Molecular Cell, Vol. 11, 671–683, March, 2003, Copyright ©2003 by Cell Press

# Proteomic Discovery of Cellular Substrates of the ClpXP Protease Reveals Five Classes of ClpX-Recognition Signals

**Julia M. Flynn,[1] Saskia B. Neher,[1] Yong-In Kim,[1–3] Robert T. Sauer,[1] and Tania A. Baker[1,2,*]**
[1]Department of Biology
[2]Howard Hughes Medical Institute
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

## Summary

**ClpXP is a protease involved in DNA damage repair, stationary-phase gene expression, and ssrA-mediated protein quality control. To date, however, only a handful of ClpXP substrates have been identified. Using a tagged and inactive variant of ClpP, substrates of *E. coli* ClpXP were trapped in vivo, purified, and identified by mass spectrometry. The more than 50 trapped proteins include transcription factors, metabolic enzymes, and proteins involved in the starvation and oxidative stress responses. Analysis of the sequences of the trapped proteins revealed five recurring motifs: two located at the C terminus of proteins, and three N-terminal motifs. Deletion analysis, fusion proteins, and point mutations established that sequences from each motif class targeted proteins for degradation by ClpXP. These results represent a description of general rules governing substrate recognition by a AAA+ family ATPase and suggest strategies for regulation of protein degradation.**

## Introduction

Protein degradation is an essential component of biological regulation and protein quality control in organisms ranging from bacteria to humans. Many cytoplasmic proteases are large multisubunit complexes in which the proteolytic active sites are sequestered within an internal chamber. Access to this chamber is controlled by axial pores that exclude native proteins and all but the smallest peptides (for review, see Lupas et al., 1997). These multimeric proteases form complexes with AAA+ ATPases, which denature and translocate substrates into the proteolytic chamber for degradation (for review, see Ogura and Wilkinson, 2001). The ClpXP, ClpAP, HslUV (ClpYQ), HflB (FtsH), and Lon proteases of bacteria share this basic mechanism with the proteasomes of eukaryotic organisms (for review, see Schirmer et al., 1996). Identifying the proteolytic targets of specific proteases is critical to any general understanding of their diverse cellular functions and provides a way to decipher the rules by which these enzymes recognize substrates.

*E. coli* ClpXP is an ATP-dependent intracellular protease. The ClpX component is a hexameric AAA+ ATPase responsible for substrate recognition, unfolding, and translocation into ClpP (Wojtkowiak et al., 1993; Wawrzynow et al., 1995; Weber-Ban et al., 1999; Kim et al., 2000). ClpX can also act independently to dismantle multimers and remodel proteins (Levchenko et al., 1995). ClpP is a 14 subunit serine peptidase (Maurizi et al., 1990a). It has a barrel-like structure comprised of two heptameric rings. Face-to-face stacking of these rings sequesters the active sites within the proteolytic chamber (Wang et al., 1997). One or two ClpX hexamers bind to $ClpP_{14}$ to form the ClpXP protease (Grimaud et al., 1998). ClpP also combines with hexamers of the ClpA ATPase to form ClpAP (Katayama et al., 1988). ClpX and ClpA generally confer distinct substrate specificities to their respective protease complexes, although these enzymes do recognize some common substrates (for review, see Gottesman, 1996; Gottesman et al., 1998).

ClpX and ClpP orthologs are found in most bacteria, mitochondria, and chloroplasts. In *E. coli*, *clpP*-defective cells show delayed recovery both from stationary phase and following a shift to nutrient poor media (Damerau and St. John, 1993). Proteolysis by ClpXP is involved in the development of competence and in sporulation in *Bacillus subtilis* and is required for viability and cell cycle progression in *Caulobacter crescentus* (Jenal and Fuchs, 1998; Msadek et al., 1998). ClpP is also important for the virulence of bacterial pathogens including *Yersinia enterocolitica*, *Streptococcus pneumoniae*, *Salmonella typhimurium*, and *Listeria monocytogenes* (for review, see Porankiewicz et al., 1999).

Despite the diverse physiological roles of ClpXP, only a few substrates have been identified. *E. coli* ClpX was originally discovered as a component required for ClpP-dependent degradation of the λO phage replication protein (Gottesman et al., 1993). Since then, four additional phage or plasmid proteins (Mu repressor, MuA transposase, RK2 replication protein TrfA, and the P1 antidote protein PhD) and three *E. coli* proteins (the stationary phase sigma factor $\sigma^S$, the SOS protein UmuD′, and a type I restriction-modification subunit HsdR) have been identified as ClpXP substrates (see Gottesman, 1996, and references therein; also see Frank et al., 1996; Konieczny and Helinski, 1997; Makovets et al., 1998). ClpXP also degrades proteins modified by addition of the ssrA tag, an 11 residue sequence added cotranslationally to the C terminus of nascent polypeptides on stalled ribosomes (Keiler et al., 1996; Gottesman et al., 1998).

ClpX interacts with peptide sequences—referred to as recognition signals—at the C termini of the ssrA tag and MuA (Levchenko et al., 1997; Gottesman et al., 1998). In contrast, signals near the N terminus of λO appear most important for ClpX recognition (Gonciarz-Swiatek et al., 1999). In addition to these examples of direct recognition, auxiliary proteins are implicated in targeting some substrates to ClpXP; UmuD confers instability to UmuD′ (Gonzalez et al., 2000), and RssB targets $\sigma^s$ to ClpXP (Muffler et al., 1996). Although progress is being made in understanding how ClpX recognizes some members of this small group of substrates, general rules governing substrate recognition have yet to emerge.

*Correspondence: tabaker@mit.edu
[3]Current address: Abbott Bioresearch Center, 100 Research Drive, Worcester, Massachusetts 01605.

Here, we report the identification of more than 50 *E. coli* proteins that are trapped in a ClpX-dependent fashion within an active-site mutant of ClpP. Analysis of these ClpXP substrates provides a more comprehensive understanding of the cellular roles of this protease and reveals five distinct classes of ClpX-recognition motifs. This study provides the general description of the sequence rules that mediate substrate recognition by an energy-dependent intracellular protease and establishes a foundation for understanding how degradation may be regulated.

## Results

### Protein Trapping by ClpXP In Vivo

To identify new substrates, we took advantage of the ability of inactivated ClpP to accept and retain proteins translocated into its chamber by the ClpX ATPase (Kim et al., 2000). A ClpP$^{trap}$ variant was constructed containing an active-site mutation (S97A) as well as a C-terminal tandem-affinity tag ($Myc_3$-TEV-$His_6$). Proteins translocated into the proteolytic chamber of ClpXP$^{trap}$ in vitro were not degraded and were only released slowly (Kim et al., 2000; Singh et al., 2000; data not shown). To test whether ClpP$^{trap}$ also captured substrates in vivo, it was coexpressed in *E. coli* with GFP-ssrA, a model ClpXP substrate. GFP-ssrA copurified with ClpP$^{trap}$ during affinity chromatography, confirming that trapping occurred in vivo (data not shown). Cellular trapping of GFP-ssrA was prevented by an ssrA-tag mutation (C-terminal A→D, data not shown) that prevents ClpXP degradation in vitro (Flynn et al., 2001), indicating that trapping requires the same ClpX-substrate interactions needed for degradation.

To determine if capture by ClpP$^{trap}$ depended on the ClpX or ClpA ATPases, experiments were performed in *clpX*$^+$*clpA*$^+$, *clpX*$^+$*clpA*$^-$, *clpX*$^-$*clpA*$^+$, and *clpX*$^-$*clpA*$^-$ strains. To avoid trapping a heterogeneous collection of ssrA-tagged proteins, we deleted the gene encoding SmpB, a protein required for ssrA tagging (Karzai et al., 1999), from the trapping strains. These strains also carried an insertion in the chromosomal copy of *clpP* and expressed ClpP$^{trap}$ under control of an IPTG-inducible promoter. Proteins that copurified with ClpP$^{trap}$ in each strain were visualized by staining after electrophoresis on 2D gels (Figure 1).

Approximately 70 proteins copurified with ClpP$^{trap}$ in the strain expressing both ClpX and ClpA (Figure 1B). A subset of approximately 50 of these proteins were trapped in the strain expressing just ClpX (Figure 1C), whereas about 30 proteins were trapped in the strain expressing just ClpA (Figure 1D). In the absence of ClpX and ClpA (Figure 1A), only a handful of polypeptides copurified with ClpP$^{trap}$; most were ClpP fragments and one was DnaK (data not shown). Because the identities of the vast majority of ClpP$^{trap}$-captured proteins depended on the presence of ClpX or ClpA, we conclude that these ATPases selectively recognize and translocate proteins into the trap. The proteins captured in a ClpX-dependent or ClpA-dependent fashion are therefore likely to be substrates for degradation by ClpXP or ClpAP. About 10 proteins were present in both the ClpX-only and ClpA-only samples (Figures 1C and 1D), suggesting that these proteins are substrates for both proteases (see Table 1). Below, we characterize many of the proteins captured by ClpP$^{trap}$ in a ClpX-dependent manner.

### Identification of ClpXP Substrates

To identify cellular proteins captured by ClpXP$^{trap}$, complexes were isolated from the strain containing ClpX but not ClpA (*clpX*$^+$*clpA*$^-$) and separated on a 1D gel. Gels slices were excised, digested with trypsin, and analyzed by tandem-mass spectrometry. This procedure identified 60 *E. coli* proteins in addition to ClpP, ClpX, and the TEV protease (Table 1). One of the most abundant trapped proteins was $\sigma^S$ (Figure 1C), the stationary-phase sigma factor that is degraded by ClpXP during exponential growth (Schweder et al., 1996). Proteins captured by ClpXP$^{trap}$ included a wide variety of regulatory proteins and biological catalysts (Table 1), including many with suggested roles in stationary phase and oxidative stress responses (see Discussion). Based on annotations, nearly all of these proteins reside in the cytoplasm with ClpXP. One outer membrane protein, OmpA, and one inner membrane protein, RseA, were apparent exceptions (see below and Discussion). Mass spectrometry of the *clpX*$^-$*clpA*$^-$ sample revealed the presence of peptides from only 2 of the 60 *E. coli* proteins trapped in the *clpX*$^+$*clpA*$^-$ strain (see Experimental Procedures), providing further evidence of the importance of ClpX for the observed capture.

Western blots confirmed ClpX-dependent trapping of five proteins and also established whether full-length proteins or fragments were captured (Figure 2A, upper panel). Antibodies against Dps, Rsd, and DksA reacted with species having molecular weights expected for each full-length protein. In contrast, protein fragments rather than the full-length RseA and LexA copurified with the ClpP$^{trap}$ (Figure 2A, upper panel). For LexA, the two antibody-reactive bands had electrophoretic mobilities expected for protein fragments generated by RecA-mediated autocleavage between $Ala^{84}$ and $Gly^{85}$ (Little et al., 1980). For RseA, the trapped fragment bound antibodies that recognize the protein's N-terminal, cytoplasmic domain. These data strongly suggest that trapping of RseA and LexA depends upon initial cleavage of these proteins by other proteases (see Discussion). None of the five proteins tested was detected in trapped complexes isolated from the *clpX*$^-$*clpA*$^-$ strain (Figure 2A, lower panel), confirming the specificity of trapping.

Degradation experiments support the hypothesis that proteins that copurify with ClpXP$^{trap}$ are substrates for ClpXP degradation. For example, Dps, a DNA binding protein induced during starvation (Almiron et al., 1992) and one of the most abundant trapped proteins, had a significantly longer half-life in *clpX*$^-$ than in *clpX*$^+$ cells during outgrowth from stationary phase (Figure 2B) and was efficiently degraded by ClpXP in vitro (see below). DksA, the *dnaK* suppressor protein (Kang and Craig, 1990), was also stabilized in the *clpX*$^-$ strain, suggesting that ClpXP participates in degradation of this protein in vivo (Figure 2B). Note, however, that other proteases must also contribute to the degradation of Dps and DksA

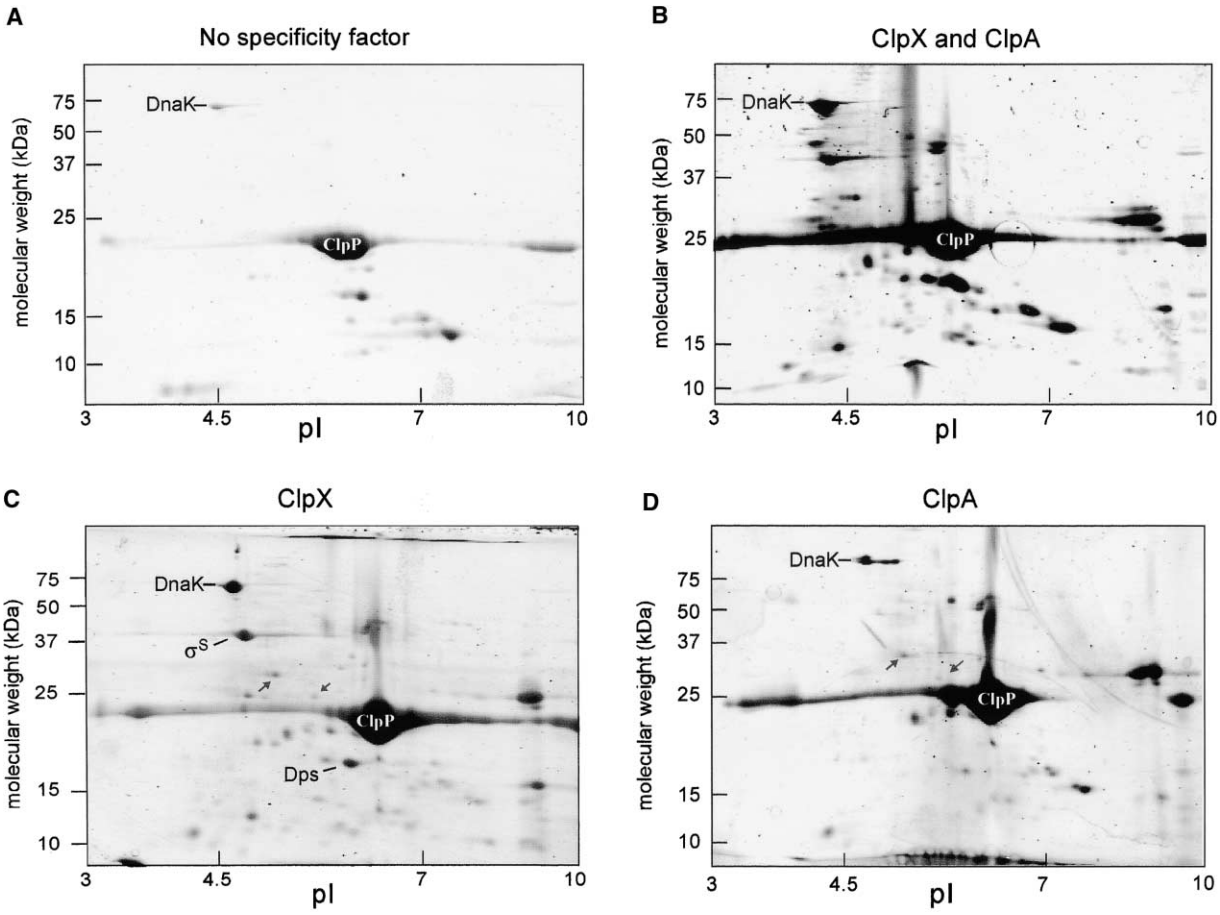



Figure 1. 2D Gel Analysis of Proteins Captured by ClpP$^{trap}$

Panels show proteins captured by ClpP$^{trap}$ in *E. coli* strains JF148 (A), JF176 (B), JF162 (C), and JF172 (D). Arrows indicate representative proteins captured by both ClpXP$^{trap}$ and ClpAP$^{trap}$.

in vivo because these proteins were still degraded in the *clpX*$^{-}$ strain (see Discussion). The N-terminal and C-terminal autocleavage fragments of LexA were also found to be degraded by ClpXP in vivo and in vitro (our unpublished data). Finally, E.L. Mettert and P.J. Kiley (personal communication) demonstrated that another trapped protein, the transcription regulator Fnr, was degraded in a ClpXP-dependent manner in vivo when cells were grown aerobically. Hence, $\sigma^S$, GFP-ssrA, the LexA N domain and C domain, Dps, DksA, and Fnr are both captured by ClpXP$^{trap}$ and appear to be substrates for ClpXP degradation. These results demonstrate the effectiveness of trapping as a method for global substrate discovery and suggest that most other captured proteins will also prove to be authentic ClpXP substrates.

### Many Trapped Substrates Have C-Terminal Degradation Signals

ClpX recognizes the C-terminal residues of certain substrates, including Leu-Ala-Ala-COOH of the ssrA tag and Arg-Arg-Lys-Lys-Ala-Ile-COOH of MuA (Levchenko et al., 1997; Flynn et al., 2001). Inspection of the C termini of the proteins trapped in a ClpX-dependent fashion revealed that 45% had sequences similar to either the ssrA tag or MuA (C motifs 1 and 2 in Figure 3A). Four trapped proteins had Ala-Ala terminal dipeptides, which in the ssrA tag is largely responsible for ClpX recognition (Flynn et al., 2001). Other trapped proteins had nonpolar C-terminal dipeptides and basic side chains in the region 3 to 6 residues before the C terminus. Positively charged residues at these positions are important for ClpX recognition of MuA (Levchenko et al., 1997).

ClpX binding to the C-terminal sequences from Crl, RibB, LldD, YdaM, and YbaQ was tested by inhibition of ClpXP degradation of GFP-ssrA (Figure 3B, inset). Synthetic peptides corresponding to the 11 C-terminal amino acids of each of these proteins inhibited degradation of GFP-ssrA (Figure 3B). Controls confirmed the specificity of this inhibition; neither an ssrA-peptide variant with Asp-Asp-COOH (Gottesman et al., 1998) nor the C-terminal peptide of Dps, which is not similar to either the ssrA or MuA tags, affected degradation of GFP-ssrA. Hence, the C-terminal residues of a number of proteins captured in a ClpX-dependent fashion bind ClpX, as expected for sequences that function as recognition signals.

Table 1. ClpXP$^{trap}$-Associated Proteins

| Gene | Swissprot Accession # | # Peptides | C-Terminal Signal | N-Terminal Signal | Gene Product or Function |
|---|---|---|---|---|---|
| **Transcriptional regulators** | | | | | |
| crl | P24251 | 9 | C-M1[a] | N-M3 ++ | Curlin genes regulatory protein |
| dksA | P18274 | 6 | C-M1 | N-M3 ++ | DnaK suppressor protein |
| fnr | P03019 | 8 | C-M1 | N-M1 ++ | Transcription regulator FNR |
| iscR | P77484 | 9 | C-M1 | N-M2 ++ | Iron-sulfur cluster regulator |
| lexA | P03033 | 3 | | N-M2 ++ | LexA repressor |
| rpoS ($\sigma^S$) | P13445 | 110 | | N-M1 + | RNA polymerase sigma factor $\sigma^S$ |
| rsd | P31690 | 1 | C-M2 | | Regulator of sigma D |
| rseA | P38106 | 2 | | | Negative regulator of sigma-E |
| **Translation** | | | | | |
| rplE | P02389 | 2 | | N-M1 + | 50S ribosomal protein L5 |
| rplJ | P02408 | 57 | C-M1 | N-M3 + | 50S ribosomal protein L10 |
| rplK | P02409 | 5 | | N-M1 + | 50S ribosomal protein L11 |
| rplN | P02411 | 17 | | | 50S ribosomal protein L14 |
| rplS | P02420 | 11 | | N-M1 + | 50S ribosomal protein L19 |
| rplU | P02422 | 2 | C-M1 | N-M3 + | 50S ribosomal protein L21 |
| tufB | P02990 | 2 | | N-M3 + | Elongation factor Ef-Tu |
| **Chaperones and degradation** | | | | | |
| clpX | P33138 | 5 | | N/A | Clp protease ATP-binding subunit |
| dnaK | P04475 | 75 | | N/A | Chaperone Hsp70 |
| gcp | P05852 | 7 | C-M1 | | O-sialoglycoprotein endopeptidase |
| groEL | P06139 | 6 | | N/A | Chaperone Hsp60 |
| lon | P08177 | 3 | | | ATP-dependent protease Lon |
| pepB | P37095 | 2 | C-M1 | | Aminopeptidase B |
| **Detoxification (protection)** | | | | | |
| dps | P27430 | 70 | | N-M1 ++ | Global regulator protein Dps |
| katE | P21179 | 1 | C-M1 | N-M3 + | Hydroperoxidase II |
| nrdH | Q47414 | 2 | C-M1 | N-M2 + | Glutaredoxin-like protein NrdH |
| tpx | P37901 | 4 | C-M1 | N-M1 + | Thiol peroxidase |
| **Cell division** | | | | | |
| ftsZ | P06138 | 3 | | N/A | Cell division GTPase |
| **Transposition** | | | | | |
| insH | P03837 | 4 | | N-M3 ++ | IS5 transposase |
| **Cell motility and transport proteins** | | | | | |
| cheW | P07365 | 2 | C-M1 | N-M1 ++ | Chemotaxis protein CheW |
| cysA | P16676 | 13 | C-M1 | N-M1 ++ | Sulfate permease A protein |
| exbB | P18783 | 5 | C-M1 | | Uptake of enterochelin |
| gatA | P37187 | 5 | | N-M1 ++ | Galactitol-specific enzyme IIA |
| ompA[b] | P02934 | 4 | | N-M2 + | Outer membrane protein 3a |
| secA | P10408 | 5 | | N-M1 + | Protein translocase protein SecA |
| **Metabolism and energy production** | | | | | |
| aceA[b] | P05313 | 12 | | N-M2 ++ | Isocitrate lyase |
| acnB | P36683 | 1 | C-M1 | | Aconitase |
| aldA | P25553 | 1 | C-M1 | | Aldehyde dehydrogenase |
| atpD | P00824 | 6 | | N-M1 ++ | $\beta$ subunit of F1 ATP synthase |
| cysD | P21156 | 2 | | N-M1 + | Sulfate adenylyltransferase |
| dadA | P29011 | 1 | | N-M2 ++ | D-amino acid dehydrogenase |
| fabB | P14926 | 1 | | N-M2 ++ | $\beta$-ketoacyl-acyl carrier protein synthase I |
| gapA[b] | P06977 | 4 | | N-M1 ++ | Glyceraldehyde 3-phosphate dehydrogenase |
| gatY | P37192 | 3 | | N-M2 ++ | Tagatose 1,6-bisphophate aldolase |
| gatZ | P37191 | 5 | | N-M2 ++ | Tagatose 6-phosphate kinase |
| glcB | P37330 | 2 | | N-M3 ++ | Malate synthase |
| glpD | P13035 | 1 | C-M1 | | Glycerol 3-phosphate dehydrogenase |
| glyA | P00477 | 2 | C-M1 | N-M1 + | Glycine hydroxymethyltransferase |
| iscS | P39171 | 1 | | N-M2 ++ | Cysteine desulferase |
| iscU | P77310 | 4 | | | IscU |
| lipA | P25845 | 3 | | N-M2 ++ | Lipoic acid synthetase |
| lldD | P33232 | 2 | C-M1[a] | | L-Lactate dehydrogenase |
| moaA | P30745 | 3 | | N-M3 + | Molybdopterin biosynthesis, protein A |
| paaA | P76077 | 1 | C-M2 | | Phenylacetic acid degradation protein |
| pncB | P18133 | 4 | C-M2 | N-M1 + | Nicotinate phosphoribosyltranferase |
| ribB | P24199 | 8 | C-M2[a] | | Riboflavin biosynthase |
| tnaA[b] | P00913 | 32 | | N-M1 + | Tryptophanase |
| udp | P12758 | 1 | | N-M1 + | Uridine phosphorylase |
| **Unknown function** | | | | | |
| ybaQ | P77303 | 1 | C-M2[a] | N-M3 + | |
| ycbW | P75862 | 5 | C-M1 | N-M2 ++ | |
| ydaM | P77302 | 3 | C-M1[a] | | |
| yebO | P76266 | 4 | | | |
| ygaT | P76621 | 7 | | ++ | |

Proteins are grouped into functional categories based on annotations from the SwissProt database (Bairoch and Apweiler, 2000) and the general literature. For each protein, the gene name, SwissProt accession number, number of peptides identified by MS/MS analysis, and protein name are listed. Proteins with C-terminal sequences similar to those of the ssrA tag (C-M1) or the MuA tag (C-M2) are marked. Proteins whose N-terminal peptides bind to ClpX strongly (++) or moderately (+) are marked. GroEL, FtsZ, ClpX, and DnaK were not tested for binding of their N-termini to ClpX (N/A). The N termini of the proteins that bind to ClpX are categorized as containing N motif 1 (N-M1), N motif 2 (N-M2) or N motif 3 (N-M3) as defined in Figure 4B. A Western blot revealed the presence of Rsd in ClpP$^{trap}$ (see Figure 2), establishing that the identity of trapped proteins can be determined reliably from a single peptide.

[a] Proteins whose corresponding C-terminal peptides inhibit ClpXP degradation of GFP-ssrA.

[b] Proteins that were also found to be captured by ClpAP$^{trap}$.

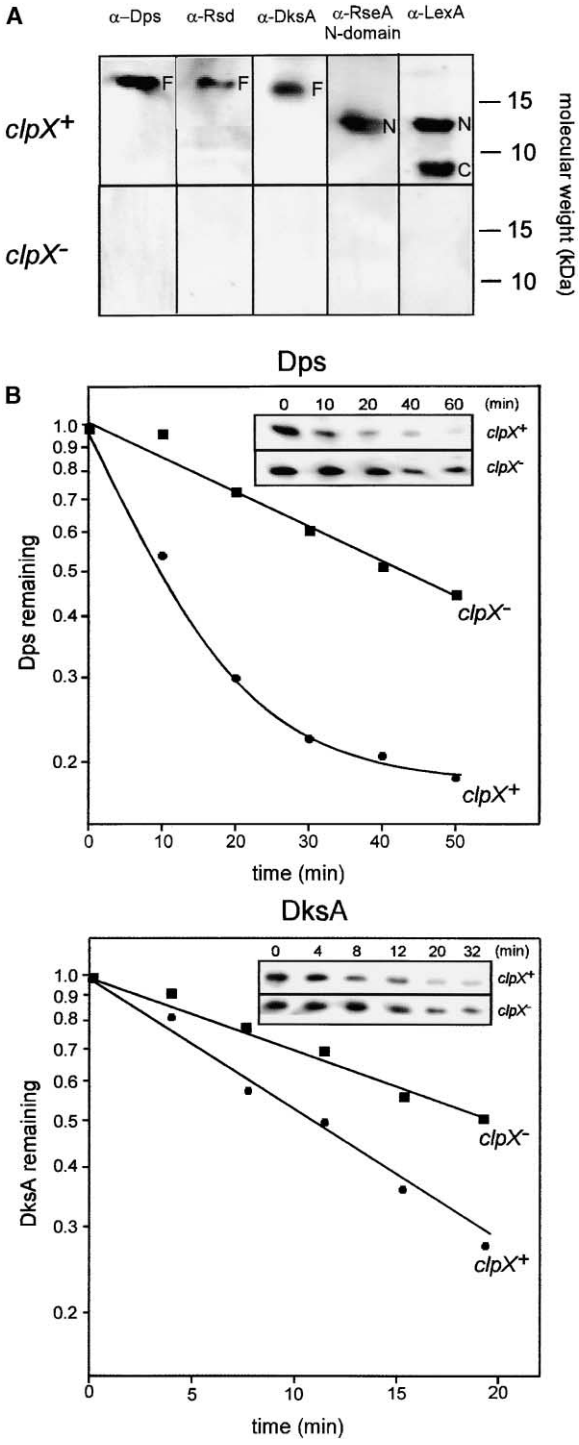



Figure 2. Western Blots of Trapped Proteins

(A) The molecular weights of bands for Dps (18.5 kDa), Rsd (18.1 kDa), and DksA (17.3 kDa) correspond to full-length proteins (F). The molecular weight of the RseA band (13 kDa) corresponds to an N-terminal fragment (N). The LexA fragments have masses (9 and 13 kDa) expected for autocleavage fragments consisting of residues 1–84 (N) and 85–202 (C). No immunoreactivity was observed in samples trapped in a *clpX*$^-$*clpA*$^-$ strain.

(B) ClpX-dependent degradation in vivo. Following dilution from a stationary phase culture, protein synthesis was inhibited with spectinomycin at an $A_{600}$ of 0.1, and samples were removed at specific time points and assayed by Western blotting with anti-Dps or anti-DksA antibodies as indicated.

To test directly for functional recognition, we fused the 10 C-terminal residues of Crl, Gcp, and YbaQ to a stable reporter protein—Arc repressor—and assayed ClpXP degradation in vitro. Each fusion protein but not the parent Arc protein was rapidly degraded (Figure 3C). Thus, these C-terminal sequences function as ClpXP-degradation signals. By extension, we suggest that most if not all of the proteins listed in Figure 3A have C-terminal peptide signals that make them substrates for ClpXP.

### Peptide Arrays Identify N-Terminal ClpX Binding Signals

To test for potential N-terminal ClpX-recognition signals, we prepared a peptide array with the N-terminal 11 residues of the ClpXP-trapped proteins and several previously identified ClpXP substrates attached covalently to a filter. This array was incubated with ClpX and ATPγS and was washed, and peptide-associated ClpX was detected with anti-ClpX antibody (Figure 4A). ClpX bound to the N-terminal peptides of about 60% of the proteins tested. The specificity of peptide binding was evident from inspection of the filter; ClpX binding ranged from very strong to undetectable. Notably, ClpX bound strongly to the N-terminal peptide of λO, a protein whose N-terminal residues are known to be important for ClpXP degradation (Gonciarz-Swiatek et al., 1999). These results suggest that ClpX may recognize many trapped proteins through N-terminal signals.

### Alignments Reveal Multiple Classes of N-Terminal Recognition Motifs

Inspection of the N-terminal sequences bound by ClpX revealed several distinct motifs. For instance, λO, Dps, and sixteen other trapped proteins contained good matches to the consensus: polar-T/$\phi$-$\phi$-basic-$\phi$, where $\phi$ indicates a hydrophobic side chain (N motif 1 in Figure 4B; also see Table 1). As an example of an N motif 1 protein, we studied Dps. Purified Dps was efficiently degraded in a reaction requiring ClpX, ClpP, and ATP (Figure 5A; data not shown). In contrast, a truncated Dps variant missing most of N motif 1 ($Dps^{6-167}$) was resistant to ClpXP degradation (Figure 5A). Thus, the N-terminal residues of Dps are required for its degradation by ClpXP. These residues are absent in the Dps crystal structure (Grant et al., 1998), suggesting that they are unstructured and would therefore be accessible to ClpX. A deletion variant of λO missing N motif 1 is also less susceptible to ClpXP degradation (Gonciarz-Swiatek et al., 1999), supporting a role for this sequence in ClpX recognition.

To establish that N motif 1 functions as a ClpX-recognition signal in vivo, we coexpressed $ClpXP^{trap}$ with either Dps or $Dps^{6-167}$. As expected, full-length Dps copurified with $ClpP^{trap}$ (Figure 5B), but the truncated variant, $Dps^{6-167}$, did not (Figure 5B). These data demonstrate that N motif 1 is essential for Dps-ClpX interactions in the cell.

To determine the sufficiency of the N motif 1 sequence for ClpXP degradation, we constructed Arc fusion proteins containing the first 12 residues of Dps or λO. Following cellular removal of the N-terminal methionines, the purified proteins produced were $Dps^{2-12}$-Arc and $\lambda O^{2-12}$-Arc. ClpXP degraded both fusion proteins in vitro at rates similar to those observed for full-length Dps

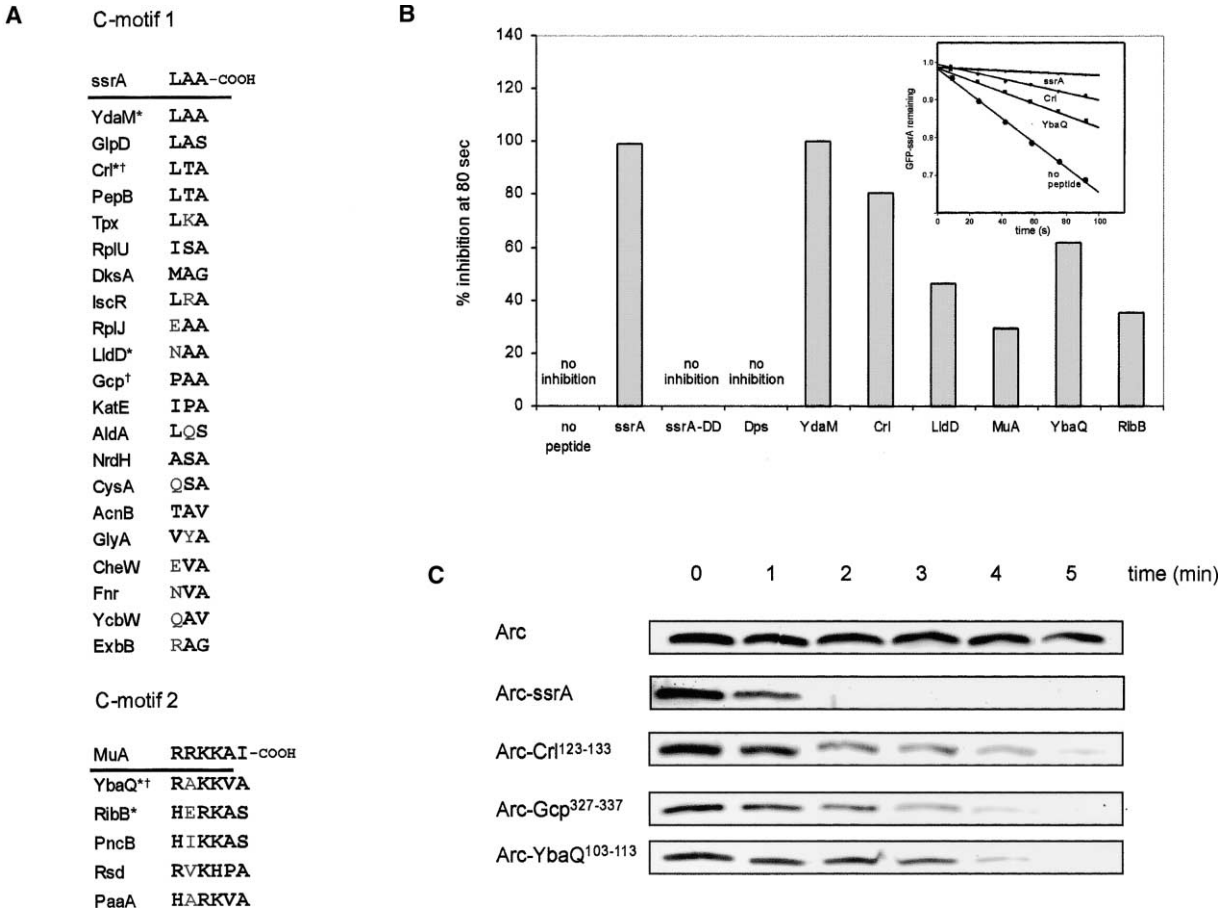



Figure 3. C-terminal Recognition Signals in Trapped Proteins

(A) Sequence similarities of trapped proteins with the ssrA tag (C motif 1) and MuA (C motif 2). Dissimilar amino acids are shadowed in gray. *, proteins whose corresponding C-terminal peptides inhibit ClpXP degradation of GFP-ssrA; †, proteins whose C-terminal peptides target Arc-fusion proteins for ClpXP degradation.

(B) ClpXP degradation of GFP-ssrA in the presence of C-terminal peptides. Bars indicate percent inhibition after 80 s of degradation from experiments like those shown in inset. Peptide sequences were ssrA (CAANDENYALAA), ssrA-DD (CAANDENYALDD), Dps (CFLWFIESNIE), YdaM (CKNDGRNRVLAA), Crl (CDFRDEPVKLTA), LldD (CALAPMAKGNAA), MuA (CILEQNRRKKAI), YbaQ (CARREERAKKVA), and RibB (CAYR QAHERKAS).

(C) ClpXP degradation of Arc fusion proteins with the ssrA tag or C-terminal residues of Crl (FRDEPVKLTA), Gcp (RWPLAELPAA), and YbaQ (RREERAKKVA) assayed by SDS-PAGE.

and λO (Figures 5A and 6A). Thus, the N-terminal regions of Dps and λO contain sequences that are both necessary and sufficient to target proteins for degradation by ClpXP. Next, we mutated conserved residues in N motif 1. $Dps^{2-12}$-Arc fusion proteins containing Asp substitutions for Thr3, Lys5, or Leu6 were degraded significantly less efficiently by ClpXP than the parental Dps-Arc fusion (Figure 6A). These data establish that several of the conserved residues in N motif 1 are important for its function as a ClpX-recognition signal.

DadA, IscS, OmpA, and nine additional proteins shared N-terminal sequences matching the pattern $NH_2$-Met-basic-$\phi$-$\phi$-$\phi$-$X_5$-$\phi$ (N motif 2 in Figure 4B; Table 1). Adding either the $OmpA^{1-11}$ or $IscS^{1-11}$ sequences to the N terminus of Arc converted it into a substrate for ClpXP degradation (Figure 6B). Mutating Lys2 or Ile5 of the $IscS^{1-12}$ sequence to Asp abolished detectable degradation of the fusion protein, showing that these residues are essential for ClpX recognition of this sequence motif (Figure 6B).

Ten other proteins, including Crl and DksA, contained N-terminal sequences that generally fit the consensus $\phi$-X-polar-X-polar-X-basic-polar (N motif 3 in Figure 4B; Table 1). When $DksA^{1-12}$, a representative sequence containing this motif, was fused to Arc, the resulting protein was degraded by ClpXP (Figure 6C), although less rapidly than fusion proteins carrying N motif 1 or N motif 2 signals. Thus, representative sequences containing each of the three N motifs were sufficient to confer susceptibility to degradation by ClpXP.

## Discussion

### Substrate Discovery through Intracellular Trapping

Targeted protein degradation in bacteria is a dynamic process in which substrates of proteases like ClpXP

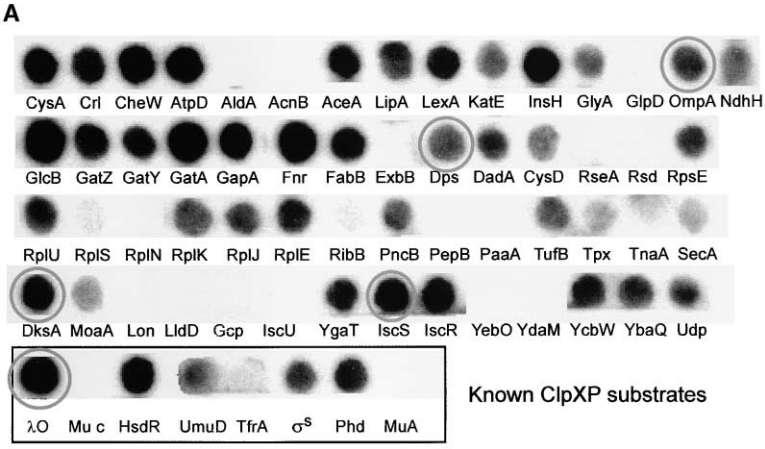
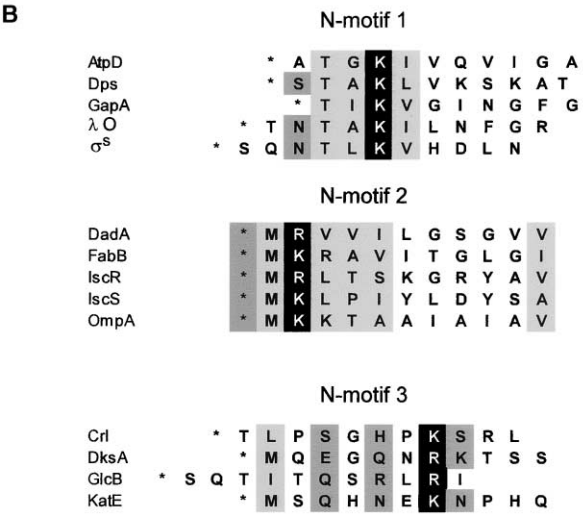



Figure 4. N-Terminal Recognition Signals

(A) A filter with covalently bound peptides corresponding to the N-terminal 11 residues of trapped proteins and known ClpXP substrates was incubated with ClpX, and bound protein was detected as in a Western blot (see Experimental Procedures). Removal of the N-terminal Met was assumed for proteins with Ala, Ser, Thr, or Gly at position 2 and peptides corresponding to residues 2–12 of the unprocessed molecule (Ben-Bassat et al., 1987). Peptides shown to target fusion proteins for ClpXP degradation are circled.

(B) Many ClpX binding sequences contain one of three motifs: N motif 1: polar-T/$\phi$-$\phi$-basic-$\phi$; N motif 2: $NH_2$-Met-basic-$\phi$-$\phi$-$\phi$-$X_5$-$\phi$; or N motif 3: $\phi$-X-polar-X-polar-X-basic-polar. Additional members of each group are listed in Table 1. Asterisks correspond to the $\alpha$-amino group.




change as cells respond to shifts in nutrients and to environmental stress. As a result, studying the full impact of degradation on the bacterial proteome requires methods for identifying protease substrates under a variety of environmental conditions. Here, we have described the use of an inactive, epitope-tagged variant of the ClpP protease as an intracellular trap for ClpXP substrates. Following capture and affinity purification, tandem-mass spectrometry identified more than 50 *E. coli* proteins. Similar strategies could be applied to identify protein targets of ClpXP under different growth conditions in *E. coli* or in other bacteria. Similar methods should also work to identify substrates of the ClpAP, HslUV, and Lon proteases.

Several observations support the conclusion that most $ClpXP^{trap}$-captured proteins are authentic ClpXP substrates. First, their capture by $ClpP^{trap}$ depended on the presence of ClpX. Second, two known ClpXP substrates—$\sigma^S$ and GFP-ssrA—were captured. Third, five newly identified trapped proteins (DksA, Dps, Fnr, and two fragments of LexA) were subsequently shown to be substrates for ClpXP degradation. Fourth, the majority of $ClpXP^{trap}$-captured proteins displayed C-terminal and/or N-terminal peptide sequences that bound to ClpX or were very similar to known recognition signals, and seven of the peptides identified in this manner were shown to target fusion proteins for ClpXP degradation. This collection of proteins captured by $ClpXP^{trap}$ represents a large increase in the number of known ClpXP substrates.

For a few $ClpXP^{trap}$-associated proteins, the relevance to ClpXP-mediated degradation was uncertain. For example, DnaK was also associated with $ClpP^{trap}$ in the absence of ClpX. Because DnaK binds unfolded proteins (Pelham, 1986), we assume that it binds denatured or unassembled $ClpP^{trap}$ subunits. Hence, we have no evidence that DnaK is a ClpXP substrate. For OmpA, questions arose because the captured protein is normally located in another compartment, the outer membrane. OmpA is highly expressed, however, and may saturate the SecA-mediated secretion pathway under some circumstances; ClpXP degradation of this cytoplasmic OmpA could play a role in protein quality control. For RseA, we found that $ClpXP^{trap}$ captured an N-terminal fragment corresponding to its cytoplasmic domain, whereas neither its C-terminal periplasmic domain nor

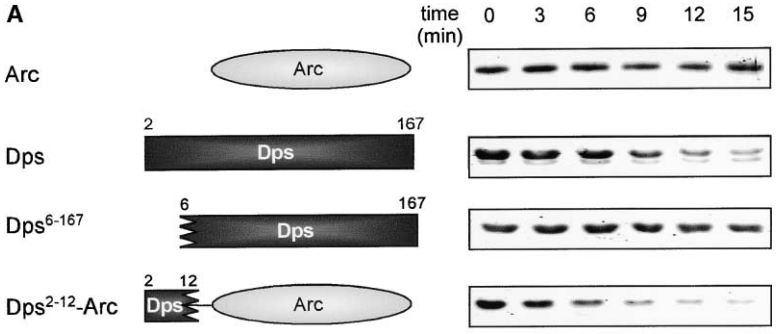
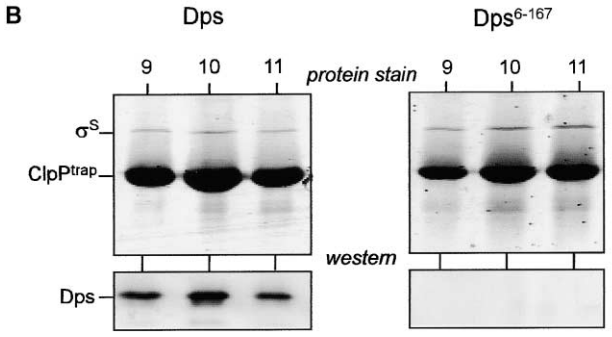



Figure 5. Dps Has an N-Terminal Degradation Signal

(A) ClpXP degradation of full-length Dps, full-length Arc, $Dps^{6-167}$, or $Dps^{2-12}$-Arc assayed by SDS-PAGE.

(B) Purification of $ClpP^{trap}$ complexes formed in strains expressing Dps or $Dps^{6-167}$. $ClpP^{trap}$ was purified by Ni-NTA followed by gel filtration. The three peak $ClpP^{trap}$ fractions (9–11) are shown: (upper panel) stained with Sypro orange; (lower panel) probed with anti-Dps antibody. Note the presence of $\sigma^S$ in the upper panel confirms that trapping occurred efficiently in both strains.




the full-length protein, which spans the inner membrane, was trapped. Specific trapping of this N-terminal RseA domain supports a model proposed by Alba et al. (2002) in which ClpXP-mediated degradation of the N-terminal domain of RseA requires prior cleavage of RseA by inner-membrane proteases.

Seven proteins captured by $ClpXP^{trap}$ had masses ranging from 50 to 102 kDa even though structural calculations suggest the ClpP chamber can only accommodate globular proteins as large as 50 kDa (Wang et al., 1997; Ortega et al., 2000). How might these larger proteins be trapped? EM images of $ClpXP^{trap}$-substrate complexes reveal substrate density both within the ClpP chamber and at the axial ends of ClpXP particles (Ortega et al., 2002), suggesting that captured proteins can be associated with $ClpP^{trap}$ with only a portion of the substrate inside the chamber.

### Molecular Definition of ClpX-Recognition Motifs

Identification of cellular proteins captured by $ClpXP^{trap}$ led to the discovery of five peptide motifs that target proteins for ClpXP degradation. Overall, nearly 90% of the proteins captured by $ClpXP^{trap}$ contain sequences that are attractive candidates for ClpX-recognition signals. Twenty-six of the captured proteins have C-terminal sequences that are plausible sites of ClpX interaction based on their similarities to known recognition signals, peptide-inhibition studies, and fusion protein analysis. These sequences fall into two classes; C motif 1 is ssrA-like, and C motif 2 is more similar to the MuA-recognition sequence. In addition, 40 of the captured proteins have N-terminal peptides that bound ClpX on a peptide array. Alignments of the N-terminal ClpX binding sequences reveal three peptide motifs. Representative sequences from each of these motifs convert an attached protein into a ClpXP substrate, demonstrating that these sequences are functional ClpX-recognition signals. Single point mutations in highly conserved motif residues also stabilize these fusion proteins, confirming the importance of these determinants for recognition. Thus, analyzing a large group of new ClpXP substrates has allowed us to define sequence rules governing substrate choice.

The ClpX-recognition motifs were clearly enriched in the trapped population of proteins compared to the entire proteome. For example, the percentage of trapped proteins terminating with the dipeptide Ala-Ala-COOH (the critical region of C motif 1) was enriched 7-fold. N motif 1 is the most defined of the three N-terminal recognition motifs. A strict consensus for this motif — $T^1$-$X^2$-$K^3$-[ILV]$^4$ located from 1 to 4 residues from the N terminus—is present in the trapped protein population at a 10-fold higher frequency than in the proteome. Despite inherent uncertainties about whether these sequences will be accessible or functional in any specific protein, the identification of five classes of defined ClpX-recognition signals provides a useful foundation for the bioinformatic identification of other likely ClpX substrates.

In bacteria, many proteins are degraded by more than one protease. For example, ssrA-tagged proteins are degraded by ClpXP, ClpAP, and FtsH, whereas SulA is

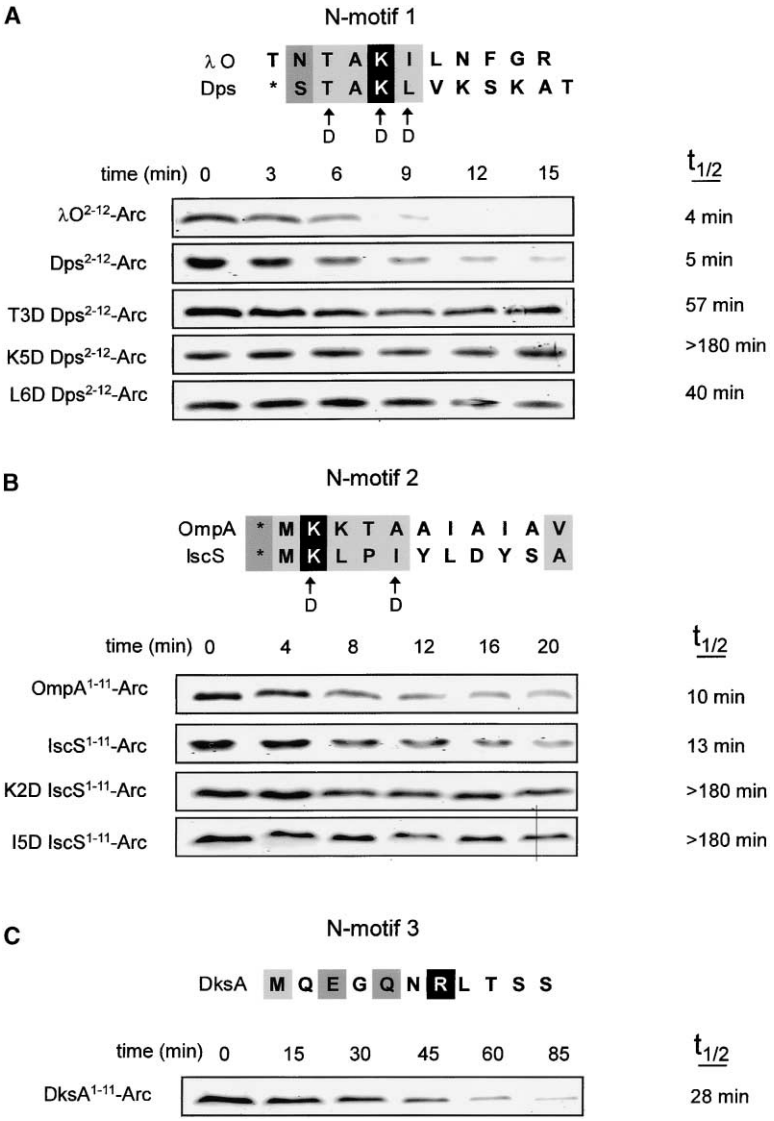



Figure 6. ClpXP Degradation of Arc-Fusion Proteins with Wild-Type or Mutant N-Terminal Recognition Signals

Degradation of each protein (5 μM) was assayed by SDS-PAGE, and half-lives ($t_{1/2}$) were determined from plots of intensity versus time.

degraded by HslUV (ClpYQ) and Lon (Gottesman et al., 1998; Herman et al., 1998; Wu et al., 1999). Some of the new ClpXP substrates identified here are also substrates for other proteases. For example, the C-terminal autocleavage fragment of LexA is degraded by ClpXP (our unpublished data) but is also a substrate for the Lon protease (Little, 1983). Likewise, both ClpXP and other proteases appear to contribute to the degradation of Dps and DksA. Finally, a preliminary analysis of the proteins captured by $ClpP^{trap}$ in a strain expressing ClpA but not ClpX indicates that ClpAP recognizes about ten proteins that are also recognized by ClpXP.

How most shared substrates are recognized by multiple proteases is not presently known. In the case of ssrA-tagged proteins, it has been established that the same 11 residue peptide targets them to ClpXP and to ClpAP, but it is also known that these proteases recognize different sets of amino acid residues within this peptide (Gottesman et al., 1998; Flynn et al., 2001). We believe that it is also likely that ClpXP and ClpAP will recognize nonidentical recognition signals in other shared substrates. Current evidence supports the idea that the precise peptide motifs that target proteins for degradation by ClpXP and ClpAP are different. For example, the shared substrates identified include proteins with N motif 1 and N motif 2 (see Table 1), but most N motif 1 or N motif 2 proteins are not common substrates. Furthermore, in vitro degradation experiments demonstrate that Dps, which is recognized by ClpXP via N motif 1, is not degraded by ClpAP (our unpublished data), indicating that this signal is not recognized by both proteases. Similarly, ClpA does not recognize C motif 1 in the ssrA tag or C motif 2 in MuA (Flynn et al., 2001; I. Levchenko and T.A.B., unpublished data), and thus it is unlikely to directly recognize similar sequence motifs in other proteins.

In some instances, a ClpX-recognition signal normally located at a protein terminus can also function at some internal positions (Hoskins et al., 2002). However, analysis of previously characterized substrates and those described here suggests that ClpX-recognition signals are most commonly found near either the N terminus or

C terminus of a protein. This localization is probably explained by the observation that these regions are frequently accessible in native proteins. Moreover, the free α-amino and α-carboxyl groups at the protein termini provide additional unique recognition determinants.

For LexA repressor, there is good evidence that an efficient ClpX binding sequence is not recognized in the context of the full-length native protein. LexA contains an N motif 2 sequence, which bound ClpX on the peptide array, but full-length LexA was neither captured by $ClpXP^{trap}$ nor degraded by ClpXP (our unpublished data). Inspection of the LexA crystal structure shows that portions of its N-terminal motif are buried in the native protein (Luo et al., 2001). In fact, for LexA and for RseA, accessible ClpXP recognition signals appear only to be produced following initial cleavage by other proteases. Recognition of cryptic peptide signals that are exposed as a result of polypeptide cleavage or protein denaturation probably represents a general strategy used by ClpX to interact with some substrates. This may explain why some captured proteins lacked recognizable N-terminal or C-terminal ClpX binding motifs (see Table 1).

About one quarter of the captured proteins contain potential ClpX-recognition signals at both the N terminus and C terminus. In these cases, both signals might be utilized for ClpXP degradation or one or the other might be more accessible in the native protein or in protein complexes and therefore be used to a greater extent. In fact, precedence for multiple signals contributing to a protein's recognition by ClpX is evident from deletion analysis of the λO protein, which reveals that information located near both its N- and C termini contributes to the efficiency of its degradation (Gonciarz-Swiatek et al., 1999). Even though some $ClpXP^{trap}$-captured proteins appear to have recognition signals at both termini, it seems unlikely that two ClpXP enzymes would ever degrade a single substrate from both ends, because the recognition signals bind rather weakly to ClpX hexamers and thus the probability that two ClpXP enzymes would simultaneously engage one substrate molecule is very low.

This study has revealed the presence of five classes of ClpX-recognition signals. In addition, one protein whose N-terminal peptide bound ClpX did not contain a recognizable motif, suggesting that there may be additional classes of signals. Why are there so many different types of signals? One attractive model is that signal diversity allows differential regulation of protein degradation. For example, proteins that bind specifically to one of the recognition motifs could specifically repress ClpXP degradation of these proteins but not those bearing other signals. As some single proteins appear to carry distinct classes of recognition signals, possibilities for combinatoral control of protein turnover are also present. It is common for multiple regulatory proteins to work together to control gene expression, and similar strategies could also help to regulate the precise composition of the proteome by degradation.

### Trapped Proteins and Roles for ClpXP-Mediated Degradation

Many of the proteins captured by $ClpXP^{trap}$ are coregulated in response to cellular stress and changes in environment. For example, our analysis suggests that ClpXP degrades a set of proteins that are active during stationary phase. Five trapped proteins (Rsd, Dps, KatE, FtsZ, and GlpD) were encoded by genes transcribed under control of the stationary-phase $\sigma^S$ factor, two additional captured proteins (Crl and DksA) have been implicated in controlling the level of $\sigma^S$, and $\sigma^S$ itself represented one of the major trapped proteins (see Hengge-Aronis, 1996, and references therein; also see Pratt and Silhavy, 1998; Jishage and Ishihama, 1998; Webb et al., 1999). ClpXP is known to regulate $\sigma^S$ levels by degrading it during exponential phase but not during stationary phase (Schweder et al., 1996). Our experiments indicate that Dps and DksA are degraded by ClpXP as cells recover from stationary phase and re-enter logarithmic growth. Hence, ClpXP appears to regulate the levels of some stationary-phase proteins by direct degradation as well as by degrading $\sigma^S$.

Many proteins trapped by ClpXP help cells cope with oxidative stress and shifts between aerobic and anaerobic growth. Nine of the trapped proteins—Fnr, AceA, AcnB, AldA, GlcB, GlpD, MoaA, Tpx, and LldD—are encoded by genes regulated by the anoxic transcriptional regulatory proteins Fnr and/or ArcA (see Lynch and Lin, 1996, and references therein; also see Kim et al., 1999; Pellicer et al., 1999a, 1999b; Anderson et al., 2000). Some oxidative stress probably occurred during our trapping experiments, as aerobic metabolism reduces $O_2$ to reactive species. Six trapped proteins—Fnr, IscR, IscU, AcnB, MoaA, and LipA—contain Fe-S centers, which can serve as sensors of oxidative stress. For example, the Fe-S cluster of Fnr is oxidized during aerobic growth (Kiley and Beinert, 1998), reducing Fnr activity and potentially enhancing its degradation by ClpXP. Based on these initial studies, ClpXP may degrade proteins whose Fe-S clusters have been damaged by oxidation as a general response to oxidative stress.

Six ribosomal proteins were captured by $ClpXP^{trap}$. Why should proteins—such as ribosomal proteins—that are generally long lived, be ClpXP substrates? Ribosome populations are reduced following a nutritional downshift (Davis et al., 1986), and ClpXP may degrade ribosomes when nutrients become limiting, releasing amino acids for new protein synthesis. It is possible that ribosome turnover had begun when cells were harvested for our trapping studies during late exponential growth. Alternatively, ClpXP may degrade unassembled ribosomal proteins or damaged subunits. In fact, we suspect that for a number of substrates, ClpXP may function to degrade only a fraction of the protein population, depending upon damage, assembly state, or growth conditions.

The definition of ClpX-recognition signals and the apparent role of ClpXP degradation in a variety of stress responses provide a foundation for understanding strategies for regulating protein turnover. Because peptide signals are critical for degradation, the use of signal binding partners that mask or enhance substrate recognition by ClpX is one useful regulatory strategy. Regulating the availability of cryptic recognition signals provides another way to control degradation in response to environmental change. For example, denaturation of proteins during heat shock or following initial cleavage by other proteases could expose latent ClpX-recognition

sequences. Identification of additional ClpXP substrates under a broad range of environmental conditions should permit further definition of the molecular mechanisms that contribute to the cellular control of targeted protein degradation.

## Experimental Procedures

### Solutions

The following solutions were used: TBS: 50 mM Tris-HCl (pH 7.5) and 125 mM NaCl; ClpX buffer: 50 mM HEPES-KOH (pH 7.5), 150 mM KCl, 5 mM $MgCl_2$, 100 $\mu$M $ZnSO_4$, and 2 mM DTT; PBS: 150 mM NaCl, 20 mM, Na-phosphate (pH 7.3); TEV buffer: 50 mM Tris-HCl (pH 8.0), 0.5 mM EDTA, and 1 mM DTT. PD buffer, S buffer, W20 buffer, Clp buffer, and W500 buffer are as described (Kim et al. 2000).

### Proteins

Dps and $Dps^{6\text{-}167}$ (Grant et al., 1998), GFP-ssrA (Yakhnin et al., 1998), ClpP (Kim et al., 2000), and Arc derivatives (Arc-st11 and the fusions) (Robinson and Sauer, 1996) were purified as described. ClpX was purified using standard chromatographic methods; the protocol is available upon request.

### Strains and Plasmids

*E. coli* strains were grown in LB broth. The W3110 *clpP::cat* Δ*smpB-1*, W3110 *clpP::cat clpA::kan* Δ*smpB*, and W3110 *clpP::cat clpX::kan* Δ*smpB* strains were derived from W3110 Δ*smpB-1* (Karzai et al., 1999). From this strain, additional protease mutations (*clpA::kan, clpX::kan, and clpP::cat*) were introduced by P1 transduction. To generate the MC4100 *clpX::kan clpP::cat clpA::kan* strain, the *clpP::cat* allele was transduced into SG22178.

A plasmid expressing ClpP without the propeptide sequence (Δ1–13) was constructed by PCR amplification of the *clpP* gene, cleavage with SphI and BglII, and cloning into the SphI-BglII fragment of QE-70. The active-site S97A mutation was introduced using Quikchange (Qiagen) and appropriate primers to generate pYK162. The $Myc_3$-TEV-$His_6$ sequence was introduced on an oligonucleotide cassette between the BglII and HindIII sites of pYK162 to produce pJF105. The C-terminal appended tag is: DSILTHRNRS HHHHHH GGEN LYFQGAYTSG EQKLISEEDL NGEQKLISEE DLNGEQKLIS EE DLN. Strains used for trapping were: JF148 (MC4100 *clpX::kan clpP::cat clpA::kan/pJF105*), JF176 (W3110 *clpP::cat* Δ*smpB-1/pJF105*), JF172 (W3110 *clpP::cat clpX::kan* Δ*smpB-1/pJF105*), and JF162 (W3110 *clpP::cat clpA::kan* Δ*smpB-1/pJF105*).

A plasmid expressing $Dps^{6\text{-}167}$ was constructed by PCR amplification from strain SK101 (Martinez and Kolter, 1997), cleavage with NdeI and BamHI, and cloning into the NdeI-BamHI fragment of pET3a (Novagen). A plasmid expressing *arc-st11* in pET-11a was constructed by PCR amplification of pET-28b-Arc-ssrA (Burton et al., 2001) and ligation into the NheI-BamHI fragment of pET-11a (Novagen) to form pET-11a-Arc-st11. The first 12 residues of Dps and λO and the first 11 residues of IscS, OmpA, and DksA were fused to Arc-st11 by using oligonucleotide cassettes. The mature N-terminal sequences of the fusion proteins are: $Dps^{2\text{-}12}$-Arc: *STAKLVKSK AS*MGK; $\lambda O^{2\text{-}12}$-Arc: *TNTAKILNF GRA*SMGK; $IscS^{1\text{-}11}$-Arc: *MKLPIYLDY S AS*MGK; $OmpA^{1\text{-}11}$-Arc: *MKKTAIAIAV AS*MGK; $DksA^{1\text{-}11}$-Arc: *MQEGQNRKTS S*MGK (Dps, λO, IscS, OmpA, and DksA are in italics; Arc is underlined). The T3D, K5D, and L6D $Dps^{2\text{-}12}$-Arc mutants and the K2D and I5D $IscS^{1\text{-}11}$-Arc mutants were constructed using oligonucleotide cassettes. The C-terminal 10 residues of Crl, Gcp, and YbaQ were fused to Arc-st11 by PCR amplification of the Arc-st11 gene with primers containing the C-terminal sequence of each respective protein and ligation into the NheI-BamHI fragment of pET11a. The sequences of the C-terminal regions of the resulting fusion proteins are: Arc-$YbaQ^{103\text{-}113}$: *QHD*RREERA KKVA; Arc-$Crl^{123\text{-}133}$: *QHD*FRDEPV KLTA; Arc-$Gcp^{327\text{-}337}$: *QHD*RWPLAE LPAA. All constructs were confirmed by DNA sequencing.

A plasmid expressing Dps under control of the arabinose promoter (pJF119) was constructed by removal of the *dps* and *araC* genes from pBAD18-dps (Martinez and Kolter, 1997) and cloning into the AvaI-HindIII fragment of pSU38. $Dps^{6\text{-}167}$-pSU38 was constructed by PCR amplification of $dps^{6\text{-}167}$ from the $dps^{6\text{-}167}$-pET3a plasmid and ligation into the EcoRI/XbaI fragment of pBAD18. The $dps^{6\text{-}167}$ and *araC* genes were cut from the resulting plasmid and cloned into the AvaI-HindIII fragment of pSU38 to form pJF121. pJF119 and pJF121 were then transformed into JF176.

### Protein Trapping In Vivo

Strains JF148, JF162, JF172, and JF176 were grown in 4 liters of LB/amp at 30°C to an $A_{600}$ of 0.4, induced with 0.5 mM IPTG, and grown for 2.5 additional hr. Cells were harvested by centrifugation and resuspended in 3 ml S buffer per gram of cells. Following lysis by French press, the lysate was centrifuged for 30 min at 25,000 × g, and the supernatant was added to 2.5 ml nickel-NTA resin (Qiagen) equilibrated in S buffer. After mixing for 2 hr at 4°C, the resin was packed into a column, washed with 200 ml S buffer, 100 ml W20 buffer, and eluted with 5 ml W500 buffer. The Myc antibody affinity resin was generated by cross-linking 9E10 antibody to protein G agarose (Invitrogen) as described (Harlow and Lane, 1988). The elutant from the nickel-NTA column was mixed with 1.5 ml of this resin equilibrated in PBS. After mixing for 2 hr at 4°C, the beads were packed into a column and washed with 60 ml PBS, followed by 60 ml PBST (PBS + 0.1% Tween 20), and finally by 20 ml TEV reaction buffer. The slurry was then mixed with 1 ml TEV reaction buffer and 400 units of TEV protease (GIBCO),and agitated at room temperature for 30 min. The released protein was collected and stored at −20°C.

### Trapping of Dps and $Dps^{6\text{-}167}$ In Vivo

Dps or $Dps^{6\text{-}167}$ was coexpressed with $ClpP^{trap}$ under the same conditions as above, by the addition of 0.2% L-arabinose at the same time as the IPTG. $ClpP^{trap}$ complexes were purified on a Ni-NTA column as above followed by filtration chromatography on a Superdex 200 PC 3.2/30 column run in Clp buffer.

### 2D Gels

Samples for 2D gel analysis were exchanged into 8 M urea and 2% CHAPS and loaded on a 7 cm Immobiline DryStrip (pH 3–10L) for focusing on a IPGphor system (Pharmacia), followed by 12.5% SDS-PAGE (Bjellqvist et al., 1993). Spots were visualized using Sypro Ruby protein stain (Molecular Probes) on a Fluorimager 595 (Molecular Dynamics).

### Mass Spectrometry

Samples for MS/MS analysis were separated by 12.5% SDS-PAGE. Gel slices (approximately 0.5–1.0 cm) were excised, digested with trypsin, and analyzed by microcapillary reverse-phase HPLC nanoelectrospray tandem mass spectrometry using a Finnigan LCQ DECA quadropole ion trap mass spectrometer (Harvard Microchemistry Facility). Control analyses performed on samples purified from the $clpX^{-}clpA^{-}$ strain yielded peptides from ClpP, TEV protease, and keratin, as well as four peptides of Dps. The presence of this small number of Dps peptides was probably an artifact due to purification of Dps in the laboratory during sample preparation; Western analysis failed to detect any Dps in this sample (see Figure 2A).

### Degradation In Vivo

Cultures of W3110 or W3110 *clpX::kan* cells were grown overnight in LB broth at 37°C ($A_{600} \approx 3$), diluted 1:100 in fresh LB broth, and allowed to grow for 50 min at 37°C ($A_{600} \approx 0.1$). At this point, 150 μg/ml of spectinomycin was added. Samples were removed at specific times and analyzed by SDS-PAGE, followed by Western blotting (see below).

### Western Blots

Western blots were performed following the guidelines of Amersham for use with the ECF substrate (Amersham) using the following primary antibodies: anti-Dps (from Richard Bugess, University of Wisconsin, Madison), anti-LexA (from John Little, University of Arizona), anti-Rsd (from Akira Ishihama, National Institute of Genetics), anti-DksA (from Diana Downs, University of Wisconsin, Madison), or anti-N domain RseA and anti-C domain RseA (from Carol Gross, UCSF).

### Degradation In Vitro

$ClpX_6$ (0.3 μM), $ClpP_{14}$ (0.8 μM), ATP (4 mM), and an ATP regeneration system (50 μg/ml creatine kinase and 2.5 mM creatine phosphate) were mixed in PD buffer and incubated for 2 min at 30°C. For all degradation experiments, 5 μM of protein was added, and samples were removed at specific times and analyzed by SDS-PAGE. For peptide-inhibition experiments, GFP-ssrA (1 μM) was added with peptide (50 μM), and degradation was monitored by fluorescence as described (Flynn et al., 2001).

### Peptide Arrays

A cellulose filter containing peptides corresponding to the 11 N-terminal residues of all the trapped proteins (except GroEL, FtsZ, ClpX, and DnaK) and known ClpXP substrates was prepared by the MIT Biopolymers facility using an Abimed instrument. Each peptide contained two additional C-terminal β-alanines and was covalently attached to the filter by a polyethylene glycol linker. The filter was soaked in ethanol, washed three times for 5 min in TBST (TBS + 0.1% Tween 20), blocked overnight in TBST plus 10% milk, and then washed twice with TBST and twice in ClpX buffer for 5 min. $ClpX_6$ (0.8 μM) and ATPγS (4 mM) (Roche) were incubated at 30°C in 5 ml ClpX buffer for 2 min and added together with 0.1% milk to the filter for 6 hr at 4°C. The filter was washed three times with ClpX buffer and ATPγS (0.5 mM) and incubated with anti-ClpX antibody in 5 ml ClpX buffer and ATPγS (1 mM) for 30 min. Next, the filter was washed three times as above and incubated with goat anti-rabbit IgG HRP-conjugated antibody (Amersham) and ATPγS (1 mM) for 20 min. After three final washes, the filter was incubated with ECL substrate (NEN) and visualized on film.

### Acknowledgments

This work was supported by NIH grant AI-16892 and HHMI. T.A.B. is an employee of HHMI. We thank Richard Burgess, John Little, Akira Ishihama, Diana Downs, and Carol Gross for antibodies, Steven Finkel for strains and advice, Patricia Kiley for communication of unpublished work, Igor Levchenko, Shari Specter, and Anthony Schwacha for technical expertise and discussions, and the Harvard Microchemistry Facility for mass spectrometry. We thank Deborah Siegele and members of the Baker and Sauer labs for help, advice, and comments on this manuscript.

Received: July 30, 2002
Revised: January 8, 2003

### References

Alba, B.M., Leeds, J.A., Onufryk, C., Lu, C.Z., and Gross, C.A. (2002). DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma(E)-dependent extracytoplasmic stress response. Genes Dev. *16*, 2156–2168.

Almiron, M., Link, A.J., Furlong, D., and Kolter, R. (1992). A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. Genes Dev. *6*, 2646–2654.

Anderson, L.A., McNairn, E., Leubke, T., Pau, R.N., and Boxer, D.H. (2000). ModE-dependent molybdate regulation of the molybdenum cofactor operon moa in *Escherichia coli*. J. Bacteriol. *182*, 7035–7043.

Ben-Bassat, A., Bauer, K., Chang, S.Y., Myambo, K., Boosman, A., and Chang, S. (1987). Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. J. Bacteriol. *169*, 751–757.

Bjellqvist, B., Pasquali, C., Ravier, F., Sanchez, J.C., and Hochstrasser, D. (1993). A nonlinear wide-range immobilized pH gradient for two-dimensional electrophoresis and its definition in a relevant pH scale. Electrophoresis *14*, 1357–1365.

Burton, R.E., Siddiqui, S.M., Kim, Y.I., Baker, T.A., and Sauer, R.T. (2001). Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine. EMBO J. *20*, 3092–3100.

Damerau, K., and St. John, A.C. (1993). Role of Clp protease subunits in degradation of carbon starvation proteins in *Escherichia coli*. J. Bacteriol. *175*, 53–63.

Davis, B.D., Luger, S.M., and Tai, P.C. (1986). Role of ribosome degradation in the death of starved *Escherichia coli* cells. J. Bacteriol. *166*, 439–445.

Flynn, J.M., Levchenko, I., Seidel, M., Wickner, S.H., Sauer, R.T., and Baker, T.A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. Proc. Natl. Acad. Sci. USA *98*, 10584–10589.

Frank, E.G., Ennis, D.G., Gonzalez, M., Levine, A.S., and Woodgate, R. (1996). Regulation of SOS mutagenesis by proteolysis. Proc. Natl. Acad. Sci. USA *93*, 10291–10296.

Gonciarz-Swiatek, M., Wawrzynow, A., Um, S.J., Learn, B.A., McMacken, R., Kelley, W.L., Georgopoulos, C., Sliekers, O., and Zylicz, M. (1999). Recognition, targeting, and hydrolysis of the λ O replication protein by the ClpP/ClpX protease. J. Biol. Chem. *274*, 13999–14005.

Gonzalez, M., Rasulova, F., Maurizi, M.R., and Woodgate, R. (2000). Subunit-specific degradation of the UmuD/D′ heterodimer by the ClpXP protease: the role of trans recognition in UmuD′ stability. EMBO J. *19*, 5251–5258.

Gottesman, S. (1996). Proteases and their targets in *Escherichia coli*. Annu. Rev. Genet. *30*, 465–506.

Gottesman, S., Clark, W.P., de Crecy-Lagard, V., and Maurizi, M.R. (1993). ClpX, an alternative subunit for the ATP-dependent Clp protease of *Escherichia coli*. Sequence and *in vivo* activities. J. Biol. Chem. *268*, 22618–22626.

Gottesman, S., Roche, E., Zhou, Y., and Sauer, R.T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. Genes Dev. *12*, 1338–1347.

Grant, R.A., Filman, D.J., Finkel, S.E., Kolter, R., and Hogle, J.M. (1998). The crystal structure of Dps, a ferritin homolog that binds and protects DNA. Nat. Struct. Biol. *5*, 294–303.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., and Maurizi, M.R. (1998). Enzymatic and structural similarities between the *Escherichia coli* ATP-dependent proteases, ClpXP and ClpAP. J. Biol. Chem. *273*, 12476–12481.

Harlow, E., and Lane, D. (1988). Antibodies, A Laboratory Manual (Plainview, NY: Cold Spring Harbor Laboratory Press).

Hengge-Aronis, R. (1996). Regulation of gene expression during entry into stationary phase. In *Escherichia coli* and *Salmonella*, F.C. Neidhardt, ed. (Washington, D.C.: American Society for Microbiology), pp. 1497–1512.

Herman, C., Thevenet, D., Bouloc, P., Walker, G.C., and D'Ari, R. (1998). Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH). Genes Dev. *12*, 1338–1347.

Hoskins, J.R., Yanagihara, K., Mizuuchi, K., and Wickner, S. (2002). ClpAP and ClpXP degrade proteins with tags located in the interior of the primary sequence. Proc. Natl. Acad. Sci. USA *99*, 11037–11042.

Jenal, U., and Fuchs, T. (1998). An essential protease involved in bacterial cell-cycle control. EMBO J. *17*, 5658–5669.

Jishage, M., and Ishihama, A. (1998). A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. Proc. Natl. Acad. Sci. USA *95*, 4953–4958.

Kang, P.J., and Craig, E.A. (1990). Identification and characterization of a new *Escherichia coli* gene that is a dosage-dependent suppressor of a dnaK deletion mutation. J. Bacteriol. *172*, 2055–2064.

Karzai, A.W., Susskind, M.M., and Sauer, R.T. (1999). SmpB, a unique RNA-binding protein essential for the peptide-tagging activity of SsrA (tmRNA). EMBO J. *18*, 3793–3799.

Katayama, Y., Gottesman, S., Pumphrey, J., Rudikoff, S., Clark, W.P., and Maurizi, M.R. (1988). The two-component, ATP-dependent Clp protease of *Escherichia coli*. Purification, cloning, and mutational analysis of the ATP-binding component. J. Biol. Chem. *263*, 15226–15236.

Keiler, K.C., Waller, P.R., and Sauer, R.T. (1996). Role of a peptide

tagging system in degradation of proteins synthesized from damaged messenger RNA. Science *271*, 990–993.

Kiley, P.J., and Beinert, H. (1998). Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. FEMS Microbiol. Rev. *22*, 341–352.

Kim, S.J., Han, Y.H., Kim, I.H., and Kim, H.K. (1999). Involvement of ArcA and Fnr in expression of Escherichia coli thiol peroxidase gene. IUBMB Life *48*, 215–218.

Kim, Y.I., Burton, R.E., Burton, B.M., Sauer, R.T., and Baker, T.A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. Mol. Cell *5*, 639–648.

Konieczny, I., and Helinski, D.R. (1997). The replication initiation protein of the broad-host-range plasmid RK2 is activated by the ClpX chaperone. Proc. Natl. Acad. Sci. USA *94*, 14378–14382.

Levchenko, I., Luo, L., and Baker, T.A. (1995). Disassembly of the Mu transposase tetramer by the ClpX chaperone. Genes Dev. *9*, 2399–2408.

Levchenko, I., Yamauchi, M., and Baker, T.A. (1997). ClpX and MuB interact with overlapping regions of Mu transposase: implications for control of the transposition pathway. Genes Dev. *11*, 1561–1572.

Little, J.W. (1983). Variations in the *in vivo* stability of LexA repressor during the SOS regulatory cycle. Paper presented at: Cellular Responses to DNA Damage, April 10–15, 1983.

Little, J.W., Edmiston, S.H., Pacelli, L.Z., and Mount, D.W. (1980). Cleavage of the *Escherichia coli* LexA protein by the RecA protease. Proc. Natl. Acad. Sci. USA *77*, 3225–3229.

Luo, Y., Pfuetzner, R.A., Mosimann, S., Paetzel, M., Frey, E.A., Cherney, M., Kim, B., Little, J.W., and Strynadka, N.C. (2001). Crystal structure of LexA: a conformational switch for regulation of self-cleavage. Cell *106*, 585–594.

Lupas, A., Flanagan, J.M., Tamura, T., and Baumeister, W. (1997). Self-compartmentalizing proteases. Trends Biochem. Sci. *22*, 399–404.

Lynch, A.S., and Lin, E.C.C. (1996). Responses to molecular oxygen. In *Escherichia coli* and *Salmonella*, N.C. Frederick, ed. (Washington, D.C.: American Society for Microbiology), pp. 1526–1538.

Makovets, S., Titheradge, A.J., and Murray, N.E. (1998). ClpX and ClpP are essential for the efficient acquisition of genes specifying type IA and IB restriction systems. Mol. Microbiol. *28*, 25–35.

Martinez, A., and Kolter, R. (1997). Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. J. Bacteriol. *179*, 5188–5194.

Maurizi, M.R., Clark, W.P., Katayama, Y., Rudikoff, S., Pumphrey, J., Bowers, B., and Gottesman, S. (1990a). Sequence and structure of ClpP, the proteolytic component of the ATP- dependent Clp protease of *Escherichia coli*. J. Biol. Chem. *265*, 12536–12545.

Msadek, T., Dartois, V., Kunst, F., Herbaud, M.L., Denizot, F., and Rapoport, G. (1998). ClpP of *Bacillus subtilis* is required for competence development, motility, degradative enzyme synthesis, growth at high temperature and sporulation. Mol. Microbiol. *27*, 899–914.

Muffler, A., Fischer, D., Altuvia, S., Storz, G., and Hengge-Aronis, R. (1996). The response regulator RssB controls stability of the sigma(S) subunit of RNA polymerase in *Escherichia coli*. EMBO J. *15*, 1333–1339.

Ogura, T., and Wilkinson, A.J. (2001). AAA+ superfamily ATPases: common structure–diverse function. Genes Cells *6*, 575–597.

Ortega, J., Singh, S.K., Ishikawa, T., Maurizi, M.R., and Steven, A.C. (2000). Visualization of substrate binding and translocation by the ATP-dependent protease, ClpXP. Mol. Cell *6*, 1515–1521.

Ortega, J., Lee, H.S., Maurizi, M.R., and Steven, A.C. (2002). Alternating translocation of protein substrates from both ends of ClpXP protease. EMBO J. *21*, 4938–4949.

Pelham, H.R. (1986). Speculations on the functions of the major heat shock and glucose- regulated proteins. Cell *46*, 959–961.

Pellicer, M.T., Fernandez, C., Badia, J., Aguilar, J., Lin, E.C., and Baldom, L. (1999a). Cross-induction of glc and ace operons of *Escherichia coli* attributable to pathway intersection. Characterization of the glc promoter. J. Biol. Chem. *274*, 1745–1752.

Pellicer, M.T., Lynch, A.S., De Wulf, P., Boyd, D., Aguilar, J., and Lin, E.C. (1999b). A mutational study of the ArcA-P binding sequences in the aldA promoter of *Escherichia coli*. Mol. Gen. Genet. *261*, 170–176.

Porankiewicz, J., Wang, J., and Clarke, A.K. (1999). New insights into the ATP-dependent Clp protease: *Escherichia coli* and beyond. Mol. Microbiol. *32*, 449–458.

Pratt, L.A., and Silhavy, T.J. (1998). Crl stimulates RpoS activity during stationary phase. Mol. Microbiol. *29*, 1225–1236.

Robinson, C.R., and Sauer, R.T. (1996). Equilibrium stability and sub-millisecond refolding of a designed single-chain Arc repressor. Biochemistry *35*, 13878–13884.

Schirmer, E.C., Glover, J.R., Singer, M.A., and Lindquist, S. (1996). HSP100/Clp proteins: a common mechanism explains diverse functions. Trends Biochem. Sci. *21*, 289–296.

Schweder, T., Lee, K.H., Lomovskaya, O., and Matin, A. (1996). Regulation of *Escherichia coli* starvation sigma factor (sigma s) by ClpXP protease. J. Bacteriol. *178*, 470–476.

Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S., and Maurizi, M.R. (2000). Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. Proc. Natl. Acad. Sci. USA *97*, 8898–8903.

Wang, J., Hartling, J.A., and Flanagan, J.M. (1997). The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. Cell *91*, 447–456.

Wawrzynow, A., Wojtkowiak, D., Marszalek, J., Banecki, B., Jonsen, M., Graves, B., Georgopoulos, C., and Zylicz, M. (1995). The ClpX heat-shock protein of *Escherichia coli*, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone. EMBO J. *14*, 1867–1877.

Webb, C., Moreno, M., Wilmes-Riesenberg, M., Curtiss, R., 3rd, and Foster, J.W. (1999). Effects of DksA and ClpP protease on Sigma S production and virulence in *Salmonella typhimurium*. Mol. Microbiol. *34*, 112–123.

Weber-Ban, E.U., Reid, B.G., Miranker, A.D., and Horwich, A.L. (1999). Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. Nature *401*, 90–93.

Wojtkowiak, D., Georgopoulos, C., and Zylicz, M. (1993). Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of *Escherichia coli*. J. Biol. Chem. *268*, 22609–22617.

Wu, W.F., Zhou, Y., and Gottesman, S. (1999). Redundant *in vivo* proteolytic activities of *Escherichia coli* Lon and the ClpYQ (HslUV) protease. J. Bacteriol. *181*, 3681–3687.

Yakhnin, A.V., Vinokurov, L.M., Surin, A.K., and Alakhov, Y.B. (1998). Green fluorescent protein purification by organic extraction. Protein Expr. Purif. *14*, 382–386.