C-terminal specific protein degradation: Activity and substrate specificity of the Tsp protease

KENNETH C. KEILER,¹ KAREN R. SILBER,¹ KEVIN M. DOWNARD,²
IOANNIS A. PAPAYANNOPOULOS,² KLAUS BIEMANN,² AND ROBERT T. SAUER¹

¹ Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

² Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(RECEIVED February 28, 1995; ACCEPTED May 4, 1995)

Abstract

The activity of Tsp, a periplasmic endoprotease of *Escherichia coli*, has been characterized by assaying the cleavage of protein and peptide substrates, determining the cleavage sites in several substrates, and investigating the kinetics of the cleavage reaction. Tsp efficiently cleaves substrates that have apolar residues and a free α -carboxylate at the C-terminus. Tsp cleaves its substrates at a discrete number of sites but with rather broad primary sequence specificity. In addition to preferences for residues at the C-terminus and cleavage sites, Tsp displays a preference for substrates that are not stably folded: unstable variants of Arc repressor are better substrates than a hyperstable mutant, and a peptide with little stable structure is cleaved more efficiently than a protein substrate. These data are consistent with a model in which Tsp cleavage of a protein substrate involves binding to the C-terminal tail of the substrate, transient denaturation of the substrate, and then recognition and hydrolysis of specific peptide bonds.

Keywords: intracellular proteolysis; mass spectrometry; periplasmic protease; substrate specificity

Proteases are ubiquitous in biological systems. The structures, properties, and mechanisms of a large number and variety of extracellular proteases are known (Beynon & Bond, 1989), but our knowledge of intracellular proteases is far less complete. For example, intracellular proteolysis is extremely selective (Goldberg, 1992; Gottesman & Maurizi, 1992) and yet the simple question of what determines the substrate specificity of an intracellular protease has in no case been clearly answered. In yeast and higher eukaryotes, some proteins are marked for degradation by modification with ubiquitin (Hershko & Ciechanover, 1992), but in bacteria there is no ubiquitin system. Factors that have been shown to be important for the proteolytic fate of specific bacterial proteins include thermal stability (Parsell & Sauer, 1989), the identity of N-terminal sequences (Tobias et al., 1991), and the identity of C-terminal sequences (Bowie & Sauer, 1989; Parsell et al., 1990). A better understanding of proteolytic selectivity requires purification of the relevant proteases and biochemical studies of their interactions with substrates.

Tsp is a periplasmic protease of *Escherichia coli* that was purified based upon its ability to selectively degrade a variant of the N-terminal domain of λ repressor with the apolar C-terminal

sequence WVAAA, but not to degrade a variant with the polar C-terminal sequence RSEYE (Silber et al., 1992). The gene encoding Tsp has been called both *tsp* and *prc* (Hara et al., 1991; Silber et al., 1992), and Tsp may be identical with protease Re (Park et al., 1988) and with a protease purified based upon its ability to cleave oxidized glutamine synthetase (Roseman & Levine, 1987). The λ repressor variant that is a substrate for Tsp in vitro is still degraded in cells deleted for the *tsp* gene, indicating that another protease must be responsible for degradation of this substrate in the deletion strain (Silber & Sauer, 1994). However, cells deleted for the *tsp* gene are defective in the C-terminal processing of two periplasmic proteins, penicillin-binding protein 3 (Nagasawa et al., 1989) and tonB (R. Larsen & K. Postle, pers. comm.), suggesting that these proteins may be natural substrates of the Tsp protease. Tsp has no sequence homology to any characterized protease and is not inhibited by a wide array of small-molecule protease inhibitors (Silber et al., 1992). *ctpA*, a gene with homology to *tsp*, has recently been identified in the cyanobacterium *Synechocystis* sp. PCC 6803 and is believed to encode a protease that is responsible for C-terminal processing of the D1 protein (Anbudurai et al., 1994).

In this paper, we probe the characteristics of peptides and proteins that make them substrates for Tsp in vitro, determine the location of cleavage sites in several protein and peptide substrates, and present a preliminary characterization of the kinetic

Reprint requests to: Robert T. Sauer, Department of Biology, Massachusetts Institute of Technology, 68-571, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139; e-mail: bobsauer@eagle.mit.edu.

Fig. 1. CD spectra of the Tsp (filled diamonds) and TspH₆ (open squares) enzymes.

properties of Tsp. These studies confirm the importance of C-terminal sequence in determining substrate preference but show that additional factors are also involved. One of these factors appears to be thermodynamic stability because experiments with a set of related substrates with differing stabilities indicate that the denatured protein is the only proteolytically susceptible form. Another factor may be the presence of appropriate cleavage sites because we find that Tsp makes endoproteolytic cuts in its substrates at a limited number of discrete bonds.

Results

Comparison of Tsp and TspH₆

To simplify the purification of Tsp, we constructed a *tsp* gene encoding six additional C-terminal histidine residues and purified

the resulting gene product, called TspH₆, using nickel chelate chromatography. The CD spectra of Tsp and TspH₆ are almost identical and consistent with approximately 35–40% α -helix (see Fig. 1). The molecular weight of TspH₆ was determined to be 78.5 ± 5.3 kDa by sedimentation equilibrium centrifugation. The polypeptide molecular weight calculated from the sequence is 75.0 kDa, indicating that TspH₆ is monomeric, as has previously been shown for Tsp (Silber et al., 1992).

As shown in Figure 2, Tsp and TspH₆ have very similar activities in cleaving several protein substrates. In addition, time courses of the cleavage of Arc repressor and the #105 variant of the N-terminal domain of λ repressor by TspH₆ and Tsp are indistinguishable (not shown). These results indicate that the histidine tag does not affect the specificity or kinetic properties of Tsp. Moreover, because Tsp and TspH₆ are purified by completely different methods and yet display the same enzymatic activity, it is extremely unlikely that the observed proteolytic activity is due to a contaminant. The experiments described in the following sections were performed using the TspH₆ enzyme.

Importance of C-terminal sequences in different protein and peptide substrates

Four sets of proteins or peptides were tested as potential substrates for Tsp (Table 1). Within each set, the amount of cleavage is sensitive to the identity of residues or functional groups at the C-terminus. For example, the two λ repressor variants and four Arc variants shown in Table 1A differ only in their C-terminal sequences and yet show dramatically different sensitivities to Tsp cleavage. In single point cleavage assays, the variants with relatively apolar C-terminal tails (λ #105, Arc, and Arc-st7) are good substrates for Tsp, whereas those with charged or highly polar tails (λ N-terminal domain, Arc-st5, and Arc-st6) are poor substrates. In addition, even though each of the Arc variants has the apolar sequence GRIGA in its C-terminal tail, only those with this sequence at the C-terminus (Arc and Arc-st7) are good substrates.

Fig. 2. Degradation of λ repressor and Arc substrates by Tsp and TspH₆. Reaction mixtures were incubated at 37 °C for 4 h, and cleavage was assayed by gel electrophoresis. At shorter incubation times the same cleavage products were observed, but at reduced levels. Arrow indicates the position of the Tsp and TspH₆ protein bands.

Table 1. Cleavage of protein and peptide substrates by Tsp

A. Protein substrates		
Protein ^a	Protein background: tail sequence	% Cleavage
λ N domain	λ 1-97: RSEYE	0
λ #105 N domain	λ 1-97: WVAAA	95
Arc repressor	Arc 1-48: GRIGA	90
Arc-st5	Arc 1-48: GRIGAKNQHE	0
Arc-st6	Arc 1-48: GRIGAHHHHHH	20
Arc-st7	Arc 1-48: GRIGAHHHHHHGRIGA	90
B. Peptide substrates		
Peptide ^b	Sequence	% Cleavage
BAS9	Ac WAAAAAARBAAB COOH	50
BAS7	Ac WAAAAAARBAAB CONH ₂	<5
BAS10	Ac WAAAAAARBGGG COOH	45
BAS8	Ac WAAAAAARBGGG CONH ₂	<5

^a Protein substrates were incubated with Tsp for 8 h and cleavage was quantified by laser densitometry of SDS gels.

^b Sequences of peptides tested as substrates of Tsp. B represents a non-natural amino acid amino: butyric acid. Peptides with the N-terminus blocked by acetylation (Ac) or the C-terminus blocked by amidation (CONH₂) are indicated. Peptides were incubated with Tsp for 5 h, chromatographed on an HPLC reverse-phase column, and the amount of cleavage was quantified by integrating the peak corresponding to intact peptide.

The presence of a free α-carboxyl group is also important in determining whether closely related peptides with apolar C-terminal sequences are cleaved efficiently by Tsp. As shown in Table 1B, peptides BAS7 and BAS9 differ only in whether the C-terminus is a carboxyl or carboxyamide group; the same is true for peptides BAS8 and BAS10. In both cases, the peptide with the α-carboxyl group is cleaved readily by Tsp, whereas the variant with the α-carboxyamide group is cleaved poorly or not cleaved at all (Table 1B; Fig. 3).

Site of proteolytic cleavage

To determine the sites at which Tsp cleaves its substrates, the digestion products of Arc, Arc-st7, the #105 variant of the N-terminal domain of λ-repressor, and peptide BAS10 were analyzed by MALDI-TOF mass spectrometry. The major products from the digestion of each substrate are listed in Table 2 and are shown schematically in Figure 4. Several conclusions emerge from these studies. First, Tsp is clearly an endoprotease. Although Tsp recognizes determinants at the C-terminus of substrates, it cleaves these substrates at a limited number of discrete sites within each polypeptide chain (Table 2; Fig. 4). These sites occur at different distances in the sequence from the C-terminus and map to diverse locations in the native secondary and tertiary structures. The P1 position of the scissile peptide bond is Ala, Ser, or Val in 14 of the 20 cleavage sites; Ile, Leu, Arg, and Lys occupy the P1 position in the remaining 6 sites. Hence, Tsp appears to prefer P1 residues that are small and uncharged or that have aliphatic side chains extending through the C_γ, C_δ, or C_ε

Fig. 3. Cleavage of peptides BAS10 (free α-carboxyl group) and BAS8 (α-carboxyamide group) by TspH₆. Reverse-phase HPLC traces (A₂₂₆ versus time) are shown after 0 and 2 h of incubation with TspH₆ (pH 8, 37 °C).

positions. The same general preferences are evident at the P1' position of the scissile peptide bond except there is no discernible preference for a small residue and the repertoire of allowed residues includes Met, Tyr, and Trp.

Kinetic properties and the importance of substrate stability

To determine approximate kinetic parameters for Tsp cleavage of a protein and a peptide substrate, initial rates of cleavage were determined for different substrate concentrations of Arc and peptide BAS9 by using CD to monitor the loss of α-helicity that occurs upon degradation. Figure 5A and B shows the rate versus substrate concentration plots for Tsp cleavage of Arc and BAS9, respectively. Fitting of these data yields the apparent kinetic parameters shown in Table 3. The apparent *K_m* for Tsp cleavage of peptide BAS9 is comparable to the *K_m* observed for Tsp cleavage of Arc, but *k_{cat}* and *k_{cat}/K_m* are almost 30-fold higher for cleavage of the peptide. Thus, Tsp-catalyzed degradation of peptide BAS9 is significantly more efficient than Tsp-catalyzed degradation of Arc. By contrast, Tsp and chymotrypsin have similar values of *k_{cat}/K_m* using Arc as a substrate (Table 3).

One factor that might influence the efficiency with which Tsp cleaves different proteins or peptides is the equilibrium between the native and denatured forms of the substrates. This could occur, for example, if the denatured substrate were actually the form in which the peptide bond was cleaved (see Discussion). To test this idea, the susceptibility to cleavage of Arc repressor

Fig. 4. Sites of Tsp cleavage are indicated by arrows in four protein and peptide substrates. Elements of secondary structure in λ repressor and Arc (Pabo & Lewis, 1982; Breg et al., 1990) are taken from the structures of the wild-type proteins. α -Helices and β -strands are labeled. The first 50 residues of λ #105 are not to scale.

Table 2. Products of Tsp cleavage of substrates determined by MALDI-TOF mass spectrometry

Substrate	Measured m/z value ^a	Assigned fragment	Assigned cleavage ^b	Difference from calculated m/z value ^c
λ #105	11,177.15	1-102	S...A	-2.77
	11,032.40	1-100	S...A ^d	-5.36
	10,967.11	1-99	S...V ^d	+1.42
	10,676.52	1-97	S...L ¹ W	-4.72
	10,563.14	1-96	S...S ¹ L	-5.04
	10,120.10	1-92	S...S ¹ M	-4.55
	10,003.07	1-91	S...V ¹ S	-4.51
	7,132.32	1-66	S...A ¹ K	+8.12
	6,762.50	1-62	S...A ¹ A	+6.78
	6,413.92	1-59	S...A ¹ Y	+6.56
	6,113.65	1-56	S...A ¹ L	+4.63
	1,577.36	67-80	A ¹ K...I ¹ A	+0.49
	1,489.77	80-91	S ¹ I...V ¹ S	+2.06
	1,205.00	82-90	A ¹ R...A ¹ V	+0.65
	1,161.54	92-102	V ¹ S...A ^c	+0.17
	1,076.18	93-102	S ¹ M...A	+1.89
	1,020.90	92-100	V ¹ S...A ^d	+0.69
	948.29	92-99	V ¹ S...V ^d	+0.16
	932.52	93-100	S ¹ M...A ^d	+0.39
	861.19	93-99	S ¹ M...V ^d	+0.14
Arc	734.06	60-66	A ¹ Y...A ¹ K ^c	-1.79*
	629.45	97-102	S ¹ L...A	-1.31
	487.78	97-100	S ¹ L...A ^d	-0.83
	6,229.22	1-53	M...A ^f	+1.95
	5,566.04	7-53	K ¹ M...A	+0.63
Arc-st7	3,570.47	23-53	V ¹ R...A ^f	+0.44
	2,678.29	1-22	M...V ¹ R ^f	+0.99
	2,137.44	23-40	V ¹ R...R ¹ V	+1.11
	7,263.96	1-61	M...R ¹ I	+1.10
BAS10	4,844.48	23-64	V ¹ R...A	+1.09
	1,753.94	26-40	V ¹ A...R ¹ V	-0.41
BAS10	1,340.61	3-17	A ¹ R...B	+0.08

^a Measured m/z values represent average (i.e., polyisotopic) values for the $[M+H]^+$ ion.

^b Arrows indicate cleavage between the residues shown.

^c The difference between measured and calculated m/z values for the assigned sequence is $\leq 0.18\%$ (except for sequence marked by *, which was confirmed independently; see footnote e).

^d In addition to 1-102, the starting material contained small amounts of 1-100 and 1-99. Thus, fragments that end at position 99 or 100 are assumed to result from these truncated substrates and not from cleavage by Tsp.

^e Sequence confirmed by an independent MS/MS experiment (Bie-mann, 1990).

^f N-terminal sequence confirmed by Edman degradation.

variants with decreased stability (Arc-NK29, lysine for asparagine at position 29) or increased stability (Arc-PL8, leucine for proline at position 8) were compared with that of wild-type Arc. Arc-NK29 and Arc-PL8 each differ from wild-type Arc at a single amino acid residue and have the same apolar C-terminal sequence as Arc. Moreover, neither of the sites of mutation is adjacent to a scissile bond. The melting temperatures of these three proteins under comparable conditions are approximately

Fig. 5. Rate versus substrate concentration plots for Tsp cleavage of peptide BAS9 and Arc. Cleavage was assayed by changes in CD ellipticity. Error bars indicate the uncertainty in determining the rate in each assay. Points below 25 μ M have errors of similar magnitude that have not been shown.

37 °C (Arc-NK29), 50 °C (Arc), and 70 °C (Arc-PL8) (M. Milla, pers. comm.). As shown in Figure 6, the hyperstable PL8 Arc variant is cleaved significantly more slowly than Arc, whereas the unstable NK29 variant is cleaved significantly more rapidly than Arc. These data suggest that substrates are cleaved only in a denatured form.

Discussion

In this work, we have investigated the substrate specificity of Tsp and probed some of its basic enzymatic properties. Previous studies had shown that Tsp cleaves a variant of λ repressor's N-terminal domain with an apolar, but not a polar, C-terminal tail sequence (Silber et al., 1992). Here we have shown that Tsp also selectively cleaves variants of Arc repressor depending upon their C-terminal sequences. Arc variants with relatively apolar C-terminal residues are good substrates for Tsp, whereas those with polar C-terminal residues are poor substrates. The studies reported here also show that a free α -carboxyl group at the C-terminus of an appropriate substrate is important in determining whether it will be cleaved by Tsp. Peptides BAS9 and BAS10, which have apolar C-terminal sequences and free

Table 3. Apparent kinetic parameters for the cleavage of Arc and the peptide BAS9 by Tsp and chymotrypsin (Chy)^a

Protease	Substrate	Apparent kinetic parameters
Tsp	Arc	$K_m = 50 (\pm 36) \mu\text{M}$ $k_{cat} = 0.19 (\pm 0.05) \text{ s}^{-1}$ $k_{cat}/K_m = 3.8 (\pm 2.9) \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$
Chy	Arc	$k_{cat}/K_m = 4.4 (\pm 0.5) \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$
Tsp	BAS9	$K_m = 35 (\pm 13) \mu\text{M}$ $k_{cat} = 3.7 (\pm 0.5) \text{ s}^{-1}$ $k_{cat}/K_m = 1.1 (\pm 0.4) \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$

^a K_m and k_{cat} values were determined by fitting the data in Figure 4 to the hyperbolic function $v = k_{cat}[E_t]\{[S]/(K_m + [S])\}$.

α -carboxyl groups, are cleaved readily by Tsp, whereas otherwise identical peptides with amidated C-termini are cleaved very poorly. The finding that C-terminal determinants are important in determining cleavage susceptibility in several unrelated sets of proteins and peptides strongly suggests that this is a general mechanism by which Tsp discriminates among potential substrates.

How does the chemical nature of a protein's or peptide's C-terminal region influence whether it will be a good substrate for Tsp? Because Tsp is an endoprotease with some cleavage sites far from the C-terminus, it is unlikely that the C-terminal region of a substrate directly affects the chemistry of catalysis. There are several plausible models by which the C-terminal sequence of a substrate could affect Tsp cleavage. First, C-terminal sequences that are not strongly polar or charged could act as a site to which Tsp can initially bind to the substrate, with the α -carboxyl group forming part of this binding site. Such binding might be important simply for tethering Tsp to the substrate but might also be required to activate the protease in some fashion. Activation does not occur in trans because neither the #105 variant of the N-terminal domain of λ repressor nor a peptide corresponding to the last nine residues of #105 stimulates

cleavage of the wild-type N-terminal domain of λ repressor (data not shown). We note that the tail sequences of Arc and the N-terminal domain of λ repressor are partially disordered and accessible in the native structures of these proteins (Weiss et al., 1987; Breg et al., 1990), which would allow them to act as binding sites for Tsp even in the native proteins. Second, the tail sequences could act indirectly by promoting unfolding or other conformational changes in the substrate that result in increased susceptibility of the substrate to Tsp cleavage. This model, however, is not consistent with experiments showing that Tsp-sensitive variants with hydrophobic tails have stabilities and conformations comparable to Tsp-insensitive proteins with hydrophilic tails in both the λ repressor and Arc proteins (Parsell & Sauer, 1989; Milla et al., 1993). Moreover, although Tsp can discriminate between variants of λ repressor or Arc differing only at their C-termini, nonspecific proteases cleave variants with hydrophobic or hydrophilic tails equally well (Silber, 1992). The peptide pairs examined here (BAS7/BAS9 and BAS8/BAS10) have little stable structure as assayed by CD, but their spectra are identical, indicating that amidation of the terminal carboxylate in peptides BAS7 and BAS8 does not lead to significant conformational changes. Because there is no evidence that apolar C-terminal sequences alter substrate conformation or stability, we favor the idea that these sequences act as binding sites for Tsp.

Mapping of the sites of Tsp cleavage in the primary structures of several protein and peptide substrates reveals a limited number of discrete sites of proteolysis that occur throughout the substrate sequences. Tsp cleaves peptide bonds with Ala, Ser, or Val (and to a lesser extent Ile, Leu, Lys, or Arg) at the P1 position and with these same side chains plus Met, Tyr, or Trp at the P1' position. These specificity preferences appear to be sufficiently lax that most proteins and peptides would be expected to contain some cleavage sites. We note, however, that many of the substrates contain peptide bonds that fit these broad rules but are not sites at which cleavage is observed. For example, peptide BAS10 is cleaved at the Ala²-Arg³ bond but not at the Ala⁷-Arg⁸ or Ala¹²-Arg¹³ bonds. These latter bonds may not be cleaved because of additional side-chain preferences at adjoining residues (P2, P2', etc.) or because other factors are operative. For example, if Tsp contains a site for binding the C-terminus of the substrate that is distinct from the active site, then the distances between these sites may introduce a requirement for some minimal distance between the C-terminus of the substrate and a cleavage site.

We find that an Arc repressor variant that is less stable than wild type is cleaved more rapidly and a variant that is more stable than wild type is cleaved more slowly. These results are consistent with a model in which denaturation of substrate is required prior to cleavage. Indeed, the positions of the cleavage sites in Arc and the N-terminal domain of λ repressor map to many different locations in the native secondary and tertiary structure, including the hydrophobic core, surface loops, unstructured regions, and the middle of α -helices. All of these sites cannot be recognized in the context of the native structures of Arc or the N-terminal domain of λ repressor because many of the scissile peptide bonds would be inaccessible to the protease in the native proteins. If global rather than local denaturation of protein substrates is a prerequisite for cleavage, then the global unfolding rate constant would need to be greater than or equal to k_{cat} for Tsp cleavage ($>0.2 \text{ s}^{-1}$ for Arc at 37 °C). In

Fig. 6. Tsp cleavage of wild-type Arc and stability variants assayed by loss of CD ellipticity at 222 nm. Melting temperatures of PL8, Arc, and NK29 are approximately 70, 50, and 37 °C, respectively.

fact, the unfolding rate constant for Arc is $\sim 3 \text{ s}^{-1}$ at 37°C (C. Waldburger, pers. comm.) and thus global unfolding of Arc could precede cleavage.

Kinetic analysis of substrate cleavage by Tsp is complicated because there are often multiple cleavage sites, and the proteolytic products may themselves be substrates for further cleavage or even inhibitors. For these reasons, only apparent values of kinetic parameters can be measured. Cleavage of Arc by Tsp at 37°C and pH 8 is relatively slow (apparent $k_{\text{cat}}/K_m = 3.8 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$) but is similar in rate to cleavage of the same substrate by chymotrypsin (apparent $k_{\text{cat}}/K_m = 4.4 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$). The cleavage of peptide BAS9 ($k_{\text{cat}}/K_m = 1.1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$) by Tsp is almost 30-fold more efficient. We presume that BAS9 is a better substrate for Tsp than Arc because it is largely unfolded at 37°C (B. Schulman, pers. comm.), although other factors may also contribute. For comparison, cleavage of pentapeptide substrates by elastase have k_{cat}/K_m values of approximately $2 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ (Thomson & Blout, 1973; Bauer et al., 1976). Thus, Tsp does not appear to be any less efficient in cleaving susceptible protein and peptide substrates than well-characterized enzymes such as chymotrypsin and elastase, but it is far more specific in its choice of substrates.

A protease such as Tsp almost certainly comes in contact with a wide variety of different proteins in the periplasm of the bacterial cell and must be able to discriminate reliably between appropriate and inappropriate substrates. Our studies suggest that the chemical nature and structural accessibility of the C-terminal sequence of potential substrates play a central role in determining whether such proteins are cleaved by Tsp. Proteins with polar C-terminal residues or with C-terminal sequences that are buried within the native structure should not be cleaved by Tsp. Similarly, even if a protein has an accessible and apolar C-terminus that allowed transient Tsp binding, it could probably still escape cleavage if its native structure were sufficiently stable. By using a combination of specificity determinants for substrate recognition, Tsp is apparently able to avoid degrading the wrong proteins in the cell.

Materials and methods

Proteins and peptides

Wild-type Tsp was purified from *E. coli* strain X90 containing the Tsp overproducing plasmid pKS6-1w (Silber et al., 1992). Cells were grown in 6–12 L of LB broth at 37°C with gentle shaking to an absorbance of 0.8 at 600 nm and induced by addition of 1 mM IPTG. The culture was grown for an additional 3 h, harvested by centrifugation, and resuspended in 10 volumes of a spheroplasting buffer containing 100 mM Tris-HCl, pH 8, 5 mM EDTA, 500 mM sucrose. After resuspension, the cells were collected again by centrifugation, resuspended in 10 volumes of cold water to burst the outer membrane and release the periplasmic proteins, and the cytoplasmic and membrane portions were removed by centrifugation. The soluble periplasmic fraction was chromatographed on a 20-mL DEAE Sephacel column equilibrated in 10 mM Tris-HCl, pH 8, 100 mM KCl. The flow-through fraction was collected, concentrated to 0.5 mL by ultrafiltration, and chromatographed on a 24-mL Superose 6 column in 10 mM Tris-HCl, pH 8, 100 mM KCl, using a Pharmacia FPLC apparatus. Fractions containing Tsp at greater than

95% purity, as assayed by SDS-PAGE, were pooled. Tsp was stored in column buffer at 0°C with no loss of activity over a period of several months.

A gene encoding a variant of Tsp (TspH₆) containing six additional histidine residues at the C-terminus was constructed by subcloning the *Eco* RI–*Hind* III fragment from pKS6-1w (which contains all of the *tsp* gene with the exception of 66 bases at the 3' end of the gene) into *Eco* RI and *Xba* I cut pBluescript II KS(+) together with a *Hind* III–*Xba* I oligonucleotide cassette designed to restore the wild-type bases at the 3' end of the gene followed by the His₆ tag and normal termination codon. The structure of the resulting plasmid (pKK101) was confirmed by restriction mapping and DNA sequencing. The TspH₆ protein was purified from *E. coli* strain KS1000, which contains a deletion of the chromosomal *tsp* gene (Silber & Sauer, 1994), transformed with pKK101. Cells were grown, induced, and harvested as described above. The cell pellet was resuspended in five volumes of 100 mM Tris-HCl, pH 8, 200 mM KCl, 1 mM EDTA, 2 mM CaCl₂, 10 mM MgCl₂, 1.4 mM 2-mercaptoethanol, lysed by sonication, and the lysate was cleared by centrifugation. The supernatant was applied to a 20-mL Ni-NTA column (Qiagen) equilibrated in 100 mM NaH₂PO₄, 10 mM Tris-HCl, pH 8. Bound proteins were eluted with a 500-mL gradient formed by linear mixing of the column buffer at pH 8 and column buffer titrated to pH 4.5 with HCl. Fractions containing the TspH₆ protein, which eluted between pH 5.3 and pH 4.8, were pooled, concentrated to 0.5 mL by ultrafiltration, and chromatographed on a 24-mL Superdex 75 column equilibrated in 10 mM Tris-HCl, pH 8, 100 mM KCl. Fractions containing TspH₆ at greater than 95% purity were pooled.

The Arc-st7 protein differs from wild-type Arc in having 11 additional C-terminal residues (HHHHHHGRIGA). The *arc-st7* gene was constructed by synthesis of an appropriate *Hind* III–*Cla* I cassette that was inserted between the *Hind* III and *Cla* I sites of plasmid pSA600 (Milla et al., 1993). Arc repressor, the Arc-st5, Arc-st6, and Arc-st7 variants, the N-terminal domain of λ repressor (residues 1–102), and the #105 variant of the N-terminal domain were purified by ion-exchange chromatography, affinity chromatography, and gel filtration as described (Lim & Sauer, 1991; Milla et al., 1993). The NK29 and PL8 mutants of Arc were gifts of Marcos Milla. Peptides BAS7–BAS10 were the generous gift of Brenda Schulman. The expected molecular weights of all peptides were confirmed by mass spectrometry.

Characterization of Tsp

The CD spectra of Tsp and TspH₆ (50 $\mu\text{g/mL}$) were recorded at 5°C in CD buffer (50 mM phosphate, pH 7.0, 100 mM KCl) using an Aviv model 60DS circular dichroism spectrometer. The α -helical content was estimated using the program Varselect (Manavalan & Johnson, 1987). Sedimentation equilibrium centrifugation was performed in 10 mM Tris-HCl, pH 8, 100 mM KCl using a Beckman XL-A analytical ultracentrifuge. Molecular weight values were determined for 0.5 μM , 1.7 μM , 3.4 μM , and 5 μM TspH₆ at rotor speeds of 15,000 rpm and 13,000 rpm by nonlinear least-squares fitting of the data to a single species molecular weight function ($C(r) = C(a)\exp[\omega^2 M(1 - \nu\rho)(r^2 - a^2)/2RT]$) (Laue et al., 1992) using the program Nonlin (Johnson & Frazier, 1985; Brenstein, 1989).

Assays for proteolysis

For assays of Tsp or TspH₆ cleavage of protein substrates by gel electrophoresis, 0.5 µg of protease and 2 µg of substrate in 30 µL of reaction buffer (10 mM Tris-HCl, pH 8.0, 200 mM KCl) were incubated at 37 °C for 4 h and the reaction products were electrophoresed on 15% polyacrylamide-SDS-Tris-tricine gels (Schagger & von Jagow, 1987). Cleavage of peptides by TspH₆ was monitored by HPLC on a reverse-phase column (gradient from 0 to 100% acetonitrile in 0.1% TFA), following incubation of 500 µg peptide with 3.4 µg TspH₆ in 1 mL of reaction buffer for 5 h at 37 °C. The extent of cleavage was quantified by loss of area in the peak corresponding to the intact peptide and by appearance of additional peaks in the chromatograph. Reverse-phase column chromatography was also used to monitor the time course of Tsp cleavage of protein substrates such as Arc and the #105 variant of the N-terminal domain of λ repressor. For these assays, 10 µg of substrate was mixed with 2 µg TspH₆ in 300 µL of reaction buffer at 37 °C, and 50-µL aliquots were removed from the reaction at 20-min intervals, quenched with 0.1% TFA, and analyzed by reverse-phase chromatography as described above.

Kinetic constants for cleavage of Arc repressor and peptide BAS9 were determined by mixing 50 nM TspH₆ with 2.5–220 µM Arc repressor or 1.6–150 µM BAS9 in reaction buffer at 37 °C and monitoring changes in CD ellipticity at 222 nm. The rate of cleavage in mdeg/s was determined from a linear fit of the ellipticity versus time data and converted to units of µM substrate cleaved/s/µM enzyme assuming that cleaved substrate has no ellipticity at 222 nm. K_m and k_{cat} values for Arc and BAS9 were determined by fitting the rate versus substrate concentration data to the hyperbolic function $v = k_{cat}[E_t]\{[S]/(K_m + [S])\}$ using the program Nonlin. The relative rates of cleavage of Arc repressor variants were determined by mixing 45 nM Tsp with Arc, Arc-PL8, or Arc-NK29 at concentrations of 8 µM, and monitoring cleavage using the CD assay described above.

Mass spectrometry

The peptides produced by Tsp cleavage of λ #105, Arc, Arc-st7, and the peptide BAS10 were analyzed by MALDI-TOF mass spectrometry (Hillenkamp et al., 1991). Typically 50–150 nmol of substrate were digested with 100 pmol of Tsp or TspH₆ for 1–6 h. In some cases, the digestion products were partially purified by reverse-phase chromatography or nickel-NTA affinity chromatography prior to mass spectrometric analysis. A sample of each digest was diluted with a 10 mM solution of 2,5-dihydroxybenzoic acid or 3,5-dimethoxy-4-hydroxycinnamic acid to a concentration of approximately 10 µM. One microliter of each matrix/analyte solution was deposited and dried on the sample probe of a VESTEC 2000 linear time-of-flight mass spectrometer (PerSeptive Biosystems). Samples were irradiated at 337 nm with a nitrogen laser (model VSL337ND, Laser Science Inc.). Measured m/z values for the $[M+H]^+$ ions detected were matched with segments of the protein using a computer program (COMPOST) that has been previously described (Papayannopoulos & Biemann, 1991). In cases where more than one peptide from a given substrate could correspond to the measured mass, sequence assignments were made based on the consistency of the peptide termini with those from other peptide digestion products. In this regard, peptides formed through a single cleavage

of the intact protein were favored over those requiring two cleavages. In several cases, sequence assignments were confirmed independently from the collision-induced dissociation mass spectra of the peptide products (see Biemann, 1990) or by N-terminal sequencing using the Edman degradation.

Acknowledgments

We thank Bronwen Brown, Brenda Schulman, and Marcos Milla for gifts of proteins and peptides. We also thank Ray Larsen and Kathleen Postle for communicating unpublished data. This study was supported by NIH grants AI-15706 and RR-00317.

References

- Anbudurai PR, Mor TS, Ohad I, Shestakov SV, Pakrasi HB. 1994. The CtpA gene encodes the C-terminal processing protease for the D1 protein of the photosystem II reaction center complex. *Proc Natl Acad Sci USA* 91:8082–8086.
- Bauer CA, Thompson RC, Blout ER. 1976. The active centers of *Streptomyces griseus* protease 3, alpha-chymotrypsin, and elastase: Enzyme-substrate interactions close to the scissile bond. *Biochemistry* 15:1296–1299.
- Beynon RJ, Bond JS, eds. 1989. *Proteolytic enzymes: A practical approach*. Oxford, UK: IRL Press.
- Biemann K. 1990. Sequencing of peptides by tandem mass spectrometry and high-energy collision-induced dissociation. *Methods Enzymol* 193:455–479.
- Bowie JU, Sauer RT. 1989. Identification of C-terminal extensions that protect proteins from intracellular proteolysis. *J Biol Chem* 264:7596–7602.
- Breg JN, van Opheusden JHJ, Burgering MJM, Boelens R, Kaptein R. 1990. Structure of Arc repressor in solution: Evidence for a family of beta-sheet DNA-binding proteins. *Nature* 346:586–589.
- Brenstein RJ. 1989. *Nonlin for Macintosh*. Carbondale, Illinois: Robelko Software.
- Goldberg AL. 1992. The mechanisms and functions of ATP-dependent proteases in bacterial and animal cells. *Eur J Biochem* 203:9–23.
- Gottesman S, Maurizi MR. 1992. Regulation by proteolysis: Energy-dependent proteases and their targets. *Microbiol Rev* 56:592–621.
- Hara H, Yamamoto Y, Higashitani A, Suzuki H, Nishimura Y. 1991. Cloning, mapping, and characterization of the *Escherichia coli* *prc* gene, which is involved in C-terminal processing of penicillin-binding protein 3. *J Bacteriol* 173:4799–4813.
- Hershko A, Ciechanover A. 1992. The ubiquitin system for protein degradation. *Annu Rev Biochem* 61:761–807.
- Hillenkamp F, Karas M, Beavis RC, Chait BT. 1991. Matrix-assisted desorption/ionization mass spectrometry of biopolymers. *Anal Chem* 63:1193A–1203A.
- Johnson ML, Frazier S. 1985. Nonlinear least-squares analysis. *Methods Enzymol* 117:301–342.
- Laue TM, Shah BD, Ridgeway TM, Pelletier SL. 1992. Computer-aided interpretation of analytical sedimentation data for proteins. In: Harding S, Rowe A, Horton J, eds. *Analytical ultracentrifugation in biochemistry and polymer science*. Cambridge, UK: Royal Society of Chemistry. pp 90–125.
- Lim WA, Sauer RT. 1991. The role of internal packing interactions in determining the structure and stability of a protein. *J Mol Biol* 219:359–376.
- Manavalan P, Johnson WC Jr. 1987. CD secondary structure analysis using variable selection. *Anal Biochem* 167:76–85.
- Milla ME, Brown BM, Sauer RT. 1993. P22 Arc repressor: Enhanced expression of unstable mutants by addition of polar C-terminal sequences. *Protein Sci* 2:2198–2205.
- Nagasawa H, Sakagami Y, Suzuki A, Suzuki H, Hara H, Hirota Y. 1989. Determination of the cleavage site involved in C-terminal processing of penicillin-binding protein 3 of *Escherichia coli*. *J Bacteriol* 171:5890–5893.
- Pabo CO, Lewis M. 1982. The operator-binding domain of lambda repressor: Structure and DNA recognition. *Nature* 298:443–447.
- Papayannopoulos IA, Biemann K. 1991. A computer program (COMPOST) for predicting mass spectrometric information from known amino acid sequences. *J Am Soc Mass Spectrom* 2:174–177.
- Park JH, Lee YS, Chung CH, Goldberg AL. 1988. Purification and characterization of protease Re, a cytoplasmic endoprotease in *Escherichia coli*. *J Bacteriol* 170:921–926.

- Parsell DA, Sauer RT. 1989. The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. *J Biol Chem* 264:7590-7595.
- Parsell DA, Silber KR, Sauer RT. 1990. Carboxy-terminal determinants of intracellular protein degradation. *Genes & Dev* 4:277-286.
- Roseman JE, Levine RL. 1987. Purification of a protease from *Escherichia coli* with specificity for oxidized glutamine synthetase. *J Biol Chem* 262:2101-2110.
- Schagger H, von Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* 166:368-379.
- Silber KR. 1992. The role of carboxyl termini in selective protein degradation [thesis]. Cambridge, Massachusetts: Massachusetts Institute of Technology.
- Silber KR, Keiler KC, Sauer RT. 1992. Tsp: A tail-specific protease that selectively degrades proteins with nonpolar C termini. *Proc Natl Acad Sci USA* 89:295-299.
- Silber KR, Sauer RT. 1994. Deletion of the *prc* (*tsp*) gene provides evidence for additional tail-specific proteolytic activity in *Escherichia coli* K-12. *Mol Gen Genet* 242:237-240.
- Thomson RC, Blout ER. 1973. Dependence of the kinetic parameters for the elastase-catalyzed amide hydrolysis on the length of peptide substrates. *Biochemistry* 12:57-65.
- Tobias JW, Shrader TE, Rocap G, Varshvsky A. 1991. The N-end rule in bacteria. *Science* 254:1374-1377.
- Weiss MA, Karplus M, Sauer RT. 1987. ¹H NMR aromatic spectrum of the operator binding domain of the lambda repressor: Resonance alignment with application to structure and dynamics. *Biochemistry* 26:890-897.