

Tsp: A tail-specific protease that selectively degrades proteins with nonpolar C termini

(*Escherichia coli*/hydrophobic tail/ λ repressor/protein sequence homology)

KAREN R. SILBER, KENNETH C. KEILER, AND ROBERT T. SAUER

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

Communicated by H. Gobind Khorana, September 27, 1991

ABSTRACT An *Escherichia coli* protease designated Tsp (tail-specific protease) has been purified, and its gene has been cloned and sequenced. Tsp specifically degrades a variant of the N-terminal domain of λ repressor in which the five C-terminal residues, which are polar in wild type, have been replaced by nonpolar residues. This substrate specificity *in vitro* parallels the previously reported selective degradation *in vivo* of N-terminal-domain variants with nonpolar C-terminal residues. The gene sequence and N-terminal protein sequence of Tsp predict a protein of 660 amino acids. The deduced protein sequence of Tsp shows no significant homology to known protease sequences but does show sequence similarity to the human and bovine interphotoreceptor retinoid-binding proteins, which bind hydrophobic ligands.

Within cells, different proteins are degraded with half-lives ranging from minutes to days (1). This heterogeneity implies that intracellular degradation is selective, but the mechanisms of this selectivity are incompletely understood. The C-terminal sequences of some proteins can have dramatic effects on their rates of degradation in *Escherichia coli* (2, 3). For example, the wild-type N-terminal domain of λ repressor (residues 1–102) has the polar C-terminal amino acid sequence RSEYE¹⁰² and has a half-life *in vivo* of >10 hr. The rapidly degraded “#105” variant of this protein has the hydrophobic C-terminal sequence WVAAA¹⁰² and has a half-life of 15 min (3). To identify possible cellular component(s) responsible for the degradation of proteins with destabilizing C termini, we have purified an *E. coli* activity that specifically degrades the #105 variant but not the wild-type N-terminal domain. This activity resides in a single, purified polypeptide (Tsp, for tail-specific protease). We have cloned and sequenced the *tsp* gene. * Based on sequence comparisons, Tsp shows similarities to the human and bovine interphotoreceptor retinoid-binding proteins (IRBPs) but does not resemble any proteases in the protein sequence data base.

MATERIALS AND METHODS

Assays for Tsp Activity. Tsp activity in crude lysates and column fractions was assayed by measuring the degradation of ³⁵S-labeled substrates (either the #105 variant or the wild-type N-terminal domain of λ repressor) to products soluble in 10% (wt/vol) trichloroacetic acid (3). Reaction mixtures contained \approx 9000 cpm of ³⁵S-labeled substrate, a sample of the *E. coli* fraction in lysis buffer or column buffer, and 0.02% Nonidet P-40 in a volume of 50 μ l. Reaction mixtures were incubated at 37°C for 1 or 2 hr and were processed as described (3).

In more purified fractions, degradation was assayed by the appearance of substrate digestion products following SDS/PAGE. Wild-type or #105 protein (2 μ g), which was purified by the method of Lim and Sauer (4), was mixed with a sample in a 10- μ l reaction volume. Reaction mixtures were incubated at 37°C for 3–6 hr and were stopped by boiling for 3 min in Laemmli sample buffer (5). Samples were electrophoresed in SDS/16.5% polyacrylamide gels with the Tricine buffer system (6), and the gels were stained with Coomassie blue R250.

Purification of Tsp from Strain X90. *E. coli* strain X90 [*ara* Δ (*lac-pro*) *nalA argEam rif thi-1/F' lacI^Q lac⁺ pro⁺*] cells were grown at 37°C to an OD₆₀₀ of \approx 2.0 in 24.5 liters of LB broth. The cells were harvested and were lysed by sonication. Nucleic acids and membranes were precipitated with 0.1% polyethyleneimine, the preparation was centrifuged at 16,000 \times g for 30 min, and a 40–80% (saturation) ammonium sulfate cut of the supernatant was collected.

The 40–80% pellet was resuspended in 100 mM sodium phosphate (pH 7), adjusted to 1 M ammonium sulfate, and loaded onto a phenyl-Sepharose CL-4B column (Pharmacia). The column was washed with loading buffer and developed with a linear gradient from 1 M to 0 M ammonium sulfate in 100 mM sodium phosphate (pH 7). In some purifications, Tsp activity was eluted at \approx 0.5 M ammonium sulfate; in others, Tsp was eluted throughout the column profile. Fractions containing Tsp activity were pooled, and protein was precipitated with ammonium sulfate. The pellet was resuspended in 10 mM Tris (pH 8.2) and dialyzed against the same buffer. This material was loaded onto a DEAE-Sephacel column (Pharmacia), which was washed with 10 mM Tris (pH 8.2). Tsp activity was eluted in the flow-through and wash fractions, which were combined and loaded onto a Bio-Rex 70 column (Bio-Rad). The column was developed with a linear gradient from 25 to 500 mM KCl. Tsp activity was eluted at \approx 250 mM KCl, and fractions containing this activity were pooled. As a final purification step, portions of the Bio-Rex 70 pool were concentrated in an Amicon Centricon 30 microconcentrator and were electrophoresed in a 7.5% Laemmli gel (5). The gel was stained with KCl, and the \approx 82-kDa band was excised and electroeluted from the gel slices.

Protein Sequencing. Protein sequencing was performed by sequential Edman degradation using an automated gas-phase sequenator. Aliquots of Bio-Rex 70 column fractions containing Tsp were precipitated with 10% trichloroacetic acid and electrophoresed in a 7.5% Laemmli gel. Protein was electroblotted from the gel onto ProBlott membrane (Applied Biosystems), and the Coomassie-stained Tsp band was excised and sequenced. To obtain internal amino acid sequence, Tsp was electrophoresed as described above, the protein was blotted onto nitrocellulose, and the Ponceau

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IPTG, isopropyl β -D-thiogalactopyranoside; IRBP, interphotoreceptor retinoid-binding protein.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M75634).

S-stained Tsp bands were excised. Tsp was digested with trypsin, fragments were separated by HPLC on a C₁₈ column, and a 26-amino acid tryptic peptide was sequenced.

Cloning and Sequencing of *tsp*. Two codon-biased oligonucleotide probes, T and N, were designed to encode the 26 amino acids of the sequenced tryptic peptide and the 15 N-terminal residues of Tsp, respectively. Probe T was labeled with [γ -³²P]ATP and T4 polynucleotide kinase and was hybridized to a nylon membrane containing plaque-lifted DNA from the Kohara phage library (7) as described (8). Overlapping Kohara phage 335 (15D5) and 336 (19H3) showed strong hybridization to the probe.

Kohara phage 335 DNA ($\approx 0.5 \mu\text{g}$) was digested to completion with individual restriction enzymes, electrophoresed in a 0.7% agarose gel, and transferred to a GeneScreenPlus nylon membrane (NEN). The 3-kilobase *EcoRI* fragment showed strong hybridization to both ³²P-labeled probes T and N (8). This fragment, which contains the *tsp* gene, was gel-purified from agarose with GeneClean II (Bio 101, La Jolla, CA) and ligated to pBluescript KS(+) (Stratagene) that had been linearized with *EcoRI*. Plasmids pKS6-1w and pKS7-1a contain the *tsp* gene running with and against the *lac* promoter of pBluescript KS(+), respectively. DNA sequencing (9) of both strands of the *EcoRI* insert, except for regions encoding the *htpX* gene (10), was performed with the single-stranded forms of pKS6-1w and pKS7-1a.

Purification of Tsp from an Overproducing Strain. X90/pKS6-1w cells were grown at 37°C to an OD₆₀₀ of ≈ 1.5 in 10 liters of LB broth, and Tsp expression was induced by the addition of isopropyl β -D-thiogalactopyranoside (IPTG) to 1 mM. After an additional 3 hr, the cells were harvested and lysed by sonication. The sonicate was centrifuged at 37,000 $\times g$ for 20 min to remove cell debris, dialyzed against 10 mM Tris (pH 8.2), and loaded onto a DEAE-Sephacel column. The column was washed with 10 mM Tris (pH 8.2) and developed with steps containing 50 mM, 100 mM, and 150 mM KCl. Fractions containing Tsp were identified by gel electrophoresis, pooled, dialyzed against 10 mM Tris, pH 8.2/25 mM KCl, and loaded onto a Bio-Rex 70 column. The column was washed with 10 mM Tris, pH 8.2/25 mM KCl and developed with a linear gradient from 25 mM to 500 mM KCl. Fractions containing Tsp activity were pooled, concentrated 100-fold, and chromatographed on a Superose 6 FPLC column. Tsp was eluted in a single, symmetric peak. For molecular weight determination, Tsp (10 μg) was chromatographed on a Sephacryl S-200 column.

Inhibition Assays. Tsp (0.2 μg) was incubated for 20 min at room temperature with an inhibitor or metal cation in 46.5 μl of buffer containing 10 mM Tris (pH 8.0) and 10 mM KCl. Unlabeled #105 (1 μg) and ³⁵S-labeled #105 (9000 cpm) were added, bringing the total volume to 50 μl . The reaction mixture was incubated at 37°C for 30 min, and the trichloroacetic acid-soluble fraction was prepared as described (3). The activity was calculated relative to a control of solvent (water or ethanol) alone.

Tsp Induction and Osmotic Shock Procedure. X90/pKS6-1w cells were grown in LB broth at 37°C to an OD₆₀₀ of ≈ 1.0 , and *lac* promoter-mediated transcription was induced in one of two cultures by adding IPTG to 1 mM. After an additional 2.5 hr, cells were harvested and osmotic shock fractions (periplasmic fractions) were prepared essentially as described (11) for cells in late exponential phase. The remaining cytoplasmic/membrane fraction and the initial whole cell preparation were boiled in Laemmli sample buffer prior to electrophoresis of the fractions in 7.5% Laemmli gels.

Pulse-Chase Experiment. X90/pKS6-1w cells were grown as described (3) and were induced with IPTG for 20 min. Cells were pulse-labeled for 15 sec with L-[³⁵S]methionine, excess unlabeled L-methionine was added, and samples were taken

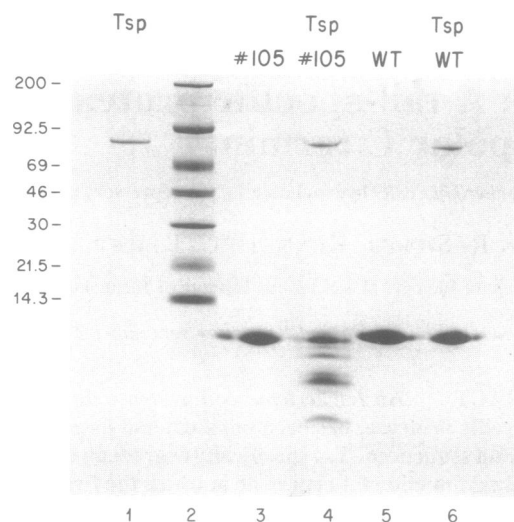


FIG. 1. Specific degradation of the #105 variant of the N-terminal domain of λ repressor by Tsp. Reaction mixtures were incubated at 37°C for 6 hr, and cleavage of the #105 or wild-type (WT) protein was assayed by gel electrophoresis. Lane 1, Tsp purified from strain X90; lane 2, protein standards (molecular mass in kilodaltons indicated at left); lanes 3 and 5, control reactions of #105 or WT (no Tsp added); lanes 4 and 6, reactions of #105 or WT with Tsp.

at 0, 20, 50, 120, 270, and 600 sec of the chase period. Lysates were prepared and electrophoresed in a 7.5% Laemmli gel.

Sequence Comparisons. In data-base searches, the BLAST (12) service at the National Center for Biotechnology Information was used to search for proteins with similarities to the Tsp sequence. The BESTFIT program of the Genetics Computer Group's sequence-analysis software package (13) was used to compare the Tsp sequence with protease active-site sequences (14).

RESULTS

Identification of Tsp and Cloning of the *tsp* Gene. A proteolytic activity that degrades the #105 variant of the N-terminal domain of λ repressor was purified from *E. coli* X90. This activity was designated Tsp, for tail-specific protease. As shown in Fig. 1, Tsp is specific in that it cleaves the #105 protein into multiple fragments (lane 4) but does not cleave the wild-type protein (lane 6). The degradation of #105 does not require addition of ATP, metal ions, or other obvious cofactors.

In SDS/polyacrylamide gels, Tsp appears to correspond to a single polypeptide of ≈ 82 kDa (Fig. 1). Sequential Edman degradation of intact Tsp gave the sequence NH₂-VEDITRADQIPVLKE. Sequencing of an internal tryptic fragment of Tsp yielded the sequence NH₂-VGVLDPG-FYVGLTDDVKVQLQKLEK-COOH. These two sequences were used to design oligonucleotide probes that were used to screen a Kohara phage library (7) and to locate the *tsp* gene on the phage clones. A 3-kilobase *EcoRI* fragment of Kohara phage 335, which maps to ≈ 40.3 min on the *E. coli* chromosome, contained the *tsp* gene.

DNA Sequence of *tsp*. The DNA sequence of the *tsp* gene and its flanking regions was determined (Fig. 2). The N-terminal sequence of Tsp is encoded by the DNA sequence starting at nucleotide 135. From this point, an open reading frame extends for 660 codons before terminating with a TAA codon at base 2115. This open reading frame also codes for the 26-residue sequence determined for the internal tryptic fragment of Tsp. The molecular mass calculated for the 660-residue protein is 74,323 Da. This is in reasonable agreement with the ≈ 82 -kDa value estimated from SDS/PAGE.

(M) (N) (M) (F) (F) (R) (L) (T) (A) (L) (A) (G) (L) (L) (A) (I) (A) (G)

1 GAATTCGGGTATGCTTTTGGATGTCGCGCCAGAACACCTGGTGTCTGAAACGGAGGCGCGGCCAGGCATGAACATGTTTTTTAGGCTTACCGCGTTAGCTGGCTGCTGCAATAGCA
 (Q) (T) (F) (A) V E D I T R A D Q I P V L K E E T Q H A T V S E R V T S R F T R S H Y R
 121 GCGAGACCTTCGCTGTAGAGATATCACCGCTGCTGATCAAAATTCGCGTATAAAGGAAAGAGAGCGAGCATGCGACCGTAAGTGAGCGGTAACTGCGCGCTTACCGCTTCTCATTATC
 Q F D L D Q A F S A K I F D R Y L N L L D Y S H N V L L A S D V E Q F A K K K
 241 GCGAGTTCGACCTCGATCAGCAATTTTCGCGCAAAATCTTTGACCGCTACCTGGAATCTGCTGATTAACGACCAACCGTGTCTGGCAAGCGATGTGAACAGTTTCGCGAAAGAAAA
 E L G D E L R S G K L D V F Y D L Y N L A Q K R R F E R Y Q Y A L S V L E K P M
 361 CCGAGTTAGCGGATGAATCGCTTCAGGCAACTCGAGCTTCTACGATCTGCGCGCAAAAGCGCCGTTTTGAGCGTACCAGTACCGTTTGTGCGTACTGGAAAGCCGA
 D F T G N D Y N L D R S K A P W P K N E A E L N A L W D S K V K F D E L S L K
 481 TGGATTTACCGGCAACGACACTTATAACCTTGACCGCAGCAAGCGCCCTGGCCGAAAAACGAGGCTGAGTTGAACCGCGCTGTGGACAGTAAAGTCAAATTCGACGAGTTAAGCCCTGA
 L T G K T D K E I R E T L T R R Y K F A I R R L A Q T N S E D V F S L A M T A F
 601 AGCTGACAGAAAAACGGATAAAGAAATTCGTGAACCTGACTCGCGCTACAATTTGCCATTCGTCTGCGCGCAAAACACAGCGAAGATTTTTCTCGCTGGCAATGACGGCGT
 A R E I D P H T N Y L S P R N T E Q F N T E M S L S L E G I G A V L Q M D D D Y
 721 TTGCGGTGAAATCGACCGCATACCACTATCTTCCCGCTAATACCGAACAGTTCACACTGAAATGATTTGTGCTGGAGGTATGGCGCAGTCTGCAATGGATGATGACT
 T V I N S M V A G G P A A K S K A I S V G D K I V G V G Q T G K P M V D V I G W
 841 ACACCGTTATCAATTCGATGGTGGCAGTGGTCCCGCAGCAAGGTAAGCTATCAGCGTTGGTGACAAATTTGCGGTGTGGTCAACAGCGAAGCGGATGGTTCAGTGTGGT
 R L D D V V A L I K G P K G S K V R L P A G K G T R T V L T R E R I
 961 GCGGCTTCGATGATGTTGCTTAAATTAAGGGCCGAAAGCGGAGTAAAGTTTCGTGGAATTTTACCTGCTGTAAAGGACCAAGACCCGCTACTGTACGTTGACCGGTGACGCTA
 R L E D R A V K M S V K T V G K E K A V G V L D I P G F Y V G L T D D D V K V Q L Q
 1081 TTCGTCGAAAGACCGCGGTAAATGTCGCGTGAAGACCGTGGTAAAGAGGATCGCGCGTGTGGATATTCGCGGCTTCTATGTTGGTTTGAACAGCATGTCAAAGTGAACACTGC
 K L E K Q N V S S V I I D L R S N G G A L T E A V S L S G L F I P A G G P I V Q
 1201 AGAACTGGAAAAACAGAAATGTCAGCAGCTCATCATCGACTCGTAGCAATGGGGGTTAACTGAAGCGGTATCGCTCTCGGCTCTGTTTATCTCCGCGTCCATTTGTT
 V R D N N G K V R E D S D T D G Q V F Y K A G P L V L V D R F S A S A S E I F A
 1321 AGTTCGCGGATAAACAGCGCAAGTTGCGTAGATAGCGATACCGGACAGGTTTCTATAAAGCCCGCTGGTGGTCTGGTTCAGCCGCTTCAGTCTTCAGAAATCGATCTTTG
 A A H Q D Y G R A L V V G E P T F G K G T V Q Q Y R S L N R I Y D Q M L R P E W
 1441 CCGCGCAATTCAGGATTCGCTGCTGGTGTGGTGAACCGGTTTGGTAAAGCCACCGTTTCAGCAATACCGTTCATTGAACCGTATTACGATCAGATTTACGCTCCTGAAT
 P A L G S V Q Y T I Q P F Y R V N G S T Q R K G V T P D I I M P T G N E E T E
 1561 GGCAGCGTGGGTTCTGTGACATACAGCATCCAGAAATCTATCGCGTTAACGGCGGAGTACGCAACGTAAGCGGTAAAGCGGTAAAGCGGACAGATCATATGCGGCGGGTAAAGAAACCG
 T G E K F E D N A L P W D S I D A A T Y V K S G D L T A F E B L K E H N A R
 1681 AAACGGGTGAGAAATTCGAAGATAACCGCTCCCGTGGGATAGCATGATGCGCGGACTTATGTGAAATCAGGAGATTAAACCGGCTTTGAAGCGGAGCTGCTGAAGGAACATAATCGCC
 I A K D P E F Q N I M K D I A R F N A M K D K R N I V S L N Y A V R E K E N N E
 1801 GTATCGGAAAGATCTGAGTTCCAGAACATCATGAAGATATCGCGCTCAACGCTATGAAGACCAAGCGCAATATCGTTTCTGTAATACCTGTGGTGGAGAAAGATAATG
 D D L A R L R L N E R F K R E G K P E L K L D D L D P K D I Y Q E P D Y L D E
 1921 AAGATGATCGACCGCTGCGCGGTTTGAACGAAACGCTTTAAACCGGAAGGTAACCGGAGTTGAAGAACTGGATGATCTACGAAAGATTACCAGGACCGGATCTTATCTGGATG
 T V N I A L D L A K L E K A R P A E Q P A P V K stop
 2041 AGACGGTGAATATCGCACTCGATCTGGCAAGCTTGAAGCCAGACCGCGGAACACCGCTCCCGTCAAGTAATCAATCAGSCACAAAGAAATTTGCTGCTGATTTTTTAAACAGCG

FIG. 2. Nucleotide sequence of the *tsp* gene and predicted amino acid sequence. Numbers refer to the nucleotide sequence, beginning at the *EcoRI* site. Amino acid symbols in parentheses represent a potential presequence. The underlined sequence is a potential transcription terminator.

The start point of protein synthesis for Tsp is uncertain. Neither the GTA codon for the N-terminal valine nor the preceding GCT codon is a known translational start codon. There are ATG codons 20 and 22 codons upstream of the N-terminal valine codon. Each is preceded by a potential Shine-Dalgarno sequence (AGGN₇ATG in the first case and GGAGGN₁₁ATG in the second). Translational initiation at either of these ATG codons would give rise to sequences that resemble signal sequences (15). Based on comparison with *E. coli* promoter consensus sequences, a promoter for *tsp* transcription could not be identified. After the TAA termination codon, there is an inverted repeat followed by a run of thymidines that probably acts as a Rho-independent transcription terminator. The DNA sequence directly following this region corresponds to the sequence of the beginning of the heat shock gene *hspX* (10).

Overexpression, Purification, and Properties of Tsp. The *tsp* gene was placed under the control of the *lac* promoter in plasmid pKS6-1w. When the *lac* promoter is induced, X90 cells carrying pKS6-1w overexpress an ≈82-kDa protein (Fig. 3, compare lanes 2 and 3). Moreover, lysates of such cells show at least a 5-fold increase in degradation of ³⁵S-labeled #105 protein as compared with X90 cells carrying the plasmid vector alone. We purified the ≈82-kDa protein from X90/pKS6-1w cells (Fig. 3, lane 1). In protease assays like those shown in Fig. 1 and by SDS/PAGE, this material was indistinguishable from the Tsp that was purified in smaller quantities from X90 cells.

The effect of inhibitors and metal cations on Tsp activity is shown in Table 1. None of the inhibitors significantly affected Tsp activity. Cysteine protease inhibitors were not studied because Tsp contains no cysteine. Stimulation of Tsp activity was seen with MnCl₂, CoCl₂, and CaCl₂ and inhibition was seen with FeCl₃, CuSO₄, and ZnSO₄.

In gel filtration chromatography, Tsp has an apparent molecular mass of ≈68 kDa. Comparison of this estimate with the ≈82 kDa obtained by SDS/PAGE and the ≈74 kDa calculated from the predicted protein sequence indicates that Tsp is a monomer in solution. Because the molecular mass of Tsp determined by gel filtration was lower than the expected 74 kDa, we used laser desorption time-of-flight mass spec-

trometry to determine whether the purified Tsp might be processed by proteolytic cleavage of C-terminal residues. The molecular mass of Tsp determined by mass spectrometry was 75 kDa, indicating that such C-terminal processing is unlikely to occur.

Since the DNA sequence near the beginning of the *tsp* gene encodes a potential signal sequence, which is not present in the purified protein, Tsp may be secreted. To examine this possibility, we performed a cellular fractionation experiment. Tsp was found in both the periplasmic and cytoplasmic/membrane fractions of X90/pKS6-1w cells (Fig. 3, lanes 5 and 7). The periplasmic location of at least some of the Tsp is consistent with the presence of a signal sequence. It should be noted that these fractionation results may have been affected by Tsp overproduction and that more detailed studies are needed to examine Tsp localization. To test whether Tsp is synthesized with a signal peptide that is later removed, we performed a pulse-chase experiment (data not shown). Following 15 sec of pulse-labeling, Tsp was present as an ≈82-kDa band in an SDS/polyacrylamide gel. Precursor forms were not evident and no changes in the quantity or mobility of this band were observed during a 10-min chase. Although this result suggests that Tsp is not processed following synthesis, the possibilities that processing occurs extremely rapidly or that precursor and mature forms might have the same mobility in SDS/polyacrylamide gels cannot be excluded.

DISCUSSION

We have described the purification of an *E. coli* protease, designated Tsp, and the cloning and sequencing of its gene. Tsp degrades the #105 protein, which has a nonpolar sequence (WVAAA) at its C terminus, but does not degrade the wild-type N-terminal domain of λ repressor, which has polar residues (RSEYE) at its C terminus. The substrate specificity of Tsp *in vitro* is that expected from the studies *in vivo*. Namely, the #105 variant is degraded in the cell much more rapidly than the wild-type protein (3). At present, we cannot be sure that Tsp is responsible for the rapid degradation of #105 *in vivo*, but this seems to be the most likely possibility.

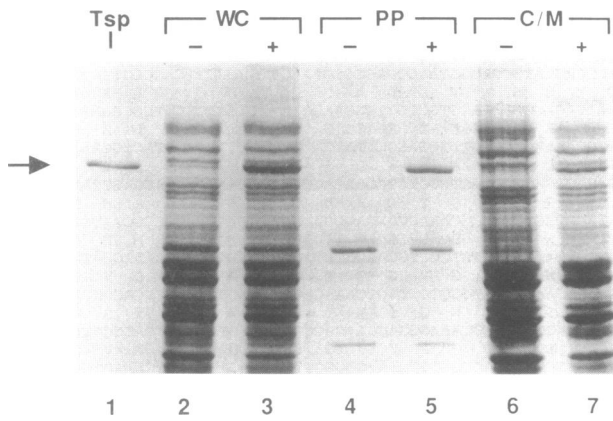


FIG. 3. Overexpression and cellular distribution of Tsp. Lane 1, 2.2 μ g of Tsp purified from X90/pKS6-1w cells. For lanes 2-7, cultures of X90/pKS6-1w cells were grown with (+) or without (-) induction of *tsp* expression. Whole cells (WC) were fractionated by osmotic shock into the periplasmic (PP) and cytoplasmic/membrane (C/M) fractions. Arrow indicates the position of the Tsp band.

If Tsp does degrade the #105 protein in the cell, then it will be important to determine where this degradation occurs, as Tsp seems to be a periplasmic protein but may also be present in the cytoplasm.

As the #105 and wild-type substrates differ only in their C-terminal five residues, these C-terminal "tails" must in some manner mediate Tsp's cleavage of #105, but not of wild type. We believe that Tsp may bind directly to the C-terminal residues of the #105 protein. These residues are known to be unstructured in the wild-type N-terminal domain (16) and thus are also likely to be unstructured and accessible to the protease in the #105 variant. The C-terminal residues of the #105 variant are unlikely to affect proteolysis indirectly—for example, by interacting with or destabilizing portions of the

Table 1. Effect of inhibitors and metal cations on Tsp activity

Inhibitors		Metal cations	
Addition*	Activity, %	Addition*	Activity, %
None	100	None	100
3,4-DCI	84	MnCl ₂	252
PMSF	98	CoCl ₂	135
TLCK	115	CaCl ₂	125
TPCK	94	MgCl ₂	103
Pepstatin	91	ZnSO ₄	53
<i>o</i> -PA	91	CuSO ₄	27
EDTA	88	FeCl ₃	20

*Concentrations are 1 mM except *o*-PA (2 mM) and EDTA (10 mM). 3,4-DCI, 3,4-dichloroisocoumarin; PMSF, phenylmethylsulfonyl fluoride; TLCK, 7-amino-1-chloro-3-tosylamido-2-hepatonone ("tosyl-L-lysine chloromethyl ketone"); TPCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone ("tosyl-L-phenylalanine chloromethyl ketone"); *o*-PA, *o*-phenanthroline.

folded protein—since thermal stabilities of the #105 protein and the wild-type N-terminal domain are identical (3). Hence, the simplest and most appealing model is that the #105 variant is a substrate for Tsp because its hydrophobic C-terminal tail provides an anchoring site to which the protease binds. By this model, the wild-type N-terminal domain is not a substrate because its C-terminal residues are not hydrophobic.

Preliminary characterization of the digestion products of the #105 variant shows that Tsp is an endoprotease. The major sites of cleavage appear to be A⁴⁹-L⁵⁰, A⁶⁶-K⁶⁷, A⁸¹-R⁸², and V⁹¹-S⁹² (unpublished work). Hence, Tsp cleaves the #105 protein at specific sites following small hydrophobic residues such as alanine and valine. Tsp could initially bind to the hydrophobic C-terminal tail of the #105 protein and then make endoproteolytic cleavages within the polypeptide chain, perhaps waiting for transient unfolding of the protein to allow access to the cleavage sites. Such a model is

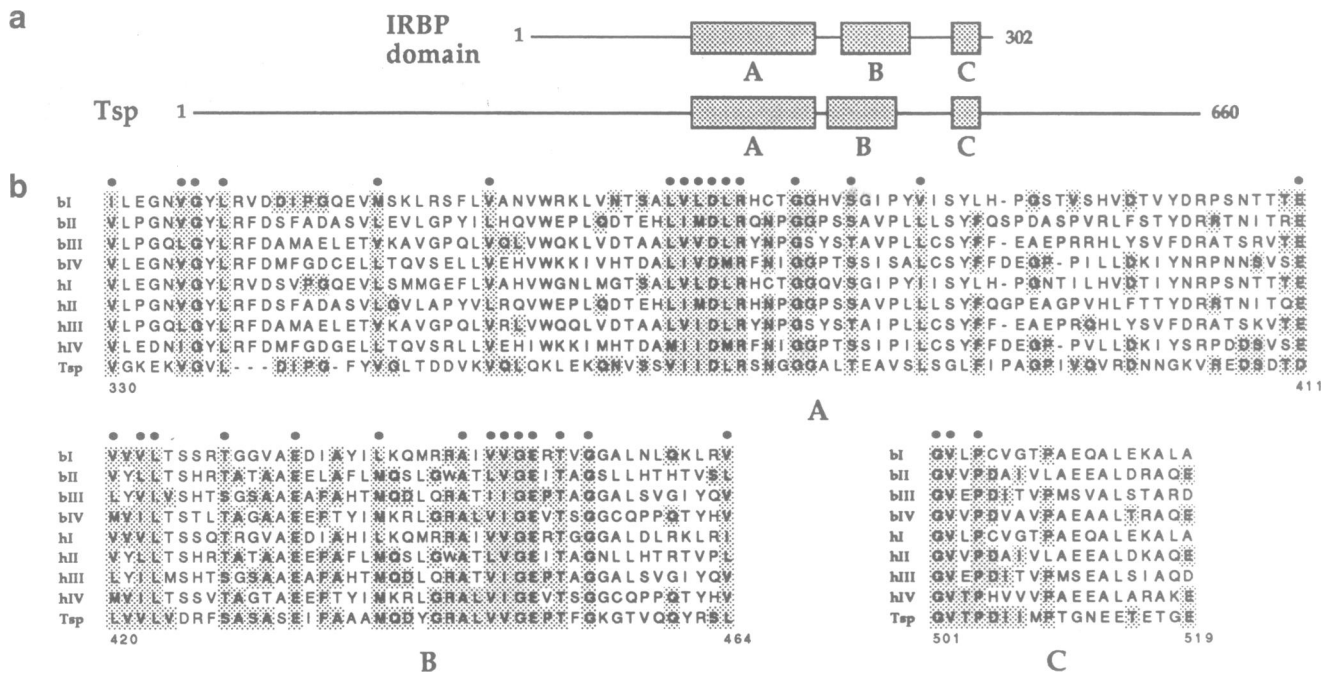


FIG. 4. Sequence homology between Tsp (660 residues) and the repeated domains of bovine and human IRBP (\approx 302 residues). (a) Schematic alignment of Tsp and an IRBP domain. (b) Alignments of Tsp with the bovine (bi-biv) and human (hi-hiv) IRBP repeats. The alignment of the IRBP domains is from ref. 17. The alignment of Tsp with the IRBP domains is based on pairwise alignments from a BLAST search. Numbers correspond to the Tsp sequence. Positions of identity between Tsp and at least one IRBP repeat, and identical or conserved positions (I/V/L/M, D/E, S/T) in all nine sequences (noted by filled circles), are shaded. To assess the significance of the similarity between Tsp and IRBP, a profile (18) of all eight IRBP domains was constructed and compared with the Tsp sequence and the sequence data base. The similarity between the IRBP profile and Tsp was 8.7 standard deviations greater than the mean similarity calculated for other sequences in the data base.

speculative and needs to be tested but predicts that for a protein to be a substrate for Tsp, several criteria (e.g., exposed nonpolar C terminus, appropriate rates of unfolding and refolding, and the presence of specific cleavage sequences) would need to be fulfilled.

In sequence searches (see *Materials and Methods*), no significant homology between Tsp and any of the known protease sequences or protease active-site sequences (14) was detected. Surprisingly, significant sequence homology between Tsp and the human and bovine IRBP was observed. IRBP consists of four tandemly-repeated, homologous segments of ≈ 300 amino acids each (for review, see ref. 17). Fig. 4 shows a sequence alignment of Tsp and the eight domains of human and bovine IRBP. There are three segments of sequence similarity, designated A, B, and C. In these segments, 51% of the Tsp residues are identical to a residue in at least one IRBP repeat, and 23% of the Tsp residues show identity or strong conservation with positions in all eight IRBP repeats. We have speculated that Tsp may bind directly to the hydrophobic C-terminal sequence of the #105 protein. IRBP binds hydrophobic ligands such as retinol, retinoic acid, cholesterol, and fatty acids (17). Thus both proteins recognize hydrophobic ligands that contain a hydroxyl or carboxyl group at one end. The homologous segments of Tsp and IRBP could form a hydrophobic binding pocket that is used in ligand binding. Such a binding pocket would also be expected to contain a positively charged residue that confers specificity for the carboxyl group. We note that arginine-371 of Tsp is conserved in all eight IRBP domains and arginine-444 is conserved in six of the eight IRBP domains.

Hara *et al.* (19) have recently reported the cloning and sequencing of the *prc* gene of *E. coli*. *prc* and *tsp* are the same gene. They map to the same location on the chromosome and, with the exception of base 1018, are identical in sequence. Most interestingly, cells containing a conditionally lethal deletion of the *prc* gene are deficient in the C-terminal processing of penicillin-binding protein 3. In wild-type strains, cleavage *in vivo* of penicillin-binding protein 3 occurs after a valine residue that is 11 residues from the C terminus (20). Since purified Tsp is a C-terminal-specific protease that can cleave after valine, it seems extremely likely that it is directly responsible for the C-terminal cleavage of penicillin-binding protein 3 that occurs in the cell.

Tsp is similar in many ways to two previously purified *E. coli* proteases, protease Re (21) and a protease that degrades oxidized glutamine synthetase (22). Each of these proteases has a similar molecular weight and is monomeric. All behave similarly during purification, are ATP-independent, and show similar patterns of inhibition. Cleavage *in vitro* of oxidized glutamine synthetase occurs after an alanine residue, 38 residues from the C terminus, which would be consistent with Tsp activity. Primary sequence information for protease Re or the protease that degrades oxidized glutamine synthetase is not available but should eventually

provide the most stringent test of the possible identity of these proteases with Tsp.

We thank J. Bowie for performing the profile analysis of the IRBP homology; C. Burkins, P. Matsudaira, and W. Lane for protein sequencing; I. Papayannopoulos for mass spectral analysis; D. Siegele for the Kohara phage lift; W. Lim for the purified wild-type N-terminal-domain protein; and members of the Grossman, Burden, and Sauer labs for help and valuable discussions. This work was supported by National Institutes of Health Grant A116892. Mass spectral experiments were supported by National Institutes of Health Grant RR00317 to K. Biemann (Massachusetts Institute of Technology Mass Spectrometry Facility). K.R.S. is a predoctoral fellow of the W. M. Keck Foundation.

1. Goldberg, A. L. & St. John, A. C. (1976) *Annu. Rev. Biochem.* **45**, 747–803.
2. Bowie, J. U. & Sauer, R. T. (1989) *J. Biol. Chem.* **264**, 7596–7602.
3. Parsell, D. A., Silber, K. R. & Sauer, R. T. (1990) *Genes Dev.* **4**, 277–286.
4. Lim, W. A. & Sauer, R. T. (1991) *J. Mol. Biol.* **219**, 359–376.
5. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
6. Schagger, H. & von Jagow, G. (1987) *Anal. Biochem.* **166**, 368–379.
7. Kohara, Y., Akiyama, K. & Isono, K. (1987) *Cell* **50**, 495–508.
8. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
9. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
10. Kornitzer, D., Teff, D., Altuvia, S. & Oppenheim, A. B. (1991) *J. Bacteriol.* **173**, 2944–2953.
11. Heppel, L. A. (1971) in *Structure and Function of Biological Membranes*, ed. Rothfield, L. I. (Academic, New York), pp. 223–247.
12. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
13. Devereux, J., Haeblerli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
14. Barrett, A. J. (1986) in *Proteinase Inhibitors*, eds. Barrett, A. & Salvesen, G. (Elsevier, New York), pp. 3–22.
15. Oliver, D. B. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ed. Neidhardt, F. C. (Am. Soc. Microbiol., Washington), pp. 56–69.
16. Weiss, M. A., Karplus, M. & Sauer, R. T. (1987) *Biochemistry* **26**, 890–897.
17. Liou, G. I. H., Geng, L. & Baehr, W. (1991) *Prog. Clin. Biol. Res.* **362**, 115–137.
18. Gribskov, M., McLachlan, A. D. & Eisenberg, D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4355–4358.
19. Hara, H., Yamamoto, Y., Higashitani, A., Suzuki, H. & Nishimura, Y. (1991) *J. Bacteriol.* **173**, 4799–4813.
20. Nagasawa, H., Sakagami, Y., Suzuki, A., Suzuki, H., Hara, H. & Hirota, Y. (1989) *J. Bacteriol.* **171**, 5890–5893.
21. Park, J. H., Lee, Y. S., Chung, C. H. & Goldberg, A. L. (1988) *J. Bacteriol.* **170**, 921–926.
22. Roseman, J. E. & Levine, R. L. (1987) *J. Biol. Chem.* **262**, 2101–2110.