Cytochrome b_562_ is not cleaved by the tail-specific protease Tsp in _in vitro_ or in the periplasm of _Escherichia coli_ but becomes a good substrate when the C-terminal sequence WVAAA is added. Following randomization of the final three residue positions of this substrate, 54 different mutants with single residue substitutions were recovered. The steady-state expression levels of cytochrome variants bearing these mutant tails were similar in an _E. coli_ strain deleted for the tsp gene but differed markedly in a strain containing Tsp. Wild-type cytochrome b_562_ and seven variants, displaying a range of intracellular expression levels, were purified. These proteins were found to have the same T_m values in thermal denaturation experiments but to be cleaved by Tsp at rates differing by as much as 30-fold. Overall, the rates of Tsp cleavage of these proteins in _in vitro_ correlate with their rates of cleavage in _in vivo_ as determined by pulse-chase experiments. These results indicate that the C-terminal sequence of the cytochrome-b_562_ variants is important in determining their proteolytic fate in the cell and show that this degradation is mediated predominantly by Tsp. There are different selectivity rules at each of the three C-terminal positions. The identity of the C-terminal residue of the substrate, where small, uncharged residues (Ala, Cys, Ser, Thr, Val) are preferred, is most important in determining the rate of substrate cleavage by Tsp. Non-polar residues are also preferred at the second and third positions, but larger and more hydrophobic side chains are also acceptable at these positions in good substrates.

The observed selectivity of intracellular proteolysis implies the existence of mechanisms that allow proteases to discriminate between correct and incorrect protein substrates (1, 2). Some substrate recognition may be possible at the level of exposure of appropriate cleavage sites. For example, a protease might cleave almost any unfolded or partially folded protein in which peptide bonds flanked by the proper P1 and P1’ residues are accessible. This type of mechanism may serve to rid the cell of misfolded or unfolded proteins but is unlikely to allow significant selectivity unless the local sequence determinants of cleavage site selection occur only rarely. Some protein substrates may be marked for degradation by covalent modification with molecular tags, which then serve as recognition determinants in subsequent steps. The ubiquitin-proteasome system of eukaryotic cells is currently the best understood example of this type, although the determinants that cause particular proteins to be modified by ubiquitin are still not well understood (3). In bacteria, there are systems in which the identity of sequences at the C-terminal or N-terminal ends of proteins appear to serve as determinants of proteolytic degradation (4–6). In these cases, such sequences may serve as secondary binding sites, which allow a protease to tether a substrate while waiting for rare unfolding events that expose the sites of primary cleavage.

Specific degradation of proteins with non-polar C-terminal sequences was first reported for several cytoplasmic proteins in _Escherichia coli_ (4, 5), but the protease that mediates this degradation has not been identified. Tsp (tail-specific protease) is a periplasmic protease of _E. coli_ that was purified based on its ability to differentially degrade two protein substrates that differed only in their C-terminal residues (7). The protein substrate cleaved by Tsp had a relatively apolar C-terminal sequence (WVAAA), while the protein resistant to Tsp cleavage had a relatively polar C-terminal sequence (REYE). Although this specificity in _in vitro_ is similar to that observed in _in vivo_ for cytoplasmic degradation (5), gene knockout experiments have shown that Tsp is not involved in cytoplasmic degradation (8). Experiments in _in vitro_ have established that Tsp is an endoprotease that cleaves substrates at discrete sites throughout the polypeptide chain in a reaction that depends upon the identity of the substrate's C-terminal sequence and requires the presence of a free α-carboxyl group (9). The precise determinants that allow some C-terminal sequences but not others to be recognized by Tsp are currently unknown. Moreover, it has not yet been established that C-terminal-specific degradation of substrates by Tsp occurs in the cell. In this paper, we address these issues by studying the susceptibility of cytochrome-b_562_ variants to Tsp-mediated proteolysis in _in vivo_ and _in vitro_. Cytochrome b_562_ is a periplasmic protein that can be readily detected in _in vitro_ lysates because binding of the protein to heme gives rise to a characteristic red absorbance spectrum (10). We show that wild-type cytochrome b_562_ is resistant to Tsp-mediated cleavage but becomes a good substrate when a WVAAA C-terminal tail is added. To investigate sequence preferences at the C-terminus, we constructed libraries of cytochrome b_562_ variants with WVAAA, with each of the last three tail positions randomized individually, and assayed for steady-state levels of the modified variants in cells containing Tsp or deleted for Tsp. These results reveal different preferences at each of the three C-terminal positions and show that Tsp is the major periplasmic protease responsible for C-terminal-specific degradation of these substrates.

**EXPERIMENTAL PROCEDURES**

Strains, Plasmids, and Mutagenesis—_E. coli_ strain X90 is ara_Δ(lac pro) gyra argE(Am) Rif’ thi-1/F’ lac-14 lac’ pro’; _E. coli_ strain K51000 is X90 Δsp(pro)3:kan ada-51:Thi10 (8). Plasmid pKK101 is a _pBluescript_-derived vector that encodes ampicillin resistance and the Tsp-His_6_ protease under control of a lac-promoter (9). Plasmid pRW-1 (a gift from Michael Hecht) is a _pEMBL-18_ derived plasmid that encodes ampicillin resistance and the cytochrome-b_562_ gene under transcrip-
tional control of a lac promoter (11).

A plasmid (pCyb2) encoding a variant of cytochrome b$_{562}$ with the C-terminal tail sequence WVAAA was constructed by ligating the PstI-BamHI backbone fragment from pRW-1 (the PstI site is near the 3'-end of the cytochrome b$_{562}$ gene; the BamHI site is roughly 150 base pairs downstream) with a double-stranded oligonucleotide cassette encoding the stop codons for the WVAAA sequence, and the wild-type stop codon and termination sequences. The structure of pCyb2 was confirmed by restriction mapping and DNA sequencing. To randomize the C-terminal codons of the cytochrome b$_{562}$-WVAAA gene, the pCyb2 construction was repeated using an oligonucleotide cassette containing an equimolar mixture of G, A, T, and C at the appropriate codons. These libraries were transformed into X90 cells, single colonies were isolated, and the C-terminal sequences of genes from 60–75 independent candidates were determined by DNA sequencing.

Protein Purification—Wild-type Tsp (using a six-histidine tag) was purified from E. coli strain K1000(pK101) using nickel-chelate chromatography and ion-exchange chromatography as described (9). Cytochrome b$_{562}$ was purified from E. coli strain K1000(pRW-1) using a protocol adapted from Ames et al. (12). Cells were grown at 37 °C with gentle shaking for 12 h in 1 liter of 2X-YT broth supplemented with 150 μg/ml ampicillin and 1 mM isopropyl-1-thio-β-D-galactopyranoside and were harvested by centrifugation. The cell pellet was resuspended in 10 ml of chloroform, incubated at room temperature for 15 min, and 100 ml of 10 mM Tris-HCl (pH 8.0) was added. This mixture was centrifuged for 20 min at 4,000 rpm, and the red aqueous supernatant containing cytochrome b$_{562}$ and other periplasmic proteins was recovered. 10 ml of a 0.5 M potassium citrate (pH 4.0) buffer was added, and the pH was brought to 4.0 by addition of 1 N HCl. After stirring for 30 min at 4°C, precipitated proteins were removed by centrifugation at 10,000 rpm for 30 min. The supernatant was loaded onto a Mono S column (Pharmacia Biotech Inc.) equilibrated in 25 mM potassium citrate (pH 4.0), and the column was developed with a linear gradient from 0 to 300 mM KCl in the same buffer. Fractions containing cytochrome b$_{562}$, which eluted between 75 and 100 mM KCl, were pooled, concentrated, and chromatographed on a Superose 12 column (Pharmacia) in buffer containing 50 mM Tris-HCl (pH 8.0), 0.1 mM EDTA. Fractions that contained cytochrome b$_{562}$ were pooled and concentrated, and heme was added by rotatory evaporation to a concentration of 1.4 mg/ml. At different times, 0.5-ml aliquots were removed from the cuvette at various time points and assayed for cleavage of cytochrome b$_{562}$ by SDS gel electrophoresis. As expected, loss of absorbance at 418 nm was found to correlate with loss of the intact cytochrome b$_{562}$ band.

Stability of Cytochrome b$_{562}$ Variants—The circular dichroism spectra of cytochrome b$_{562}$ variants were measured using 25 μM protein in buffer containing 50 mM Tris-HCl (pH 8.0), 0.1 mM EDTA. Thermal melts were performed using the same protein concentration and buffer composition by monitoring the ellipticity at 222 nm as a function of temperature. Melting curves were fit to a two-state transition between native and denatured protein by non-linear least squares fitting using the program NONLIN for Macintosh (13, 14).

RESULTS

Wild-type cytochrome b$_{562}$ is not cleaved by Tsp in vitro (Fig. 1A) and is expressed at comparable levels in tsp+ and tsp− cells as determined by heme absorbance (Fig. 1B). Neither finding is surprising. The Tsp protease is thought to prefer non-polar tails that are accessible in the folded protein (7, 9), whereas cytochrome b$_{562}$ has a very polar C-terminal sequence (HQKYR), which is relatively inaccessible in the crystal structure (15, 16). In an attempt to make cytochrome b$_{562}$ a substrate for Tsp, we constructed a gene with the cytochrome coding sequence followed by codons for the C-terminal pentapeptide WVAAA. This sequence was chosen because purified Tsp is known to cleave variants of the N-terminal domain of λ repressor and of Arc repressor, which have the WVAAA tail in...
As shown in Fig. 1A, the purified cytochrome-\textit{b}_{562}\textendash WVAAA protein is also cleaved by Tsp in vitro. The cytochrome-\textit{b}_{562}\textendash WVAAA protein is expressed at a much lower steady-state level in cells containing Tsp than in cells lacking Tsp (Fig. 1B), and pulse-chase experiments show that the reduced steady-state level is caused by increased intracellular degradation, which is Tsp dependent (Table I).

Steady-state Expression Levels Depend on the Identity of the C-terminal Residues—Each of the three C-terminal residues of the cytochrome-\textit{b}_{562}\textendash WVAAA substrate was individually randomized, and from 16 to 20 different amino acids were recovered at each position. The resulting variants were assayed for steady-state expression levels in \textit{tsp}^{+} and \textit{tsp}^{-} cells as shown in Fig. 2. In this figure, open bars represent expression levels in the absence of Tsp, and closed bars represent expression levels in the presence of Tsp. In the cells containing Tsp, there are significant differences in expression levels depending on the chemical identity of the three C-terminal residues. These differences are discussed in greater detail below. In the cells lacking Tsp, all of the variants are expressed at similar levels. This suggests that Tsp is the major periplasmic protease responsible for C-terminal-specific protein degradation of the cytochrome-\textit{b}_{562} substrates.

Cytochrome variants with Ala, Cys, Val, Ser, and Thr at the C terminus (position 1) are expressed at the lowest steady-state levels in cells containing Tsp (Fig. 2), indicating that Tsp prefers substrates that have small, uncharged side chains at this position. Variants with polar side chains, large side chains, Pro, or Gly at the C terminus are expressed at reasonably high levels.

| Variant | Half-life in vivo \textit{tsp}^{+} | Half-life in vivo \textit{tsp}^{-} | Relative resistance to cleavage by Tsp in vitro |
|---------|-----------------|-----------------|-------------------------------|
| AAA     | <2              | >120            | 0.01                          |
| AYA     | <2              | >120            | 0.01                          |
| LAA     | <2              | >120            | 0.01                          |
| AAV     | 6               | >120            | 0.02                          |
| AGA     | 36              | >120            | 0.10                          |
| QAA     | 23              | >120            | 0.13                          |
| AAK     | >120            | >120            | 0.30                          |

\(^{a}\) Rate of cleavage of wild type/rate of cleavage of variant.

\(^{1}\) K. C. Keiler and R. T. Sauer, unpublished data.

Fig. 2. Steady-state levels of cytochrome-\textit{b}_{562} variants in the \textit{tsp}^{+} strain X90 (filled bars) and the otherwise isogenic \textit{tsp}^{-} strain KS1000 (open bars). The error bars indicate the standard deviation from the mean for three different X90 cultures.
levels even in the presence of Tsp. In many cases, the expression levels of these variants are as high as in cells deleted for Tsp (Fig. 2). This suggests that the identity of the C-terminal residue is extremely important in determining whether a protein will be a good or poor substrate for Tsp.

At the penultimate amino acid residue (position 2), variants with Ala, Tyr, Ile, and Trp are expressed at the lowest levels in the presence of Tsp, and variants with Arg, Lys, and Gln are expressed at the highest levels. In general, hydrophobic residues at position 2 appear to be preferred by Tsp relative to hydrophilic residues. Moreover, only a few side chains at this position increase expression to levels comparable to those seen in tsp cells. This suggests that position 2 is less important than position 1 in determining resistance to Tsp cleavage.

At the third position from the C terminus, Tsp prefers Ala, Leu, Val, and Ile. The least preferable side chains are Asn, Gin, and Met. It is somewhat surprising that Leu and Met, which are often considered to be conservative substitutions for each other, have such different effects at this position. No side chains at position 3 increase steady-state expression to levels observed in the absence of Tsp, suggesting that this position is less important than either position 1 or 2 in determining resistance to Tsp cleavage.

Comparison of Cleavage Rates in Vivo and in Vitro—We purified cytochrome-b_{562} variants with at least one stabilizing and one destabilizing residue at each of the three C-terminal positions and assayed rates of cleavage by Tsp in vitro (Table I). Tsp rapidly cleaved variants with the AAA, AAY, AYA, and LAA tails but did not cleave or only slowly cleaved variants with the AAK, AGA, and QAA tails. Pulse-chase experiments were used to determine half-lives for these same variants in the cell (Fig. 3). As expected, variants that had low steady-state expression levels (AAA, AAY, LAA) had the shortest half-lives, and variants that had high steady-state expression levels (AAK, AGA) had considerably longer half-lives (Fig. 2, Table I). As shown in Table I, the resistance of the cytochrome-b_{562} variants to cleavage by Tsp in vitro correlates reasonably well (R^2 ~ 0.9) with the half-lives of the variants in a tsp strain in vivo. Thus, these data support the model that differential degradation of these proteins by Tsp in vivo is responsible for the observed differences in their half-lives and steady-state levels.

Effects of Tail Sequence on Protein Stability—In principle, the different C-terminal sequences of the cytochrome-b_{562} variants might affect susceptibility to Tsp by reducing the thermodynamic stability of the protein. To test for this possibility, the stabilities of the purified cytochrome-b_{562} variants were measured by temperature denaturation monitored by CD spectroscopy (Fig. 4). The melting curve of each variant showed a T_m between 68 and 69 °C, indicating that the C-terminal mutations do not affect the global stability of cytochrome b_{562}.

**DISCUSSION**

The work presented here has established the importance of specific amino acids at each of the three C-terminal residues in determining whether a protein is efficiently cleaved by Tsp in vitro and in vivo. In previous studies based on screening of a small number of potential protein and peptide substrates, we had concluded that Tsp appeared to recognize substrates with non-polar or hydrophobic C-terminal residues and not to recognize substrates with polar C-terminal residues (7, 9). The data summarized in Fig. 2 reveal that this view is an oversimplification. While no good substrates have highly polar tails and most good substrates do have non-polar residues at the three C-terminal positions, there is considerable fine specificity. For example, small non-polar residues are preferred relative to larger hydrophobic side chains at the C-terminal position. At the other two positions, there is no simple correlation between the size of non-polar side chains and effects on Tsp cleavage. For example, alanine and tyrosine are the most destabilizing side chains at the penultimate residue, while valine has a significantly smaller effect at this position.

How do the C-terminal residues of a substrate affect its degradation by Tsp? Unlike systems in which sequence signals act to target substrates to subcellular compartments specialized for degradation (17), the C-terminal sequences that mediate degradation by Tsp appear to be recognized by the protease itself. This is shown most clearly by the strong correlation between the half-lives of the cytochrome-b_{562} variants in vivo and the resistance of these purified variants to cleavage by purified Tsp in vitro. Since other macromolecules are not required for Tsp-mediated degradation of substrates with non-polar tails, then Tsp must either recognize these C-terminal sequences directly or recognize indirect effects of these sequences on stability or structure. Several experiments suggest that Tsp recognizes the tail sequences directly. As shown here, C-terminal sequences that make cytochrome b_{562} a good substrate for Tsp do not alter the thermodynamic stability of the protein. The same is true of C-terminal sequences that make Arc repressor and the N-terminal domain of λ repressor good substrates for Tsp cleavage in vitro (7, 9). Moreover, it seems unlikely that destabilizing C-terminal tails act indirectly by allowing other sequence or structural determinants to be recognized because the same tail sequence (e.g., WVAAA) can make Arc repressor, λ repressor, and cytochrome b_{562} (proteins that differ in primary, secondary, tertiary, and quaternary structure) good substrates for Tsp cleavage. We believe that Tsp directly recognizes protein and peptide substrates by using a binding site that requires a free α-carboxyl group and side chains of the appropriate size, shape, and hydrophobicity at the last three residue positions of the substrate. Such a binding site
would serve to tether the substrate to the enzyme, and binding of an appropriate C-terminal tail at this site might also function to activate the enzyme.

Our studies have shown that Tsp can efficiently degrade a periplasmic protein with a WVAAA C-terminal tail. An independent protease in the cytoplasm of E. coli is also capable of rapidly degrading proteins with WVAAA tails (5, 8). Why does E. coli use the Tsp system in the periplasm and an independent system in the cytoplasm to degrade proteins with certain C-terminal sequences? It seems unlikely that this is a general mechanism for removing unfolded or misfolded proteins from the cell. First, most damaged or misfolded proteins would not be expected to have the proper C-terminal sequence to allow degradation by a C-terminal-specific pathway. Second, Tsp and its cytoplasmic counterpart are not limited to degrading unfolded proteins. Both the cytochrome-b_{562} variants studied here and the λ-repressor variants used to study cytoplasmic C-terminal-specific degradation are stably folded (5). Recent studies suggest that one function of tail-specific proteases is to work in conjunction with a peptide-tagging system that marks certain proteins in E. coli for degradation (18).² In this system, proteins translated from damaged mRNA are modified by C-terminal addition of a peptide with the sequence AANDENYA-LAA. The C-terminal residues of this peptide tag then render the tagged protein susceptible to degradation by Tsp or by its cytoplasmic counterpart.

Acknowledgments—We thank Michael Hecht for the gift of pRW-1 and Cliff Robinson for helpful advice.

REFERENCES
1. Gottesman, S., and Maurizi, M. R. (1992) Microbiol. Rev. 56, 592–621
2. Goldberg, A. L., and St. John, A. C. (1976) Annu. Rev. Biochem. 45, 747–803
3. Ciechanover, A. (1994) Cell 79, 13–21
4. Bowie, J. U., and Sauer, R. T. (1989) J. Biol. Chem. 264, 7586–7602
5. Parsell, D. A., Silber, K. R., and Sauer, R. T. (1990) Genes & Dev. 4, 277–286
6. Tobias, J. W., Shrader, T. E., Rocap, G., and Varshavsky, A. (1991) Science 254, 1374–1377
7. Silber, K. R., Keiler, K. C., and Sauer, R. T. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 295–299
8. Silber, K. R., and Sauer, R. T. (1994) Mol. & Gen. Genet. 242, 237–240
9. Keiler, K. C., Silber, K. R., Downard, K. M., Bienmann, K., and Sauer, R. T. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 2718–2722
10. K.C. Keiler, P.R.H. Waller, and R.T. Sauer, submitted for publication.

² K.C. Keiler, P.R.H. Waller, and R.T. Sauer, submitted for publication.