Substrate Recognition through a PDZ Domain in Tail-Specific Protease†

Kirk D. Beebe, Jinu Shin, Jing Peng, Chaity Chaudhury, Jena Khera, and Dehua Pei*

Ohio State Biochemistry Program and Department of Chemistry, The Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210

Received November 24, 1999; Revised Manuscript Received January 11, 2000

ABSTRACT: Tail-specific protease (Tsp) is a periplasmic enzyme that selectively degrades proteins bearing a nonpolar C-terminus. Its substrate specificity suggests that Tsp may contain a substrate recognition domain, which selectively binds to the nonpolar C-termini of substrate proteins, separate from its catalytic site. In this work, we show that substrate recognition of Tsp is mediated by a PDZ domain, a small protein module that promotes protein−protein interactions by binding to internal or C-terminal sequences of their partner proteins. Partial proteolysis by V8 protease at a single peptide bond immediately N-terminal to the PDZ domain resulted in two distinct and relatively stable fragments and complete loss of catalytic activity. Photoaffinity labeling with a fluorescent nonpolar peptide caused the covalent attachment of the peptide to a single site on the Tsp protein. Systematic deletion mutagenesis of Tsp localized the binding site to amino acids 206−307, a region that completely encompasses the putative PDZ domain (217−301). The isolated PDZ domain (amino acids 206−334) is capable of folding into a well-behaved structure and binds to a nonpolar peptide with a dissociation constant (K_D) of 1.9 μM, similar to that of the intact Tsp protein. Site-directed mutagenesis of a surface residue at the peptide binding site of the PDZ domain, valine 229, to Glu or Gln resulted in an increase in the K_M value but had no effect on the k_{cat} value. The use of a separate substrate recognition domain such as a PDZ domain may be a general mechanism for achieving selective protein degradation.

Intracellular proteolysis occurs in all living organisms, providing a variety of functions including the elimination of abnormal proteins, the maintenance of amino acid pools during starvation, the generation of protein fragments that act as hormones, antigens, and other effectors, and the regulation of the intracellular concentrations of certain proteins. Because protein degradation must be highly selective in order for the cell not to cannibalize itself, the proteases or protease complexes need to recognize specific signals embedded in their cognate substrates. It has been recognized that the sequence identity at the N- and C-termini of a protein serves as a determinant of proteolytic degradation (*1*, *2*). The effect of the N-terminal residue on the metabolic stability of a protein, the N-end rule, has been relatively well studied (*3*). The mechanism by which the C-terminal sequence of a protein determines its proteolytic susceptibility has only recently started to be elucidated. An important advance in this area was the identification and cloning of a tail-specific protease (Tsp)[1] in *Escherichia coli* periplasm, which selectively degrades proteins with nonpolar C-terminal sequences (*4*−*6*). Tsp and its cytoplasmic counterpart (*7*) may also be involved in the elimination of aberrant polypeptides synthesized from defective mRNAs using a peptide tagging system (*8*).

Tsp represents a new class of protease. It was first identified as an activity that degrades a λ repressor variant with a nonpolar C-terminal sequence, WVAAA, but not the wild-type repressor, which has a polar C-terminal sequence, RSEYE (*4*). Cloning of the *tsp* gene showed that it had no sequence homology to any of the well-characterized members of other protease families and was not inhibited by several common protease inhibitors (*4*). Site-directed mutagenesis study has established that serine 430 and lysine 455, both in the C-terminal portion of the protein, are essential for catalytic activity (*9*). On the basis of this observation, it has been proposed that Tsp belongs to the family of proteases that use a serine-lysine dyad for catalysis; the members of this family include LexA, type I signal peptidase, and the class A β-lactamases (*10*). The unusual substrate specificity of Tsp suggests that it may contain a separate binding site for the hydrophobic tails of its substrate proteins that is distinct from the catalytic site (*5*, *11*). Amino acid sequence analysis has revealed an internal region with some sequence similarity to the PDZ domain (*12*). PDZ domains are ∼100 aa modules initially found in a number of eukaryotic cell-junction-associated proteins, where they serve to localize ion channel proteins within a particular region of the membrane by binding to the C-termini of the channel proteins (*13*−*16*). PDZ domains were subsequently found in many other proteins such as phosphatases and kinases (*12*, *17*−*19*). Crystal structural studies of several PDZ domains have revealed that the C-termini of their partner proteins bearing specific sequences bind to a surface cleft in the PDZ domain, with the C-terminal carboxylate group bound by the highly conserved GLGF motif of the PDZ domain through hydrogen bonds to the backbone amides (*20*−*22*). In addition to the

† This work was supported by funds from the National Institutes of Health (AI 40575) and The Ohio State University. J.K. was a summer undergraduate researcher supported by the REU program of the National Science Foundation (CHE 9619937).

* To whom correspondence should be addressed: Department of Chemistry, The Ohio State University, 100 West 18th Avenue, Columbus, OH 43210. Telephone (614) 688-4068; Fax (614) 292-1532; E-mail pei.3@osu.edu.

[1] Abbreviations: Tsp, tail-specific protease; EDANS, 5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid.

10.1021/bi992709s CCC: $19.00
Published on Web 02/23/2000

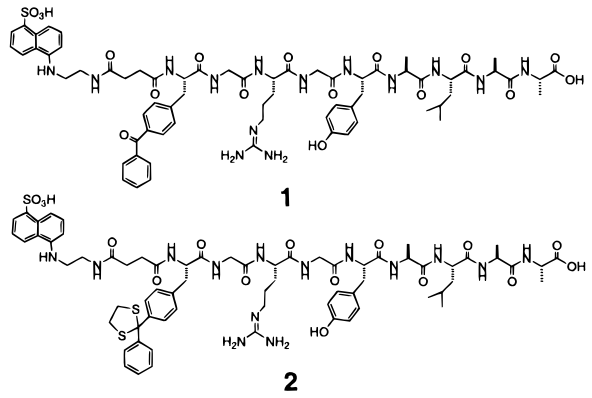

FIGURE 1: Structures of the PDZ-binding peptides.

carboxylate group, the C-terminal three or four residues are also involved in intricate interactions with the PDZ domain (*23*−*27*). The substrate specificity of Tsp is consistent with having a PDZ domain as the separate substrate binding site, recognizing proteins with the proper C-terminal sequences for degradation. In this work, we show that Tsp indeed contains a PDZ domain as the substrate recognition domain through a combination of partial proteolysis, photoaffinity labeling, and site-directed mutagenesis studies.

MATERIALS AND METHODS

Materials. All amino acids, Wang resin, and other peptide synthesis reagents were purchased from Advanced ChemTech (Louisville, KY). V8 protease, 5-[(2-aminoethyl)amino]-naphthalene-1-sulfonic acid (EDANS), and biotin were from Sigma Chemical Co. Fmoc-*p*-benzoyl-L-phenylalanine (Bpa) was purchased from Peninsula Laboratories. Oligonucleotide primers were ordered from Integrated DNA Technologies (Coralville, IA).

Synthesis of Photoaffinity Labels. Peptides were synthesized on Wang resin on 0.2-mmol scale with a homemade peptide synthesis apparatus and fluorenylmethoxycarbonyl-(Fmoc-) protected amino acids (*28*). Three equivalents of amino acids were used during coupling reactions and double coupling was performed if necessary (as judged by ninhydrin test). Coupling of the biotin moiety was accomplished by the addition of 3 equiv of biotin, *N*-hydroxybenzotriazole monohydrate (HOBt), *O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HBTU), and *N*-methylmorpholine. The EDANS group was added as previously described (*11*). Cleavage and deprotection of peptides were performed for 1 h at room temperature in a cocktail containing 5 mL of trifluoroacetic acid (TFA), 0.13 mL of ethanedithiol, 0.25 mL of thioanisole, 0.10 mL of anisole, 0.375 g of crystalline phenol, and 0.25 mL of water. The solvent was evaporated and the crude peptides were triturated with diethyl ether (3 × 15 mL). The crude peptides were purified by reversed-phase HPLC (C-18 column) and analyzed by matrix-assisted laser desorption ionization (MALDI) mass spectrometry. Peptide **1** (Figure 1): calculated mass 1377.0, found 1379.3. Peptide **2**: calculated mass 1453.1, found 1455.4.

Construction of Tsp Mutants. Site-directed mutants were constructed by the Quik-Change mutagenesis method (Stratagene). Plasmid pKK101 (*5*), which contains the entire *tsp* gene plus six histidine residues at the C-terminus, was used as the template. The primers used were 5′-GGAAGGTATTGGCGCCGAGCTGCAAATGGATGATG-3′ (primer 1) and 5′-GGAAGGTATTGGCGCACAGCTGCAAATGGATGATG-3′ (primer 2), and their complements for the V229E and V229Q mutations, respectively. A new *Pvu*II restriction site was introduced into the plasmid DNA by the V229Q mutation, whereas an *Nar*I site was created by the V229E mutation, providing convenient markers for screening for mutants.

Deletion mutants were generated by the polymerase chain reaction (PCR) using plasmid pKK101 (*5*) as template. The following primers were used to construct the various mutants: primer 3, 5′-GGGAATTCTTACTTGACGGGAGCGGGTTG-3′; primer 4, 5′-GGAAGGCCATATGAGCTATCTTTCCCCGCGT-3′; primer 5, 5′-GGGAATTCTTAACGGGTCTTGGTCCCT-3′; primer 6, 5′-CATGCCATGGCATATGAGCGAAGATGTTTTCTCGCT-3′; primer 7, 5′-GGGAATTCTTATTTCTCTTTACCGACGGTCTT-3′; and primer 8, 5′-CATGCCATGGCATATGGTAGAAGATATCACGCGTG-3′). Primers 3 and 4 were used to construct the N-terminal deletion mutant, ΔNT Tsp (amino acids 206−660). The PCR product was digested with restriction endonucleases *Nde*I and *Eco*RI and inserted into the corresponding sites of plasmid pET22b (Novagen). The resulting plasmid was digested with *Nde*I and *Sca*I, and the Tsp coding fragment was ligated to the *Nde*I/*Sca*I-linearized plasmid pET15b (Novagen). This procedure added an N-terminal histidine tag to ΔNT Tsp. Primers 6 and 7 were employed to isolate the PDZ domain, PDZ-a (amino acids 185−334). Primers 5 and 6 were used to construct PDZ-b (amino acids 185−307). Primers 4 and 7 were used to construct PDZ-c (amino acids 206−334). Primers 4 and 5 were used to construct PDZ-d (amino acids 206−307). Primers 7 and 8 were used to construct the C-terminal truncation mutant, ΔCT Tsp (amino acids 1−334). The PCR products from all above reactions were digested with *Nde*I and *Eco*RI and ligated with *Nde*I/*Eco*RI-linearized plasmid pET-28a (Novagen). This procedure resulted in the addition of a six-histidine tag to the N-terminus of each construct. Amino acid numbering of all mutants used in this study and positions referred to are based on that used in previous work (*9*).

Purification of Wild-Type and Mutant Tsp. Escherichia coli BL21 (DE3) cells carrying the appropriate expression plasmid were grown in 3 L of LB medium containing 75 mg/L ampicillin at 37 °C. When OD_{600} reached 0.7, the cells were induced by the addition of isopropyl β-D-thiogalactopyranoside to a final concentration of 140 μM and grown for an additional 3−6 h at 28 °C. The cells were harvested by centrifugation at 4 °C (6000 rpm for 10 min), resuspended in 60 mL of ice-cold column binding buffer (5 mM imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, pH 7.9), and lysed by sonication. The crude cell lysate was centrifuged for 15 min at 17 000 rpm (Sorvall SS-34 rotor), and the clear supernatant (65 mL) was loaded onto a 10-mL His·Bind column (Novagen) (*29*) that had been equilibrated with the binding buffer. The bound protein was eluted with 60 mL of an elution buffer (1 M imidazole, 0.5 M NaCl, and 20 mM Tris-HCl, pH 7.9) and the fractions containing homogeneous Tsp were pooled. After being concentrated in a Centriprep 10 concentrator (Amicon), the protein sample was mixed with glycerol (33% final) and stored at −80 °C. Protein concentrations were determined by the Bradford

assay with bovine serum albumin as standard (*30*). Wild-type Tsp (without histidine tag) was purified as previously described (*11*).

Tsp Assay. All assays were performed with 2-*N*-[[(1-sulfono-5-naphthyl)amino]ethyl]aminosuccinyl-Ala-Ala-Arg-Ala-Ala-[N^{ϵ}-[4-[4-(dimethylamino)phenylazo]benzoyl]lysyl]-(6-aminocaproyl)$_2$-Glu-Asn-Tyr-Ala-Leu-Ala-Ala as substrate (*11*). Cleavage of the substrate at the Ala−Arg peptide bond by Tsp causes an increase in fluorescence yield of the EDANS group, which was monitored on a Perkin-Elmer LS-5 spectrofluorometer. A typical reaction (1.5 mL) contained 50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 0−20 μM substrate, and 44.4 nM of Tsp enzyme. To determine the rate of ΔNT Tsp, 1.4 μM of ΔNT Tsp was incubated with 50 μM substrate and the reaction was carried out and monitored similarly as described above.

Photoaffinity Labeling. All labeling reactions were carried out at 4 °C in a Rayonet photochemical reactor equipped with two RPR-3500 lamps (350 nm) (Southern New England Ultraviolet Co.). A reaction mixture (total volume of 30 μL) containing 25 mM Tris-HCl (pH 8.0), 100 mM NaCl, wild-type or mutant Tsp protein (typically 50 μg), and 100 μM peptide **1** was irradiated for 20 min under the 350 nm lamps and then analyzed by electrophoresis on SDS−PAGE gels. The labeled protein bands in the gel were visualized on a Fisher Biotech 302 nm transluminator (Model FBTI 614) and photographed with a Polaroid camera. In competition experiments, a reaction mixture containing 25 mM Tris-HCl (pH 8.0), 100 mM NaCl, 50 μg of Tsp, 100 μM fluorescently labeled peptide **1**, and 0−1200 μM unlabeled competing peptide (peptide **2**) was incubated under the 350 nm lamps for 10 min at 4 °C, followed by SDS−PAGE analysis.

To identify the site of photolabeling, PDZ-a domain (3 mg, 187 nmol of PDZ domain) was irradiated in the presence of 40 μM peptide **1** for 3 h at 4 °C (total volume of 10 mL). To remove the excess photoaffinity labels, the sample was loaded onto a His·Bind metal affinity column (Novagen). After extensive washing, the bound protein was eluted with 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 1 M imidazole and exhaustively dialyzed against 100 mM ammonium bicarbonate buffer (pH 7.9). The resulting protein was digested with 1:25 (w/w) trypsin at 37 °C for 3 days. The peptides generated were separated on a reversed-phase HPLC equipped with an analytical C_{18} column. A linear gradient of 0−50% acetonitrile in water plus 0.05% trifluoroacetic acid was carried out in 1 h at a flow rate of 1.0 mL/min. Individual peaks were collected, lyophilized to dryness, and analyzed on a Perkin-Elmer electrospray ionization mass spectrometer (model API-300) in the positive ion mode.

V8 Proteolysis of Tsp. Tsp (6 μg) was digested with 0.3 μg of V8 protease in 0.1 M Tris-HCl, pH 7.4, for varying lengths of time (0−60 min) at room temperature (total reaction volume of 12 μL). The reactions were quenched by the addition of 12 μL of 2× SDS−PAGE gel loading buffer and heating for 10 min at 95 °C and analyzed on an SDS−10% polyacrylamide gel. For mass spectrometric analysis, the reaction was similarly carried out (60 μg of Tsp in a 120-μL reaction) and quenched by the addition of TFA to 0.5% final concentration and stored at −80 °C until use. The resulting sample was passed through a C-8 reversed-phase HPLC column (eluted with acetonitrile plus 0.1% trifluoroacetic acid) before being injected into a Perkin-Elmer electrospray ionization mass spectrometer (model API-300). Further confirmation of the site of cleavage was achieved by digestion of a Tsp variant containing a C-terminal six-histidine tag (*5*), denaturation with 6 M urea, and passing the sample through a Ni^{2+} metal affinity column (Novagen) (*29*). After being washed with 6 column volumes of the binding buffer (20 mM Tris-HCl, pH 7.9, 500 mM NaCl, 5 mM imidazole, and 6 M urea), the column was eluted with 20 mM Tris-HCl, pH 7.9, 500 mM NaCl, 1 M imidazole, and 6 M urea. The eluted proteins (∼2 mL) were precipitated with 4 volumes of cold acetone and analyzed on an SDS−8% polyacrylamide gel.

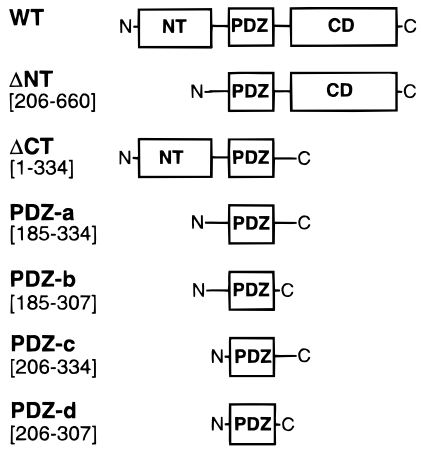

FIGURE 2: Schematic representation of the structures of wild-type and mutant Tsp used in this work. The numbers in brackets indicate the starting and ending points within the Tsp protein used for each construct, according to the numbering scheme of Keiler and Sauer (*9*). Each construct also contained an N-terminal histidine tag. NT, N-terminal domain; CD, catalytic domain.

Binding Assay. The dissociation constants (K_D) of peptide **1** to native Tsp and the isolated PDZ-c domain were determined by measuring fluorescence resonance energy transfer from the tryptophan(s) in the protein to the EDANS group on the peptide (*31*). A 1.5-mL solution containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 250 nM peptide **1**, and 0.1−8.0 μM wild-type Tsp or PDZ-c in a quartz cuvette was placed in a Perkin-Elmer spectrofluorometer. The excitation wavelength was set at 290 nm while the emission at 520 nm was monitored. The observed fluorescence yield at 520 nm, F, was fitted to the equation $F = F_{max}[P]/(K_D + [P])$, where [P] is the concentration of Tsp or PDZ-c added and F_{max} is the maximum fluorescence intensity at 520 nm. Inner filtering effects were minimal over the concentration range of 0.1−8.0 μM PDZ-c; in the absence of peptide **1**, the fluorescence yield of PDZ-c at 370 nm increased essentially linearly with protein concentration. Experiments were carried out in triplicate.

RESULTS

V8 Proteolysis of Tsp. Partial proteolysis of Tsp by V8 protease resulted in two fragments of estimated molecular masses of 55 and 25 kDa (Figure 3), accompanied by the loss of Tsp activity. Since the intact Tsp has an apparent molecular mass of 75−80 kDa on an SDS−polyacrylamide gel, this indicates selective cleavage of Tsp at a single site by V8 protease. To determine whether the cleavage occurred near the N-terminus or C-terminus of Tsp, the V8 digestion

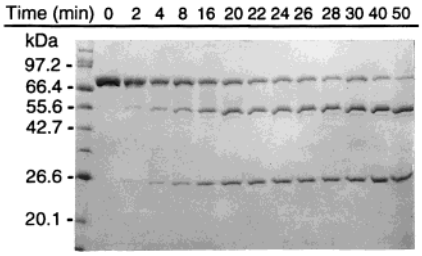

FIGURE 3: SDS−10% polyacrylamide gel showing the cleavage of Tsp by V8 protease. At the indicated time points, the reaction was quenched by the addition of equal volume of 2× SDS−PAGE loading buffer and heating at 95 °C for 5 min.

was performed on a Tsp variant that carries a C-terminal histidine tag (*5*). The resulting fragments were denatured by 6 M urea and loaded onto a Ni^{2+} affinity column (*29*), upon which the histidine-tagged C-terminal fragment should bind, whereas the N-terminal fragment would not. SDS−PAGE analysis revealed that the bound fragment had a molecular mass of 55 kDa (data not shown). Thus, the single V8 cleavage occurred near the N-terminus of Tsp sequence. To locate the precise site of cleavage, the reaction products were analyzed by electrospray ionization mass spectrometry (ESI MS). The molecular masses of intact Tsp (wild type) and the two cleavage fragments were found to be 75 168, 49 491, and 25 689 Da, respectively. This is most consistent with cleavage of the peptide bond between N211 and T212 (calculated masses of 49 506 and 25 708 Da for the two fragments). Surprisingly, all of our attempts to separate the two fragments under nondenaturing conditions (gel-filtration and ion-exchange chromatography) failed, suggesting that the two fragments remain tightly associated even after proteolysis. We also noted that prolonged storage of the protein at 4 °C produced a similar cleavage pattern.

Photoaffinity Labeling. Previous studies with both small peptide and large protein substrates suggest that Tsp contains a separate substrate binding site that is responsible for recognizing the hydrophobic C-terminus of a substrate (*5*, *11*). To test for the presence of such a substrate recognition site, we synthesized a fluorescent photoaffinity label (peptide **1**, Figure 1), in which an EDANS group and a photolabile benzophenone group were attached to the N-terminus of octapeptide GRGYALAA, a sequence known to act as a competitive inhibitor of Tsp (*11*). Binding of peptide **1** to Tsp would bring the benzophenone group to the vicinity of the substrate recognition site; upon exposure to 350 nm light, the fluorescent peptide would be covalently attached to the protein via the benzophenone group (*32*) and the protein would become fluorescent. As expected, when Tsp (50 μg) was incubated with peptide **1** (100 μM) under 350 nm light for 25 min, the peptide became covalently attached to Tsp, as evidenced by the appearance of a fluorescent band at ∼75 kDa on SDS−polyacrylamide gels (Figure 4, panels A and B, lane 2). The extent of modification was quantitated by examining the remaining protease activity of Tsp. The maximal amount of labeling was found to be 64% after 200 min of exposure to the UV light. When a control experiment was carried out with phosphatase SHP-1 (*33*), a 68-kDa protein readily available in this laboratory, no labeling was observed under the same set of conditions (Figure 4, panels A and B, lane 5). Further, successful labeling required the presence of all three components: Tsp protein, the peptide,

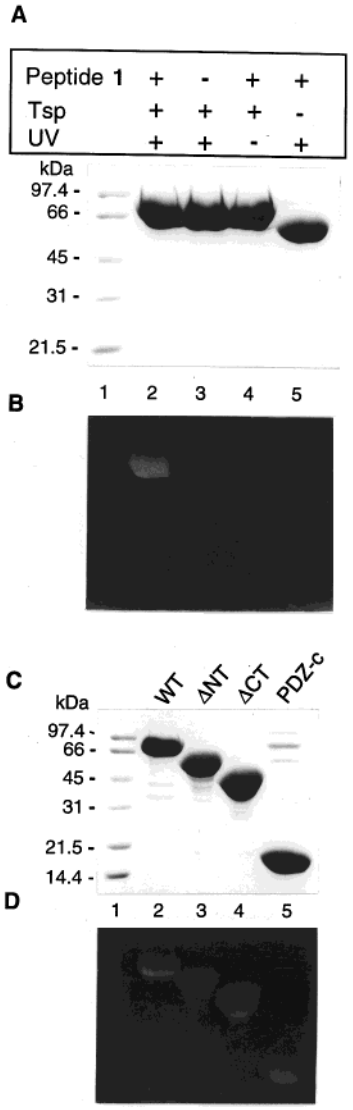

FIGURE 4: Photoaffinity labeling of wild-type Tsp and its deletion mutants. (A) Coomassie blue-stained SDS−polyacrylamide gel showing wild-type Tsp labeled with peptide **1**. Lane 1, molecular mass markers; lane 2, positive reaction with all components present; lane 3, no peptide; lane 4, no UV light; and lane 5, negative control with SHP-1 protein (no Tsp). (B) Fluorescent image of the same gel shown in panel A. (C) Coomassie blue-stained SDS−polyacrylamide gel showing Tsp deletion mutants that had been labeled by peptide **1**. Lane 1, molecular mass markers; lane 2, wild-type Tsp; lane 3, ΔNT Tsp; lane 4, ΔCT Tsp; and lane 5, PDZ-c domain. (D) Fluorescent image of the same gel shown in panel C.

and UV light. Peptide **1** was not cleaved by Tsp under the assay conditions despite the high enzyme concentration. This result demonstrates that Tsp indeed contains a recognition site for the nonpolar C-terminus of a substrate, which may or may not be the same as the catalytic site.

To locate the binding site for peptide **1**, photoaffinity labeling experiments were carried out with various deletion mutants of Tsp. Primary sequence analysis has previously identified a region (amino acids 217−301) that exhibits some sequence similarity to the PDZ domain (*12*). Since PDZ domains are known to bind to the C-termini of their partner proteins, the putative PDZ domain in Tsp may be the substrate recognition site. Consistent with this notion, when photoaffinity labeling experiments were carried out with ΔNT Tsp (amino acids 206−660), a truncation mutant lacking the

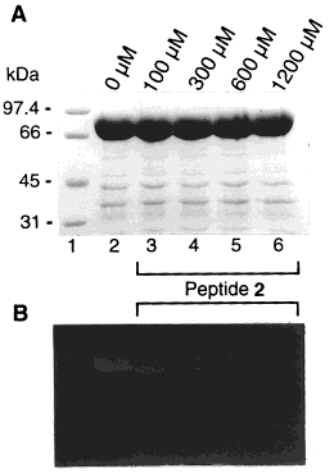

FIGURE 5: Competition between photoaffinity label (peptide **1**) and peptide **2** for binding to Tsp. (A) Coomassie blue-stained SDS−polyacrylamide gel showing the labeled Tsp in the presence of 0−1200 μM peptide **2** (lanes 2−6). Lane 1 shows molecular mass markers. (B) Fluorescent image of the same gel shown in panel A.

sequence N-terminal to the putative PDZ domain, efficient labeling was observed (Figure 4, panels C and D, lane 3). Similarly, a C-terminal truncation mutant, ΔCT Tsp (amino acids 1−334), which lacks all of the sequences C-terminal to the putative PDZ domain, was also efficiently labeled by peptide **1** (Figure 4, panels C and D, lane 4). These results suggest that the substrate recognition site is located between amino acids 206 and 334, the region where the putative PDZ domain is located (Figure 2). Therefore, the putative PDZ domain is the likely substrate recognition site of Tsp.

To isolate the PDZ domain, several truncations were performed by the PCR method (PDZ-a−d, Figure 2). We found that the PDZ domain core (PDZ-d, amino acids 206−307) as defined by sequence homology (*12*) failed to produce any soluble protein. Extension of the PDZ core by 27 amino acids at the C-terminus but not at the N-terminus afforded a soluble protein (PDZ-c, Figure 2). When PDZ-c (124 μM) was incubated with peptide **1** (100 μM) in the presence of 350 nm UV light, the PDZ domain was efficiently labeled with the fluorescent probe (Figure 4, panels C and D, lane 5). This labeling was both UV- and peptide-dependent. PDZ-a (amino acids 185−334) was also purified as soluble protein and was efficiently labeled by peptide **1** (data not shown). PDZ-b (amino acids 185−307) was purified only in very small quantities as the majority was present in inclusion bodies; however, data suggested that PDZ-b cross-links to peptide **1** as well.

To further rule out the possibility that the observed labeling may be due to nonspecific modification at multiple sites, we performed competition experiments, in which photoaffinity labeling by peptide **1** was carried out in the presence of a competing peptide (peptide **2**, Figure 1). Peptide **2** was obtained as a byproduct from the synthesis of peptide **1** and was formed by the addition of one molecule of ethylenedithiol, presumably to the benzophenone group, during the peptide deprotection step. Peptide **2** binds Tsp and inhibits Tsp activity but does not covalently cross-link to Tsp due to the loss of the chromophore. As shown in Figure 5, when increasing amounts of peptide **2** were added to the reaction between Tsp and peptide **1**, the amount of fluorescently labeled Tsp decreased. The SDS−polyacrylamide gel stained with Coomassie blue showed that there were equal amounts of Tsp in all reactions (Figure 5, panel A). Significant competition was observed when the concentration of peptide **2** (e.g., 300 μM) was in only slight excess of peptide **1** (100 μM), despite irreversible labeling by peptide **1**, suggesting that both peptides compete for binding to a specific site instead of a series of nonspecific sites. These results taken together demonstrate that PDZ is indeed the domain through which Tsp recognizes and binds its substrates.

To locate the site of covalent attachment on the PDZ domain, PDZ-a was incubated with excess peptide **1** in the presence of 350 nm light and the unreacted peptides were removed by dialysis. The covalently modified protein was digested with trypsin and the resulting peptides were separated by reversed-phase HPLC. All major peaks were collected and checked for the presence of fluorescence. Out of the 20 peaks collected, peaks 11 and 15 showed strong absorbance at 214 nm and were intensely fluorescent. These two fractions were further analyzed by electrospray ionization mass spectrometry (ESI MS). Peak 11 gave a peak at m/z = 840.7. Although this closely matches the calculated molecular mass of an unmodified peptide from the PDZ domain, L_{295}-EILPAGK_{302} (calcd MH^+ = 840.52), this peptide should not be fluorescent. Therefore, we assume that the peak at m/z = 840.7 carried two positive charges, derived from a single peptide that had a molecular mass of (840.7 × 2) − 2 = 1679.4 Da. After trypsin digestion at the arginyl residue, the N-terminal fragment of peptide **2**, which contained both the EDANS group and the benzophenone moiety, had a molecular mass of 830.4 Da as determined by ESI MS under the same conditions. Thus, the unknown peptide fragment derived from the PDZ domain should have a molecular mass of 1679.4 − 830.4 = 849.0. This value matches closely the mass of peptide M_{325}SVKTVGK_{332} (calcd M = 848.5), which is located at the extreme C-terminus of the PDZ-a protein. No other peaks matched the observed molecular mass (=10 Da deviation). Peak 15 produced a peak at m/z = 1072.0, which did not match either a fragment from the unmodified PDZ-a or any fragment plus the fluorescent probe. One possibility is that a modified peptide might have undergone other structural changes during the experimental process. Experiments with an N-terminally biotinylated peptide (which has similar sequence to peptide **1**) also resulted in labeling of residues in the C-terminus of PDZ-a.

*Direct Binding of Peptide **1** to Tsp and PDZ Domain.* Binding of peptide **1** to Tsp was also demonstrated by the presence of fluorescence resonance energy transfer from the tryptophan residue(s) in Tsp (excitation wavelength = 290 nm) to the EDANS group in the peptide (emission wavelength = 520 nm). By varying the PDZ domain (PDZ-c) concentration (0.1−8.0 μM) and examining the fluorescence emission from the EDANS group at 520 nm (peptide concentration fixed at 250 nM), a dissociation constant (K_D) of 1.9 ± 0.3 μM was observed. The binding affinity of peptide **1** to Tsp was similarly determined, resulting in a K_D value of 1.8 ± 0.3 μM.

Catalytic Property of Tsp Mutants. All PDZ domains contain a GLGF loop motif or its homologue (*12*, *20*). This sequence is critical for recognizing the C-terminus of its partner protein by forming a series of hydrogen bonds between the PDZ backbone amides and the carboxylate of the partner protein (*20*). In Tsp, the corresponding sequence

Table 1: Kinetic Constants for Wild-Type and Mutant Tsp Enzymes[a]

enzyme	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\times 10^3$ M^{-1} s^{-1})
wild type	4.4 ± 0.6	0.050 ± 0.000	11
V229Q	9.5 ± 1.3	0.050 ± 0.010	5.3
V229E	11.1 ± 1.4	0.050 ± 0.002	4.5
ΔNT			<0.003
ΔCT			NA

[a] Assay reactions were performed in 50 mM Tris-HCl, pH 8.0, 250 mM NaCl, and 44.4 nM WT Tsp or full-length mutants in the presence of 0−20 μM substrate (*11*). The ΔNT construct (1.4 μM) was assayed in the presence of 50 μM substrate and the rate was determined by following the reaction over 45 min. Data reported are the mean ± SD of three independent experiments. NA, no detectable activity.

is GIGA and valine 229 is located immediately C-terminal to the GIGA motif (*4*). With the PDZ domain of PSD-95 as a model (*20*), Val-229 of Tsp should be solvent-exposed and is likely involved in hydrophobic interactions with a substrate. With mutation of this residue it was expected that the binding affinity of Tsp to a hydrophobic peptide would change, but major changes in the global structure of Tsp would not occur. To this end, valine 229 was mutated to a glutamine and glutamate, respectively, with the expectation that the hydrophilic side chains of Gln and Glu would lower the affinity of Tsp to substrates containing hydrophobic C-termini. Both V229Q and V229E mutants produced soluble proteins, which were efficiently labeled by peptide **1** upon UV irradiation (data not shown) and were catalytically active (Table 1). While these mutations had no detectable effect on the k_{cat} value (0.05 s^{-1}), they caused a 2−3-fold increase in the K_M value when assayed against a peptide substrate (*11*). This is consistent with the notion that changes in the PDZ domain only affect the overall binding affinity of a substrate but do not affect the catalytic power of the active site, which is located at the C-terminal region of Tsp and is physically remote from the PDZ domain.

The deletion mutants were also evaluated for catalytic activity. Although ΔNT Tsp contains both the PDZ domain and all of the catalytic residues as identified by mutagenesis studies (*9*), it had little catalytic activity, with an apparent k_{cat}/K_M value of 3.0 ± 0.3 M^{-1} s^{-1} toward the fluorescent peptide substrate (*11*) (Table 1). We believe that ΔNT Tsp is catalytically inactive and the observed residual activity is likely due to protease contamination, since the ΔNT Tsp preparation exhibited the same amount of activity toward a control substrate, which contains a glutamate at the C-terminus but is otherwise identical to the fluorescent peptide substrate (*11*). We also noted that ΔNT Tsp was somewhat less stable than wild-type Tsp, undergoing denaturation when the ionic strength of the solution was low. The C-terminal truncation mutant, ΔCT Tsp, was also catalytically inactive.

DISCUSSION

Tsp is different from the other known proteases in that it recognizes both the C-terminus and the sequences flanking the scissile bond of a substrate. The substrate specificity of Tsp is primarily determined by the sequence at the C-terminus, with protein containing small hydrophobic C-terminal residues being the preferred substrates (*5*). On the other hand, the sequences flanking the known cleavage sites

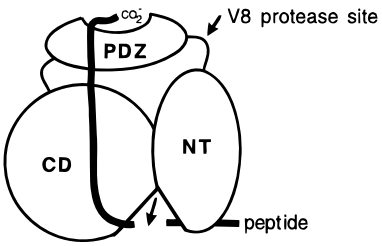

FIGURE 6: Proposed model of substrate recognition and catalysis by Tsp. The PDZ domain recruits a peptide or protein substrate via the C-terminus of the substrate. Once the substrate is recruited, the catalytic domain (CD) performs catalysis while the N-terminal domain (NT) assists either directly or indirectly in an unknown manner.

in several peptide as well as protein substrates are quite diverse (*5*). This unusual specificity allows Tsp to selectively degrade any proteins with nonpolar C-termini but not those proteins that are needed for proper cellular functions and which necessarily do not contain nonpolar C-termini. One of the functions of Tsp (or its homologue) is to degrade polypeptides synthesized from damaged mRNAs, which are potentially toxic to the cell and are targeted for destruction even before they are released from the ribosome by tagging a nonpolar tail to their C-termini (*8*). Because the cleavage site in a substrate and its C-terminus have been found to be distinct and typically separated by significant distance (*5*), it is unlikely that a single active site can interact with both motifs. Therefore, a separate binding domain has been proposed to be responsible for recognizing the C-terminal sequence of a potential substrate (*5*). In this work, we have shown that this C-terminal binding site is mediated by a PDZ domain, based on several lines of evidence. First, sequence alignment shows that amino acids 217−301 share significant sequence homology to several well-known PDZ domains (*12*). Second, proteolysis by V8 protease, which cleaves after aspartyl and glutamyl residues, resulted in cleavage at a single site immediately before the PDZ domain. This suggests the presence of individual domains in Tsp. Third, deletion mutagenesis locate the C-terminal binding site to amino acids 206−307, the same region where the PDZ domain resides. Fourth, the Tsp fragment containing amino acids 206−334 forms a soluble protein and is capable of binding to a peptide with a nonpolar C-terminus. Finally, mutagenesis in the PDZ domain, which is remote from the catalytic site, decreased the catalytic activity of Tsp by reducing its affinity to the substrate.

We propose that Tsp contains a minimum of three separate domains: an N-terminal domain (amino acids 1−205), a PDZ domain (amino acids 206−307), and a C-terminal catalytic domain (amino acids 308−660) (Figure 6). We speculate that the PDZ domain first recognizes a cognate substrate by binding to its nonpolar C-terminus. By doing so, it recruits a substrate to the catalytic site, which subsequently cleaves the substrate at multiple sites. This would predict that the catalytic site is a relatively poor enzyme that has very low affinity toward its substrates so that it cannot form a productive E·S complex without the help from the PDZ domain. Indeed, addition of peptide **1** or other peptides with nonpolar C-termini strongly inhibited the catalytic activity of Tsp (*11*). An alternative mechanism is that binding of the C-terminus of a substrate to the PDZ domain causes a conformational change that activates the

catalytic domain for peptide cleavage. This latter mechanism is less likely to be the case as it would predict that peptide **1** or other nonpolar peptides act as activators instead of inhibitors of Tsp. The function of the N-terminal domain remains unclear. It is clearly important for catalysis, despite containing none of the residues identified as critical for catalysis (*9*), as its removal resulted in a protein that is capable of binding to a substrate but has essentially no catalytic activity. Even a single cleavage in the apparent linker region between the N-terminal domain and the PDZ domain inactivated the enzyme. It is possible that the active site is formed by residues from both the N- and C-terminal domains. Note that after V8 proteolysis, the two fragments remained tightly associated.

Since the identification of Tsp in 1992, several Tsp homologues have subsequently been discovered in both prokaryotic and eukaryotic organisms (*9*, *34*). Sequence analysis shows that members of the HtrA protease family, which are heat-shock-induced serine proteases found in bacteria, yeast, and human, also contain one or multiple putative PDZ domains (*12*, *35*). Furthermore, the two-component, ATP-dependent proteases ClpAP, ClpXP, and ClpYQ as well as the Lon protease have recently been shown to contain a 100-amino acid sensor and substrate discrimination (SSD) domain in their ATPase component, which recognizes specific sequences within a substrate protein (*36*, *37*). Therefore, it appears that the use of a separate substrate recognition domain (e.g., PDZ or SSD domains) is a common mechanism by which specific proteases discriminate their target protein(s) from other nonspecific proteins.

REFERENCES

1. Bachmair, A., Finley, D., and Varshavsky, A. (1986) *Science 234*, 179−186.
2. Parsell, D. A., Silber, K. R., and Sauer, R. T. (1990) *Genes Dev. 4*, 277−286.
3. Varshavsky, A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12142−12149.
4. Silber, K. R., Keiler, K. C., and Sauer, R. T. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 295−299.
5. Keiler, K. C., Silber, K. R., Downard, K. M., Papayannopoulos, I. A., Biemann, K., and Sauer, R. T. (1995) *Protein Sci. 4*, 1507−1515.
6. Keiler, K. C., and Sauer, R. T. (1996) *J. Biol. Chem. 271*, 2589−2593.
7. Silber, K. R., and Sauer, R. T. (1994) *Mol. Gen. Genet. 242*, 237−240.
8. Keiler, K. C., Waller, P. R. H., and Sauer, R. T. (1996) *Science 271*, 990−993.
9. Keiler, K. C., and Sauer, R. T. (1995) *J. Biol. Chem. 270*, 28864−28868.
10. Paetzel, M., and Dalbey, R. E. (1997) *Trends Biochem. Sci. 22*, 28−31.
11. Beebe, K. D., and Pei, D. (1998) *Anal. Biochem. 263*, 51−56.
12. Ponting, P. C. (1997) *Protein Sci. 6*, 464−468.
13. Cho, K. O., Hunt, C. A., and Kennedy, M. B. (1992) *Neuron 9*, 929−942.
14. Woods, D. F., and Bryant, P. J. (1993) *Mech. Dev. 44*, 85−89.
15. Itoh, M., Nagafuchi, A., Yonemura, S., Kitani-Yasuda, T., and Tsukita, S. (1993) *J. Cell Biol. 121*, 491−502.
16. Ponting, C. P., and Philips, C. (1995) *Trends Biochem. Sci. 20*, 102−103.
17. Kornau, H. C., Seeburg, P. H., and Kennedy, M. B. (1997) *Curr. Opin. Neurobiol. 7*, 368−373.
18. Sato, T., Irie, S., Kitadi, S., and Reed, J. C. (1995) *Science 268*, 411−415.
19. Hata, Y., Butz, S., and Sudhof, T. C. (1996) *J. Neurosci. 16*, 2488−2494.
20. Doyle, A. D., Lee, A., Lewis, J., Kim, E., Sheng, M., and MacKinnan, R. (1996) *Cell 85*, 1067−1076.
21. Cabral, J. H. M., Petosa, C., Sutcliffe, M. J., Raza, S., Byron, O., Poy, F., Marfatia, S. M., Chisti, A. H., and Liddington, R. C. (1996) *Nature 382*, 649−652.
22. Schultz, J., Hoffmuller, U., Krause, G., Ashurst, J., Marcias, M. J., Schmieder, P., Schneider-Mergener, J., and Oschkinat, H. (1998) *Nat. Struct. Biol. 5*, 19−24.
23. Kornau, H.-C., Schenker, L. T., Kennedy, M. B., and Seeburg, P. H. (1995) *Science 269*, 1737−1740.
24. Kim, E., Nlethammer, M., Rothschild, A., Jan, Y. N., and Sheng, M. (1996) *Nature 378*, 85−88.
25. Songyang, Z., Fanning, A. S., Fu, C., Xu, J., Marfatia, S. M., Chishti, A. H., Crompton, A., Chan, A. C., Anderson, J. M., and Cantley, L. C. (1997) *Science 275*, 73−77.
26. Stricker, N. L., Christopherson, K. S., Yi, B. A., Schatz, P. J., Raab, R. W., Dawes, G., Bassett, D. E., Jr., Bredt, D. S., and Li, M. (1997) *Nat. Biotech. 15*, 336−342.
27. Schepens, J., Cuppen, E., Wieringa, B., and Hendricks, W. (1997) *FEBS Lett. 409*, 53−56.
28. Bodansky, M. (1993) *Principles of Peptide Synthesis*, 2nd ed., Springer-Verlag, Germany.
29. Hichuli, E., Dobeli, H., and Schacher, A. (1987) *J. Chromatogr. 411*, 177−184.
30. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
31. Selvin, P. R. (1994) *Methods Enzymol. 246*, 300−334.
32. Dorm'an, G., and Prestwich, G. D. (1994) *Biochemistry 33*, 5661−5673.
33. Pei, D., Neel, B. G., and Walsh, C. T. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1092−1096.
34. Oelmuller, R., Herrmann, R. G., and Pakrasi, H. B. (1996) *J. Biol. Chem. 271*, 21848−21852.
35. Pallen, M. J., and Wren, B. W. (1997) *Mol. Microbiol. 26*, 209−221.
36. Smith, C. K., Baker, T. A., and Sauer, R. T. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 6678−6682.
37. Van Melderen, L., and Gottesman, S. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 6064−6071.

BI992709S