Role of the PDZ Domains in *Escherichia coli* DegP Protein

Jack Iwanczyk,1 Daniela Damjanovic,1 Joel Kooistra,1 Vivian Leong,1 Ahmad Jomaa,1 Rodolfo GHIRLando,2 and Joaquín Ortega1*

Department of Biochemistry and Biomedical Sciences, McMaster University, 1200 Main Street West, Hamilton, Ontario L8N 3Z5, Canada,1 and Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, DHHS, Bethesda, Maryland 20892-05402

Received 27 November 2006/Accepted 29 January 2007

PDZ domains are modular protein interaction domains that are present in metazoans and bacteria. These domains possess unique structural features that allow them to interact with the C-terminal residues of their ligands. The *Escherichia coli* essential periplasmic protein DegP contains two PDZ domains attached to the C-terminal end of the protease domain. In this study we examined the role of each PDZ domain in the protease and chaperone activities of this protein. Specifically, DegP mutants with either one or both PDZ domains deleted were generated and tested to determine their protease and chaperone activities, as well as their abilities to sequester unfolded substrates. We found that the PDZ domains in DegP have different roles; the PDZ1 domain is essential for protease activity and is responsible for recognizing and sequestering unfolded substrates through C-terminal tags, whereas the PDZ2 domain is mostly involved in maintaining the hexameric cage of DegP. Interestingly, neither of the PDZ domains was required for the chaperone activity of DegP. In addition, we found that the loops connecting the protease domain to PDZ1 and connecting PDZ1 to PDZ2 are also essential for the protease activity of the hexameric DegP protein. New insights into the roles of the PDZ domains in the structure and function of DegP are provided. These results imply that DegP recognizes substrate molecules targeted for degradation and substrate molecules targeted for refolding in different manners and suggest that the substrate recognition mechanisms may play a role in the protease-chaperone switch, dictating whether the substrate is degraded or refolded.

PDZ domains represent a common protein interaction motif, and their name was derived from the first three proteins in which such domains were identified, namely PSD-95, *Drosophila melanogaster* Disc large protein, and zonula occludens protein 1. Bacterial PDZ domains are homologous to the metazoan PDZ domains (23, 24, 26); however, their topology is different (20), and thus they are designated “PDZ-like” domains.

PDZ domains are approximately 90 residues long and have a common structure consisting of six β-strands and two α-helices, which fold in an overall six-stranded β-sandwich. The C-terminal ends of a protein substrate (7) usually bind in a groove of the domain formed between one of the α-helices and the adjacent β-strand, which thus serves as an extra β-strand added to the β-sheet (8). In this manner, the C-terminal peptide backbone participates in an extensive hydrogen-bonding pattern with the main chain atoms of the PDZ domain β-strand. The terminal carboxylate group is also stabilized by a series of hydrogen bonds with the highly conserved “carboxylate-binding loop” (3). However, the side chain of the C-terminal residue and the side chain of the residue at position −2 are the structures that are most critical for the specificity of substrate recognition by the PDZ domain, rather than the extensive hydrogen bonds with the main chain of the PDZ domain β-strand (29). In this context, the C-terminal residue of the substrate molecule is referred to as P0, and upstream residues are designated P−1, P−2, etc.

The *Escherichia coli* DegP protein (also called HtrA or protease Do) is an essential periplasmic protein (21) that functions as both a chaperone and a protease in a temperature-dependent fashion. DegP contains two PDZ domains, PDZ1 and PDZ2, following the N-terminal protease domain (2). One flexible loop links the protease domain to the PDZ1 domain, and another flexible loop links the PDZ1 domain to the PDZ2 domain (Fig. 1A and B) (19). X-ray crystallography studies have shown that DegP oligomerizes into a hexamer with a central chamber containing the proteolytic sites. The PDZ domains extend out from the ends of the cage, but they can also adopt a closed conformation blocking the lateral routes of access to the internal chamber (Fig. 1C) (19).

The PDZ1 domain of DegP contains a deep binding cleft formed by β-strand 14, the “carboxylate-binding loop,” and helix H (Fig. 1D, left panel) (19). This domain is predicted to be involved in substrate binding through recognition of the C-terminal residues of the substrate molecule. Hydrophobic residues form the pockets for the P0 and P−2 residues of the substrate (Fig. 1D, right panel). In this manner, the PDZ1 domain is proposed to bind substrates with hydrophobic C-terminal residues.

Interestingly, residues 358 to 362 in the loop separating the two PDZ domains fill the corresponding binding cleft of PDZ2 (Fig. 1E) (19) and adopt an extended conformation identical to the conformation exhibited by C-terminal peptide substrates bound to other PDZ domains. Thus, this substrate-like segment may block substrates from entering the groove. However, based on the X-ray structure of DegP, the carboxylate-binding...
loop of PDZ2 (Fig. 1E, left panel) is highly flexible, and the possible reorientation of the substrate-like segment may allow binding of substrates.

In addition to the structural evidence presented previously, there is also experimental evidence demonstrating that other proteases bind substrates through their PDZ domains. For instance, the periplasmic *E. coli* tail-specific protein (Tsp) binds to the C-terminal end of substrates through its PDZ domain. This recruits the substrate to the catalytic site, which is located in a different domain, resulting in substrate cleavage at multiple sites (1, 16, 27, 30). Moreover, the X-ray structure of the *E. coli* DegS protein in complex with the C-terminal tail of the OmpC protein (35) shows that the DegS PDZ domain is used for substrate binding, also demonstrating how peptide binding to the PDZ domain modulates protease activity by inducing a series of reversible conformational changes that activate the protease (33, 35). Finally, a previous study (30) showed that the isolated PDZ domains of *Salmonella enterica* serovar Typhimurium DegP are able to bind the 11-amino-acid SsrA peptide. Interestingly, in this study *E. coli* DegP bound to the SsrA peptide, but the isolated PDZ domains of the *E. coli* protein were unable to bind to it. The SsrA peptide is usually appended to the C terminus of incompletely translated proteins, targeting them for proteolytic degradation and thus avoiding the buildup of ribosomes stalled on defective mRNA molecules (4–6, 15, 17). It therefore appears that substrate recognition by PDZ domains is a common mechanism in proteases.

In this work, we found that the PDZ1 domain of *E. coli* DegP is essential for protease activity and that both PDZ domains are dispensable for chaperone activity. Our results suggest that the two PDZ domains in DegP perform different functions; PDZ1 is involved in targeting substrates for proteolysis by recognizing the last three C-terminal residues of the substrate, whereas PDZ2 is required for maintenance of the hexameric cage of DegP. Additionally, we found that the loops between the protease and the PDZ1 domain and between the two PDZ domains are essential for protease activity in the hexameric DegP protein. The implications of these results for the protease-chaperone switch of DegP are discussed below.

**FIG. 1.** Structure of DegP and peptide-binding clefts of the PDZ1 and PDZ2 domains. (A) Ribbon representation of the DegP monomer (PDB identification no. 1KY9), showing the protease domain (blue), the PDZ1 domain (yellow), and the PDZ2 domain (red). The loop between the protease and the PDZ1 domain is cyan, and the loop between the two PDZ domains is dark gray. Residues involved in generating the DegP deletion mutants in this study are indicated. (B) Linear diagram of the DegP protein. Residues involved in generating the DegP deletion mutants are indicated. (C) Side views of the open conformation (left) and closed conformation (right) of the molecular surfaces of the DegP hexamer. The PDZ2 domain is not defined in the open conformation of the original structure of DegP. (D and E) Ribbon representations (left) and molecular surface representations (right) of the PDZ1 (D) and PDZ2 (E) domains, showing the substrate-binding cleft. In the ribbon representations elements important for substrate binding, including the carboxylate-binding loop (CBL), are indicated by different colors. The carboxylate-binding loop is not defined in the X-ray structure of the PDZ2 domain, as indicated by the red dashed line. The hydrophobic pockets of the binding cleft are orange in the surface representations. Residues 355 to 358 in the loop separating the PDZ domains are indicated by a ball-and-stick representation in panel E, filling the binding cleft of the PDZ2 domain.
MATERIALS AND METHODS

Plasmids and mutagenesis. The pET21b-DegP, pET21b–PDZD1 263-366 DegP, and pET21b–PDZD2 DegP plasmids were obtained as described previously (10). The pET21b–PDZD1 +2 DegP plasmid was generated by PCR using the pET21b–DegP plasmid as the template, and it was subcloned into the expression vector pET21b. The pET21b–PDZD1 263-354 DegP plasmid was obtained from pET21b–DegP by engineering Nhel and XbaI restriction sites between residues 262 and 263 and between residues 354 and 355, respectively, using the QuiChange site-directed mutagenesis method (Stratagene). The plasmid was digested with Nhel and XbaI, purified from the Nhel-XbaI fragment by agarose gel electrophoresis, and ligated with T4 DNA ligase (Invitrogen). The QuiChange method was used to remove the remaining nucleotides. The pET21b plasmids used to express the Δ258-260 and Δ357-364 DegP mutants were obtained by the QuiChange method from the pET21b–DegP plasmid. The pET21b–Δ258-260–ΔPDZD2 plasmid was constructed from the pET21b–ΔPDZD2 DegP plasmid also by using the QuiChange method. The same method was used to construct all the proteolytically inactive variants (S210A) of the mutants.

The enhanced green fluorescent protein (EGFP) gene was amplified from the pEGFP-C2 vector (Clontech) by PCR and subcloned into the expression vector pPROEX-HTb (Invitrogen) using NcoI and XbaI sites. A stop codon was introduced after the XbaI site by the QuiChange method to obtain the pPROEX-HTb-EGFP expression plasmid. The pPROEX-HTb-EGFP-SraA plasmid expressing EGFP-SraA was constructed in several steps. First, two complementary oligonucleotides coding for the 11-residue SraA tag followed by a stop codon and an Nhel site were synthesized with XbaI- and HindIII-compatible ends. Second, the pPROEX-HTb vector was digested with XbaI and HindIII and ligated with T4 DNA ligase to the annealed oligonucleotide coding for the SraA tag to obtain the pPROEX-HTb-SraA plasmid. Finally, the EGFP gene was obtained from the pPROEX-HTb-EGFP plasmid as an NcoI-XbaI insert and ligated into the pPROEX-HTb-SraA plasmid digested with the NcoI and XbaI enzymes, producing the pPROEX-HTb-EGFP-SraA plasmid that was used to express EGFP-SraA.

The QuiChange method was used to obtain the pPROEX-HTb-EGFP-SFS plasmid from the pPROEX-HTb-EGFP plasmid by replacing nucleotides coding for the last three residues of EGFP (ISR). This plasmid was used to produce EGFP-SFS, and all the subsequent mutations in the SFS tag were generated by the QuiChange method from the pPROEX-HTb-EGFP-SFS plasmid.

Protein expression and purification. Wild-type DegP and all the DegP mutants used in the assays were expressed as C-terminally His-tagged proteins. The procedure used to express and purify the proteins with a HiTrap metal chelating column was performed as described previously (10).

To express EGFP, EGFP-SraA, EGFP-SFS, and several variants of EGFP-SFS as N-terminally His-tagged proteins, the expression vector pPROEX-HTb containing one of the EGFPs was transformed into E. coli BL21(DE3) competent cells. The cells were grown in LB medium at 37°C to optical density at 600 nm of 0.5, and expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 5 h at 25°C. Lysis was performed in 20 ml lysis buffer (50 mM Tris-HCl [pH 7.3], 0.1 M NaCl, 5% glycerol) by adding 256 µl of lysosyme (50 mg/ml) and incubating the preparation for 1 h at 4°C, followed by sonication on ice. The lysate was cleared by centrifugation at 39,000 × g for 40 min. NaCl was added to obtain a concentration to 0.5 M, and the lysate was filtered through a 0.45-µm filter and added to a HiTrap metal chelating column (GE Healthcare Life Sciences) equilibrated with 50 mM Tris-HCl (pH 7.3)–0.5 M NaCl–5% glycerol. Nonspecifically bound proteins were washed with increasing concentrations of imidazole up to 45 mM. EGFP (or one of the EGFP variants) was eluted with 240 mM imidazole. Fractions containing EGFP were pooled, diluted fivefold in buffer A (50 mM Tris-HCl [pH 7.3], 0.1 M NaCl), and loaded into a HiTrap Q HP column (GE Healthcare Life Sciences) equilibrated in buffer A. The column was washed with 50 ml of the same buffer, and the protein was eluted by increasing the concentration of NaCl to 200 or 400 mM depending on the EGFP mutant. Fractions containing the EGFP were pooled and loaded into a HiPrep 26/10 desalting column (GE Healthcare Life Sciences) equilibrated with 50 mM Tris-HCl (pH 7.3)–0.15 M NaCl. The eluted protein was concentrated to obtain a concentration of 8 μg/ml with a concentrator (Vivaspin 6; 10,000 molecular-weight cutoff; Vivascience), and glycerol was added to a final concentration of 20% (vol/vol).

Size exclusion chromatography. A Superdex 200 10/300 GL column (GE Healthcare Life Sciences) equilibrated in 50 mM HEPES (pH 7.3) containing 100 mM NaCl at 4°C was used for the gel filtration chromatography experiments. Protein samples (100 µl) at concentrations between 0.5 and 3.5 mg/ml depending on the experiment were applied to the column. Samples were incubated at 4, 37, or 45°C before the protein was loaded into the column. A gel filtration calibration kit (high molecular weight; GE Healthcare Life Sciences) was used for column calibration.

Analytical centrifugation. Sedimentation equilibrium and velocity experiments were carried out as described previously (10). Samples analyzed in the sedimentation equilibrium experiments were examined at three loading concentrations, as shown in Table 1. Data collected at different speeds and different loading concentrations were analyzed globally in terms of various species analysis models using SEDPHAT 4.0 (P. Schuck; http://www.analyticalultracentrifugation.com/sedphat/sedphat.htm). Solution densities were determined at 20°C with a Mettler-Toledo DE51 density meter and were corrected using values obtained at 4°C. Partial specific volumes were calculated based on the amino acid composition using SEDNTERP (J. Philo; http://www.cephalica.mailway.com). Sedimentation coefficient distributions [given in the form c(s)] were obtained by analysis of the sedimentation velocity data with SEDFIT (25) (P. Schuck; http://www.analyticalultracentrifugation.com/default.html). All sedimentation coefficients below are the values determined at 4°C in 50 mM HEPES (pH 7.3).

Chaperone and protease activity assays. Refolding assays using citrate synthase (Roche) as a substrate were performed as described previously (10). Similarly, protease assays with malate dehydrogenase (MDH) from porcine heart (Roche and Sigma), bovine milk β-casein (Sigma), and egg white lysozyme (Bioshop) were performed using previously described methods (10).

EGFP refolding assay. Prior to the experiment EGFP and the EGFP-SraA, EGFP-SFS, and EGFP-SFS variants were denatured by incubating 1 mg/ml of protein in 6 M guanidine-HCl containing 25 mM Tris-HCl (pH 7.3), 10 mM diithiothreitol, and 1 mM EDTA for 2 h at 25°C. To perform the assay, 200 pmol of the guanidine-HCl-denatured EGFP (or EGFP variant) was diluted into 500 µl of EGFP refolding buffer (50 mM Tris-HCl, pH 7.3) with or without 800 µmol of a specific DegP S210A mutant. Fluorescence was determined at 37°C with excitation at 489 nm and emission at 509 nm by using a fluorescence spectrophotometer (Cary Eclipse; Varian) (9, 28). The amount of refolded EGFP (or EGFP variant) was expressed as a percentage of the fluorescence intensity obtained for the same native EGFP at an identical concentration. Where indicated below, the EGFP refolding assay was done in the presence of the different DegP mutants. In these cases, a fluorescence value of 100% was considered the intensity of fluorescence produced by the same concentration of the native EGFP (or EGFP variant) in the presence of the specific DegP mutant at the concentration used in the assay. This control was necessary as we noticed a small increase in the fluorescence of the native EGFPs in the presence of DegP or DegP mutants.

**RESULTS**

**PDZ domain requirements for protease and chaperone activities of DegP**. In agreement with previous reports (31), we showed that deletion of both PDZ domains (Fig. 2A) or the PDZ1 domain (10) resulted in a dramatic reduction in the...
protease activity of DegP with all three substrates tested (MDH, β-casein, and lysozyme). Conversely, we also showed (10) that the ΔPDZ2 DegP mutant was able to hydrolyze the substrates tested at rates that were similar to (MDH and β-casein) or higher than (lysozyme) the rates observed with wild-type DegP.

Interestingly, none of the PDZ deletions had a significant effect on the chaperone activity of DegP (Fig. 2C) (10), as shown by analysis of the chaperone activity of the S210A proteolytically inactive variants of the mutants (in which serine 210 in the catalytic triad was changed to alanine).

We hypothesized that the lack of PDZ2 involvement in protease activity may be a consequence of residues 358 to 362 filling its binding cleft (Fig. 1E). As the ΔPDZ2 DegP mutant analyzed in the present study had very low proteolytic activity (Fig. 2B), we chose to examine the ΔPDZ1 263-354 DegP mutant. In this mutant residues 358 to 362 were deleted. Consistent with previous results, the proteolytically inactive S210A variant of this mutant exhibited full chaperone activity (Fig. 2C). To investigate whether the substrate-like segment occupied the binding cleft blocks the access of potential substrates, we looked at our previously published data for a different ΔPDZ1 DegP mutant (ΔPDZ1 263-366) (10) in which both the PDZ1 domain and the substrate-like segment from the loop connecting the PDZ domains were removed. This should have facilitated access of the substrate to the C-terminal binding groove of PDZ2. Interestingly, the ΔPDZ1 263-366 DegP mutant exhibited no (MDH and β-casein) or very mild (lysozyme) protease activity. Therefore, the hypothesis that the substrate-like segment occupies the binding cleft and blocks the binding of potential substrates is an unlikely explanation for the dispensability of the PDZ2 domain for the protease activity of DegP. These results suggest that despite the flexible carboxylate-binding loop in PDZ2 that could allow reorientation of the substrate-like segment and entrance of substrates, this domain is not actually required for protease activity.

Together, these results indicate that neither of the PDZ domains is required for the chaperone activity of DegP and that only the PDZ1 domain is essential for its protease activity.

**Both PDZ domains are essential for maintaining the hexameric cage of DegP.** In the “closed” conformation of the DegP hexamer, the PDZ2 domain of each monomer interacts with the PDZ1 domain of the opposing trimer in a zipper-like arrangement (Fig. 1C) (19). Our previous studies showed that residues 39 to 78 in the LA loops, as well as the PDZ2 domains, are essential for maintenance of the DegP hexameric cage (10). In our previous study, we characterized the oligomeric state of the ΔPDZ2 DegP mutant (and its S210A variant) by size exclusion chromatography and analytical ultracentrifugation and showed that this mutant is a trimer. Similarly, the ΔPDZ1 263-366 DegP mutant (and its S210A variant) (Fig. 1A and B) were shown to be trimers. We concluded that the PDZ2 domain is essential for the maintenance of the hexameric cage in DegP, as is the PDZ1 domain, which acts as a spacer allowing the PDZ2 domain to reach the PDZ1 and PDZ2 domains of the opposing monomer rather than the adjacent monomer (10).

To rule out the possibility that the previously characterized ΔPDZ1 263-366 mutant (10) is a trimer because the PDZ2 domain is not adequately folded due to the absence of the substrate-like segment filling the binding cleft, we extended our analysis by characterizing the oligomeric state of the ΔPDZ1 263-354 mutant (Fig. 1A and B). This protein includes not only the entire PDZ2 domain but also residues 358 to 362 of the substrate-like segment (Fig. 1E). This mutant was incubated at 4, 37, or 43°C for 1 h before it was loaded onto a size exclusion chromatography column. In all cases this mutant and its S210A variant (Fig. 1A and B) eluted at a position very close to the expected elution volume for a different ΔPDZ1 DegP mutant (ΔPDZ1 263-366) (10) in which both the PDZ1 domain and the substrate-like segment from the loop connecting the PDZ domains were removed. This should have facilitated access of the substrate to the C-terminal binding groove of PDZ2. Interestingly, the ΔPDZ1 263-366 DegP mutant exhibited no (MDH and β-casein) or very mild (lysozyme) protease activity. Therefore, the hypothesis that the substrate-like segment occupies the binding cleft and blocks the binding of potential substrates is an unlikely explanation for the dispensability of the PDZ2 domain for the protease activity of DegP. These results suggest that despite the flexible carboxylate-binding loop in PDZ2 that could allow reorientation of the substrate-like segment and entrance of substrates, this domain is not actually required for protease activity.

To evaluate this explanation for the dispensability of the PDZ2 domain we characterized the oligomeric state of the ΔPDZ1 263-354 mutant (ΔPDZ1 263-354 DegP S210A) in the “open” conformation of the DegP hexamer (Fig. 2C) (10). To rule out the possibility that the previously characterized ΔPDZ1 263-366 mutant (10) is a trimer because the PDZ2 domain is not adequately folded due to the absence of the substrate-like segment filling the binding cleft, we extended our analysis by characterizing the oligomeric state of the ΔPDZ1 263-354 mutant (Fig. 2C) (10). This protein includes not only the entire PDZ2 domain but also residues 358 to 362 of the substrate-like segment (Fig. 1E). This mutant was incubated at 4, 37, or 43°C for 1 h before it was loaded onto a size exclusion chromatography column. In all cases this mutant and its proteolytically inactive variant (ΔPDZ1 263-354 DegP S210A) eluted at a position very close to the expected elution volume for the trimeric form of the protein (Fig. 3A and B).

Similar experiments with the ΔPDZ1+2 and ΔPDZ1+2 DegP S210A mutants yielded identical results. The ΔPDZ1+2 DegP S210A mutant was also characterized by analytical ultracentrifugation, and data consistent with the results of size exclusion chromatography were obtained. Sedimentation velocity data analyzed in terms of a continuous c(s) distribution resulted in excellent fits consistent with the presence of one well-resolved species having a sedimentation coefficient of 3.1 S (Fig. 3C). Sedimentation equilibrium data analyzed in terms of a single ideal solute also resulted in excellent fits; these data showed that the molecular mass was 78.9 ± 0.8 kDa (expected.....
monomer mass, 27,003.4 Da; stoichiometry, n = 2.92 ± 0.03), which indicated that this DegP mutant is a monodisperse trimer in solution (Fig. 3D).

These results confirmed the essential role of the PDZ1 and PDZ2 domains in maintaining the hexameric cage. Furthermore, they showed the stability of the ΔPDZ1 263-354 and ΔPDZ1 + 2 mutants and their proteolytically inactive variants at the temperatures used in the protease and chaperone assays performed.

C-terminal tags enhance trapping of unfolded proteins by DegP. We next wanted to study the mechanism of substrate recognition by DegP. The presence of two PDZ domains in this protein led us to investigate whether DegP recognizes substrates through their C-terminal residues. To this end, we used...
EGFP and constructed derivatives of EGFP with different C-terminal residues. EGFP was chosen because it fluoresces in its native state but not in nonnative conformations (9) and also because it has the ability to refold spontaneously into its native conformation upon incubation in refolding buffer.

As a control, we first tested whether DegP$_{S210A}$ could trap guanidine-HCl-denatured EGFP in its nonnative form by chemically denaturing EGFP and then diluting it in the refolding reaction mixture with or without DegP$_{S210A}$. Without DegP$_{S210A}$, unfolded EGFP rapidly regained fluorescence, and 1 min after the reaction was initiated 68% of the EGFP reached a folded and fluorescent state. In contrast, only 50% of the initial fluorescence of EGFP was recovered after addition of DegP$_{S210A}$ to the refolding mixtures (Fig. 4A), indicating that DegP$_{S210A}$ trapped some of the unfolded EGFP before it refolded.

To test whether a C-terminal tag in EGFP influences sequestration of unfolded proteins by DegP$_{S210A}$, we denatured EGFP-SsrA, which contained the 11-amino-acid SsrA protease recognition tag (AANNENYALAA) fused to the carboxyl terminus of EGFP. Dilution of the denatured protein in a refolding reaction mixture without DegP$_{S210A}$ resulted in recovery of about 76% of the original fluorescence, whereas in the presence of DegP$_{S210A}$ only 33% of the initial fluorescence was recovered (Fig. 4B), suggesting that the 11-residue SsrA tag enhances sequestration of unfolded EGFP by DegP$_{S210A}$.

It has been shown that the last three residues of the C-terminal regions of substrates are the most critical residues for binding by PDZ domains (8). To determine whether this is the case with the PDZ domains of DegP, we obtained EGFP with its last three residues (ISR) changed to SFS (EGFP-SFS). Serine, phenylalanine, and serine are the three carboxyl-terminal amino acids of the PapG pilin, which is a natural substrate for DegP (11, 12). The refolding assay was performed, and we observed that only 57% of the unfolded EGFP-SFS was refolded in the presence of DegP$_{S210A}$, compared to the 95% of the protein that was refolded in the absence of DegP$_{S210A}$ (Fig. 4C).

Altogether, these data suggest that C-terminal tags enhance trapping of unfolded substrates by DegP. Even though the tags are different lengths, the sequence of the last three carboxyl-terminal amino acids appears to be sufficient to enhance sequestration.

Residue at P$_{-1}$ of the SFS C-terminal tag influences sequestration of unfolded EGFP by DegP. To determine which of the last three residues of the substrate C-terminal tag influences sequestration of unfolded EGFP by DegP, we generated point mutations in the original EGFP-SFS substrate. In particular, we were interested in determining whether changes at P$_{-1}$ in the C-terminal tag alter substrate trapping by DegP. The X-ray structures of DegP (19) and other PDZ domain-containing proteins (3) indicated that the side chain of the P$_{-1}$ residue points away from the surface of the binding cleft and is thus expected to have little influence on the binding specificity. We initially generated three EGFP-SFS mutants. The highly hydrophobic phenylalanine residue at P$_{-1}$ was changed to a small nonpolar residue (alanine), a negatively charged residue (glutamate), or a positively charged residue (lysine). The change to a nonpolar residue (EGFP-SAS) or to a negatively charged residue (EGFP-SES) significantly increased the amount of EGFP sequestered in an unfolded state in the refolding assay with DegP$_{S210A}$ (Fig. 5). Conversely, the change to a positively charged lysine residue (EGFP-SKS) decreased the amount of EGFP sequestered, demonstrating that the residue at P$_{-1}$ in
the substrate molecule has a great influence on the trapping ability of DegP.

Additionally, we changed the residues at P₀ and P₋₂ in the C-terminal tag of EGFP-SFS from the polar amino acid serine to the nonpolar amino acid alanine either individually or simultaneously. The three mutants, EGFP-AFS, EGFP-SFA, and EGFP-AFA, showed a slight decrease in the amount of substrate sequestered by DegP S²₁₀₀₄ (Fig. 5). This result was surprising because, in principle, the hydrophobic residues forming the pockets for P₀ and P₋₂ in the PDZ₁ domain of DegP are better suited to bind alanine than to bind serine. However, additional mutations in EGFP-SFS are required to understand the binding requirements at P₀ and P₋₂.

The PDZ₁ domain is responsible for sequestering the unfolded substrate through the C-terminal tag. Having shown that the PDZ₁ domain is essential for protease activity, we hypothesized that the PDZ₁ domain is involved in sequestering the unfolded substrate through the C-terminal tag. To investigate this possibility, we examined the proteolytically inactive variants (DegP S²₁₀₀₄) of our PDZ deletion mutants in an EGFP fluorescence assay using denatured EGFP-SFS as a substrate. The ΔPDZ₁ + 2, ΔPDZ₂, ΔPDZ₁ 263-354, ΔPDZ₁ 263-366 DegP S²₁₀₀₄, and full-length DegP S²₁₀₀₄ mutants (Fig. 1A and B) were purified and used in the refolding reactions. We found that fluorescence recovery was not affected by the presence of the ΔPDZ₁ + 2 DegP S²₁₀₀₄ mutant or either of the two ΔPDZ₁ DegP S²₁₀₀₄ mutants. However, the level of fluorescence recovery significantly decreased upon incubation with ΔPDZ₂ DegP S²₁₀₀₄ or DegP S²₁₀₀₄, which was used as a control (Fig. 6A). This suggests that the amount of sequestered EGFP-SFS in an unfolded state increased and is consistent with the hypothesis that the PDZ₁ domain is responsible for binding the three C-terminal SFS residues. Interestingly, ΔPDZ₂

FIG. 5. Importance of the residue at P₋₂ of the substrate C-terminal tag for sequestration by DegP. Guanidine-HCl-denatured EGFP variants with the C-terminal tags indicated (amino acids indicated by one-letter abbreviations) were refolded in the absence or presence of DegP S²₁₀₀₄ protein. For each EGFP, plots similar to those in Fig. 4 were generated, and then the difference between the percentage of protein refolded in the absence of DegP S²₁₀₀₄ and the percentage of protein refolded in the presence of DegP S²₁₀₀₄ protein at 8 min was calculated (at this time fluorescence had reached a plateau). Experiments were done in triplicate, and averages and standard deviations of the percentages of fluorescence reductions were calculated. Thus, the values reflect the percentage of each EGFP sequestered by DegP S²₁₀₀₄ protein in an unfolded state. ISR, results obtained for wild-type EGFP (isoleucine, serine, and arginine are the three C-terminal residues of EGFP).

FIG. 6. PDZ₁ domain mediates sequestration of the unfolded substrate through the C-terminal tag. (A) Guanidine-HCl-denatured EGFP-SFS was allowed to refold either alone or in the presence of DegP S²₁₀₀₄ or one of the PDZ mutants. The intensity of fluorescence in each reaction was determined after 8 min of incubation at 37°C and was expressed as a percentage of the intensity of fluorescence of native EGFP-SFS at the same concentration in the presence of the DegP S²₁₀₀₄ mutant. The experiment was done in triplicate, and averages and standard deviations of the percentages of fluorescence intensities were calculated. Identical experiments were performed with EGFP-SsrA (B), EGFP-SES (C), and EGFP (D).
DegP$_{S210A}$ was not as efficient as at sequestering EGFP-SFS, full-length DegP$_{S210A}$, suggesting that the PDZ2 domain plays a role in stabilizing the SFS C-terminal tag bound to the PDZ1 domain.

To confirm these results, we repeated the experiment using unfolded EGFP-SsrA (Fig. 6B) and unfolded EGFP-SES (Fig. 6C). Consistent with our previous results, neither the PDZ1 DegPS210A mutant nor the two PDZ1 DegP$_{S210A}$ mutants affected the refolding of EGFP-SsrA and EGFP-SES. Similarly, only the PDZ2 DegP$_{S210A}$ and DegP$_{S210A}$ mutants showed significant sequestration of the unfolded substrates (Fig. 6B and C). As a negative control, we measured the effects of the mutants on the refolding of untagged EGFP. As expected, only the DegP$_{S210A}$ mutant exhibited some ability to sequester the unfolded form of EGFP. The differences between the fluorescence observed in the refolding reaction with EGFP in the presence of the mutants and the fluorescence observed in the reaction in the absence of DegP were not significant (Fig. 6D).

Loops connecting the protease and the PDZ1 domain and connecting the two PDZ domains are essential for protease activity. The PDZ1 domain is connected to the protease domain by a loop (Fig. 1A and B). To determine whether this loop is essential for DegP protease and chaperone activities, we generated the 258-260 DegP mutant and a proteolytically inactive variant of this mutant (258-260 DegPS210A), which lacked most of the loop-forming residues. Both mutants eluted at the elution volume expected for a hexamer when they were loaded onto a size exclusion chromatography column after incubation at 4, 37, or 43°C for 1 h (Fig. 7A). Data obtained in sedimentation equilibrium experiments with the 258-260 DegPS210A mutant exhibited excellent fits when they were analyzed in terms of two noninteracting ideal solutes, returning molecular masses for predominantly hexamers of 281.3 ± 0.7 kDa (expected monomer mass, 46,544.8 Da; stoichiometry, n = 6.0 ± 0.7) and for some dodecamers of 549.4 ± 45 kDa (stoichiometry, n = 11.8 ± 1.0). The hexamer was the predominant species, an observation confirmed by the results of a sedimentation velocity analysis showing the presence of two well-resolved sedimenting species having uncorrected c($s$) values of approximately 7 and 10.2 S. Integration of the c($s$) distribution revealed the presence of approximately 10% dodecamers in this sample (Fig. 7B). Moreover, the c($s$) distribution reproducibly returned a sedimentation coefficient identical to that observed for the DegPS210A mutant. This shows that the overall shapes of the two proteins are very similar and that the structural integrity of the 258-260 DegPS210A mutant is maintained. Thus, we concluded that both the 258-260 DegP and 258-260 DegPS210A mutants are stable hexamers similar to wild-type DegP.

The chaperone activity of the 258-260 DegPS210A mutant as a continuous sedimentation coefficient distribution [c($s$)] obtained using the program SEDFIT. (C) Plot showing the averages and standard deviations for the hydrolysis rates for the 258-260 DegP mutant for three substrates (MDH, β-casein, and lysozyme). WT, wild type.

FIG. 7. Characterization of the 258-260 DegP mutant. (A) Analysis of the oligomeric state of the 258-260 DegP mutant and its proteolytically inactive S210A variant by size exclusion chromatography. The plots at the top show the elution profiles of the mutants. The dashed lines indicate the expected elution volumes (Ve) for a hexameric form (11.5 ml) and a trimeric form (13 ml) of the mutant. The gels below the plots contained samples from each peak resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and stained with Coomassie brilliant blue. OD 280, optical density at 280 nm. (B) Sedimentation velocity data for the 258-260 DegP$_{S210A}$ mutant expressed
was not affected (data not shown); however, the mutation had a dramatic effect on the protease activity with MDH and β-casein, which was completely eliminated. Hydrolysis of lysozyme was not completely eliminated, but it was significantly decreased compared with the activity of wild-type DegP (Fig. 7C).

Interestingly, a mutant lacking the entire PDZ2 domain and the flexible loop connecting the protease and PDZ1 domains (Δ258-260+ΔPDZ2 DegP mutant) exhibited protease activity (Fig. 8A, right panel). This mutant eluted as a trimer in size exclusion chromatography (Fig. 8A, left and middle panels). Therefore, the loop appears to be essential for protease activity only when DegP forms a hexameric cage.

We hypothesized that the loop between the protease and the PDZ1 domain is essential only in the context of a hexameric DegP because the flexibility that this loop may confer allows exposure of the binding cleft of PDZ1 to substrate molecules, which may otherwise be buried within the hexameric cage. However, the hexameric Δ258-260 DegP S210A mutant was able to trap unfolded EGFP-SFS (Fig. 7D) and EGFP-SES (data not shown) in the fluorescence assay, suggesting that the PDZ1 binding cleft in this mutant is readily accessible to the C-terminal residues of the substrate molecules.

We next examined whether the flexible loop between the PDZ1 and PDZ2 domains is required for DegP activities by generating the Δ357-364 DegP mutant (Fig. 1A and B) and its inactive S210A variant. Analysis of the oligomeric state showed that these two mutants were stable dodecamers as determined by size exclusion chromatography (Fig. 8B, left and right panels) and sedimentation equilibrium analysis (Fig. 8C) (experimental molecular mass, 566 ± 38 kDa, corresponding to a

![Image of Figure 8](https://example.com/figure8.png)

**FIG. 8.** Characterization of the Δ258-260+ΔPDZ2 and Δ357-364 DegP mutants. (A) The plot on the left shows the elution profile of the Δ258-260+ΔPDZ2 DegP S210A mutant obtained with a Superdex-200 column after incubation at 4°C for 1 h. The dashed lines indicate the estimated elution volumes (Ve) for a hexameric form (11.8 ml) and a trimeric form (13.3 ml) of the mutant. A sample from the peak was resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and stained with Coomassie brilliant blue (middle panel). The averages and standard deviations for the proteolysis rates for the Δ258-260+ΔPDZ2 DegP mutant with MDH, β-casein, and lysozyme are shown on the right. WT, wild type; OD 280, optical density at 280 nm. (B) Plots showing the elution profiles for full-length (FL) and truncated (TF) forms of the Δ357-364 DegP mutant. The elution profile for the proteolytically inactive variant of this mutant (S210A) is also shown. The dashed lines in the elution profile on the left indicate the estimated elution volumes for a dodecameric form (10 ml) and a hexameric form (11.5 ml) of the mutant. The gels below the plots contained fractions of each peak resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. (C) Sedimentation equilibrium profiles expressed in terms of A280 versus the radius for the Δ357-364 DegP S210A mutant at a loading A280 of 0.58. Data were collected at 4,000 rpm (blue), 6,000 rpm (green), and 8,000 rpm (red) at 4.0°C and were analyzed in terms of a single ideal solute. Best fits are indicated by black lines through the experimental points, and the corresponding distributions of the residuals are shown. (D) Graph showing the averages and standard deviations for the proteolysis rates of the Δ357-364 DegP mutant with different substrates. (E) Plot showing the percentages of unfolded EGFP-SFS refolded in the absence and presence of the Δ357-364 DegP S210A mutant compared to the results obtained for the DegP S210A mutant. The plot was generated as described in the legend to Fig. 6.
stoichiometry of 12.3 ± 0.8). Two hexameric cages presumably associate to form the dodecamers observed. Consistent with our previous results, the mutation did not alter the chaperone activity of the protein (data not shown). However, the Δ357-364 DegP mutant was unable to hydrolyze MDH and β-casein (Fig. 8D). As observed for the mutant lacking the loop between the protease and PDZ1 domains, the PDZ1 binding cleft in the Δ357-364 DegP mutant was also accessible to the substrate molecules, as shown by the EGFP fluorescence assay (Fig. 8E), indicating that the loop between the PDZ domains in DegP is also essential for proteolysis but is not required for substrate access to the PDZ1 binding cleft.

Surprisingly, the Δ357-364 DegP mutant hydrolyzed lysozyme at rates similar to the rates observed for the wild type. These data, together with the slow lysozyme hydrolysis observed with the DegP protein lacking the loop between the protease and PDZ1 domains and the ΔPDZ1 DegP mutants, suggest that lysozyme may be able to access the catalytic triad of DegP in the protease domain without assistance from the PDZ1 domain. Consistently, we noticed that in our experimental conditions degradation of β-casein or MDH occurred with incubation times on the order of minutes, while the incubation times in the lysozyme hydrolysis assays had to be extended to 4 h in order to observe significant degradation (10). This observation also suggests that the targeting of lysozyme for degradation by DegP may be different from the targeting of β-casein and MDH.

Interestingly, the Δ357-364 DegP mutant showed partial cleavage during expression and purification, and an approximately 43-kDa truncated polypeptide was formed (Fig. 8B, left panel). A sample of this mutant containing mainly the truncated form of the protein was obtained by inducing expression for 3 h in the presence of 7 mM dithiothreitol. This mutant eluted as a dodecamer when it was loaded on a size exclusion chromatography column (Fig. 8B, middle panel), and it hydrolyzed lysozyme at rates similar to those observed for the predominantly nontruncated protein (data not shown). Therefore, it appears that truncation occurred without affecting the oligomeric structure of the protein, and we concluded that the lack of protease activity exhibited by this mutant in the presence of MDH and β-casein was due to the absence of the loop between the PDZ1 and PDZ2 domains rather than to the presence of ~40% of the protein in the truncated form.

**DISCUSSION**

Our data showing that the PDZ1 domain of E. coli DegP is essential for protease activity and that the last three C-terminal residues of the substrate molecules enhance sequestration of unfolded proteins provide new insights into the mechanism of substrate recognition for proteolytic degradation by DegP. More generally, because our study showed that both PDZ domains are dispensable for the chaperone activity of DegP, the data imply that DegP recognizes substrate molecules targeted for degradation and substrate molecules targeted for refolding in different ways. Thus, we hypothesized that the mechanism through which DegP recognizes the substrate may play a role in the protease-chaperone switch, dictating whether the substrate is degraded or refolded.

Our results are consistent with a mechanism in which the DegP protease first recognizes its substrates by binding to exposed C-terminal residues, much like other PDZ-containing proteins (e.g., Tsp [1, 16]). By doing this, it recruits a substrate to the catalytic site and subsequently cleaves it at multiple sites. In this context, the C-terminal residues allow DegP to tether the substrate and attack the primary cleavage sites (12). It has been shown that in other proteases, such as DegS, binding of the C terminus of a substrate to the PDZ domain induces a conformational change that activates the catalytic triad (35). Whether binding of the C-terminal residues of the substrate to the PDZ1 domain of DegP also causes such allosteric activation of the proteolytic domain is currently unknown. In this study, we found that the loop connecting the protease domain and the PDZ1 domain is essential for protease activity. In a manner similar to that in DegS, this loop in DegP may be required to transmit the conformational change that makes the catalytic site active. However, based on the protease activity observed with the Δ258-260+ΔPDZ2 DegP mutant that was also missing this loop but was proteolytically active, we concluded that this is probably not the case. Alternatively, we hypothesized that residues 258 to 260 between the protease and the PDZ1 domain and residues 357 to 364 between the two PDZ domains (Fig. 1A and B) are flexible loops required to expose the binding cleft of PDZ1 and allow substrate binding, but the EGFP fluorescence assay showed that the PDZ1 binding cleft in these mutants seems to be readily accessible to the C-terminal residues of the substrate molecules. Interestingly, the loop between the protease and the PDZ1 domain was found to be dispensable for the protease activity in a trimeric mutant of DegP. Additional experiments are necessary to determine why these loops are essential for the protease activity of hexameric DegP but dispensable in a trimeric mutant of this protein.

The X-ray structure of DegP suggests that the hydrophobic binding pockets for P₀ and P₋₁ in the binding cleft in the PDZ1 domain of DegP confer the binding specificity. Several crystal structures, including the structure of PSD-95 PDZ3 bound to its peptide ligand, initially suggested that the residue at P₋₁ points away from the interaction surface of the binding cleft, which correlates with the relatively low specificity of this PDZ domain with respect to substrate residues at P₋₁ (3). However, subsequent studies indicated that substitutions at this site could affect the binding preference for PSD-95 PDZ3 (22) and other PDZ domains, including PDZ1 in the NHERF (13, 14, 34) and InaD (18) proteins. Our results indicate that point mutations in the residue at P₋₁ of the SFS C-terminal tag greatly influenced the ability of DegP to sequester the unfolded substrate. In fact, changes in a negatively charged or small nonpolar residue, such as alanine, dramatically increased trapping by PDZ1, which correlated well with the presence of the positively charged residue arginine 325 in the vicinity of the –1 binding pocket in the PDZ1 domain (Fig. 1D).

Our results showed that the C-terminal SsrA tag and other C-terminal tags increased the sequestration of unfolded EGFP. Although the EGFP refolding assay used in this study does not measure binding directly, the increased sequestration of unfolded substrate is consistent with changes in the C-terminal tag binding to the E. coli PDZ1 DegP domain. Moreover, the results obtained with the SsrA tag are in agreement with the results of a previous report showing the ability of the
PDZ domains of S. enterica serovar Typhimurium DegP to bind the SsrA peptide (30). Interestingly, in this previous report, isolated E. coli DegP PDZ domains were unable to bind the SsrA peptide.

Previous studies have shown that PDZ domains are highly modular domains and that their structure is usually not compromised by mutations within the PDZ domain itself (32) or in neighboring domains. Many of our results confirmed that the mutants in this study were properly folded, that all of the mutants tested showed chaperone activity, and that the oligomeric structures were stable even upon incubation for 1 h at 43°C. Further proof that the structural integrity of these mutants was maintained was obtained from sedimentation velocity analyses; mutants with small deletions (e.g., Δ258-260 DegPΔN210A mutant) had a c(s) distribution very similar to that of the corresponding full-length DegP, whereas mutants with larger deletions had sedimentation values that appeared to correlate with their masses. It is therefore unlikely that the results obtained with these mutants reflected incorrect folding or protein instability.

ACKNOWLEDGMENTS

We thank Alba Guarne and Cecelia Trainor for insightful comments and critical reading of the manuscript. We thank Mochan Li for preliminary work on the Δ357-364 DegP mutant and Saeid Darvish-Kazem for constructing the EGFP-SFS mutants.

This work was supported by grants from the Canadian Institutes of Health Research, the Canada Foundation for Innovation, and the Ontario Innovation Trust. J.O. was a recipient of a Canadian Institutes of Health Research salary award. R.G. was supported by the Intramural Research Program of the NIH National Institute of Diabetes and Digestive and Kidney Diseases.

REFERENCES