Structural and functional analysis of the PDZ domains of human HtrA1 and HtrA3

STEVEN T. RUNYON,[1] YINGNAN ZHANG,[2] BRENT A. APPLETON,[2] STEPHEN L. SAZINSKY,[3] PING WU,[2] BORLAN PAN,[2] CHRISTIAN WIESMANN,[2] NICHOLAS J. SKELTON,[1] AND SACHDEV S. SIDHU[2]

[1]Department of Medicinal Chemistry, Genentech, Inc., South San Francisco, California 94080, USA
[2]Department of Protein Engineering, Genentech, Inc., South San Francisco, California 94080, USA
[3]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(RECEIVED June 3, 2007; FINAL REVISION July 20, 2007; ACCEPTED July 27, 2007)

Abstract

High-temperature requirement A (HtrA) and its homologs contain a serine protease domain followed by one or two PDZ domains. Bacterial HtrA proteins and the mitochondrial protein HtrA2/Omi maintain cell function by acting as both molecular chaperones and proteases to manage misfolded proteins. The biological roles of the mammalian family members HtrA1 and HtrA3 are less clear. We report a detailed structural and functional analysis of the PDZ domains of human HtrA1 and HtrA3 using peptide libraries and affinity assays to define specificity, structural studies to view the molecular details of ligand recognition, and alanine scanning mutagenesis to investigate the energetic contributions of individual residues to ligand binding. In common with HtrA2/Omi, we show that the PDZ domains of HtrA1 and HtrA3 recognize hydrophobic polypeptides, and while C-terminal sequences are preferred, internal sequences are also recognized. However, the details of the interactions differ, as different domains rely on interactions with different residues within the ligand to achieve high affinity binding. The results suggest that mammalian HtrA PDZ domains interact with a broad range of hydrophobic binding partners. This promiscuous specificity resembles that of bacterial HtrA family members and suggests a similar function for recognizing misfolded polypeptides with exposed hydrophobic sequences. Our results support a common activation mechanism for the HtrA family, whereby hydrophobic peptides bind to the PDZ domain and induce conformational changes that activate the protease. Such a mechanism is well suited to proteases evolved for the recognition and degradation of misfolded proteins.

Keywords: structure/function studies; chaperonins; NMR spectroscopy; docking proteins; PDZ domain; ligand specificity

Supplemental material: see www.proteinscience.org

Reprint requests to: Sachdev S. Sidhu, Department of Protein Engineering, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA; e-mail: sidhu@gene.com; fax: (650) 225-3734.

Abbreviations: BSA, bovine serum albumin; ELISA, enzyme-linked immunosorbent assay; Erbin-PDZ, the PDZ domain of human Erbin; GST, glutathione S-transferase; HtrA, high-temperature requirement A; HtrA1-PDZ, the PDZ domain of HtrA1; HtrA2-PDZ, the PDZ domain of HtrA2; HtrA3-PDZ, the PDZ domain of HtrA3; IGF, insulin-like growth factor; IGFBP, IGF binding protein; IPTG, isopropyl-β-D-thiogalactoside; NiNTA, Nickel-nitrilotriacetic acid; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; PBS, phosphate-buffered saline; PDZ, PSD-95/Discs-large/ZO-1; PEG, polyethylene glycol; TGF-β, transforming growth factor β; ZO, zonula occludens; ZO1-PDZ1, the first PDZ domain of human ZO-1.

Article and publication are at http://www.proteinscience.org/cgi/doi/10.1110/ps.073049407.

In multicellular organisms, PDZ domains are common modular protein domains that mediate a wide range of specific protein–protein interactions by binding in a sequence-specific manner to the C termini of their biological partners or, in some instances, to internal sequence motifs (Harris and Lim 2001; Sheng and Sala 2001). Specificity of PDZ domains for their ligands was originally classified in two groups based on the presence of a Ser/Thr residue (type I) or a hydrophobic residue (type II) at position^{-2} (in the standardized PDZ ligand nomenclature, the ligand C terminus is designated position^{0} and the preceding positions are numbered −1, −2, etc., and the corresponding binding sites on the PDZ domain are numbered site^{0}, site^{-1}, etc.) (Songyang et al. 1997). More recent studies have suggested that the determinants of selectivity for most PDZ domains are considerably more complex, with binding depending on all four C-terminal ligand side chains and being influenced by residues further upstream (Fuh et al. 2000; Kozlov et al. 2000, 2002; Karthikeyan et al. 2001, 2002; Im et al. 2003a,b; Kang et al. 2003; Skelton et al. 2003; Appleton et al. 2006; Zhang et al. 2006). For example, the binding profile of the Erbin-PDZ domain (Erbin-PDZ) is extremely specific for the last four ligand residues ([D/E][T/S]WV_{COOH}) and that of the first PDZ domain of zonula occludens-1 (ZO1-PDZ1) is similar ([R/K/S/T][T/S][W/Y][V/I/L]$_{COOH}$) but exhibits increased promiscuity for three of the last four residues. Both Erbin-PDZ and ZO1-PDZ1 also employ auxiliary ligand interactions upstream of position^{-3} that modulate binding affinity (Appleton et al. 2006). Moreover, it has become apparent that the highly specific nature of these protein–protein interactions is important for the biological function of scaffolding proteins that contain PDZ domains (Zhang et al. 2006).

The high-temperature requirement A (HtrA) family represents an interesting group of PDZ-containing proteins that are produced by both bacterial and mammalian cells. HtrA proteins are characterized by the presence of a serine protease domain followed by one or two PDZ domains. The prototypical prokaryote *Escherichia coli* contains three periplasmic HtrA family members: the founding member HtrA (also know as DegP), DegS, and DegQ. DegP and DegQ act as chaperones at normal temperatures and are essential for survival at high temperatures, where the proteolytic function mediates degradation of denatured proteins (Kolmar et al. 1996; Krojer et al. 2002). The membrane-anchored DegS is responsible for initiating a proteolytic cascade that activates expression of periplasmic proteases/chaperones in response to unfolded proteins (Wilken et al. 2004).

Four human genes encoding for proteins with significant homology with bacterial HtrA family members have been identified. The protein products of three of these genes have been studied, and, of these, HtrA2/Omi is the best characterized. Under normal conditions, HtrA2/Omi resides in the mitochondrial intermembrane space and is believed to be responsible for determining the fate of misfolded proteins in a role analogous to that of bacterial HtrA proteins in the periplasm (Jones et al. 2003; Ekert and Vaux 2005). Under conditions of cellular stress, HtrA2/Omi is released into the cytoplasm, along with other mitochondrial proteins, where it participates in the apoptotic cellular death process (Ekert and Vaux 2005). The other two human HtrA family members (HtrA1 and HtrA3) are less well characterized, and they are also structurally more complex as, in addition to the defining protease and PDZ domains, both also contain a secretion signal and a region homologous to human mac25, which is characterized by the presence of a domain with homology with insulin-like growth factor binding protein (IGFBP) and a Kazal-type serine protease inihibitor domain (Zumbrunn and Trueb 1996; Hu et al. 1998; Nie et al. 2003b).

HtrA1 was originally identified as a gene that is down-regulated in a human fibroblast cell line after transfection with the oncogenic simian virus 40 (SV40) (Zumbrunn and Trueb 1996), and has since been implicated in a number of malignancies. The HtrA1 gene is down-regulated in malignant melanoma and ovarian cancer (Baldi et al. 2002; Chien et al. 2004), and overexpression results in the inhibition of tumor growth and proliferation (Baldi et al. 2002) and enhanced chemotherapy-induced cytotoxicity (Chien et al. 2006). In contrast, the gene is up-regulated in cartilage of osteoarthritic joints and may contribute to the development of arthritic diseases (Hu et al. 1998). HtrA1 has also been implicated in amyloid precursor protein processing (Grau et al. 2005), and thus may play a role in Alzheimer's disease. Recently, a single nucleotide polymorphism that leads to increased expression has been identified in the promoter region of the gene in patients with the wet form of age-related macular degeneration, suggesting a key role for HtrA1 in the pathogenesis of this disease (DeWan et al. 2006; Yang et al. 2006). HtrA1 is also up-regulated in the placenta during pregnancy, suggesting a role in normal placental development and function (De Luca et al. 2004). It is noteworthy that all of these diverse biological processes are, to some extent, dependent upon proteolytic degradation and modification of the extracellular matrix. HtrA3 is also up-regulated in the placenta and was originally identified as a ''pregnancy-related'' serine protease (Nie et al. 2003b). However, detailed expression profiling revealed that, aside from coincident up-regulation in the placenta, HtrA1 and HtrA3 show very different expression patterns, suggesting that, despite high structural similarity, the two proteases mediate different tissue-specfic functions (Nie et al. 2003a).

Yeast-two-hybrid studies have revealed that the PDZ domain of HtrA1 binds to the C termini of procollagen C-propeptides, and denatured procollagen is a substrate for the protease in vitro (Murwantoko et al. 2004). HtrA1 has also been shown to degrade fragments of amyloid precursor protein in vitro, and inhibition of the protease leads to the accumulation of β-amyloid in cell culture supernatants (Grau et al. 2005). Within the synovial fluids of patients with arthritic disease, elevated HtrA1 levels have been shown to cause degradation of fibronectin and subsequent induction of matrix metalloproteases, suggesting that HtrA1 contributes to the destruction of extracellular matrix through direct and indirect mechanisms (Grau et al. 2006). Both HtrA1 and HtrA3 bind to and inhibit members of the TGF-β family (Oka et al. 2004; Tocharus et al. 2004; Grau et al. 2005), and HtrA1 has been shown to cleave IGFBP-5 (Hou et al. 2005). These findings suggest that these proteases regulate TGF-β and IGF signaling, and both of these pathways are intimately associated with production and maintenance of the extracellular matrix. Taken together, these examples suggest HtrA1 and HtrA3 play a role in protein quality control in the extracellular matrix, in a manner analogous to the function of HtrA2/Omi in mitochondria and, in addition, may modulate specific signaling pathways. Furthermore, destructive joint pathology in arthritic disease and invasive choroidal neovascularization in wet age-related macular degeneration may partly be caused by aberrant overproduction of HtrA1 (Grau et al. 2006; Yang et al. 2006).

Sequence comparison of the human HtrA family (Fig. 1) reveals a high degree of sequence conservation, with greater than 30% identity among the PDZ domains. While this level of sequence homology is expected to translate into structural conservation, key sequence differences within the ligand-binding site may impart differences in specificity, which, in turn, may lead to divergent function.

In this report, we have used phage display to explore the ligand specificities of the PDZ domains of HtrA1 (HtrA1-PDZ) and HtrA3 (HtrA3-PDZ). Together with previous studies of the PDZ domain of HtrA2/Omi (HtrA2-PDZ) (Zhang et al. 2007), our results reveal distinct differences in the ligand preferences of the different domains. To better understand the details of ligand specificity, we used affinity measurements of synthetic peptide analogs of the optimal ligands of HtrA1-PDZ and HtrA3-PDZ to quantitate the energetic contributions of individual ligand side chains. An efficient combinatorial alanine scanning approach has also been utilized to

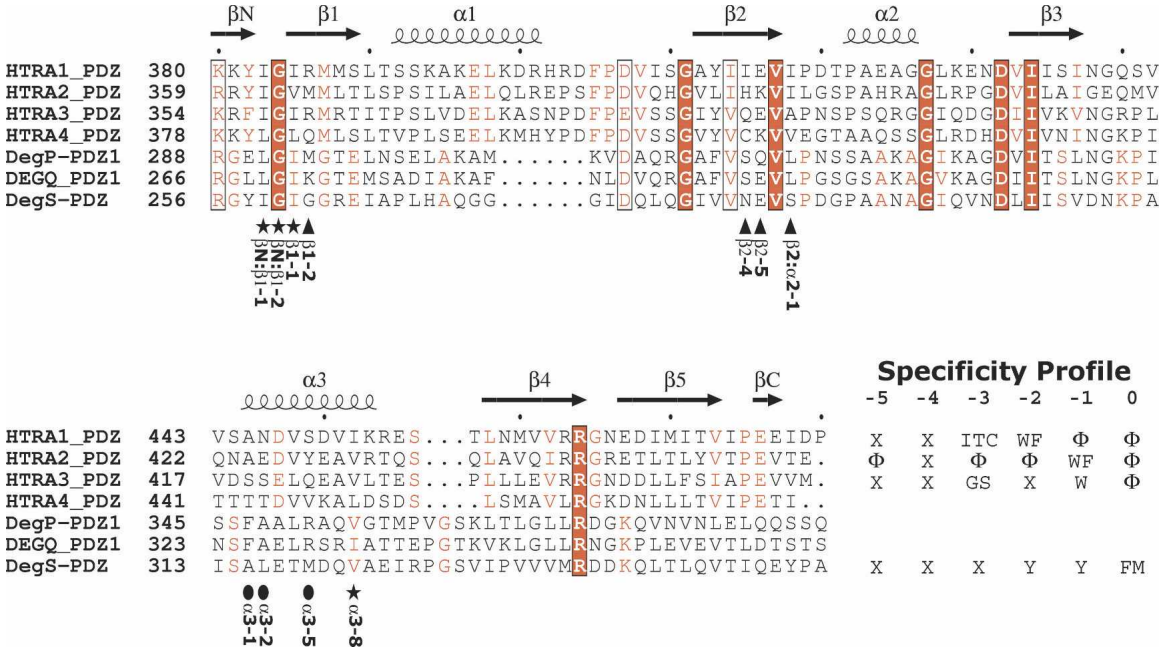

Figure 1. Sequence alignment of the PDZ domains of human and *E. coli* HtrA family members. Elements of regular secondary structure in the human domains are labeled and depicted *above* the alignment. Boxes indicate highly conserved residues between both human and bacterial HtrA family members, and red boxes indicate 100% identity. Residues that are highly conserved within either human or bacterial family members are shown in red or green font, respectively. Residues that are directly involved in conferring ligand specificity for the human domains are labeled and indicated *below* the sequences as stars (site^{0}), triangles (site^{-1} and site^{-3}), or ovals (site^{-2}). The optimal ligand-binding profiles for human HtrA1, -2, and -3 (Zhang et al. 2007; this work) and DegS (Walsh et al. 2003) are tabulated at the *bottom right*, using the single letter amino acid code with "X" indicating no preference and "Φ" indicating hydrophobic amino acids. Alignments were made using ClustalW, and the figure was created using ESPript (Gouet et al. 2003).

identify residues within HtrA1-PDZ that contribute energetically to ligand binding (Weiss et al. 2000). We have combined these mutagenic analyses with structural studies of HtrA1-PDZ by NMR spectroscopy and HtrA3-PDZ by X-ray crystallography. The high-resolution structures of the PDZ domains in complex with optimal peptide ligands provide a framework within which to understand the affinity and selectivity of peptide binding. These results are also analyzed in the light of studies of the ligand specificities of HtrA2-PDZ, Erbin-PDZ, ZO1-PDZ1, and the PDZ domains of DegS and DegP. Taken together, these findings provide a detailed understanding of the molecular interactions that are responsible for the differing specificities and biological functions of the various members of the PDZ domain family.

Results

Binding specificity profiles for HtrA1-PDZ and HtrA3-PDZ

We used phage-displayed peptide libraries to define binding specificity profiles for HtrA1-PDZ and HtrA3-PDZ. The domains were screened separately against libraries of completely random decapeptides or dodecapeptides displayed in a high valency format by fusion to the C or N terminus of the M13 major coat protein, respectively. The absence of a free peptide C terminus in the N-terminal library permits the identification of internal peptide sequences capable of binding to the PDZ domains. We were successful in obtaining specific binding clones for each PDZ domain, and DNA sequencing revealed many unique sequences that were aligned and analyzed for homology (Figs. 2, 3).

For ligands with free C termini, both HtrA1-PDZ and HtrA3-PDZ prefer an overall hydrophobic character (Fig. 2), and this is similar to the previously defined specificity of HtrA2-PDZ (Zhang et al. 2007). HtrA1-PDZ prefers hydrophobic residues at each of the last four ligand positions, while HtrA3-PDZ exhibits a strong preference for hydrophobic residues at only the last two positions. At position^{0} both domains resemble HtrA2-PDZ and other PDZ domains, in that they prefer aliphatic side chains. At position^{-1} the selection yielded Trp exclusively for HtrA3-PDZ, while HtrA1-PDZ appears to be more promiscuous, since many different hydrophobic residues, and even some hydrophilic residues, were selected. In contrast, HtrA1-PDZ is highly selective for Trp/Phe at position^{-2}, where HtrA3-PDZ shows no clear preference. At position^{-3}, HtrA1-PDZ prefers Ile but also tolerates Cys/Thr, while HtrA3-PDZ shows a preference for small Gly/Ser residues. Neither domain exhibits any strong preferences upstream of position^{-3}.

Binding selections for internal ligands were also successful for both HtrA1-PDZ and HtrA3-PDZ, and sequencing revealed 16 and 8 unique sequences, respectively (Fig. 3). In this case, alignment of the sequences was more difficult than for the C-terminal ligands, which was facilitated by the C-terminal residue functioning as an anchor position. Thus, we aligned the sequences on the basis of homology, but did not assign position numbering, since it is unclear how these ligands bind to the PDZ domains. For HtrA1-PDZ (Fig. 3A), 8 of the 16 sequences exhibited significant homology with a hydrophobic motif that stretched across almost the entire length of the peptide sequence ([G/S/A][V/L][T/S]WG[E/D]ϕ[L/V]Xϕ[L/V/I], where "ϕ" is a hydrophobe). We hypothesize that the acidic residue could serve as a surrogate C terminus, placing a Trp at position^{-2} in this subset of the sequences. The remaining sequences did not exhibit significant homology with this motif but did exhibit overall hydrophobic character. In the case of HtrA3-PDZ, all of the internal ligands exhibited striking homology with a motif that stretched across the entire length of the sequence (G[V/L][V/L]VDEWϕL[S/N]LL). This motif is predominantly hydrophobic but also contains conserved acidic residues that could mimic a free C terminus, but it is not clear how the ligand would be oriented in the binding cleft. These results show that both HtrA1-PDZ and HtrA3-PDZ can recognize hydrophobic peptides lacking a free C terminus, but in the absence of structural information, we cannot speculate on the details of molecular recognition. In the case of HtrA1-PDZ, the peptides may not adopt a single-binding mode, and in both cases, it is possible that the peptides adopt a secondary structure, since they exhibit homology across the entire length of the dodecapeptide sequence.

Affinity assays with synthetic peptides

The phage-selected ligands were used to guide the design of synthetic peptides for solution-phase competition-binding assays to compare and contrast the binding specificities of HtrA1-PDZ and HtrA3-PDZ (Table 1). For HtrA1-PDZ, four unique peptides were tested (H1-C1, -2, -3, and -4), and the optimal peptide (H1-C1) exhibited an IC_{50} of 0.9 μM, which is in good agreement with the dissociation constant determined from isothermal titration calorimetry (K_d = 1.1 μM; Supplemental Fig. S1). Blocking the C terminus by amidation (peptide H1-C1a) abolished binding, indicating that the terminal carboxylate group is required for binding of this peptide. Alanine scanning of peptide H1-C1 revealed that substitution of Val^{0} or Trp^{-1} results in modest fourfold or sevenfold reductions in binding, respectively (peptide H1-C1b and -c), whereas alanine substitution for Trp^{-2} or Ile^{-3} results in larger 44-fold or 14-fold decreases, respectively (peptides H1-C1d and -e). Alanine substitutions upstream of position^{-3} did not affect affinity

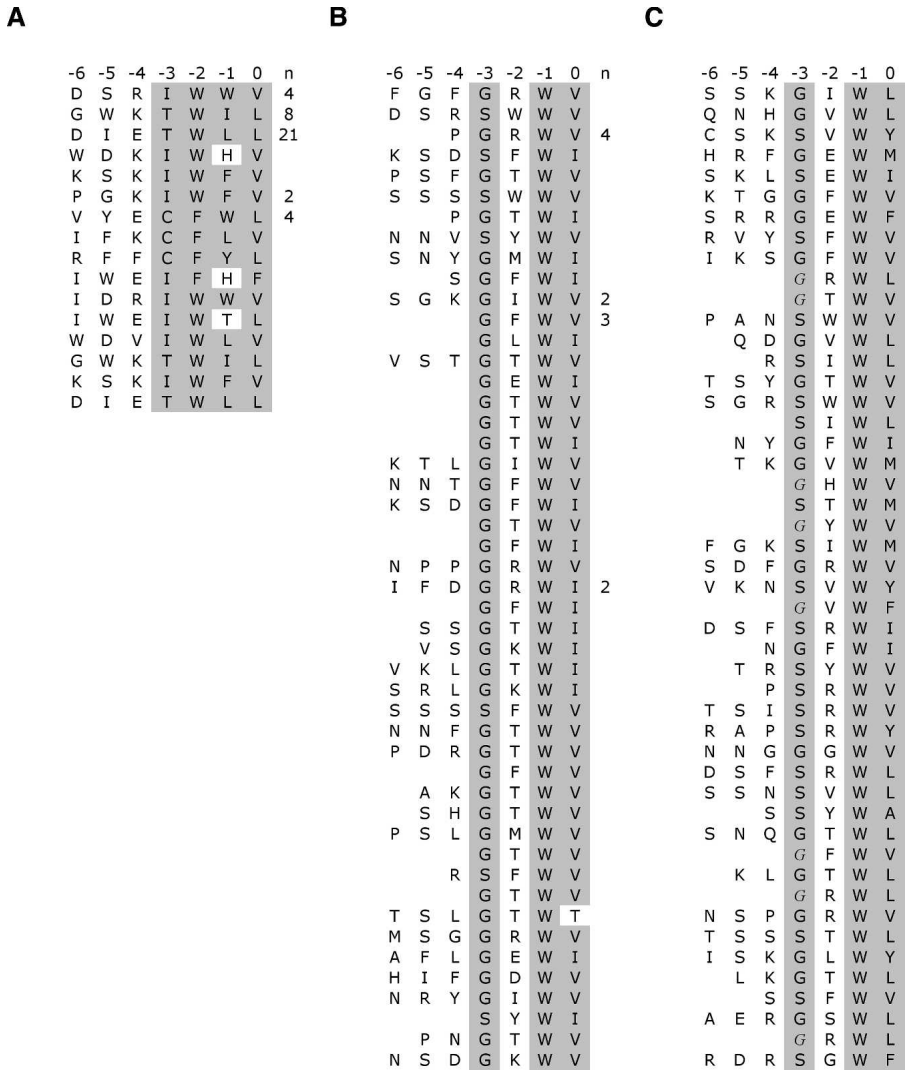

Figure 2. C-terminal peptide ligands for human HtrA PDZ domains. Sequences are shown for peptides selected from a C-terminal phage display library screened against HtrA1-PDZ (*A*), HtrA3-PDZ (*B*), and HtrA1(I415Q/I418A)-PDZ (*C*). The number of occurrences of each nonunique clone (*n*) is shown to the *right* of each sequence. Gray shading indicates sequences that match the optimal binding profile for HtrA1-PDZ (*A*) or HtrA3-PDZ (*B*,*C*), as defined in Figure 1.

significantly (peptides H1-C1f, -g, and -h). Taken together, these results confirm the optimal binding profile defined by phage display, which shows that HtrA1-PDZ prefers hydrophobic residues at each of the last four ligand residues.

Previous reports have suggested that mouse HtrA1 binds to the C termini of fibrillar procollagen proteins such as Type III collagen α3 C-propeptide (Col3a1) as well as to a Golgi matrix protein (GM130) (Murwantoko et al. 2004). We measured affinities for synthetic peptides corresponding to the C termini of Col3a1 and GM130 and compared these to our phage-derived peptides (Table 1). The Col3a1 and GM130 peptides bind with reasonable affinities in the double-digit micromolar range, which are an order of magnitude weaker than the affinity of the optimal phage-derived peptide (H1-C1). While our results confirm that HtrA1-PDZ can indeed bind to these peptides, the affinities measured in our assay are considerably weaker than those reported by Murwantoko et al. (2004) (0.3 μM and 6.0 nM for Col3a1 and GM130, respectively). It is notable that the previous affinities were determined using assays with trimeric HtrA1 immobilized on solid surfaces, which can lead to significant overestimation of affinities due to avidity effects induced by polyvalent binding (Harris and Lim 2001; Harris et al. 2001; Laura et al. 2002). In contrast, our solution-phase competition assays give a more accurate estimate of monomeric binding affinities.

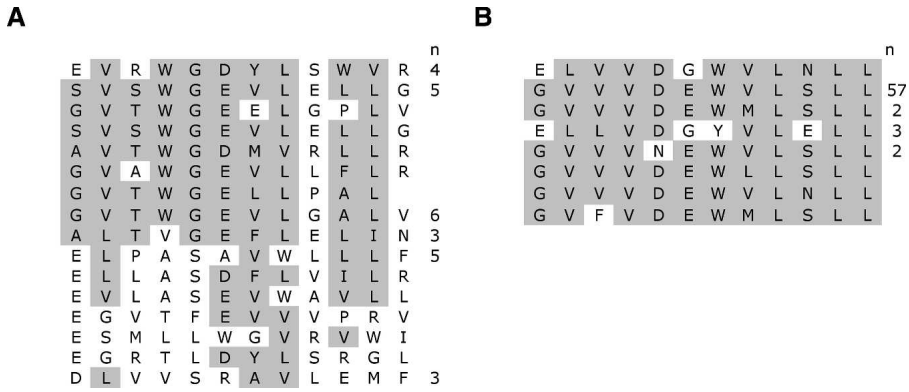

Figure 3. Internal peptide ligands for human HtrA PDZ domains. Sequences are shown for peptides selected from an N-terminal phage display library screened against HtrA1-PDZ (*A*) and HtrA3-PDZ (*B*). Gray shading indicates sequences that match the following consensus: HtrA1-PDZ ([G/S/A][V/L][T/S]WG[E/D]ϕ[L/V]Xϕ[L/V/I]), HtrA3-PDZ (G[V/L][V/L]VDEWϕL[S/N]LL).

For HtrA3-PDZ, two different peptides were tested (peptide H3-C1 and H3-C2) and both bound tightly with affinities in the submicromolar range (Table 2). While Val is preferred at position^0, substitution by other aliphatic residues had only modest effects on binding (compare peptide H3-C1 to H3-C1a and -b), and even a bulky Phe^0 substitution only reduced affinity by ~13-fold (peptide H3-C1c). In contrast, blocking of the carboxylate group by amidation (peptide H3-C1d) drastically reduced affinity by ~120-fold. Alanine scanning showed that only substitution of the Trp^{-1} residue reduced binding significantly (H3-C1f), and, surprisingly, truncation of residues from the N terminus of peptide H3-C1 revealed that the C-terminal dipeptide alone (peptide H3-C1k) is sufficient for high affinity binding to HtrA3-PDZ. Furthermore, even the C-terminal Val^0 of the dipeptide can be replaced by either Ala or Gly (peptides H3-C1l and -m) without drastically reducing binding. Taken together, these results demonstrate that a C-terminal carboxylate and a Trp^{-1} residue are necessary and sufficient to mediate binding to HtrA3-PDZ, and other side chains make only minor contributions.

To investigate the recognition of internal peptide motifs, we also measured affinities for synthetic peptides corresponding to sequences selected from library N (Fig. 3). For HtrA1-PDZ, binding was detected for only one of the two peptides tested (H1-N1), and the affinity was dramatically weaker than the affinities of the C-terminal

Table 1. *IC_{50} values for synthetic peptides binding to HtrA1-PDZ*

Peptide	Sequence												IC_{50}[a] (μM)
					−7	−6	−5	−4	−3	−2	−1	0	
H1-C2						G	W	K	T	W	I	L	7.7 ± 0.6
H1-C3						D	I	E	T	W	L	L	23 ± 3
H1-C4						W	D	K	I	W	H	V	2.8 ± 0.3
H1-C1						D	S	R	I	W	W	V	0.9 ± 0.1
H1-C1a						D	S	R	I	W	W	**V**[b]	>500
H1-C1b						D	S	R	I	W	W	**A**	3.5 ± 0.9
H1-C1c						D	S	R	I	W	**A**	V	6 ± 1
H1-C1d						D	S	R	I	**A**	W	V	40 ± 5
H1-C1e						D	S	R	**A**	W	W	V	13 ± 1
H1-C1f						D	S	**A**	I	W	W	V	2.5 ± 0.4
H1-C1g						D	**A**	R	I	W	W	V	1.3 ± 0.1
H1-C1h						**A**	S	R	I	W	W	V	2.8 ± 0.3
Col3a1					D	I	G	P	V	C	F	L	16 ± 3
GM130						E	V	K	I	M	V	V	24 ± 8
H1-N1	E	V	R	W	G	D	Y	L	S	W	V	R[b]	240 ± 170
H1-N2	S	V	S	W	G	E	V	L	E	L	L	G[b]	>500

[a]The IC_{50} value is the mean concentration of peptide that blocked 50% of PDZ domain binding to an immobilized high affinity peptide ligand. The N termini of the peptides were acetylated. Deviations from the sequence of peptide H1-C1 are in bold text.
[b]The C terminus was amidated.

Table 2. *IC_{50} values for synthetic peptides binding to HtrA3-PDZ*

	Sequence												
Peptide					−7	−6	−5	−4	−3	−2	−1	0	IC_{50}[a] (μM)
H3-C2								R	S	W	W	V	0.6 ± 0.1
H3-C1								F	G	R	W	V	0.6 ± 0.1
H3-C1a								F	G	R	W	**I**	1.0 ± 0.1
H3-C1b								F	G	R	W	**L**	2.9 ± 0.3
H3-C1c								F	G	R	W	**F**	7.7 ± 0.8
H3-C1d								F	G	R	W	**V**[b]	70 ± 10
H3-C1e								F	G	R	W	**A**	3.5 ± 0.3
H3-C1f								F	G	R	**A**	V	270 ± 110
H3-C1g								F	G	**A**	W	V	0.9 ± 0.1
H3-C1h								F	**A**	R	W	V	1.1 ± 0.2
H3-C1i									G	R	W	V	1.0 ± 0.1
H3-C1j										R	W	V	1.3 ± 0.1
H3-C1k											W	V	4.7 ± 0.4
H3-C1l											W	A	14 ± 1
H3-C1m											W	G	22 ± 3
H3-N1	E	L	V	V	D	G	W	V	L	N	L	L[b]	19 ± 6
H3-N2	G	V	V	V	D	E	W	V	L	S	L	L[b]	200 ± 50

[a]The IC50 value is the mean concentration of peptide that blocked 50% of PDZ domain binding to an immobilized high affinity peptide ligand. The N termini of the peptides were acetylated. Deviations from the sequence of peptide H3-C1 are in bold text.
[b]The C terminus was amidated.

peptides. For HtrA3-PDZ, binding was detectable for both peptides tested, and binding of one peptide (H3-N1) was only 30-fold weaker than that of the optimal peptides with free C termini. Thus, HtrA1-PDZ and HtrA3-PDZ are able to bind to internal peptide motifs, but C-terminal ligands are preferred.

Structural analysis of HtrA1-PDZ and HtrA3-PDZ

By making the appropriate PDZ-ligand fusion, as has been described previously (Appleton et al. 2006), a high-precision structure of HtrA3-PDZ in complex with peptide H3-C1 was obtained by X-ray crystallography (Fig. 4B, Table 3). Similar crystallization efforts with HtrA1-PDZ were unsuccessful, but NMR data for the domain interacting with peptide H1-C1 were sufficient to clearly define the structure of the complex (Fig. 4A, Table 4). In both structures, the PDZ fold (Fig. 1) consists of a five-stranded β-sandwich (β1–β5) capped by two α-helices (α1,α3), as has been seen in other PDZ domain structures (Sheng and Sala 2001). In addition, there are short β-strands at the N and C termini (βN and βC). Like the PDZ domains of HtrA2/Omi (Zhang et al. 2007) and bacterial DegP and DegS (Krojer et al. 2002; Wilken et al. 2004), HtrA1-PDZ and HtrA3-PDZ have a cyclically permuted fold compared to the canonical PDZ fold as exemplified by the PDZ domain of Erbin (Skelton et al. 2003), in which the first β-strand of the canonical fold corresponds to β5 of HtrA1-PDZ (Fig. 4C,D). In the structures of all three HtrA family members, the β1–β2 loop of the PDZ domain forms a well-defined α-helix, but the orientation of the helix relative to the rest of the domain varies, suggesting that this region may be conformationally dynamic. Indeed, for the HtrA1-PDZ complex, heteronuclear NOE measurements are consistent with subnanosecond scale dynamics of this region (Supplemental Fig. S2). To facilitate comparison of residues within each PDZ domain, residues will be referred to by their location within each of the secondary structure elements or loops (Aasland et al. 2002). Thus, for example, β1-1 is the first residue in strand β1 and β1:β2-1 is the first residue in the loop between strands β1 and β2.

In the structures of all three human HtrA family members, the peptide binds in an extended conformation in the cleft between strand β1 and helix α3, adding an additional strand to the antiparallel β-sheet formed by strands β1 and β2 (Fig. 5). The C-terminal carboxylate group is coordinated by the three main chain amides immediately preceding strand β1. In the case of HtrA1-PDZ, these amides are protected from solvent exchange in the complex, as determined by hydrogen/deuterium exchange NMR experiments. The amide hydrogens of residues βN:β1-1 (Ile383) and β1-1 (Ile385) both point directly at the carboxyl group, but at slightly longer distances than those usually considered to define a hydrogen bond (3.92 ± 0.4 Å and 3.52 ± 0.6 Å heavy atom distances). The amide hydrogen of βN:β1-2 (Gly384) does not point directly toward the carboxylate group and may be involved in a water-mediated hydrogen bond.

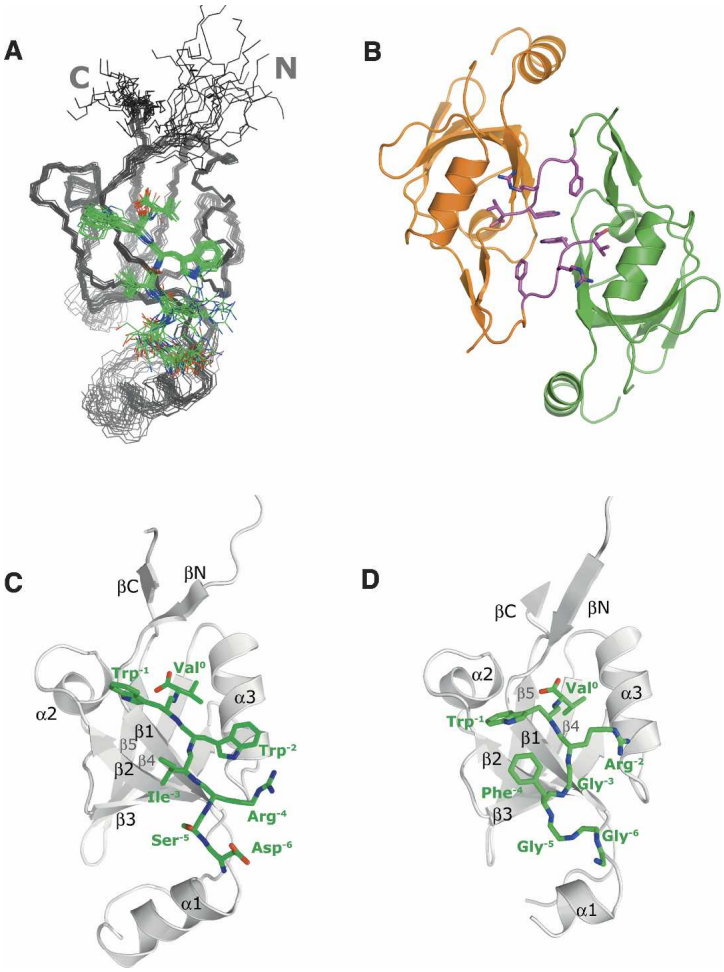

Figure 4. Structures of HtrA1-PDZ and HtrA3-PDZ bound to phage-derived peptide ligands. (*A*) Ensemble of structures for HtrA1-PDZ (gray) in complex with peptide H1-C1 colored according to atom type (carbon, green; oxygen, red; nitrogen, blue). Only the backbone N, Cα, and C atoms are shown as lines. Selected peptide side chain heavy atoms are included. Root mean square deviation to the mean structure = 0.522 ± 20.1 Å for the backbone heavy atoms of residues 378–389, 411–463, and 468–475. No distance or dihedral angle restraints were violated by more than 0.1 Å or 1°, respectively (Table 3). (*B*) Dimer of the HtrA3-PDZext crystal structure. The PDZ domains are colored green or orange. The pentapeptide ligand main chains are colored magenta; the side chains and C-terminal carboxylate are shown as sticks and colored according to atom type (carbon, magenta; oxygen, red; nitrogen, blue). The peptide is linked to the C terminus of the PDZ domain through a triglycine linker. The HtrA3-PDZext structure contains two molecules per asymmetric unit, and both copies are well-defined in the electron density with the exception of a disordered region between strands β1 and β2. Ribbon views of (*C*) HtrA1-PDZ bound to peptide H1-C1 and (*D*) HtrA3-PDZext. Peptide ligands are shown in stick representation and colored according to atom type (carbon, green; oxygen, red; nitrogen, blue). Elements of regular secondary structure and peptide residues are labeled. Structural figures were produced with PyMOL (DeLano Scientific).

While the hydrogen/deuterium exchange experiments and down-field chemical shifts observed for the three amide hydrogens suggest that hydrogen bonds are formed, the inherent lack of NOE restraints prevents precise definition of the ligand carboxylate and the carboxylate binding loop of the PDZ domain or the use of specific hydrogen bond restraints in the structure calculation. In the case of HtrA3-PDZ, all three amide groups point directly at the carboxylate oxygen atoms at distances typical of hydrogen bonds found in proteins (2.9, 2.8, and 3.2 Å heavy atom distances for βN:β1-1, βN:β1-2, and β1-1, respectively). Another similarity between the two structures is the recognition of the Val^0 side chain, which is oriented within the shallow hydrophobic pocket of the PDZ ligand-binding groove. In each structure, the hydrophobic binding pocket presents aliphatic side chains that provide a complementary surface for the valine side chain to sit in; however, there is clearly enough room in this pocket to accommodate a variety of aliphatic peptide side chains at position^0. Although two rotameric conformations of the Val^0 side chain are observed in the NMR ensemble of the HtrA1-PDZ complex, the aliphatic side chains forming

Table 3. *Summary of crystallographic data collection and refinement statistics for HtrA3-PDZext*

Data collection	
Resolution (Å)[a]	50–1.7
Space group	$P4_12_12$
Cell parameters (Å)	$a,b = 73.0$, $c = 80.1$
Unique reflections	24,442
Redundancy[a]	7.6 (5.6)
Completeness (%)[a]	99.9 (99.6)
R_{sym} (%)[a,b]	0.049 (0.461)
$\langle I\rangle/\langle\sigma(I)\rangle$[a]	39.2 (3.3)
Refinement	
Resolution	30–1.7
R_{cryst},R_{free}[c]	0.186, 0.221
Number of reflections	23,159
Number of non-hydrogen atoms	1849
Number of residues	416
Number of waters	202
RMSD bonds (Å)	0.012
RMSD angles (°)	1.4
Ramachandran plot (%)[d]	95.3, 4.7, 0, 0

[a] Values in parentheses refer to data in the highest resolution shell.
[b] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where $\langle I\rangle$ is the average intensity of symmetry-related observations of a unique reflection.
[c] $R_{cryst} = R_{free} = \Sigma|F_o - F_c|/\Sigma F_o$, where R_{free} represents 5% of the data selected randomly.
[d] Values represent the percentage of residues in the most favored, additionally allowed, generously allowed, and disallowed regions of the Ramachandran plot, respectively (Laskowski et al. 1993).

the PDZ ligand-binding pocket adopt a single conformation in all members of the ensemble.

The remaining protein–ligand interactions are notably different between HtrA1-PDZ and HtrA3-PDZ. In particular, the most striking difference between the two complex structures involves recognition of Trp^{-1}. In the case of HtrA3-PDZ (Fig. 5C), Trp^{-1} adopts a conformation nearly identical to that seen with other PDZ domains that prefer Trp at position^{-1} such as HtrA2-PDZ (Fig. 5B; Zhang et al. 2007) and Erbin-PDZ (Skelton et al. 2003). The indole ring of this residue extends across the backbone of strand β1, and inserts between the side chains of β2-5 (Glu390) and β2:α2-1 (Ala392). The orientation of the Trp^{-1} side chain with respect to strand β2 is reminiscent of the interstrand Trp contacts observed in studies of peptide β-hairpin stability (Cochran et al. 2001). Thus, the preference for Trp at position^{-1} suggests a general and somewhat nonselective contribution to binding mediated by aromatic interaction with the protein backbone of strand β1. In the case of HtrA1-PDZ (Fig. 5A), Trp^{-1} also extends across the backbone of strand β1, but the orientation of the indole ring is significantly different from that of HtrA3-PDZ. Trp^{-1} is positioned between the side chains of βN-3 (Tyr382) and β1-2 (Arg386) and abuts the Ile side chain at position β2:α2-1 (Ile418). The orientation of the indole ring does not adopt a unique conformation across the ensemble (Fig. 4A). The lack of NOE restraints (three intermolecular NOEs to aromatic protons) is consistent with the indole ring not adopting a unique conformation but, rather, being conformationally dynamic. This is in contrast to the complex of Erbin-PDZ bound to a peptide with Trp at position^{-1}, in which the indole ring occupies a single conformation directly above

Table 4. *Summary of input restraints and structural statistics for NMR ensemble of HtrA1-PDZ bound to peptide H1-C1*

Parameter	Ensemble
Input restraints:	
NOE total	1352
Intraresidue	179
Sequential	340
Medium range	252
Long range	504
Intermolecular	77
Dihedral angles total	178
ϕ	77
ψ	76
χ_1	21
Violations:	
RMSD from experimental restraints:	
NOE distance (Å)	0.0041 ± 0.0003
Dihedral (°)	0.056 ± 0.016
NOE distance violations	
Number > 0.01 Å	34.1 ± 4.8
Number > 0.1 Å	0.0
Mean maximum violations (Å)	0.07 ± 0.01
Dihedral violations:	
Number > 0.1°	6.6 ± 1.9
Mean maximum violations (°)	0.5 ± 0.19
RMSD from idealized geometry	
Bonds (Å)	0.0009 ± 0.0001
Angles (°)	0.27 ± 0.01
Impropers (°)	0.12 ± 0.01
Energies:	
Energy components (kcal/mol^{-1})	
NOE	1.24 ± 0.18
CDIH	0.04 ± 0.02
Bonds	1.5 ± 0.2
Angles	35.1 ± 1.4
Impropers	2.1 ± 0.37
van der Waal's	16.3 ± 2.4
Stereochemistry:	
Ramachandran (%)	
Favored	74.39
Allowed	24.18
Generous	0.97[a]
Disallowed	0.46[a]
Structural precision:	
Mean RMSD to mean structure	
Backbone (N,Cα,C)	0.52 ± 0.12[b]
Heavy (N*,C*,O*,S*)	0.71 ± 0.35[b]

[a] Those residues falling into generous and disallowed region were in the ill-defined N terminus of HtrA1-PDZ. No more than three members of the ensemble had any one residue in these regions.
[b] RMSD was determined for the ordered regions of HtrA1-PDZ encompassing residues 378–389, 411–463, and 468–475.

the backbone of the PDZ domain and induces large ring current shifts in protons in the immediate vicinity of the aromatic ring (Skelton et al. 2003). No such indole-induced ring current shifts of residues proximal to Trp^{-1} are observed in the HtrA1-PDZ bound to peptide H1-C1.

HtrA1-PDZ does show a strong preference for Trp at position^{-2}. Trp^{-2} is particularly well-defined in the ensemble by 20 unique intermolecular NOEs assigned to its aromatic protons. The indole ring packs against helix α3 and makes favorable hydrophobic interactions with α3-1 (Ala445), β1-3 (Met387), and the methylene of α3-2 (Gln446) (Fig. 5A). The indole ring is positioned directly above the alpha proton of Gln446, inducing a large (−1 ppm) chemical shift perturbation due to the ring current effect. Although the binding profile and synthetic peptide binding assays for HtrA3-PDZ suggest that position^{-2} is unimportant for ligand binding, the crystal structure shows that Arg^{-2} in peptide H3-C1 is positioned to form a hydrogen bond to the glutamine side chain at position α3-4 (Gln423) (Fig. 5C). A nearly identical interaction has been observed in the structure of the PDZ domain of the human calcium/calmodulin-dependent serine protein kinase (CASK) (Daniels et al. 1998).

The ligand-binding profile and synthetic peptide-binding assays for HtrA1-PDZ suggest that it prefers an Ile residue at position^{-3}. The NMR structure shows that Ile^{-3} packs into a hydrophobic patch created by β2-4 (Ile415) and β1-2 (Arg386) (Fig. 5A). The packing of Ile^{-3} of the peptide and β2-4 (Ile415) of the protein against the side chain methylenes of β1-2 (Arg386) are likely responsible for the protrusion of the Arg386–Glu416 salt bridge that affects the accessibility of site^{-1}. Residues at position^{-4} through position^{-7} do not appear to be involved in the interaction with HtrA1-PDZ and are poorly defined in the NMR ensemble. In the case of the HtrA3-PDZ complex structure, Gly^{-3} of peptide H3-C1 does not contact HtrA3-PDZ directly but adopts a positive ϕ angle positioning the carbonyl oxygen of Phe^{-4} to make a hydrogen bond with the side chain at position β2-2 (Arg360) and also allowing the side chain of Phe^{-4} to interact via π–π stacking with the side chain of Arg360 (Fig. 5C). While this is clearly an energetically favorable conformation and the phage selection shows preference for Gly or Ser at position^{-3} (Fig. 2B), the synthetic peptide binding assays for HtrA3-PDZ do not suggest an energetic advantage for a Gly at position^{-3} or for an aromatic residue at position^{-4} (Table 2). There is a type I reverse turn stabilized by a hydrogen bond between the carbonyl oxygen of Gly^{-6} and the amide of Gly^{-3}; however, the synthetic peptide binding assays do not suggest that this conformation is important for ligand binding. We speculate that the discrepancies in the apparent importance of upstream ligand residues for binding to HtrA3-PDZ, as assessed by phage display or binding assays with synthetic peptides, may stem from an additional conformational requirement for ligands fused to a large protein (such as the phage coat or a natural binding partner). A turn induced by Gly^{-3} and residues further upstream may be necessary for binding to large protein ligands, but not for binding to small peptides.

Shotgun alanine scanning of HtrA1-PDZ

We used combinatorial shotgun alanine scanning to assess the contributions of individual residues of HtrA1-PDZ to ligand binding. Three phage-displayed libraries were constructed in which 61 positions in and around the peptide binding site were represented by degenerate codons that encode either the wild-type amino acid or Ala, although two additional mutations were encoded at some positions due to the degeneracy of the genetic code (Vajdos et al. 2002). These libraries were then selected for binding to peptide H1-C1 or H1-C2, and binding clones were sequenced after two rounds of selection. The number of clones with the wild-type residue at each position was compared to the number with Ala to give an indication of the preference for the wild-type residue over Ala. To control for variations in expression or display levels for different library members, the libraries were also selected for binding to an antibody capable of recognizing an epitope tag fused to the N terminus of all library members. The ratio of wild-type to Ala in the peptide selection was then scaled by the ratio in the antibody selection to give a normalized frequency of occurrence (F, Table 5). This normalized frequency of occurrence was used to organize mutations into three categories: those that reduce binding ($F \geq 5$), do not affect binding significantly ($0.5 < F < 5$), or increase binding to peptide ($F \leq 0.5$).

The effects of Ala substitutions on binding to peptide H1-C1 or H1-C2 were mapped onto the structure of HtrA1-PDZ (Fig. 6). For peptide H1-C1, some of the substitutions that significantly reduce binding are located at positions that are in direct contact with the peptide ligand (Tyr382, Ile383, Gly384, Met387, Ser389), while others are at positions that are greater than 4.5 Å from the ligand (Lys380, Lys381, Gly411, Tyr413, Ile414, Val417, Thr421, ProP422). In the case of residues that do not directly contact the peptide ligand, it is likely that replacing these side chains with Ala introduces a local perturbation of the PDZ domain structure. For example, Ile414 (β2-3) packs into the hydrophobic interior of the domain, and replacing the bulky aliphatic side chain with a single methyl group would likely perturb the packing arrangement of the hydrophobic interior and affect the β2–β3 antiparallel β-sheet, thereby perturbing the ligand-binding site. At some positions, Ala is preferred over the wild type (Asp400, Ile418, Val442, Asn446,

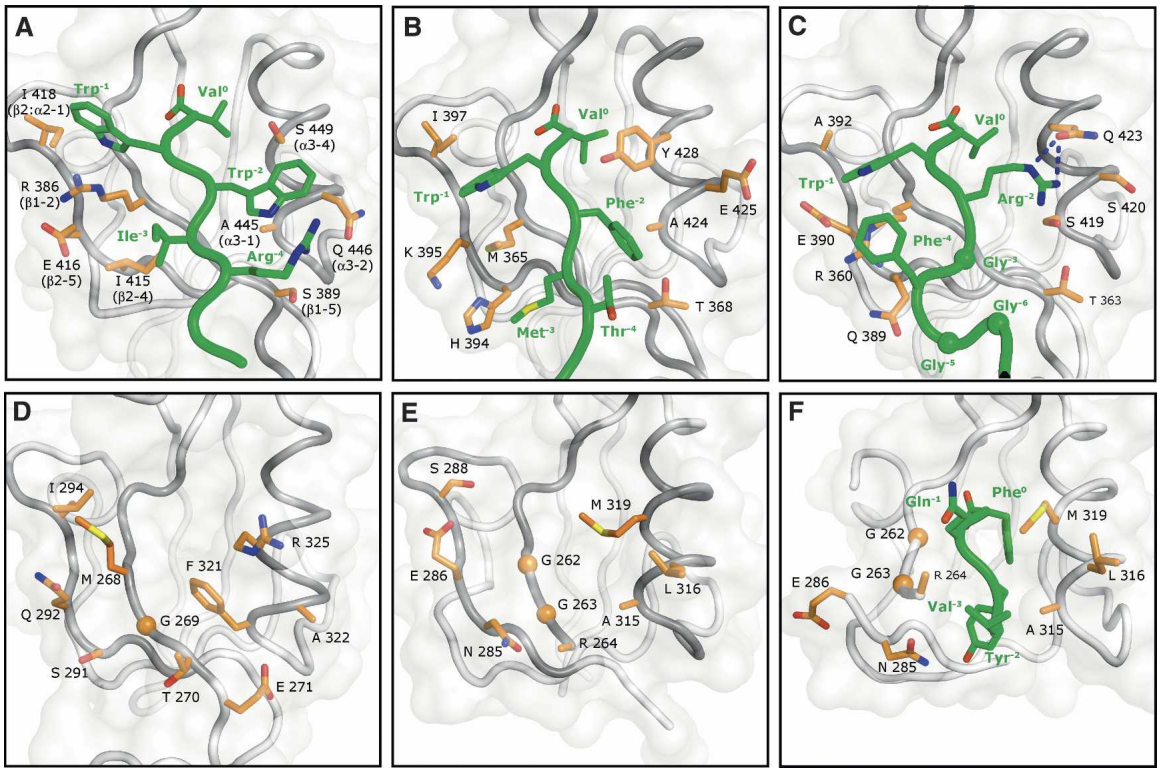

Figure 5. Binding sites of the PDZ domains of human and *E. coli* HtrA family members. In each panel, the structures are shown in the same relative orientation. Main chain traces are rendered as gray or green tubes for the PDZ domain or peptide ligand, respectively. Structurally equivalent side chains (labeled in *A*) are displayed as sticks (or spheres for Gly Cα atoms) for HtrA1-PDZ bound to peptide H1-C1 (*A*), HtrA2-PDZ bound to peptide H2-C1 ($WTMFWV_{COOH}$) (*B*), HtrA3-PDZ bound to peptide H3-C1 (*C*), unliganded DegP-PDZ1 (*D*, PDB entry 1KY9), unliganded DegS-PDZ (*E*, PDB entry 1SOT), and DegS-PDZ bound to OmpC peptide ($VYQF_{COOH}$) (*F*, PDB entry 1SOZ). Side chain carbons are colored orange or green for the PDZ domain or peptide ligand, respectively, and other side chain atoms are colored as follows: oxygen, red; nitrogen, blue; sulfur, yellow. Peptide and PDZ domain residues are labeled in green or black, respectively.

Asp447, Val448, Ser449, Leu458). Of these, only residues Ile418, Asn446, and Ser449 are within 4.5 Å of the peptide.

The effects of Ala substitutions on binding to peptide H1-C2 are similar to those for H1-C1, but there are some notable differences. The most striking differences are for residues Arg386 (β1-2), Ile415 (β2-4), and Ile418 (β2:α1-1), which are clearly important for binding to peptide H1-C2, and the structural model suggests that these side chains support favorable hydrophobic packing interactions with the Ile^{-1} side chain of the ligand (Fig. 6A). In contrast, binding of peptide H1-C1 shows a slight preference for Ala at positions 386 and 418, and the wild type is only modestly favored at position 415 (Fig. 6B). This is likely related to the nonoptimal interactions with Trp^{-1} inferred from the structural analysis (Fig. 5A), which are presumedly improved upon truncation of these side chains by Ala substitutions. Another notable difference is observed at position 389 (β1-5) where the wild-type Ser is strongly preferred for binding to H1-C1 but not for binding to H1-C2. While this residue does not appear to directly contact the peptide in the NMR structure, it is on the periphery of the binding site, and the hydroxyl group of the Ser389 side chain points directly toward the peptide in all members of the ensemble. In the NMR ensemble, the side chains of the three N-terminal peptide residues are poorly defined due to a lack of NOE restraints, and a direct involvement (or lack thereof) of Ser389 with the Arg^{-4} side chain cannot be unambiguously defined.

An HtrA1-PDZ mutant with HtrA3-PDZ-like specificity

While both HtrA1-PDZ and HtrA3-PDZ prefer hydrophobic ligands, the specificities of the two domains differ in detail. HtrA3-PDZ derives significant binding energy from interactions with a Trp side chain at $position^{-1}$, while, in contrast, HtrA1-PDZ does not interact strongly with Trp^{-1} but, instead, derives binding energy from interactions with hydrophobic side chains at the −2 and −3 positions. Our structural and functional analyses

Table 5. *Shotgun alanine scan for HtrA1-PDZ*

	Wt/Ala Ratios[a]			F[b]	
Residue	H1-C1	H1-C2	display	H1-C1	H1-C2
K380	4.2	5.4	0.4	**11**	**14**
K381	5.5	4.2	0.9	**6**	**5**
Y382	8.4	19	0.5	**17**	**38**
I383	61	>58	2.8	**>22**	**>21**
G384	82	>80	2.2	**37**	**>36**
I385	4.6	9.7	2.5	2	4
R386	1.6	12	2.0	0.8	**6**
M387	10	0.4	1.5	**7**	0.3
M388	5.2	0.7	1.4	4	0.5
S389	80	3.7	1.7	**47**	2
L390	3.6	7.2	2.9	1	2
T391	2.7	3.2	0.6	**5**	**5**
S392	1.9	1.6	1.4	1	1
S393	1.6	1.8	0.8	2	2
K394	1.3	0.6	0.5	3	1
K396	1.8	0.4	1.0	2	0.4
E397	0.4	0.4	0.4	1	1
L398	2.9	2.9	1.3	2	2
K399	1.2	1.4	0.9	1	2
D400	0.3	0.4	0.7	0.4	0.6
R401	1.7	1.7	0.7	2	2
H402	1.0	0.6	1.1	0.9	0.6
R403	0.8	1.0	0.7	1	1
D404	0.8	1.3	0.5	2	3
F405	2.8	1.9	3.0	1	0.6
P406	1.3	1.3	1.6	0.8	0.8
D407	0.6	1.0	1.2	0.5	0.8
V408	3.9	0.8	1.8	2	0.4
I409	0.6	1.1	0.5	1	2
S410	1.6	3.0	1.4	1	2
G411	89	90	2.7	**>33**	**33**
Y413	28	45	2.9	**10**	**16**
I414	28	>36	1.4	**20**	**>25**
I415	4.8	21	1.2	4	**18**
E416	1.5	1.1	1.1	1	1
V417	89	>91	2.7	**33**	**>34**
I418	0.6	>86	1.2	0.5	**>75**
P419	3.2	22	2.5	1	**9**
D420	2.3	2.3	1.7	1	1
T421	6.4	>91	2.3	3	**>40**
P422	21	>91	2.5	**8**	**>36**
Q440	2.0	1.9	1.7	1	1
S441	2.2	3.2	2.5	0.9	1
V442	5.2	1.5	44	0.1	0.03
V443	1.7	1.6	1.0	2	2
S444	22	44	12	2	4
N446	0.1	0.1	0.2	0.5	0.5
D447	0.5	1.4	2.5	0.2	0.6
V448	2.8	14	7.3	0.4	2
S449	0.3	0.5	2.3	0.1	0.2
D450	0.5	0.4	0.8	0.6	0.5
V451	1.0	0.4	0.5	2	0.8
I452	18	18	6.0	3	3
K453	0.8	0.2	0.4	2	0.5
R454	1.9	1.6	1.5	1	1
E455	0.5	0.2	0.5	1	0.4
S456	2.9	3.7	3.1	0.9	1
T457	0.8	1.1	1.2	0.7	0.9
L458	8.8	6.8	31	0.3	0.2
N459	0.7	1.4	1.6	0.4	0.9
M460	12	>83	32	0.4	**>3**

[a] The wt/Ala ratios were determined from the sequences of binding clones isolated after selection for binding to either a high affinity peptide ligand (H1-C1 or H1-C2) or an anti-gD-tag antibody (display).
[b] A normalized frequency of occurrence (F) was derived by dividing the function selection wt/Ala ratio by the display selection wt/Ala ratio. In the cases where a particular mutation was not observed among the function selection sequences, only a lower limit could be defined for the wt/Ala ratio and the F-value (indicated by a greater than sign). Bold numbers indicate Ala substitutions predicted to have a significant deleterious effect on binding ($F \geq 5$).

suggest that unfavorable steric interactions between Trp^{-1} and side chains in HtrA1-PDZ are responsible for the differences in specificity at site^{-1}. In particular, the β-branched side chains of Ile415 (β2-4) and Ile418 (β2:α1-1) in HtrA1-PDZ pack in a manner that appears to favor the formation of a salt bridge between Arg386 (β1-2) and Glu416 (β2-5), and this creates a crowded environment that prevents Trp^{-1} from forming favorable interactions with the backbone of the β1-strand. To test this hypothesis, we investigated the effects of mutations that convert

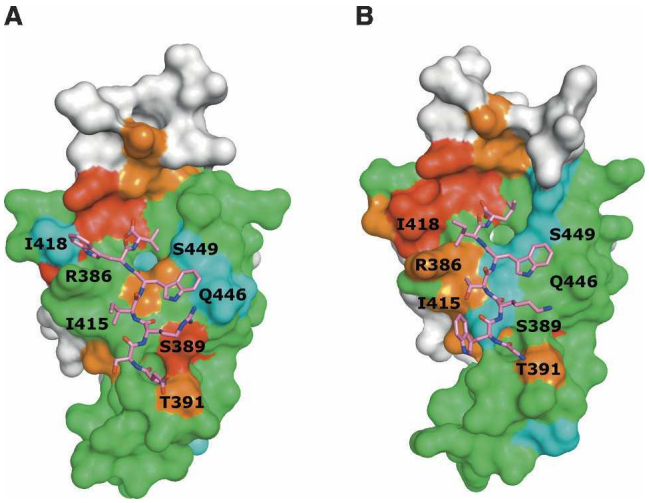

Figure 6. Results of shotgun alanine scanning for binding to peptide H1-C1 (*A*) or H1-C2 (*B*) mapped onto the structure of HtrA1-PDZ. HtrA1-PDZ is shown as a surface and the peptide ligand is shown as sticks colored according to atom type (carbon, magenta; oxygen, red; nitrogen, blue). PDZ domain residues are colored red ($F \geq 16$), orange ($16 > F \geq 5$), green ($5 > F > 0.5$), cyan ($F \leq 0.5$), or white (not scanned). The structural model of HtrA1-PDZ bound to peptide H1-C2 was generated with the homology module of InsightII (Accelrys, Inc) using the NMR ensemble of HtrA1-PDZ bound to peptide H1-C1.

the sequence of HtrA1-PDZ to that of HtrA3-PDZ at these positions. Binding selections with a C-terminal peptide library were unsuccessful for mutant HtrA1-PDZ proteins containing single substitutions (I415Q or I418A), indicating that these mutations resulted in binding sites that were incapable of recognizing peptides with high affinity. In contrast, binding selections were successful against the HtrA1-PDZ double mutant (I415Q/I418A), in which both positions were simultaneously converted to the sequence of HtrA3-PDZ (Fig. 2C). Gratifyingly, alignment of the selected peptides revealed that position^{-1} was occupied exclusively by Trp, as was the case for HtrA3-PDZ. Furthermore, and somewhat surprisingly, specificity at the −2 and −3 positions was also altered with respect to the wild type and, instead, strongly resembles that of HtrA3-PDZ, being promiscuous at position^{-2} and preferring Gly/Ser at position^{-3}. In addition, site^{0} of the mutant accepts a broad range of hydrophobic ligand residues and, thus, appears to be more promiscuous than the wild type, which prefers the aliphatic residues Val or Ile. These analyses reveal that the residues at positions β2-4 and β2:α2-1 are especially critical for determining ligand specificity in a cooperative manner, as simultaneous mutations at these positions alter specificity across the binding cleft. A similar situation was found previously in a comparison of Erbin-PDZ and ZO1-PDZ1, where structural analyses suggested that differing specificities in the −1 and −3 sites of these domains were primarily determined by the residue at a key position within the β-strand that is structurally equivalent to the β2-strand of HtrA1-PDZ (Appleton et al. 2006). Taken together, these results show that a limited number of changes at key positions within the PDZ domain fold can dramatically alter binding specificity, and this may in part explain how PDZ domains can readily evolve different specificities to provide complex scaffolding functions in multicellular organisms.

Discussion

We analyzed the PDZ domains of the human HtrA family using a combination of phage-displayed peptide libraries, synthetic peptide-binding assays, high-resolution structural analyses, and shotgun alanine scanning. The results of this systematic approach applied herein to HtrA1-PDZ and HtrA3-PDZ, and previously to HtrA2-PDZ (Zhang et al. 2007), reveal how subtle changes in PDZ domain sequence and structure impact ligand specificity. Previously, we applied the same approaches to the study of Erbin-PDZ and ZO1-PDZ1, which are typical domains involved in assembling intracellular protein complexes (Appleton et al. 2006; Zhang et al. 2006). Taken together, these studies provide an extremely detailed view of how different types of PDZ domains use a common protein fold to adapt to different cellular tasks.

As Erbin and ZO-1 are scaffold proteins that assemble intracellular complexes, the PDZ domains appear to have evolved to mediate highly specific ligand recognition. To achieve this, the domains recognize the last six or seven ligand side chains with a binding cleft that uses not only hydrophobic contacts but also provides specific electrostatic and hydrogen-bonding interactions (Appleton et al. 2006; Zhang et al. 2006). Furthermore, both domains depend critically on a C-terminal aliphatic residue for high affinity binding, as either blocking of the C-terminal carboxylate or substitution by Ala severely reduces binding. Consequently, these domains are restricted to only bind C-terminal sequences in a highly sequence-specific manner. In contrast, the binding clefts of the HtrA family PDZ domains are lined with residues that do not provide for specific electrostatic or hydrogen-bonding interactions with ligands but, rather, mainly provide hydrophobic van der Waals contacts, which can be favorable or unfavorable depending on the particular ligand sequence (Fig. 5). In addition, while these domains prefer free C termini, they can also accommodate internal sequences. Consequently, the PDZ domains of the HtrA family are promiscuous in terms of ligand specificity, as they recognize C-terminal and internal stretches of hydrophobic sequence.

It appears that, in part, the promiscuous nature of the HtrA-PDZ domain family in comparison with Erbin-PDZ and ZO1-PDZ is determined by differences in the organization of the protein fold. Erbin-PDZ and ZO1-PDZ are structurally permuted relative to the HtrA-PDZ domains, and as a result, the carboxylate binding loop is preceded by a core beta strand. This change acts to tether the loop and may limit the recognition of ligands not containing a C-terminal aliphatic side chain and carboxylate, since only a limited repertoire of conformations are readily accessible. In contrast, the carboxylate-binding loop of the HtrA PDZ domain family is preceded by a flexible tether from the protease domain (Li et al. 2002; Wilken et al. 2004; Zeth 2004). The greater conformational freedom afforded by this change relaxes the strict requirement for a particular C-terminal residue and allows relatively tight binding to be achieved with a range of C-terminal residues and also internal peptide ligands. It is noteworthy that, aside from the small HtrA family, the hundreds of human PDZ domains all adapt a fold similar to that of Erbin-PDZ and ZO1-PDZ, and we speculate that these types of domains are inherently better suited for high-specificity recognition of C-terminal sequences and, thus, have been recruited for tasks that require accurate assembly of protein–protein complexes for signaling and cellular architecture. In contrast, the more promiscuous specificity of the HtrA PDZ domain family for hydrophobic stretches appears ideal for recognition of misfolded proteins, and,

indeed, proteins resembling the DegP protease/chaperone are used ubiquitously by prokaryotes and eukaryotes for protein quality control (Wilken et al. 2004).

HtrA family members also derive specificity from domains other than the PDZ domains. In particular, the protease domain itself possesses inherent substrate specificity (Martins et al. 2003), and, thus, the overall specificity of the HtrA family members depends on a combination of the specificities of the protease and PDZ domains. Furthermore, in the case of HtrA1 and HtrA3, these mammalian proteases have acquired additional protein–protein interaction domains (i.e., the IGFBP-like domain and the Kazal-type serine protease inhibitor domain) that may serve to further fine-tune biological specificity. In the case of HtrA1, binding and inhibition of TGF-β family members is mediated by the protease domain and the preceding linker region (Oka et al. 2004), while in collagen degradation, it appears that the PDZ domain is directly involved in substrate recognition (Murwantoko et al. 2004). Thus, to fully understand the biological function of the different human HtrA family members, it will be necessary to dissect the specificities of the individual domains and, also, to determine how the domains function together in the full-length protein.

Our results also shed light on the likely binding specificities of the bacterial HtrA PDZ domains. The structures of DegS and DegP reveal that the PDZ domain of DegS (DegS-PDZ) (Fig. 5E; Wilken et al. 2004) and the first PDZ domain of DegP (DegP-PDZ1) (Fig. 5D; Krojer et al. 2002) are similar to the human HtrA PDZ domains. Therefore, we would expect that these PDZ domains would also prefer hydrophobic ligands that contain Trp or other aromatic residues at position^{-1}. Furthermore, it is likely that the bacterial HtrA PDZ domains can also recognize internal ligands, since they share a common fold topology with the human HtrA PDZ domains, and it has been noted that the carboxylate-binding loop and β1-strand of DegP are fairly flexible (Krojer et al. 2002). Our predictions are in agreement with investigations of the ligand specificity of DegS-PDZ, which revealed a hydrophobic binding profile (YY[F/M]$_{COOH}$) (Walsh et al. 2003). It is noteworthy that the peptide library used in these studies did not include Trp residues, and, based on our results, we would expect that Trp would also be accepted at site^{-1} and may even be preferred over Tyr $^{-1}$.

Enzymatic studies of HtrA2/Omi (Martins et al. 2003), HtrA1 (Murwantoko et al. 2004), DegP (Jones et al. 2002), and DegS (Walsh et al. 2003) have all shown that binding of peptides to the PDZ domains results in protease activation, and the degree of activation is correlated with the affinity of the peptide ligands for the free PDZ domain. Crystallographic studies of human HtrA2/Omi and *E. coli* DegP have revealed trimeric or hexameric structures, respectively (Krojer et al. 2002; Li et al. 2002), and recent studies of *E. coli* DegS have revealed the molecular details of the mechanism whereby PDZ domain ligands activate the protease (Wilken et al. 2004). The structure of a peptide bound to DegS-PDZ (Fig. 5F) shows the Gln^{-1} side chain pointing away from the binding cleft and making contact with a flexible loop in the protease domain. Thus, it was hypothesized that peptides activate the DegS protease by binding to the PDZ domain and acting as allosteric bridges, which interact with both the PDZ and protease domains and, in so doing, induce conformational changes that activate the protease. The same study revealed that the peptide with Gln at position^{-1} was only a moderately effective activator, and, in fact, the best activators contain aromatic residues (Trp, Tyr, or Phe) at position^{-1}, and these results were consistent with previous data showing that the DegS-PDZ binds with much higher affinity to a peptide containing Tyr^{-1} in place of Gln^{-1}. It was proposed that aromatic side chains at position^{-1} may interact with the protease loop through hydrophobic interactions rather than the polar interactions mediated by Gln^{-1}, but the end result may be a similar, activating conformational change.

Our studies of the human HtrA PDZ domain family suggest that different HtrA PDZ domains may recognize the ligand side chain at the critical position^{-1} in different ways. In the case of HtrA1-PDZ, the Trp^{-1} side chain does not interact strongly with the PDZ domain, and, thus, may be well positioned for interactions with the protease domain in a manner similar to that observed for DegS-PDZ. In contrast, HtrA2-PDZ and HtrA3-PDZ interact much more strongly with the Trp^{-1} side chain, and it is less clear whether the ligand side chain would be available for interaction with the protease domain. Further structural studies will be required to clarify these details, but it seems likely that the entire HtrA family shares a common activation mechanism, whereby peptides bind to the PDZ domain and induce conformational changes that are transmitted to the protease domain. An important aspect of this activation mechanism is that, in general, peptides with longer hydrophobic stretches are likely to bind with higher affinity to the PDZ domains and, consequently, are likely to be better activators of the proteases. Such a correlation would be well suited to proteases evolved for the recognition and degradation of misfolded proteins.

Materials and Methods

Enzymes were from New England Biolabs. Maxisorp immunoplates and 384-well assay plates were from Nalge NUNC International. *E. coli* XL1-Blue, *E. coli* BL21, and M13-VCS were from Stratagene. Plasmid pET15b was from Novagen.

Thrombin was from Calbiochem. Bovine serum albumin (BSA) and Tween 20 were from Sigma. HRP/anti-M13 antibody conjugate, pGEX6P-3, glutathione-sepharose-4B, Superdex-75, and MonoQ anion exchange resin were from Amersham Pharmacia Biotech. Nickel-nitrilotriacetic acid-agarose (NiNTA) was from Qiagen. 3,3′, 5,5′-Tetramethyl-benzidine/H_2O_2 (TMB) peroxidase substrate was from Kirkegaard and Perry Laboratories, Inc. AlphaScreen reagents and plate reader were from PerkinElmer Life Sciences. Plasmid pET15b was from Novagen. D_2O was from Cambridge Isotope Labs, Inc., and other chemicals for isotopic labeling experiments were from Spectra Stable Isotopes.

Synthetic peptides

Peptides were synthesized using standard 9-fluorenylmethoxycarbonyl (Fmoc) protocols, cleaved off the resin with 2.5% triisopropylsilane and 2.5% H_2O in trifluoroacetic acid (TFA), and purified by reversed-phase high performance liquid chromatography (HPLC). The purity and mass of each peptide was verified by liquid chromatography/mass spectrometry (LC/MS).

Isolation of peptide ligands for HtrA1-PDZ and HtrA3-PDZ

Previously described procedures were used to isolate phage-displayed peptides that bound to GST-HtrA1-PDZ or GST-HtrA3-PDZ fusion protein, using a library of random decapeptides or dodecapeptides fused to either the C (library C) or N terminus (library N) of the M13 gene-8 major coat protein (Held and Sidhu 2004), respectively. Each library contained more than 10^{10} unique members. After three rounds of selection, individual clones were grown in a 96-well format in 500 μL of 2YT broth supplemented with carbenicillin and M13-KO7, and the culture supernatants were used directly in phage ELISAs to detect peptides that bound specifically to HtrA1-, HtrA3-, or HtrA1(I415Q/I418A)-PDZ. The peptide sequences were determined from the sequences of the encoding DNA.

Shotgun alanine scanning of HtrA1-PDZ

HtrA1-PDZ was displayed on the surface of M13 bacteriophage by modifying a previously described phagemid (pS2202b) (Skelton et al. 2003). Standard molecular biology techniques were used to replace the fragment of pS2202b encoding Erbin-PDZ with a DNA fragment encoding HtrA1-PDZ. The resulting phagemid (p8HtrA1) contained an open reading frame that encoded the maltose binding protein secretion signal, followed by an epitope tag (amino acid sequence: SMADPNRFRGK DLGS), followed by HtrA1-PDZ and ending with the mature M13 gene-8 minor coat protein. *E. coli* harboring p8HtrA1 were coinfected with M13-KO7 helper phage and grown at 37°C, resulting in the production of phage particles that encapsulated p8HtrA1 DNA and displayed HtrA1-PDZ.

Libraries were constructed using previously described methods (Sidhu et al. 2000) with appropriately designed ''stop template'' versions of p8HtrA1. For each library, we used a stop template that contained TAA stop codons within each of the regions to be mutated. The stop template was used as the template for the Kunkel mutagenesis method (Kunkel et al. 1987) with mutagenic oligonucleotides designed to simultaneously repair the stop codons and introduce mutations at the desired sites. Wild-type codons were replaced with corresponding degenerate codons (Vajdos et al. 2002), which ideally encoded for only alanine and the wild type, although two additional substitutions were allowed at some positions due to the degeneracy of the genetic code. Three libraries were constructed, and each library mutated a discrete region of HtrA1-PDZ as follows: library 1, positions 380–400; library 2, positions 401–422; library 3, positions 440–460. Libraries 1, 2, and 3 contained 3.0×10^{10}, 2.5×10^{10}, or 2.3×10^{10} unique members, respectively.

Phage from the libraries were propagated in *E. coli* XL1-Blue with the addition of M13-KO7 helper phage. After overnight growth at 37°C, phage were concentrated by precipitation with PEG/NaCl and resuspended in PBS, 0.5% BSA, 0.1% Tween 20, as described previously (Sidhu et al. 2000). Phage solutions (10^{12} phage/mL) were added to 96-well Maxisorp immunoplates that had been coated with capture target and blocked with BSA. Two different targets were used; for the display selection the target was an immobilized antibody that recognized the epitope tag fused to the N terminus of HtrA1-PDZ, while for the function selection a biotinylated peptide that binds to HtrA1-PDZ with high affinity (biotin-GWKTWIL or biotin-DSRIWWV) (Laura et al. 2002) was immobilized on NeutrAvidin-coated plates. Following a 2-h incubation to allow for phage binding, the plates were washed 10 times with PBS, 0.05% Tween 20. Bound phage were eluted with 0.1 M HCl for 10 min and the eluent was neutralized with 1.0 M Tris base. Eluted phage were amplified in *E. coli* XL1-Blue and used for further rounds of selection.

Individual clones from the second round of selection were grown in a 96-well format in 500 μL of 2YT broth supplemented with carbenicillin and M13-KO7, and the culture supernatants were used directly in phage ELISAs (Sidhu et al. 2000) to detect phage-displayed HtrA1-PDZ variants that bound to either biotin-GWKTWIL, biotin-DSRIWWV, or anti-tag antibody. More than 50% of the clones exhibited positive phage ELISA signals at least twofold greater than signals on control plates coated with BSA. These positive clones were subjected to DNA sequence analysis. One 96-well plate was sequenced for each selection.

The sequences were analyzed with the program SGCOUNT as described previously (Weiss et al. 2000). SGCOUNT aligned each unique DNA sequence against the wild-type DNA sequence by using a Needleman-Wunch pairwise alignment algorithm, which translated each aligned sequence of acceptable quality, and tabulated the occurrence of each natural amino acid at each position.

Affinity assays

The binding affinities of peptides for HtrA1- or HtrA3-PDZ were estimated as IC_{50} values using a competition ELISA (Fuh et al. 2000). The IC_{50} value was defined as the concentration of peptide that blocked 50% of PDZ domain binding to immobilized peptide. Assay plates were prepared by immobilizing an N-terminally biotinylated peptide (biotin-GWKTWIL or biotin-RSWWV for HtrA1-PDZ or HtrA-3PDZ, respectively) on Maxisorp plates coated with NeutrAvidin and blocked with BSA. A fixed concentration of GST-HtrA1-PDZ (200 nM) or GST-HtrA3-PDZ fusion protein (200 nM) in PBS, 0.5% BSA, 0.1% Tween 20 (PBT buffer) was preincubated for 1 h with serial dilutions of peptide and then transferred to the assay plates. After a 1-h incubation, the plates were washed with PBS, 0.05% Tween 20, incubated for 30 min with HRP/anti-GST

antibody (1:10,000) in PBT buffer, washed again, and detected with TMB peroxidase substrate.

Protein expression and purification

For use in peptide-phage selection experiments, DNA fragments encoding for HtrA1-PDZ (380–480) or HtrA3-PDZ (354–453) were cloned into BamHI/XhoI sites of pGEX6P-3 plasmid, creating open reading frames encoding for GST-HtrA1-PDZ or GST-HtrA3-PDZ fusion proteins. *E. coli* BL21 cultures harboring the appropriate expression plasmids were grown to mid-log phase ($A_{600} = 0.8$) at 37°C in 500 mL of LB broth supplemented with carbenicillin (50 μg/mL), induced with 0.4 mM isopropyl-β-D-thiogalactoside (IPTG) and grown an additional 16 h at 30°C. The bacteria were pelleted by centrifugation at 4000*g* for 15 min, washed with PBS twice, and frozen at −80°C for 8 h. The pellet was resuspended in 50 mL of PBS and lysed by passing through the Microfluidizer Processing Equipment. The GST fusion proteins were purified from cell lysate with affinity chromatography on 2 mL of glutathione-Sepharose-4B according to the manufacturer's instructions.

For X-ray crystallography, a DNA fragment encoding for residues 354–453 of human HtrA3 was cloned into the NdeI/BamHI sites of the pET22d expression vector, creating an open reading frame encoding for HtrA3-PDZ with an N-terminal His tag and a thrombin cleavage site. In addition, standard molecular biology techniques were used to fuse extensions to the C terminus of HtrA3-PDZ to produce open reading frames encoding for HtrA3-PDZ-ext (extension, GGGFGRWV). *E. coli* BL21(DE3) cultures harboring the expression plasmid were grown at 37°C to mid-log phase ($A_{600} = 0.8$). Protein expression was induced with 0.4 mM IPTG and the culture was grown at 16°C for 16 h. The bacteria were pelleted by centrifugation at 4000*g* for 15 min, washed twice with 20 mM Tris-HCl (pH 8.0), and frozen at −80°C for 8 h. The pellet was resuspended in 100 mL of buffer A (50 mM Tris-HCl at pH 8.0 and 500 mM NaCl), and the bacteria were lysed by passing through the Microfluidizer Processing Equipment. The cell lysate was loaded onto a NiNTA column. The column was washed with buffer A plus 20 mM imidazole, and the protein was eluted with 250 mM imidazole in buffer A. Fractions containing the protein of interest were pooled, thrombin was added (1 unit/mg of protein), and the sample was dialyzed overnight against PBS at 4°C. The protein sample was concentrated and further purified over a Superdex-75 column in 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 5 mM β-mercaptoethanol.

For NMR spectroscopy, a DNA fragment encoding for residues 380–480 of human HtrA1 was cloned into the NdeI/BamHI sites of the pET15b expression vector, creating a fusion with an N-terminal His tag followed by a thrombin cleavage site. *E. coli* BLR(DE3)pLysS cells harboring the expression plasmid were grown in M9 minimal media supplemented with ^{15}N-ammonium chloride (>99%) and ^{12}C- and/or $^{13}C_{c6}$-D-glucose (>99%). The cells were grown to mid-log phase ($OD_{600} = 0.8$) at 37°C, induced with 1.0 mM IPTG, and grown at 22°C for an additional 12 h. Bacteria were pelleted by centrifugation at 4000*g* for 15 min, resuspended in buffer A with 1 mM phenylmethanesulfonyl fluoride, and lysed by passing through the Microfluidizer Processing Equipment three times. The cell lysate was clarified by centrifugation at 21,000*g* for 45 min, filtered through a 0.45-μm filter, and loaded onto a NiNTA Superflow column. The column was washed with 10 mM imidazole in buffer A and eluted with 500 mM imidazole in buffer A. Elution fractions were pooled, thrombin was added (1 unit/mg of protein), and the sample was dialyzed overnight at 4°C against buffer B (50 mM TrisHCl at pH 7.5, 100 mM NaCl) with 2 mM $CaCl_2$. The protein sample was concentrated and purified over a Superdex-75 column in 25 mM TrisHCl (pH 7.5), 300 mM NaCl. The sample was further purified over a MonoQ anion exchange column in TrisHCl (pH 7.5) buffer with a 0.1–1.0 M NaCl gradient. Protein samples were concentrated to ~2 mM in 25 mM sodium phosphate (pH 6.0) containing 10% deuterium oxide (D_2O) and 1.0 mM sodium azide. ''One hundred percent'' D_2O samples were prepared by lyophilizing 10% D_2O samples and dissolving in 99.996% D_2O.

NMR spectroscopy and structure determination

All NMR spectra were acquired at 25°C on either a Bruker DRX600 MHz or DRX800 MHz spectrometer equipped with triple resonance, triple axis actively shielded gradient probe. All NMR data were processed using NMRPipe (Delaglio et al. 1995) and analyzed using the program Sparky (version 3.11) (Goddard and Kneller 2007). Complex formation was directly monitored by measuring ^{15}N-HSQC of ^{15}N, ^{13}C-labeled HtrA1-PDZ with stepwise titration of peptide H1-C1. The appearance of new sharp peaks and the disappearance of some peaks corresponding to free HtrA1-PDZ are consistent with slow exchange on the NMR timescale. $^{1}H_N$, ^{15}N, ^{13}Cα, ^{13}Cβ, and $^{13}C'$ assignments were aided by the program Monte (Hitchens et al. 2003) using data from 3D HNCA, HNCOCA, CBCACONH, CBCANH, HNCO, and HNCACO experiments (Cavanagh et al. 1995). Side chain assignments were made by manual analysis of 3D-HCCH-TOCSY in D_2O. Peptide resonances were assigned by analysis of 2D NOESY and 2D TOCSY with ^{13}C and ^{15}N filter in F1 (Zwahlen et al. 1997; Iwahara et al. 2001). Initial structures and distance restraints were obtained by analysis of 3D NOESY-^{15}N-HSQC, 3D NOESY-^{13}C-HSQC, and 3D ^{13}C, F1-filtered, F3-edited-NOESY-HSQC spectra using automated NOE assignment with the program CYANA (version 2.0) (Herrmann et al. 2002;Güntert 2004). ϕ, ψ, and χ1 dihedral restraints were obtained by analysis of HNHA, HNHB, and TOCSY- ^{15}N-HSQC (35 ms mixing time) experiments, according to established Karplus relationships. Additional loose backbone dihedral angle restraints were obtained from analysis of backbone chemical shifts with the program TALOS (Cornilescu et al. 1999). Dihedral restraints were applied for good fits to the chemical shifts (as defined by the program) with the allowed range being the TALOS-defined mean ± the larger of 30° or three times the TALOS-calculated standard deviation. Backbone dynamics were investigated by analyzing the steady state ^{1}H-^{15}N-NOE as described (Farrow et al. 1994). One hundred structures were calculated using the simulated annealing program CNX (Accelrys, Inc.,) using distance, dihedral, and hydrogen-bond restraints starting from random protein and peptide conformations. Twenty structures with the lowest restraint violation energy were selected to represent the solution structure of the complex.

Crystallization and data collection

The HtrA3-PDZ protein (residues 354–453) was crystallized in complex with a high-affinity pentapeptide ($FGRWV_{COOH}$) identified by phage display. For crystallization of the PDZ-ligand complex (designated as HtrA3-PDZext), we took advantage of a previously described strategy (Appleton et al. 2006) and fused the five-residue peptide sequence to the C terminus of the

PDZ domain via a tri-glycine linker. Crystals were obtained by vapor diffusion in sitting drops at 19°C by mixing equal volumes of protein (10 mg/mL) with 0.1 M Bis-Tris (pH 6.5), 0.2 M $MgCl_2$, and 25% PEG 3350; the crystals were transferred to a cryobuffer containing 0.1 M Bis-Tris (pH 6.5), 0.2 M $MgCl_2$, and 35% PEG 3350 prior to flash freezing in liquid N_2. A complete data set was collected at beamline 5.0.1 of the Advanced Light Source (Berkeley, CA).

Crystallographic data processing, structure determination, and refinement

All data were processed using Denzo and Scalepack from the HKL Suite (Otwinowski and Minor 1997). The HtrA3-PDZext structure was solved by molecular replacement using AMoRe (Navaza 1994) and a search model that was generated by SWISS-MODEL (Schwede et al. 2003) using the PDZ domain from the HtrA2/Omi crystal structure (PDB entry 1LCY) (Li et al. 2002). HtrA3-PDZext crystallized in space group $P4_12_12$ with two molecules per asymmetric unit. Each PDZ domain in the asymmetric unit forms a crystallographic dimer with the ligand from an apposing molecule that is related by crystallographic symmetry. This packing arrangement creates two nonequivalent PDZ-peptide dimers, which are structurally very similar, as evidenced by a RMSD of 0.3 Å over 102 Cα atoms. Atomic models were built with COOT (Emsley and Cowtan 2004) and refined with REFMAC (Murshudov et al. 1997).

Protein Data Bank accession numbers

The atomic coordinates and structure factors (codes: 2joa and 2p3w) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, New Jersey (http://www.rcsb.org/).

Acknowledgments

We thank Clifford Quan and Thuy Tran for peptide synthesis, Yan Wu for supplying PDZ domain proteins, and Jianjun Zhang for bioinformatics support. Portions of this research were carried out at the Advanced Light Source, supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory.

References

Aasland, R., Abrams, C., Ampe, C., Ball, L.J., Bedford, M.T., Cesareni, G., Gimona, M., Hurley, J.H., Jarchau, T., and Lehto, V.-P. 2002. Normalization of nomenclature for peptide motifs as ligands of modular protein domains. *FEBS Lett.* **513:** 141–144.

Appleton, B.A., Zhang, Y., Wu, P., Yin, J.P., Hunziker, W., Skelton, N.J., Sidhu, S.S., and Wiesmann, C. 2006. Comparative structural analysis of the Erbin-PDZ domain and the first PDZ domain of ZO-1: Insights into determinants of PDZ domain specificity. *J. Biol. Chem.* **281:** 22312–22320.

Baldi, A., De Luca, A., Battista, T., Felsani, A., Baldi, F., Catricala, C., Amantea, A., Noonan, D.M., Albini, A., Natali, P.G., et al. 2002. The HtrA1 serine protease is down-regulated during human melanoma. *Oncogene* **21:** 6684–6688.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G., and Skelton, N.J. 1995. *Protein NMR spectroscopy, principles and practice*. Academic Press, New York.

Chien, J., Staub, J., Hu, S.I., Erickson-Johnson, M.R., Couch, F.J., Smith, D.I., Crowl, R.M., Kaufmann, S.H., and Shridhar, V. 2004. A candidate tumor suppressor HtrA1 is down-regulated in ovarian cancer. *Oncogene* **23:** 1636–1644.

Chien, J., Aletti, G., Baldi, A., Catalano, V., Muretto, P., Keeney, G.L., Kalli, K.R., Staub, J., Ehrmann, M., Cliby, W.A., et al. 2006. Serine protease HtrA1 modulates chemotherapy-induced cytotoxicity. *J. Clin. Invest.* **116:** 1994–2004.

Cochran, A.G., Tong, R.T., Starovasnik, M.A., Park, E.J., McDowell, R.S., Theaker, J.E., and Skelton, N.J. 2001. A minimal peptide scaffold for β-turn display: Optimizing a strand position in disulfide-cyclized β-hairpins. *J. Am. Chem. Soc.* **123:** 625–632.

Cornilescu, G., Delaglio, F., and Bax, A. 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13:** 289–302.

Daniels, D.L., Cohen, A.R., Anderson, J.M., and Brunger, A.T. 1998. Crystal structure of the hCASK PDZ domain reveals the structural basis of class II PDZ domain target recognition. *Nat. Struct. Mol. Biol.* **5:** 317–325.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6:** 277–293.

De Luca, A., De Falco, M., Fedele, V., Cobellis, L., Mastrogiacomo, A., Laforgia, V., Tuduce, I.L., Campioni, M., Giraldi, D., Paggi, M.G., et al. 2004. The serine protease HtrA1 is up-regulated in the human placenta during pregnancy. *J. Histochem. Cytochem.* **52:** 885–892.

DeWan, A., Liu, M., Hartman, S., Zhang, S.S.-M., Liu, D.T.L., Zhao, C., Tam, P.O.S., Chan, W.M., Lam, D.S.C., Snyder, M., et al. 2006. HTRA1 promoter polymorphism in wet age-related macular degeneration. *Science* **314:** 989–992.

Ekert, P.G. and Vaux, D.L. 2005. The mitochondrial death squad: Hardened killers or innocent bystanders? *Curr. Opin. Cell Biol.* **17:** 626–630.

Emsley, P. and Cowtan, K. 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60:** 2126–2132.

Farrow, N.A., Zhang, O., Forman-Kay, J.D., and Kay, L.E. 1994. A heteronuclear correlation experiment for simultaneous determination of 15N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4:** 727–734.

Fuh, G., Pisabarro, M.T., Li, Y., Quan, C., Lasky, L.A., and Sidhu, S.S. 2000. Analysis of PDZ domain-ligand interactions using carboxyl-terminal phage display. *J. Biol. Chem.* **275:** 21486–21491.

Goddard, T.D. and Kneller, D.G. 2007. SPARKY 3. University of California, San Francisco, CA.

Gouet, P., Robert, X., and Courcelle, E. 2003. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31:** 3320–3323.

Grau, S., Baldi, A., Bussani, R., Tian, X., Stefanescu, R., Przybylski, M., Richards, P., Jones, S.A., Shridhar, V., Clausen, T., et al. 2005. Implications of the serine protease HtrA1 in amyloid precursor protein processing. *Proc. Natl. Acad. Sci.* **102:** 6021–6026.

Grau, S., Richards, P.J., Kerr, B., Hughes, C., Caterson, B., Williams, A.S., Junker, U., Jones, S.A., Clausen, T., and Ehrmann, M. 2006. The role of human HtrA1 in arthritic disease. *J. Biol. Chem.* **281:** 6124–6129.

Güntert, P. 2004. CYANA: Combined assignment and dynamics algorithm for NMR applications. www.las.jp/prod/cyana/eg/.

Harris, B.Z. and Lim, W.A. 2001. Mechanism and role of PDZ domains in signaling complex assembly. *J. Cell Sci.* **114:** 3219–3231.

Harris, B.Z., Hillier, B.J., and Lim, W.A. 2001. Energetic determinant of internal motif recognition by PDZ domains. *Biochemistry* **40:** 5921–5930.

Held, H.A. and Sidhu, S.S. 2004. Comprehensive mutational analysis of the M13 major coat protein: Improved scaffolds for C-terminal phage display. *J. Mol. Biol.* **340:** 587–597.

Herrmann, T., Guntert, P., and Wuthrich, K. 2002. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319:** 209–227.

Hitchens, T.K., Lukin, J.A., Zhan, Y., McCallum, S.A., and Rule, G.S. 2003. MONTE: An automated Monte Carlo based approach to nuclear magnetic resonance assignment of proteins. *J. Biomol. NMR* **25:** 1–9.

Hou, J., Clemmons, D.R., and Smeekens, S. 2005. Expression and characterization of a serine protease that preferentially cleaves insulin-like growth factor binding protein-5. *J. Cell. Biochem.* **94:** 470–484.

Hu, S.-I., Carozza, M., Klein, M., Nantermet, P., Luk, D., and Crowl, R.M. 1998. Human HtrA, an evolutionarily conserved serine protease identified as a differentially expressed gene product in osteoarthritic cartilage. *J. Biol. Chem.* **273:** 34406–34412.

Im, Y.J., Lee, J.H., Park, S.H., Park, S.J., Rho, S.-H., Kang, G.B., Kim, E., and Eom, S.H. 2003a. Crystal structure of the Shank PDZ-ligand complex

reveals a class I PDZ interaction and a novel PDZ-PDZ dimerization. *J. Biol. Chem.* **278:** 48099–48104.

Im, Y.J., Park, S.H., Rho, S.-H., Lee, J.H., Kang, G.B., Sheng, M., Kim, E., and Eom, S.H. 2003b. Crystal structure of GRIP1 PDZ6-peptide complex reveals the structural basis for class II PDZ target recognition and PDZ domain-mediated multimerization. *J. Biol. Chem.* **278:** 8501–8507.

Iwahara, J., Wojciak, J.M., and Clubb, R.T. 2001. Improved NMR spectra of a protein–DNA complex through rational mutagenesis and the application of a sensitivity optimized isotope-filtered NOESY experiment. *J. Biomol. NMR* **19:** 231–241.

Jones, C.H., Dexter, P., Evans, A.K., Liu, C., Hultgren, S.J., and Hruby, D.E. 2002. *Escherichia coli* DegP protease cleaves between paired hydrophobic residues in a natural substrate: The PapA pilin. *J. Bacteriol.* **184:** 5762–5771.

Jones, J.M., Datta, P., Srinivasula, S.M., Ji, W., Gupta, S., Zhang, Z., Davies, E., Hajnoczky, G., Saunders, T.L., Van Keuren, M.L., et al. 2003. Loss of Omi mitochondrial protease activity causes the neuromuscular disorder of mnd2 mutant mice. *Nature* **425:** 721–727.

Kang, B.S., Cooper, D.R., Devedjiev, Y., Derewenda, U., and Derewenda, Z.S. 2003. Molecular roots of degenerate specificity in syntenin's PDZ2 domain: Reassessment of the PDZ recognition paradigm. *Structure* **11:** 845–853.

Karthikeyan, S., Leung, T., and Ladias, J.A.A. 2001. Structural basis of the Na^+/H^+ exchanger factor PDZ1 interaction with the carboxyl-terminal region of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* **276:** 19683–19686.

Karthikeyan, S., Leung, T., and Ladias, J.A.A. 2002. Structural determinants of the Na^+/H^+ exchanger regulatory factor interaction with the β2 adrenergic and platelet-derived growth factor receptors. *J. Biol. Chem.* **277:** 18973–18978.

Kolmar, H., Waller, P.R.H., and Sauer, R.T. 1996. The DegP and DegQ periplasmic endoproteases of *Escherichia coli*: Specificity for cleavage sites and substrate conformation. *J. Bacteriol.* **178:** 5925–5929.

Kozlov, G., Gehring, K., and Ekiel, I. 2000. Solution structure of the PDZ2 domain from human phosphatase hPTP1E and its interactions with C-terminal peptides from the Fas receptor. *Biochemistry* **39:** 2572–2580.

Kozlov, G., Banville, D., Gehring, K., and Ekiel, I. 2002. Solution structure of the PDZ2 domain from cytosolic human phosphatase hPTP1E complexed with a peptide reveals contribution of the β2-β3 loop to PDZ domain-ligand interactions. *J. Mol. Biol.* **320:** 813–820.

Krojer, T., Garrido-Franco, M., Huber, R., Ehrmann, M., and Clausen, T. 2002. Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. *Nature* **416:** 455–459.

Kunkel, T.A., Roberts, J.D., and Zakour, R.A. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154:** 367–382.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26:** 283–291.

Laura, R.P., Witt, A.S., Held, H.A., Gerstner, R., Deshayes, K., Koehler, M.F., Kosik, K.S., Sidhu, S.S., and Lasky, L.A. 2002. The Erbin-PDZ domain binds with high affinity and specificity to the carboxyl termini of δ-catenin and ARVCF. *J. Biol. Chem.* **277:** 12906–12914.

Li, W., Srinivasula, S.M., Chai, J., Li, P., Wu, J.W., Zhang, Z., Alnemri, E.S., and Shi, Y. 2002. Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi. *Nat. Struct. Biol.* **9:** 436–441.

Martins, L.M., Turk, B.E., Cowling, V., Borg, A., Jarrell, E.T., Cantley, L.C., and Downward, J. 2003. Binding specificity and regulation of the serine protease and PDZ domains of HtrA2/Omi. *J. Biol. Chem.* **278:** 49417–49427.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53:** 240–255.

Murwantoko,, Yano, M., Ueta, Y., Murasaki, A., Kanda, H., Oka, C., and Kawaichi, M. 2004. Binding of proteins to the PDZ domain regulates proteolytic activity of HtrA1 serine protease. *Biochem. J.* **381:** 895–904.

Navaza, J. 1994. AMoRe: An automated package for molecular replacement. *Acta Crystallogr. A* **50:** 157–163.

Nie, G.Y., Hampton, A., Li, Y., Findlay, J.K., and Salamonsen, L.A. 2003a. Identification and cloning of two isoforms of human high-temperature requirement factor A3 (HtrA3), characterization of its genomic structure and comparison of its tissue distribution with HtrA1 and HtrA2. *Biochem. J.* **371:** 39–48.

Nie, G.Y., Li, Y., Minoura, H., Batten, L., Ooi, G.T., Findlay, J.K., and Salamonsen, L.A. 2003b. A novel serine protease of the mammalian HtrA family is up-regulated in mouse uterus coinciding with placentation. *Mol. Hum. Reprod.* **9:** 279–290.

Oka, C., Tsujimoto, R., Kajikawa, M., Koshiba-Takeuchi, K., Ina, J., Yano, M., Tsuchiya, A., Ueta, Y., Soma, A., Kanda, H., et al. 2004. HtrA1 serine protease inhibits signaling mediated by Tgfbeta family proteins. *Development* **131:** 1041–1053.

Otwinowski, Z. and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276:** 307–326.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. 2003. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **31:** 3381–3385.

Sheng, M. and Sala, C. 2001. PDZ domains and the organization of supramolecular complexes. *Annu. Rev. Neurosci.* **24:** 1–29.

Sidhu, S.S., Lowman, H.B., Cunningham, B.C., and Wells, J.A. 2000. Phage display for selection of novel binding peptides. *Methods Enzymol.* **328:** 333–363.

Skelton, N.J., Koehler, M.F., Zobel, K., Wong, W.L., Yeh, S., Pisabarro, M.T., Yin, J.P., Lasky, L.A., and Sidhu, S.S. 2003. Origins of PDZ domain ligand specificity. Structure determination and mutagenesis of the Erbin-PDZ domain. *J. Biol. Chem.* **278:** 7645–7654.

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. Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* **275:** 73–77.

Tocharus, J., Tsuchiya, A., Kajikawa, M., Ueta, Y., Oka, C., and Kawaichi, M. 2004. Developmentally regulated expression of mouse HtrA3 and its role as an inhibitor of TGF-β signaling. *Dev. Growth Differ.* **46:** 257–274.

Vajdos, F.F., Adams, C.W., Breece, T.N., Presta, L.G., de Vos, A.M., and Sidhu, S.S. 2002. Comprehensive functional maps of the antigen-binding site of an anti-ErbB2 antibody obtained with shotgun scanning mutagenesis. *J. Mol. Biol.* **320:** 415–428.

Walsh, N.P., Alba, B.M., Bose, B., Gross, C.A., and Sauer, R.T. 2003. OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell* **113:** 61–71.

Weiss, G.A., Watanabe, C.K., Zhong, A., Goddard, A., and Sidhu, S.S. 2000. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc. Natl. Acad. Sci.* **97:** 8950–8954.

Wilken, C., Kitzing, K., Kurzbauer, R., Ehrmann, M., and Clausen, T. 2004. Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. *Cell* **117:** 483–494.

Yang, Z., Camp, N.J., Sun, H., Tong, Z., Gibbs, D., Cameron, D.J., Chen, H., Zhao, Y., Pearson, E., Li, X., et al. 2006. A variant of the HtrA1 gene increases susceptibility to age-related macular degeneration. *Science* **314:** 992–993.

Zeth, K. 2004. Structural analysis of DegS, a stress sensor of the bacterial periplasm. *FEBS Lett.* **569:** 351–358.

Zhang, Y., Yeh, S., Appleton, B.A., Held, H.A., Kausalya, P.J., Phua, D.C.Y., Wong, W.L., Lasky, L.A., Wiesmann, C., Hunziker, W., et al. 2006. Convergent and divergent ligand specificity among PDZ domains of the LAP and zonula occludens (ZO) families. *J. Biol. Chem.* **281:** 22299–22311.

Zhang, Y., Appleton, B.A., Wu, P., Wiesmann, C., and Sidhu, S.S. 2007. Structural and functional analysis of the ligand specificity of the HtrA2/Omi PDZ domain. *Protein Sci.* **16:** 1738–1750.

Zumbrunn, J. and Trueb, B. 1996. Primary structure of a putative serine protease specific for IGF-binding proteins. *FEBS Lett.* **398:** 187–192.

Zwahlen, C., Legault, P., Vincent, S.J.F., Greenblatt, J., Konrat, R., and Kay, L.E. 1997. Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage λ N-peptide/*boxB* RNA complex. *J. Am. Chem. Soc.* **119:** 6711–6721.