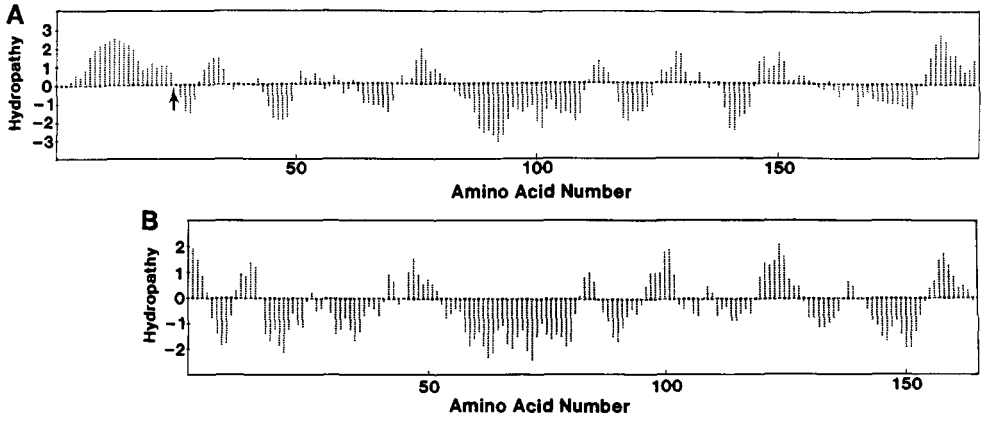


FIGURE 6: Hydropathy profiles predicted for *E. coli* PPIases calculated by using the *HPLOT* program (see text). Hydropathic index is plotted along the deduced amino acid sequence. (A) PPIase *a*; the processing site is shown by an arrow. (B) PPIase *b*.

Table I: Presence of PPIase Activity in Periplasmic and Cytoplasmic Fractions[a]

| | cytoplasmic fraction | periplasmic fraction |
|---|---|---|
| total protein (mg) | 620 | 34 |
| β-galactosidase act. (units/mg) | 16 | 1.2 |
| PPIase act. (units/mg) | 467 | 567 |
| total PPIase act. (units) | 289540 | 19278 |

[a] The values were obtained as an average of the experiments repeated 3 times.

criteria, it is strongly suggested that the amino-terminal hydrophobic stretch serves as the signal peptide to transport the PPIase *a* into the periplasmic space of the *E. coli* cell. On the other hand, PPIase *b* is not thought to have a signal sequence for the following reasons: (1) no hydrophobic sequence was present around its amino-terminal region; (2) the stop codon appears in the frame upstream of the codon for a putative initiator methionine in its gene sequence (Figure 3b); and (3) the amino acid sequence of the purified PPIase *b* is completely identical with that starting at the putative initiator methionine deduced from the gene sequence.

*Forms a and b Are Present in the Periplasmic Space and Cytoplasm, Respectively.* Because the purified form *a* lacked the amino-terminal hydrophobic stretch of 24 residues which is predicted from the nucleotide squence of the PPIase *a* gene and since it has structural features characteristic of signal peptides, we examined whether form *a* is secreted into the periplasmic space of the *E. coli* cell. For this purpose, the periplasmic fraction was prepared by osmotic shock treatment of *E. coli* cells. Table I indicates that the PPIase activity was present in the periplasmic fraction as well as the cytosolic fraction at almost equal activity per milligram of total protein. We measured simultaneously the enzymatic activity of β-galactosidase as a marker enzyme, which is naturally present in cytoplasm. This measurement allowed us to estimate that the contamination of the cytoplasmic components in the periplasmic fraction was less than 10%. This level of contamination in the periplasmic fraction was also confirmed by other cytosolic proteins, such as superoxide dismutase, and is comparable to that reported previously for many cytosolic proteins in osmotically shocked *E. coli* cells (Oliver, 1987). Furthermore, the chromatographic patterns of the two fractions on a DEAE-5PW column were quite different from each other (Figure 7). The major proteins purified from the periplasmic fraction were identified by the combination of protein sequencing and database search, and this fraction was found to contain galactose binding protein, asparaginase, and periplasmic oligopeptide binding protein (data not shown). All

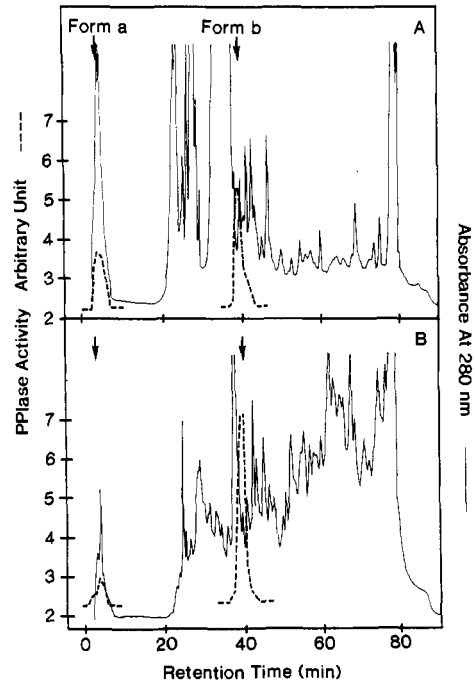


FIGURE 7: Comparison of the contents of PPIases *a* and *b* in the periplasmic (A) and cytoplasmic (B) fractions. Each fraction (2 mg of protein) was applied to a DEAE column. The elution positions of the two forms are indicated by an arrow in each chromatogram.

of the identified proteins were known to be released from the periplasm by osmotic shock treatment (Oliver, 1987) In contrast, only traces of these proteins were detected in the cytoplasmic fraction. To investigate the presence of forms *a* and *b* in each fraction, the PPIase activity in the eluent of the DEAE column was measured (Figure 7A,B). By this method, the two purified forms *a* and *b* were completely separated and eluted at the retention times of 2 and 39 min, respectively, as indicated by arrows in Figure 7. Thus, the contents of the two forms in each fraction could be estimated by measuring the enzymatic activity in the eluent from the column. As shown in Figure 7, the cytoplasmic fraction contained mainly form *b* (*a*:*b* = 1:20), whereas the ratio of form *a* over form *b* in the periplasmic fraction increased about 20-fold (*a*:*b* = 1:1). On the basis of these results and the measurement of the total PPIase activity in the periplasmic fraction (Table I), the specific activity of form *a* in the periplasmic fraction was estimated to be 294 units/mg and that in cytoplasm to be 16 units/mg. This result thus indicates that form *a* is exclusively translocated into the periplasmic space of the *E. coli* cell. The chromatographic analysis and enzyme assay revealed that

Table II: Summary of the Features of the PPIases Known

| PPIase (superfamily) | type | sensitivity to CsA | sensitivity to FK506 | sp act.[a] | mol wt | species found | content (%) | no. of genes: mammalian | no. of genes: others | cellular localization |
|---|---|---|---|---|---|---|---|---|---|---|
| FKBP[b] | FK506 | − | + | 1/20 | ~11 000 | human, bovine | 0.1–0.4 | 1–2 | ~1 (yeast) | cytoplasm |
| cyclophilin[c] | CsA | + | − | 1 | ~18000 | human, pig, rat, bovine, yeast, fungi | 0.1–0.4 | >20 | 1–2 (yeast, fungi) | cytoplasm, mitochondria (fungi) |
| | non-CsA | − | − | 1 | ~21000 | *E. coli* | 0.04–0.4 | unknown | 2 (*E. coli*) | cytoplasm, periplasm |

[a] The specific activity over that of CsA-type PPIase was given. [b] Harding et al. (1989); Siekierka et al. (1989); Maki et al. (1990). [c] Danielson et al. (1988); Fischer et al. (1989); Haendler et al. (1987); Koletsky et al. (1986); Takahashi et al. (1989); Tropschug et al. (1989).

PPIase *b* can be recovered in the periplasmic fraction upon osmotic shock treatment in levels far beyond those observed for other cytosolic proteins. However, the specific activity of form *b* in the cytosolic fraction was estimated to be about 1.5 times higher than that in the periplasmic fraction. Since form *b* is not expected to have a signal sequence, it is probably leaked from the cytoplasm by osmotic shock treatment; it is known that some of the nonperiplasmic proteins, such as elongation factor Tu and uridine phoshorylase, are released during osmotic shock treatment (Oliver, 1987). These observations are thus consistent with the idea that processed form *a* is exclusively translocated into the periplasmic space whereas form *b* is present in the cytoplasm of the *E. coli* cell. After this result was obtained, we found that Liu and Walsh (1990) reported the overexpression and isolation of an ORF gene product which is identical with the *E. coli* PPIase *a* from the periplasmic fraction of *E. coli*. Their result confirms that *E. coli* PPIase *a* is translocated into the periplasmic space of the *E. coli* cell.

## Discussion

PPIase is thought to be essential for protein folding during protein synthesis in the cell based on its refolding accelerating effect on several denatured proteins in vitro (Fischer et al., 1984; Lang et al., 1987). Despite its presumed signifiance in protein folding, its biological function is totally unknown. A conviction that PPIase plays a critical part in certain biological events came from the findings that cyclophilin and FK506 binding protein, which were thought to mediate the effects of the immunosuppressants in T-cells, possess PPIase activity and that each PPIase is inhibited by its respective ligand (Takahashi et al., 1989; Fischer et al., 1989; Siekierka et al., 1989; Harding et al., 1989). This conviction was strengthened further by the finding that cyclophilin mediated the cytotoxic CsA effect in *N. crassa* and yeast (Tropschug et al., 1989). Thus, on the assumption that the effect of CsA is mediated through inhibition of PPIase activity, PPIases are suggested not to be restricted to immunoresponses but to be involved ubiquitously across several species in many cellular events.

In order to determine the biologial functions of PPIase, we looked for a much simpler cellular system and focused on analyzing *E. coli*, since its cellular events are the best studied. Then, we examined PPIase activity in *E. coli* cells and isolated the enzyme from them guided by the activity. This approach led us to the finding that *E. coli* PPIase exists as at least two distinct forms. Because two kinds of PPIases, cyclophilin and FKBP, were found in mammalian cells (Takahashi et al., 1989; Siekierka et al., 1989), their relationships to the *E. coli* PPIases were examined in terms of sensitivity to the immunosuppressants, CsA and FK506. However, neither *E. coli* activity (*a* or *b*) was affected by these two drugs, indicating that there is a third type of PPIase in regard to drug sensitivity. The structural basis of the insensitivity of *E. coli* PPIases to the immunosuppressants was studied by sequencing the purified enzymes and the corresponding genes isolated. Such sequence analyses indicated that both *E. coli* PPIases had structures homologous to those of cyclophilins from many other species. In this regard, both *E. coli* PPIases are classified in the cyclophilin superfamily. We categorize them as non-CsA-sensitive type, emphasizing the relationship to cyclophilin in their amino acid sequences, but indicating their insensitivity to CsA. Table II summarizes the physicochemical features of the three known PPIase types, FK506-, CsA-, and non-CsA-sensitive PPIases. Currently, the non-CsA-sensitive PPIase is found only in *E. coli*; however, this does not exclude the possibility that non-CsA-sensitive PPIases may also be present in eukaryotic cells.

The probable existence of two cyclophilin-like proteins in *Drosophila* was postulated based upon the results of Northern blot analysis with the *ninaA* gene, which is a known visual transduction gene (Schneuwly et al., 1989; Shieh et al., 1989). However, the existence of the two distinct isoforms was still in question judging from the finding that in *N. crassa* two mRNA species for the cytosolic and the mitochondrial forms of cyclophilin were transcribed from a single gene and eventually produced an identical molecule by posttranslational processing of the transit signal sequence from the precursor of the mitochondrion targeted protein (Tropschug et al., 1988). In contrast, our current results definitively show that *E. coli* contains two different genes for PPIase and that the two distinctive forms are expressed in a single cell. Thus, the finding that multiple cyclophilin-related DNA sequences are present in the mammalian genome (Haendler et al., 1987; Danielson et al., 1988; Maki et al., 1990) suggests that there may be many forms of cyclophilin-like PPIase within a single species or even in a single cell, each of which probably has its specific role in the cell. The number of gene copies for PPIases found in the human genome was estimated as over 20 for cyclophilin-like PPIase and 1–2 for FKBP-like PPIase by Southern blot analysis (Table II) (Haendler et al., 1987; Maki et al., 1990), whereas in phylogenically lower organisms, such as *Neurospora* and yeast, it is estimated as 1–2 for cyclophilin-like PPIase (Tropschug et al., 1989). Although yeast genome was found to contain one copy of DNA sequence which hybridized with a human FKBP cDNA, we were unable to detect any FKBP-hybridized DNA sequence in the *E. coli* genome (data not shown). However, this does not exclude the existence of FKBP-like PPIase in an *E. coli* cell. Thus, these findings suggest that the cyclophilin gene has diverged extensively with the progressing complexity of cellular metabolism during evolution.

Our determination that one of the two forms of *E. coli* is present in cytoplasm whereas the other form is transported into the periplasmic space of the *E. coli* cell may give a clue to how cyclophilin-like PPIases have functionally diverged in the cell. Namely, it indicates that proteins with the ability

to assist folding of various polypeptides are located on both sides of the inner membrane. Many proteins are transported through this inner membrane, and during translocation, the polypeptides are thought to be in the unfolded state (Rothman, 1989), so that some mechanism is needed to modulate folding and unfolding on both sides of the membrane. We propose that the two PPIases may be involved in such a mechanism. There are several other proteins such as DnaK, SecB, and GroE which have been known to assist protein folding in the *E. coli* cell; however, these factors are all present only in *E. coli* cytosol. Thus, the presence of proteins with the ability to assist polypeptide folding in both the periplasm and cytoplasm suggests that the folding of some exported proteins may be catalyzed by the periplasmic proline isomerase, which, in turn, implies that the isomerization of prolines in some proteins destined for export may be prevented in the cytoplasm. Of course, in cases where the exported protein is translocated through the inner membrane during translation, the prevention of the isomerization in the cytoplasm may not be necessary. However, for the case that translocation is not coupled with translation, some mechanism must be postulated by which the isomerization is prevented in the cytoplasm and is subsequently effected in the periplasm, for the translocation of proteins through the inner membrane. One possible concept for such a mechanism is that proteins destined for export cannot be substrate for the cytoplasmic PPIase, but only for the periplasmic. Another mechanism might involve chaperonins which have an effect on proline isomerization, thus preventing the action of cytoplasmic PPIase on exported protein within a cytosol. Thus, our finding that the two distinct forms of PPIase are localized separately in the cytoplasmic and periplasmic compartments of *E. coli* cells raises many important questions on the understanding of the functional roles of various forms of PPIase on protein translocation, and indicates that *E. coli* offers a much simpler cellular system with which to analyze these multiple functions of PPIases on protein folding in the cell.

## Acknowledgments

We thank S. Kajie for providing *E. coli* strain ST249, K. Watanabe and T. Kobayashi for technical assistance, and K. Urakami for synthesis of the DNA probe. We are grateful to K. Ito and T. Komano of Kyoto University for suggestions and valuable discussions.

## References

Aiba, H., Adhya, S., & de Crombrugghe, B. (1981) *J. Biol. Chem. 256*, 11905–11910.

Danielson, P. E., Forss-Petter, S., Brow, M. A., Calavetta, L., Douglas, J., Milner, R. J., & Sutcliffe, J. G. (1988) *DNA 7*, 261–267.

Dietmeier, K., & Tropschug, M. (1990) *Nucleic Acids Res. 18*, 373.

Elliott, J. F., Lin, Y., Mizel, S. B., Bleackley, R. C., Harnish, D. G., & Paetkau, V. (1984) *Science 226*, 1439–1441.

Emmel, E. A., Verweij, C. L., Durand, D. B., Higgins, K. M., Lacy, E., & Crabtree, G. R. (1989) *Science 246*, 1617–1620.

Fischer, G., Bang, H., & Mech, C. (1984) *Biomed. Biochim. Acta 43*, 1101–1111.

Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T., & Schmid, F. X. (1989) *Nature 337*, 467–478.

Freedman, R. B. (1989) *Nature 341*, 692.

Goad, W. B., & Kanehisa, M. (1982) *Nucleic Acids Res. 10*, 247–263.

Haendler, B., Hofer-Warbinek, R., & Hofer, E. (1987) *EMBO J. 6*, 947–950.

Haendler, B., Keller, R., Hiestad, P. C., Kocher, H. P., Wegmann, G., & Movva, N. R. (1989) *Gene 83*, 39–46.

Harding, M. W., Handschumacher, R. E., & Speicher, D. W. (1986) *J. Biol. Chem. 262*, 8547–8555.

Harding, M. W., Galat, A., Uehling, D. E., & Schreiber, S. L. (1989) *Nature 341*, 758–760.

Kajie, S., & Anraku, Y. (1986) *Eur. J. Biochem. 154*, 457–463.

Kawamukai, M., Matsuda, H., Fujii, W., Utsumi, R., & Komano, T. (1989) *J. Bacteriol. 171*, 4525–4529.

Koletsky, A. J., Harding, M. W., & Handschumacher, R. E. (1986) *J. Immunol. 137*, 1054–1059.

Kronke, M., Leonard, W. J., Depper, J. M., Arya, S. K., Wong-Staal, F., Gallo, R. C., Waldmann, T. A., & Greene, W. C. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5214–5218.

Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Lang, K., & Schmid, F. X. (1988) *Nature 331*, 453–455.

Lang, K., Schmid, F. X., & Fischer, G. (1987) *Nature 329*, 268–270.

Lin, L.-N., Hasumi, H., & Brandts, J. F. (1988) *Biochim. Biophys. Acta 956*, 256–266.

Liu, J., & Walsh, C. T. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4028–4032.

Maki, N., Sekiguchi, F., Nishimaki, J., Miwa, K., Hayano, T., Takahashi, N., & Suzuki, M. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5440–5443.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marmur, J. (1961) *J. Mol. Biol. 3*, 208–218.

Oliver, D. B. (1987) Periplasm and Protein Secretion, in *Escherichia coli and Salmonella typhimurium 1* (Neidhardt, F. C., Ed.) pp 56–69, American Society for Microbiology, Washington, DC.

Pain, R. H. (1987) *Nature 328*, 298.

Rothman, J. E. (1989) *Cell 59*, 591–601.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schneuwly, S., Shortridge, R. D., Larrivee, D. C., Ono, T., Ozaki, M., & Pak, W. L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5390–5394.

Shieh, B.-H., Stamnes, M. A., Seavello, S., Harris, G. L., & Zuker, C. S. (1989) *Nature 338*, 67–70.

Siekierka, J. J., Hung, S. H. Y., Poe, M., Lin, C. S., & Sigal, N. H. (1989) *Nature 341*, 755–757.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.

Takahashi, N., Takahashi, Y., Blumberg, B. S., & Putnam, F. W. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4413–4417.

Takahashi, N., Hayano, T., & Suzuki, M. (1989) *Nature 337*, 473–475.

Thomson, A. W. (1989) *Immunol. Today 10*, 6–9.

Tropschug, M., Nicholson, D. W., Hartl, F.-U., Kohler, H., Pfanner, N., Wachter, E., & Neupert, W. (1988) *J. Biol. Chem. 263*, 14433–14440.

Tropschug, M., Barthelmess, I. B., & Neupert, W. (1989) *Nature 342*, 953–955.