8 Dunigan, D. D., Dietzgen, R. G., Schoelz, J. E., and Zaitlin, M., Tobacco mosaic virus particles contain ubiquitinated coat protein subunits. Virology *165* (1988) 310–312.
9 Finley, D., and Varshavsky, A., The ubiquitin system: functions and mechanisms. Trends biochem. Sci. *10* (1985) 343–346.
10 Fujiwara, T., Tanaka, K., Orino, E., Yoshimura, T., Kumatori, A., Tamura, T., Chung, C. H., Nakai, T., Yamaguchi, K., Shin, S., Kakizuka, A., Naknishi, S., and Ichihara, A., Proteasomes are essential for yeast proliferation. J. biol. Chem. *265* (1990) 16604–16613.
11 Glotzer, M., Murray, A. W., and Kirschner, M. W., Cyclin is degraded by the ubiquitin pathway. Nature *349* (1991) 132–138.
12 Goebl, M. G., Yochem, J., Jentsch, S., McGrath, J. P., Varshavsky, A., and Byers, B., The yeast cell cycle gene *CDC34* encodes a ubiquitin-conjugating enzyme. Science *241* (1988) 1331–1335.
13 Goff, S. A., and Goldberg, A. L., Production of abnormal proteins in *E. coli* stimulates transcription of *lon* and other heat shocks genes. Cell *41* (1985) 587–595.
14 Gonda, D. K., Bachmair, A., Wünning, I., Tobias, J. W., Lane, W. S., and Varshavsky, A., Universality and structure of the N-end rule. J. biol. Chem. *264* (1989) 16700–16712.
15 Heinemeyer, W., Kleinschmidt, J. A., Saidowsky, J., Escher, C., and Wolf, D. H., Proteinase yscE, the yeast proteasome/multicatalytic-multifunctional proteinase: mutants unravel its function in stress induced proteolysis and uncover its necessity for cell survival. EMBO J. *10* (1991) 555–562.
16 Hershko, A., Ubiquitin-mediated protein degradation. J. biol. Chem. *263* (1988) 15237–15240.
17 Hochstrasser, M., and Varshavsky, A., In vivo degradation of a transcriptional regulator: the yeast α2-repressor. Cell *61* (1990) 697–708.
18 Jentsch, S., McGrath. J. P., and Varshavsky, A., The yeast DNA repair gene *RAD6* encodes a ubiquitin-conjugating enzyme. Nature *329* (1987) 131–134
19 Jentsch, S., Seufert, W., Sommer, T., and Reins, H.-A., Ubiquitin-conjugating enzymes: novel regulators of eukaryotic cells. Trends biochem. Sci. *15* (1990) 195–198.
20 Johnson, E. S., Gonda, D. K., and Varshavsky, A., Cis-trans recognition and subunit-specific degradation of short-lived proteins. Nature *346* (1990) 287–291.
21 Klausner, R. D., and Sitia, R., Protein degradation in the endoplasmic reticulum. Cell *62* (1990) 611–614.
22 Leung, D. W., Spencer, S. A., Cachianes, G., Hammonds, R. G., Collins, C., Menzel, W. J., Barnard, R., Waters, W. J., and Wood, W. I., Growth hormone receptor and serum binding protein: purification, cloning and expression. Nature *330* (1987) 537–543.
23 Matthews, W., Tanaka, K., Driscoll, J., Ichihara, A., and Goldberg, A. L., Involvement of the proteasome in various degradative processes in mammalian cells. Proc. natl Acad. Sci. USA *86* (1989) 2597–2601.
24 Rechsteiner, M., Ubiquitin-mediated pathways for intracellular proteolysis. Ann. Rev. cell. Biol. *3* (1987) 1–30.
25 Rogers, S., Wells, R., and Rechsteiner, M., Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science *234* (1986) 364–368.
26 Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J., and Howley, P. M., The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell *63* (1990) 1129–1136.
27 Seufert, W., and Jentsch, S., Ubiquitin-conjugating enzymes UBC4 and UBC5 mediate selective degradation of short-lived and abnormal proteins. EMBO J. *9* (1990) 543–550.
28 Seufert, W., and Jentsch, S., Nucleotide sequence of two $tRNA^{Arg}$-$tRNA^{ASP}$ tandem genes linked to duplicated *UBC* genes in *Saccharomyces cerevisiae*. Nucl. Acids Res. *18* (1990) 1638
29 Seufert, W., McGrath, J. P., and Jentsch, S., *UBC1* encodes a novel member of an essential subfamily of yeast ubiquitin-conjugating enzymes involved in protein degradation. EMBO J. *9* (1990) 4535–4541.
30 Shanklin, J., Jabben, M., and Vierstra, R. D., Red light-induced formation of ubiquitin-phytochrome conjugates: identification of possible intermediates of phytochrome degradation. Proc. natl Acad. Sci. USA *84* (1987) 359–363.
31 Siegelman, M., Bond, M. W., Gallatin, W. M., St. John, T., Smith, H. T., Fried, V. A., and Weissman, I. L. Cell surface molecule associated with lymphocyte homing is a ubiquitinated branched-chain glycoprotein. Science *231* (19876) 823–829.
32 Wilkinson, K. D., Lee, K., Deshpande, S., Duerksen-Hughes, P., Boss, J. M., and Pohl, J., The neuron-specific PGP 9.5 is a ubiquitin carboxyl-terminal hydrolase. Science *246* (1989) 670–673.
33 Wu, R. S., Kohn, K. W., and Bonner, W. M., Metabolism of ubiquitinated histones. J. biol. Chem. *256* (1981) 5916–5920.
34 Yarden, Y., Escobedo, J. A., Kuang, W. J., Yang-Feng, T. L., Daniel, T. O., Tremble, P. M., Chen, E. Y., Ando, M. E., Harkins, R. N., Francke, U., Fried, V. A., Ullrich, A., and Williams, L. T., Structure of the receptor for platelet-derived growth factor helps define a family of closely related growth factor receptors. Nature *323* (1986) 226–232.

0014-4754/92/020172-07$1.50 + 0.20/0

Proteases and protein degradation in *Escherichia coli*

M. R. Maurizi

Laboratory of Cell Biology, National Cancer Institute, Bethesda (Maryland 20892, USA)

Abstract. In *E. coli*, protein degradation plays important roles in regulating the levels of specific proteins and in eliminating damaged or abnormal proteins. *E. coli* possess a very large number of proteolytic enzymes distributed in the cytoplasm, the inner membrane, and the periplasm, but, with few exceptions, the physiological functions of these proteases are not known. More than 90% of the protein degradation occurring in the cytoplasm is energy-dependent, but the activities of most *E. coli* proteases in vitro are not energy-dependent. Two ATP-dependent proteases, Lon and Clp, are responsible for 70–80% of the energy-dependent degradation of proteins in vivo. In vitro studies with Lon and Clp indicate that both proteases directly interact with substrates for degradation. ATP functions as an allosteric effector promoting an active conformation of the proteases, and ATP hydrolysis is required for rapid catalytic turnover of peptide bond cleavage in proteins. Lon and Clp show virtually no homology at the amino acid level, and thus it appears that at least two families of ATP-dependent proteases have evolved independently.
Key words. ATP-dependent; degradation; protease; Lon; Clp.

Introduction

Living organisms have a remarkable capacity to degrade their own proteins. One measure of this capacity is the sheer number of proteolytic enzymes found within cells. In *E. coli*, for instance, 24 different endoproteases have been identified either biochemically or genetically (table 1). Of these, at least 12 proteases are found in the cytoplasm, and the remainder are in the membrane or in the periplasmic space. Although the intracellular activities of most of these proteases have not been defined, the biochemical properties of the proteases in vitro suggest that they are capable of degrading a broad range of proteins. In addition to endopeptidases, cells possess numerous exopeptidases which further degrade the peptides generated by endoproteolytic degradation of proteins. *E. coli* has at least 12 exopeptidases[106] which serve to regenerate amino acids from the peptides released by endoproteolytic cleavage of proteins[108] and possibly to protect the cell from inhibitory effects of partial degrada-

Table 1. Proteolytic enzymes of *Escherichia coli*

Protease	Gene (map position)	M_r ($\times 10^{-3}$) subunit (native)	Substrates in vivo (in vitro)	Protease type	Remarks	References
Cytoplasmic proteases						
Lon protease (Protease La)	*lon* (10 min)	87 (450)	Sul A; λ N protein; Rcs A; Tn903 transposase	serine	ATP-dependent; Degrades λN in vitro	18, 25, 32, 94, 109, 169, 176, 177, 178
Clp protease (Protease Ti)	*clpP* (10 min)	21 (240)	pro-Clp P; Clp A; LacZ-fusions	serine	ATP-dependent activation by Clp A ATPase activity	56, 58, 70, 71, 78, 79, 97, 98
	clpA (19 min)	83 (> 450)				
RecA/LexA	*recA* (58 min)	44	(λ cl repressor; UmuD) LexA	serine	ATP-dependent; autolysis of substrates	29, 90, 144, 155
	lexA (93 min)	23				
Hfl A	*hflK* (94.5 min)	46 (200)	λ cII repressor	serine	Low activity against cII in vitro	7, 20
	hflC	37				
Protease Do		54 (300; 500)	(Casein)	serine	Degrades Ada	51, 85, 167
Protease Re		82 (82)	(Casein)	serine	Degrades oxidized GS	51, 86, 126, 145
Protease Fa		110 (110)	(Casein)	serine		51
Protease So		77 (140)	(Casein)	serine	Degrades Ada and oxidized GS	26, 51, 85
Protease Ci		(120)	(Insulin)	metallo		51
ISP-L-Eco		55 (55)	(Z-AAL-PNA)	serine		165
Protease II		58	(BAEE)	serine	Very low activity against proteins	122
(Alp-related)	(*alpA*) (57 min)		Sul A; Rcs A		Alp affects expression of ATP-dependent protease	170
Periplasmic or membrane-associated proteases						
Protease Mi		110	(Casein)	serine		51
Protease Pi (Protease III)	*ptr* (61 min)	110 (110)	(Insulin)	metallo	Insulinase superfamily	21, 28, 42, 138
Signal peptidase I	*lep* (55 min)	37 (37)	Precursors of exported proteins		Inner and outer membranes	183, 184
Signal peptidase II	*lsp* (1 min)	18	Precursors of lipoproteins		Outer membrane	171, 173, 173
Deg P (HtrA)	*degP* (3.5 min)	48	Pho A-fusions	serine	σE Heat shock protein	88, 89, 160, 161
Protease I		21 (43)	(NAPNE)	serine	Very low activity against proteins	121
Protease IV	*sppA* (38.5 min)	34 (34)	Signal peptides after processing	serine	Signal peptide peptidase	72, 73, 123
Protease IVa		(300)				140
Protease V			(Z-Phe-ONP)	serine	Inner and outer membranes	123
Protease VI		43 (43)	Membrane proteins	serine		125
Protease VII	*ompT* (12.5 min)	36 (180)	Ferric enterobactin receptor protein	serine	Cleaves between basic residues	60, 166
Protease peri7		(> 2000)	(Casein)	metallo	Activated by ATP	174
Protease peri8		(> 2000)	(Casein)	serine		174

tion products. Considering the total number of proteases and peptidases identified in *E. coli*, one is faced with the remarkable fact that > 3% of the enzymatic activities present in *E. coli* at any given time are proteolytic.

A second measure of degradative capacity in *E. coli* is the rate of intracellular protein degradation. A number of naturally unstable proteins have half-lives as short as 1–2 min (≤ 5% of the doubling time of the cells) and many mutant or otherwise abnormal proteins have half-lives in the range of 5–20 min. While the bulk of proteins that accumulate in *E. coli* are stable over several generations and have apparent half-lives of 5–20 h, a significant fraction of newly synthesized proteins and polypeptides are highly unstable and are turned over within one generation. Thus, both the number of proteases and the amount of protein turnover demonstrate that intracellular protein degradation is an active metabolic process and serves important if not essential physiological functions.

The functions of protein degradation can be conveniently divided into two categories, *housekeeping* and *regulatory*, depending on whether the protein eliminated is non-functional or functional. Housekeeping functions include the degradation of such non-functional proteins as those resulting from errors in transcription or translation, oxidative or other chemical damage, improper folding, or thermal denaturation. Degradation of such proteins would eliminate potentially harmful interactions between the non-functional and other functional proteins and would release amino acids for the synthesis of new proteins.

There are numerous examples from both prokaryotic and eukaryotic organisms of protein degradation which serves important regulatory roles. Regulatory degradation may involve the turnover of functional proteins either continuously or in response to specific metabolic signals. The selective degradation of proteins in cells under starvation conditions[93,94a,136], the initiation of the lytic life cycle of phage λ by proteolytic cleavage of the CI repressor[142], and the degradation of cyclin at the onset of mitosis in developing embryos[111] are a few of the examples of regulatory degradation that occur in response to specific physiological or environmental signals. Gottesman[54,55] has referred to targeted proteins that are otherwise stable in the absence of a signal as *conditionally unstable* proteins. Other proteins are, by contrast, *constitutively unstable*, that is, they are degraded under most metabolic conditions. These proteins are maintained at very low steady state levels, because degradation occurs rapidly following biosynthesis. Such proteins can either be eliminated from the cell very rapidly when biosynthesis is turned off, as with the SulA protein[109], or can be rapidly increased by transient inhibition of degradation, as with the *E. coli* heat shock sigma factor[161].

Even from the above brief discussion of protein degradation and its regulatory effects, it is clear that many different proteins are targeted for degradation in cells. In addition, most proteins will become targets for degradation if they are chemically or structurally perturbed, but the overwhelming majority of proteins in cells are indefinitely stable. Perhaps the most intriguing and challenging question regarding intracellular proteolysis, then, pertains to the remarkable selectivity of protein degradation. How are intracellular proteases able to discriminate with high fidelity between appropriate and inappropriate protein substrates? The distinction made above between housekeeping and regulatory degradation refers to the consequences of intracellular proteolysis and may not extend to the biochemical mechanism of degradation. Some proteases (e.g. Lon protease) degrade both abnormal proteins and specific regulatory proteins. What chemical or structural features are common to substrates for a given protease and which ones are different in substrates for different proteases? As will become apparent later in the review, we are only beginning to be able to address this question.

A second intriguing and challenging property of intracellular protein degradation is its energy-dependence. As much as 90% of intracellular protein degradation is dependent on metabolic energy, presumably supplied by ATP. Considerable progress has been made in the last 10 years in identifying enzymatic systems involved in energy-dependent protein degradation. Both prokaryotic and eukaryotic cells contain proteases whose activities in vitro are dependent on or are highly activated by ATP[9,55]. One family of such proteases, represented by Clp protease from *E. coli*, appears to have been conserved throughout evolution, and thus ATP-dependent proteases are probably a universal feature of energy-dependent proteolytic systems[58]. In addition, eukaryotic, but not prokaryotic, organisms possess an ATP-dependent system for marking proteins for degradation by the covalent attachment of ubiquitin to potential target proteins[64,139]. The ubiquitin-conjugated proteins are then degraded in an energy-dependent manner by at least one protease that appears to be ATP-dependent[139].

In *E. coli* degradation by ATP-dependent proteases involves direct interaction between the protease and the target protein, and, here again, we are only beginning to understand the mechanism by which the energy of ATP hydrolysis is used in activating the cleavage of proteins by ATP-dependent proteases. ATP hydrolysis appears to be used to promote conformational changes in the protease[177], but it may also have a role in unfolding the target protein or in moving the protein within the binding sites on the protease. Finally, we might consider that the question of the specificity of degradation and the question of the role of ATP hydrolysis in protein degradation may be related. Protein substrates must have features that allow them to be cleaved at the proteolytic active site and must also have features that can activate ATP hydrolysis by the protease[177]. Increasing the number of

recognition sites on the target protein would increase the selectivity of protein substrates exponentially.

This review will focus on the proteases found in *E. coli*, and primarily on the ATP-dependent proteases that have been identified either biochemically or genetically. *E. coli* furnished the first examples of degradation of regulatory proteins as an initiating event for development (bacteriophage λ repressor)[142] or for the induction of an emergency response (SOS induction by LexA cleavage)[92]. *E. coli* also provided the first case in which a specific protease (Lon protease) responsible for the intracellular degradation of a specific protein (Sul A) was demonstrated[109]. The Lon protease was also the first ATP-dependent protease to be purified and characterized[18,25]. Recently, a second ATP-dependent protease, Clp protease, was purified from *E. coli*[70,79]. This protease has structural features that are provocatively similar to the multicatalytic protease from eukaryotic cells. Lon and Clp proteases are responsible for 60% of the total and 70% of the energy-dependent protein degradation in growing cells. The functions of these proteases will be considered in the light of the 24 or more other endoproteolytic enzymes found in *E. coli*.

Proteins degraded in E. coli cells

Reviews by Goldberg and St. John[50] and by Miller[106] have excellent discussions of earlier experimental data that have led to our current understanding of the extent and the regulation of protein turnover in growing and non-growing cells. These results will be briefly summarized below, along with recent experimental findings that have helped identify classes of unstable proteins.

Proteins degraded in growing cells

Turnover of bulk protein in growing cells. A considerable amount of protein degradation goes on in growing *E. coli* cells, although we do not know exactly how much, and the amount that occurs depends on the environmental conditions and on the genetic background of the cells. Experiments with batch[135] or continuous[157] cultures of *E. coli* indicate that the bulk of the steady-state protein pool in *E. coli* cells are degraded at ~ 1% per hour. About 40% of the protein mass of *E. coli* is stable indefinitely, and most of the remaining protein mass has an aggregate half-life of 15–30 h. In a general survey of protein degradation in vivo, Moesteller et al.[110a] measured the degradation rates of 184 whole cell proteins and found 10 proteins with half-lives of 2–5 h and another 37 with half-lives of 5–23 h; the remainder of the proteins were stable. Schroer and St. John[146] surveyed membrane proteins in a similar way, and found only 6 out of 125 proteins that were more unstable than the average and these apparently had half-lives of several hours. An obvious but important conclusion from the results of these surveys is that the most abundant proteins found in *E. coli* are stable.

Table 2. Protein half-lives in *E. coli*

Protein	Half-life	References
Bulk protein during growth:		
~ 40% of total	> 70 h	110, 134, 135
~ 35% of total	25–70 h	
~ 10% of total	5–25 h	
~ 5% of total	2–5 h	
Bulk protein during starvation:		
~ 50% of total	1–5 h	50, 136
Abnormal proteins	20–60 min	49, 100, 134
Mutant proteins	5–60 min	53, 137
Unstable *E. coli* proteins:		
Sul A; Rcs A; sigma 32	1–2 min	32, 109, 161, 168
Tn903 transposase		
Unstable lambda proteins:		
N protein; CII; CIII;	1–2 min	7, 57, 81
O protein		
Conditionally unstable proteins:		
λ Cl; LexA; Ada	2–10 min	87, 92, 142

The bulk rate turnover measurements fail to give accurate estimates of the rate and extent of protein degradation, because short-lived proteins and polypeptides make up only a few percent of the steadystate mass of intracellular protein. When short periods of time are used for pulse-labeling, between 5 and 20% of the radioactive amino acids incorporated into newly synthesized polypeptides are released in less than one generation[180]. Even the lower range of this estimate indicates that intracellular proteolysis is a continuous and important cellular function and that *E. coli* must require a considerable degradation apparatus to account for this high rate of protein turnover. What are the proteins and polypeptides that make up this pool of rapidly degraded proteins?

The general and specific degradation rates of *E. coli* proteins are summarized in table 2. The specific proteins known to be unstable can account for only a small percentage of the pool of rapidly degraded protein. The remainder of the pool is thought to represent improperly synthesized proteins and proteins that have been damaged, denatured, or deprived of normal associations within the cell.

Abnormal proteins and mutant proteins. The first examples of rapid protein degradation in *E. coli* came from studies on the stability of abnormal proteins[49,134] produced either by the incorporation of amino acid analogs or puromycin into newly synthesized proteins. Incorporation of amino acid analogs interferes with proper folding of proteins, producing proteins with abnormal conformations, whereas puromycin causes premature termination of growing polypeptide chains and the resulting incomplete proteins would also not be expected to have stable folded conformations. Most of the abnormal proteins of both types are rapidly degraded (half-lives of 20–40 min) whereas normal proteins present in the same cells are stable. In one experiment, when canavanyl proteins were separated on an SDS gel, about 80% of the

protein bands appeared to be susceptible to degradation, although not all proteins were degraded at the same rate[96].

The first indications of the extremely short half-lives possible for individual proteins in *E. coli* came from measurements of the turnover of mutant proteins. The degradation of the X90 nonsense mutant of β-galactosidase ($t_{1/2} = 7$ min)[53] and the L1 mutant of Lac repressor ($t_{1/2} \sim 40$ min)[137], demonstrated not only that proteins can be degraded very rapidly but also that the intracellular proteases are selective. Since then, many examples of mutant proteins, including nonsense and missense mutants as well as various fusion proteins or chimeras, have been found to be rapidly turned over in vivo. Various temperature-sensitive proteins[81] have been shown to be degraded in vivo, suggesting that an inability to fold properly may render a protein susceptible to degradation. In one of the few detailed studies of the effects of mutations on the degradation of an individual protein, mutations in the lambda Cro protein that decreased the conformational stability of the protein resulted in increased degradation in vivo[124]. However, the sites or structures of the mutant proteins recognized by the degradative enzymes are unknown.

The marked instability of many mutant proteins serves to demonstrate the high capacity for protein degradation within cells and to illustrate the discriminatory ability of the *E. coli* proteases. Recognition of mutant and other abnormal proteins by the degradative system must parallel the recognition of misformed or damaged *E. coli* proteins in normal cells. What should be emphasized is the enormous range of proteins that can be perceived as abnormal and be targeted for degradation. Transcriptional and translational errors can give rise to abnormally truncated proteins or proteins that have misincorporated amino acids. Oxidative damage[30] or other chemical modifications, the binding of inhibitory ligands, the disruption of macromolecular complexes, and heat or other stress-induced denaturation are additional mechanisms by which the native structure of proteins can be affected. Since most proteins in the cell are stable but can exist in many aberrant states that are highly unstable, all proteins must have the potential to adopt structures or must have determinants masked within their primary sequences that can be recognized by intracellular proteases.

Naturally unstable proteins. A number of wild-type proteins from either *E. coli* or bacteriophage λ have been shown to be subject to rapid degradation in vivo (table 2). The rapidly degradable proteins identified to date tend to be regulatory proteins or proteins involved in critical metabolic processes such as cell division and DNA replication. Because of the functional consequences of changes in the degradation rate, proteases that perform the rate-limiting step in the degradation of some of the proteins have been identified genetically. At least 5 different proteolytic systems are involved in the degradation of the individual proteins in table 2, indicating that degradation of these proteins is highly selective. There is no apparent overlap in specificity among the proteases carrying out the initial cleavage, even for constitutively unstable proteins. Thus, Lon is needed for rapid turnover of SulA, RcsA, Tn903 transposase, and λ N protein, but has no influence on the degradation of σ32, or the λ proteins O and CII.

The entire population of a specific protein labeled during a pulse is degraded at a relatively uniform rate, indicating that there are not separate pools of stable and unstable forms of the protein. As expected, rapidly degraded proteins are present in cells in very low amounts; in a few cases the proteins are barely detectable even with very short pulses[158]. The efficiency of the degradative enzymes in finding and degrading even very small amounts of a protein suggest that the proteases have high affinity for appropriate substrates or a mechanism exists to bring the proteases and their respective targets together. Although sequence or structural information about the unstable proteins should help identify the chemical and physical features that are recognized by intracellular proteases, a comparison of the proteins degraded by a single protease, e.g., Lon protease, reveals few obvious common features. The specificity of Lon will be discussed in a later section. More information about the actual structures and other physical properties of these unstable proteins will be required to identify common determinants of protease susceptibility.

Some 'naturally unstable' proteins are not mutant proteins but may arise because of mutations in other cellular components with which they interact or because they are expressed under abnormal conditions. Degradation of the subunits of multimeric proteins such as integration host factor[112], the histone-like protein, HU[10], ribosomal subunits[31,115], and RNA polymerase[75] occurs in cells in which one of the components of the complex has been deleted, thereby preventing proper assembly of the multimer, or in cells in which one of the components is over-expressed. Presumably, the free subunits cannot fold properly in the absence of the other component or the unsatisfied bonding domains of the protein contain sites that are recognized by *E. coli* proteases.

In normal cells, the degradation of the free subunits of a multimeric enzyme would provide a post-translational means of ensuring coordinate assembly and expression of activity of the complex[84], especially in cases where expression of partial activities might have an unwanted effect. It is interesting that all of the unstable proteins listed in table 2 function in cells by interacting with other proteins or macromolecules, often as part of multicomponent complexes. Possibly, the protein in vivo is in equilibrium between a free and complexed state, and the free protein is rapidly degraded to maintain the intracellular level at the minimum functional concentration at all times. This model suggests an analogy between 'naturally unstable' proteins and those proteins mentioned

above that are unstable when cloned in excess of the proteins with which they are normally associated or when they are deprived of their partners by mutation[54, 55].

Processing of signal peptides. Most proteins of *E. coli* found in the inner or outer membrane or in the periplasmic space are synthesized with signal peptides that are cleaved from the original translation products by one of the signal peptidases found in the membrane. Processing of signal peptides proceeds with a half-life < 1 min. Signal peptides are then degraded to amino acids by Protease IV (signal peptide peptidase) and other peptidases. Since the signal peptides are usually between 15 and 30 amino acids long (approximately 10% of the length of the precursor[105] and membrane and periplasmic proteins constitute 10–15% of the protein mass of the cell[119], degradation of signal peptides amounts to turnover of perhaps 1% of the total cellular protein per generation, a significant but minor fraction of the rapidly degraded pool of polypeptides.

There are few examples of processing of cytoplasmic proteins in *E. coli*, other than the removal of the amino terminal methionine from a majority of newly synthesized proteins[8]. A histone-like protein called HLP-1, which is the product of the *skp* gene, has the first 20 amino acids removed after synthesis[65]. ClpP, a cytoplasmic protein, is synthesized with a 14-amino acid long precursor peptide, which is rapidly cleaved from the enzyme (half-time < 2 min)[97]. ClpP is the proteolytic component of Clp protease and processing appears to be auto-catalytic. Details of this process will be discussed below. There may be other instances of processing of cytoplasmic proteins in *E. coli*, but it appears unlikely that this type of cleavage reaction contributes significantly to the pool of rapidly degraded protein.

Proteins degraded after nutritional shifts

Protein degradation increases in many organisms in response to nutritional changes and especially starvation for essential nutrients[50]. The degradation of specific proteins is thought to help supply energy and amino acids for new proteins synthesis and other metabolic activities. In some cases, e.g. in spore-forming organisms[99], degradation of specific proteins may be part of the initial events in major developmental changes. Relatively little has been reported in the last 10 years regarding degradation in *E. coli* in response to nutritional changes or to starvation. Readers are directed to the comprehensive review by Goldberg and St. John[50] and a more recent one by Miller[106].

Bulk protein turnover. Protein synthesized during exponential growth begins to be degraded when cells are starved for ammonia, carbon, amino acids, phosphate, sulfate, or other nutrients. The apparent rate of protein degradation is 3–5% per hour, measured as a percent of the stable exponential phase protein that is present in the cells at the end of growth[136, 155]. Since about 50–70% of the original protein present at the end of growth is degraded after 15 h of starvation (Miller[106] and M. R. Maurizi, unpublished), the actual degradation rate for the labile proteins is ~ 8% per hour (aggregate half-life of 9 h). Newly synthesized protein in starving cells may be degraded even more rapidly than pre-existing protein[50], so that the total rate of protein degradation may be somewhat higher than given above.

Degradation of individual proteins. Most reports of changes in enzyme levels during starvation have described losses in enzymatic activities, rather than degradation of the proteins[50]. Several specific enzymes, including indole glycerolphosphate synthetase[110b] have been shown to be degraded under starvation conditions. When a survey of 20 enzymes was made in cells starved for ammonia, the activities of most of the enzymes were unchanged after 16 h (M. R. Maurizi, unpublished). A few enzyme activities (aspartate transcarbamylase, glutamate synthase, glutamine synthetase) were lost at about 8–12% per hour. Two enzymes were lost rapidly, aspartokinase III ($t_{1/2}$ = 2.5 h) and glutamate dehydrogenase ($t_{1/2}$ = 1.5 h). Degradation of both proteins was confirmed by pulse labeling and immunoprecipitation. A similar pattern of enzyme loss was observed in glucose-starved cells. These results confirmed that degradation during starvation is selective, a conclusion at variance with that of Miller in his review[106] which was based on unpublished findings from R. Moesteller. The physiological significance of the differences in stabilities of the individual proteins is not yet understood, and it remains possible that the half-lives of the proteins in starving cells merely reflect the intrinsic stability of the different proteins and not a response peculiar to the starvation conditions. In the latter case, since glutamate dehydrogenase is stable during growth, the degradation in starving cells would reflect the appearance of a new proteolytic system under those conditions. Although a number of protease activities are known to increase at the end of exponential growth (see below), no one has yet identified a proteolytic system that is unique to stationary phase or to starving cells.

Physiological regulation of protein degradation

Much of the information about enzymatic systems involved in the degradation of abnormal or mutant proteins has come from direct biochemical measurements under different physiological conditions. Factors as diverse as the availability of metabolic energy and the heat shock response have led to the identification of enzymatic systems that are directly involved with protein stability and degradation. Since little has been reported recently regarding the stringent response and its effects on starvation-induced proteolysis, the reader is directed to earlier reviews for discussions of this topic[50, 106].

Metabolic energy and ATP

In *E. coli*, as in most cells, inhibition of energy metabolism, for example by addition of cyanide in the absence of glucose, drastically reduces the rate of protein degradation. Most of the studies on the energy-dependence of degradation have measured the turnover of abnormal proteins during growth [118] or the degradation of normal proteins during starvation [100]. About 90% of protein degradation measured in this way is energy-dependent. The inhibitory effect of cyanide is also observed for the degradation of individual proteins, such as SulA [14] or glutamate dehydrogenase (M. R. Maurizi, unpublished). In vivo studies from Olden and Goldberg [118] and various in vitro studies of proteolytic enzymes [18, 25, 79] indicate that the metabolic energy for protein degradation is required in the form of ATP.

The biochemical basis of the requirement for ATP in intracellular protein degradation has become clearer in recent years. Probably all cells possess proteases that have binding sites for ATP and require the binding and hydrolysis of ATP to express proteolytic activity. In *E. coli* two such proteases, Lon and Clp [18, 25, 79] have been identified and characterized. Both proteases are abundant ($\sim 0.1-0.2\%$ of cellular protein) and mutational studies indicate that these proteases account for a more than 70% of energy-dependent degradation in growing cells. In addition, the RecA protein binds ATP as an allosteric effector and is activated to promote auto-proteolysis in a few specific target proteins. In vivo and in vitro experiments suggest that there should be other ATP-dependent proteases in *E. coli*; for example, neither Lon nor ClpA appears to be involved in starvation-induced proteolysis [78]. The majority of proteases isolated from *E. coli* cells have proteolytic activity in the absence of ATP. Some proteases, e.g. signal peptidases, may have a narrow selection of possible substrates. The other non-energy-dependent proteases may only degrade proteins that have already been cleaved by the ATP-dependent proteases or have been acted on by non-proteolytic elements of the energy-dependent degradative system.

There are two other ATP-dependent enzymatic systems that are involved in the recognition and degradation of unstable proteins: (1) heat shock proteins or molecular chaperones, and (2) the ubiquitin-conjugating systems found in eukaryotes. Since the ubiquitin system is not found in *E. coli*, it will only be discussed in a later section with respect to the different strategies employed by *E. coli* and other organisms to recognize targets for degradation. The effect of molecular chaperones will be discussed in the next section. The activities of ATP-dependent proteases will be discussed in a later section.

Heat shock and other stresses

Essentially all cells respond to sudden increases in temperature by increasing the synthesis of a set of proteins called heat shock proteins [113]. Several of the heat shock proteins are among the most abundant proteins in cells grown at physiological temperatures (37 °C for *E. coli*) and appear to have important functions in cells, since mutations in a positive regulator of heat shock, $\sigma 32$, or in some of the heat shock proteins themselves drastically reduce the viability of cells, especially above 30 °C. The cellular functions of heat shock proteins in vivo are not known, although in *E. coli* they are needed for bacteriophage replication [45]. Some heat shock proteins may bind to unfolded proteins and function as molecular chaperones [130]. Heat shock and other stresses could cause an increase in protein denaturation in vivo and the increased amounts of heat shock proteins may be needed to aid in refolding proteins.

Proteolytic regulation of the heat shock response. Synthesis of heat shock proteins in *E. coli* is under control of the heat shock sigma factor, $\sigma 32$ [62, 63]. $\sigma 32$ is required for growth even at lower temperatures [113] but is normally maintained at very low concentrations in cells, because it is rapidly degraded ($t_{1/2} = 1$ min) [161]. The protease that degrades $\sigma 32$ is not known, but degradation requires functional heat shock proteins, DnaK, DnaJ, and GrpE [161, 163, 167]. When cells are shifted to 42 °C, $\sigma 32$ degradation is transiently inhibited, but after about 4 min rapid degradation of $\sigma 32$ begins again. The inhibition of degradation along with increased translation of $\sigma 32$ mRNA allows the intracellular concentration to increase sufficiently to induce synthesis of heat shock proteins [161]. Resumption of rapid degradation of $\sigma 32$ requires DnaK and GrpE [163, 167], although there is no evidence that either protein itself is a protease.

In addition to a temperature shift, which would be expected to result in thermal denaturation of some intracellular proteins, the heat shock response can be induced in *E. coli* and other cells by the over-expression of abnormal proteins, such as canavanyl proteins or certain foreign proteins [48]. One possibility, then, is that an excess of abnormal proteins in the cell act as competitive substrates for the protease that degrades $\sigma 32$, leading to a transient increase in $\sigma 32$. Alternatively, since DnaK, DnaJ, and GrpE are also involved in abnormal protein degradation, the abnormal proteins could bind to these heat shock proteins and prevent them from mediating the degradation of $\sigma 32$.

Effects of heat shock proteins on protein degradation. Mutations in $\sigma 32$, which lead to lower expression of DnaK, DnaJ, GrpE, and GroEL, and mutations in the heat shock proteins themselves, block the degradation of many abnormal proteins, indicating that heat shock proteins are needed for the degradation of abnormal proteins such as canavanyl proteins, puromycyl fragments, and some mutant proteins [5, 47]. On the other hand, heat shock proteins may actually prevent the degradation of some abnormal proteins, such as temperature-sensitive mutants [80, 162] and some fusion proteins (M. R. Maurizi, unpublished). Since the heat shock proteins are thought to bind to unfolded proteins and to have protein-folding activities, the effect of heat shock

proteins on the stability of a particular protein or class of proteins may depend on whether the proteins are capable of being folded into stable conformations before the unfolded structures can be recognized by the intracellular proteases. In this model, heat shock proteins would increase the degradation of some abnormal proteins by keeping them in an extended conformation and preventing them from aggregating into protease-resistant inclusion bodies[80], but would decrease degradation of other proteins by accelerating their folding into protease-resistant structures similar to those of the wild-type protein.

Protease induction during heat shock. Heat shock induces synthesis of several proteases in *E. coli*, including the cytoplasmic ATP-dependent proteases, Lon and Clp[47,82,133], and the periplasmic protease, DegP (or HtrA)[89]. ClpA increases slightly in cells grown at high temperatures but synthesis is not dependent on $\sigma 32$[78]. Recently, it has been found that ClpB, a close homolog of ClpA found in *E. coli*, is also a heat shock protein[80a,154a]. The presence of increased amounts of these proteases presumably helps remove denatured proteins that could interfere with the functions of normal proteins. Nevertheless, Lon and Clp do not appear to be essential for survival of cells at moderately high temperature or even to allow cells to display a normal heat shock response and adaptation. ClpB mutants have lower variability at 50 °C[126a]. DegP mutants are unable to grow under acute heat shock conditions[88,159].

Specificity and selectivity of proteolysis in vivo

The critical question concerning intracellular protein degradation is how rapid and efficient degradation of appropriate proteins is accomplished without damage to other proteins – the selectivity of protein degradation. It is useful to distinguish between the specificity of proteases and the selectivity of proteolysis by saying that *specificity* refers to the nature of the binding interactions between the protease and a polypeptide, while *selectivity* depends on the availability of such interaction sites within protein populations or within a protein itself. Selectivity might depend not only on the existence of cleavable sites within the proteins, but on the structural stability of the protein[127], the presence of effectors of the protein or the protease, the cellular locations of the protein and protease, and the relative abundance of alternative substrates for the protease.

In the simplest case, selectivity can be accomplished by imposing strict specificity requirements for proteolysis, as in the rare tetrapeptide recognition site for Factor X[114] or in the requirement for a cysteinyl glyceride at the cleavage site for *E. coli* signal peptidase II[172]. These are both examples of limited proteolysis, but similar specificity might apply to the rate-limiting steps in degradation of particular proteins. Subsequent steps in cleavage of proteins could also be guided by strict specificity requirements. Evidence for a degradative system that recognizes the amino terminal amino acids of proteins has been presented by Varshavsky and coworkers[2,3] and others[140]. Only a subset of amino acids at the amino termini of proteins allow degradation by this system, a principle called the N-end rule. If the proteases carrying out the initial degradation reactions cut at positions that tend to leave amino termini with 'destabilizing' amino acids, the N-end rule degradative system would take over. *E. coli* proteases appear to discriminate between proteins on the basis of the carboxy terminal amino acid compositions as well. Hydrophobic amino acids near the carboxy terminus lead to increased degradation of some proteins[128]. Specific carboxy-terminal extensions added to some unstable proteins can protect the proteins from degradation in vivo[11,32].

Compartmentalization of proteases also influences selectivity of proteolysis. In *E. coli* the cytoplasm and the periplasmic space each have proteases that carry out particular functions, but there is no evidence of translocation of proteins between these compartments for the purpose of degradation. The co-localization of proteases and protein substrates within macromolecular assemblies may be a significant mechanism for controlling degradation, as in the cleavage of signal peptides by membrane associated signal peptidases[170,182]. Cytoplasmic proteases tend to be very large, and it is possible that the proteases have binding sites for interactions with other cellular constituents, which might allow them to come into contact with appropriate protein substrates. In this regard, one mechanism by which heat shock proteins might participate in general protein turnover would be by binding to both proteases and to unfolded regions of proteins. Such a combined action of proteases and heat shock proteins could explain why mutations in heat shock proteins produce defects in the degradation of the same abnormal proteins degraded by Lon[162]. Selectivity may also involve specific binding between the proteases and potential protein substrates not only at the proteolytic active site but also at other sites on the protease. As described below for Lon protease, the requirement for multiple sites of interaction between the protease and protein targets allows a protease with rather broad specificity towards cleavage sites to discriminate effectively between different proteins.

Lastly, the turnover rate for a particular protein will depend on the relative concentrations of other protein substrates in the cell and on the degree of saturation of the protease. There is no precise data on the saturation of proteases in vivo. If the estimate that 5–20% of newly synthesized protein is rapidly degraded is correct, there should be a large intracellular pool of rapidly degradable protein (0.5–2% of cellular protein) at all times during growth. Miller[106] has suggested that the decrease in this rapidly degradable pool in amino acid starved cells may free portions of the degradative apparatus to carry out increased degradation of normal proteins. Competition between the specific Lon substrates, SulA and RcsA, is evident only when one of the substrates is overproduced

in large excess[33], suggesting that Lon is normally not saturated with respect to these individual substrates, but the degree of saturation of Lon with the total pool of its physiological substrates (including abnormal proteins) is not known. When cells have high levels of abnormal proteins, increasing the amount of Lon by providing *lon* on a multicopy plasmid[46] increases the rate of abnormal protein turnover. One interesting possibility regarding the induction of the heat shock response is that the transient reduction in degradation of σ32 is the result of competition from proteins denatured by the temperature shift for the protease that degrades σ32[48].

Protease inhibitors

The only classical protease inhibitor isolated and characterized from *E. coli* is a 16,000 M_r periplasmic protein called Ecotin[27,101]. Ecotin is a potent inhibitor of several pancreatic proteases. Biochemical studies of Ecotin and analysis of the amino acid sequence have located the reactive site of the inhibitor within a disulfide linked region in which the scissile bond is between two methionines[101]. Ecotin shows no inhibitory activity against a number of *E. coli* proteases[27,125], although not all *E. coli* proteases have been tested. The physiological function of Ecotin is not known.

Two putative protease inhibitors found in *E. coli* are made by bacteriophage and their activities appear to be related to blocking the degradation of phage proteins. The presence of the CIII protein of bacteriophage λ increases the stability of the CII repressor, which has an extremely short half-life (< 2 min) in vivo[7,57]. CIII has been proposed to be an inhibitor of HflA, an *E. coli* protease required for the rate-limiting degradation of the CII protein, but it may inhibit other proteases that degrade CII as well[7,68]. In vitro, inhibition of degradation of CII by purified HflA required very high levels of CIII, suggesting that CIII may target other proteases[20]. Bahl et al.[4] showed that overproduction of CIII resulted in the stabilization of σ32 in vivo, with the concomitant induction of heat shock protein synthesis. Since the protease that degrades σ32 has not yet been identified, CIII might provide a useful biochemical probe for the isolation of the protease. Recently, it was shown that the activity of CIII resides in a 22-amino acid region of the protein that can form an amphipathic helix[81]. This segment, when expressed as an internal peptide region in an α-complementing fragment of β-galactosidase, can stabilize CII and induces the heat shock response.

Simon and his colleagues identified the T4 *pin* function which is required for the unusual stability of amber fragments of bacteriophage T4 proteins[150]. T4 *pin* behaves in vivo as expected for a gene coding for an inhibitor of Lon protease. Cells carrying the plasmid-encoded *pinA* gene show the several *lon*$^-$ phenotypes, such as production of capsular polysaccharide and sensitivity to ultraviolet light[151]. Degradation of Lon substrates in vivo, such as canavanyl proteins, is severely reduced in the presence of a functional *pinA* gene. The purified PinA protein has been shown to bind to Lon with high affinity and to inhibit proteolytic and ATPase activity in vitro (J. Hilliard, L. Simon, M. Maurizi, unpublished).

Caution must be used in interpreting inhibitory effects caused by mutations or overproduction of proteins. For unstable *E. coli* proteins that are found as part of complexes with other proteins, degradation can be partially blocked by over-expression of the associated proteins. Stabilization of SulA by FtsZ[77] and stabilization of RcsA by RcsB[158] are such examples. Modification or mutation of either the unstable protein or the associated protein could lead to altered affinity between the proteins and, hence, changes in the stability of the target protein. Inhibition of degradation by stabilization of the target protein should be easily distinguished from inhibition by interaction of an inhibitor with the protease when there is more than one substrate known for the protease.

ATP-dependent proteases

Lon protease (protease La)

Goldberg and colleagues isolated protease La from *E. coli* extracts as a protease that required ATP for the degradation of casein[25]. Protease La was subsequently identified by Charette et al.[18] and Chung et al.[25] as the product of the *lon* gene, and it will be referred to in this review as Lon protease or simply Lon. The *lon* gene had earlier been identified as the locus for mutations that made *E. coli* cells UV-sensitive and mucoid; *lon* mutations had been shown to be the same as *degR* or *degT* mutations, which interfered with the degradation of mutant forms of β-galactosidase (reviewed by Gottesman[50]). It is now clear that Lon is an ATP-dependent protease that is responsible in vivo for energy-dependent degradation of both abnormal proteins and specific regulatory proteins.

Structure and properties of Lon. The amino acid sequence of Lon protease derived from the DNA sequence of the *lon* gene indicates that Lon has 783 amino acids and a subunit M_r of 87,000[23]. The native M_r of Lon under moderately low salt conditions is 450,000[175]. Goldberg and colleagues have interpreted structural and binding studies as indicating that the native form of Lon is a tetramer[102,175], but there have been no detailed physical studies to explain the apparent discrepancy in the experimentally measured and the calculated sizes of the tetramer. Lon is inhibited by diisopropylfluoro phosphate and by peptide chloromethyl ketones[175,176], suggesting that Lon has a classical serine protease active site. The sequence around Ser679 in Lon resembles that found at the active site of the trypsin family, and mutation of that serine leads to loss of Lon activity in vivo[1]. Lack of similar conservation around histidine residues, however, indicates that Lon represents a unique family of serine proteases[23]. Lon has an intrinsic ATPase activity that is stimulated in the presence of protein substrates[177]. The

amino acid sequence of Lon reveals a single well-conserved ATP-binding consensus sequence [23] and tetramers of Lon would be expected to bind up to 4 molecules of ATP.

ATP-Dependent proteolysis by Lon protease. In the presence of ATP, Lon protease has the ability to degrade both large and small proteins in vitro, making endoproteolytic cleavages to generate peptide products of between 10 and 20 amino acids [94b, 104]. Generally, large proteins become better substrates when they are denatured [104, 177], suggesting that proteins must have an extended polypeptide structure or flexibility to interact with the substrate sites on the enzyme. Lon also cleaves the amide bond of certain short fluorogenic peptides [176]. The rate of peptide bond cleavage in the presence of ATP varies depending on the substrate [52], but as can be seen in table 4 the turnover number with most substrates is low – 1 peptide bond cleaved per 8–10 s per tetramer of Lon. The physiological substrate, λ N protein, is degraded at a somewhat faster rate, about 1 peptide bond per 1 second [94b].

With many protein substrates, degradation by Lon requires the continuous hydrolysis of ATP. However, cleavage of smaller proteins, such as oxidized insulin B chain, and fluorogenic peptides proceeds quite readily in the presence of non-hydrolyzable analogs, such as AMPPNP (table 3) [94b, 177]. The need for hydrolysis of ATP tends to be greater for higher molecular weight substrates [104], but λ N protein, M_r 12,000, was degraded in the presence of AMPPNP [94b]. Fluorogenic peptides are cleaved by Lon in the complete absence of nucleotide, and various effectors can stimulate peptide cleavage as well as ATP [176, 177]. Since the proteolytic activity of Lon does not require ATP hydrolysis and can even be expressed in the absence of ATP, ATP cannot be involved in the catalytic mechanism of peptide bond cleavage at the active site. The binding of ATP or other effectors must produce a conformational change at the proteolytic site that makes it more accessible to short regions of the polypeptide or increases the catalytic efficiency of the enzyme. The requirement for ATP hydrolysis in the degradation of larger proteins must reflect a second mode of interaction between proteins and Lon protease [177].

Table 3. Catalytic rate of peptide bond cleavage by Lon and Clp with various substrates

Substrate	k_{cat}[a] Subunit (min^{-1})	Holoenzyme	AMPPNP/ ATP[b]
1) Lon protease			
BSA[c]	2	8	0.08
α-Casein[c]	2	8	0.14
Globin[c]	4	16	0.23
λ N protein[d]	60	240	0.25
Oxidized insulin B chain[d]	20	80	1.0
Succinyl-Ala-Ala-Phe-NMC[e]	< 2	< 8	1.0
2) Clp protease			
Succinyl-Leu-Tyr-NMC[f]	> 10	> 120	–
α-Casein[f]	12	144	0.0

[a] k_{cat} was calculated from the V_{max} and the enzyme concentration, using either the subunit or the holoenzyme as the catalytic unit (tetramer of Lon or dodecamer of ClpP). The V_{max} was determined experimentally or was estimated from data obtained at partial saturation. [b] Ratio of activities in the presence of AMPPNP and ATP. [c] Calculated from data in Goldbergg and Waxman [52]. [d] Maurizi [94]. [e] Waxman and Goldberg [177]. [f] M. Maurizi, unpublished.

For several different proteins, degradation in the presence of ATP results in a stimulation of ATP hydrolysis, such that, at saturating levels of both ATP and protein substrate, there is an increase of 2.0–2.5 molecules of ATP hydrolyzed for each peptide bond cleaved [104]. This ratio is remarkably similar for proteins, even when the absolute rates of degradation differ, which could indicate a tight coupling between ATP hydrolysis and proteolysis, perhaps reflecting work done in each catalytic cycle to translocate the substrate protein within the active site. It should be emphasized, however, that Lon has a high basal ATPase activity and, since the stimulation of ATPase activity by substrates is only 2-fold for most proteins [104, 175], there are actually as many as 4 molecules of ATP consumed per peptide bond cleaved. Also, half-maximal stimulation of ATPase activity occurs with lower concentrations of proteins (about 25%) than those required for half-maximal proteolysis [177]. Thus, the acceleration of ATP hydrolysis is not linked to peptide bond cleavage per se but to some other aspect of the interaction between protein substrates and Lon protease.

Lon subunits behave cooperatively with respect to nucleotide binding and hydrolysis [102]. Up to 4 molecules of ATP can be bound per tetramer of Lon, but 2 subunit sites have very high affinity ($K_d \sim 0.1$ μM) and 2 have somewhat lower affinity ($K_d \sim 15$ μM). Assuming that the Lon subunits are homogeneous and active, these data indicate strong negative cooperativity in ATP binding. Since maximum peptide hydrolysis occurs when only the first 'high affinity' sites are occupied [52, 103, 176], Lon subunits appear to exist in at least two different states in the presence of ATP, one with the active site 'open' and one with the active site 'closed'. Lon hydrolyzes ATP to ADP quite readily, and ADP is a potent inhibitor of the enzyme. ADP binding promotes the closed active site, because ADP inhibits peptide hydrolysis independently of the presence of ATP [52]. ADP also competes with ATP for binding to the nucleotide binding site, and thus Lon active sites cycle between open and closed conformations with each round of ATP hydrolysis (described as 'active' and 'inactive' states in Goldberg and Waxman [52]).

Lon appears to have a binding site for proteins that lies outside the active site. Evidence for such an allosteric binding site for proteins comes from two findings: protein substrates, but not peptide substrates, promote the release of tightly bound ADP from Lon [103], and protein substrates activate the peptidase activity of Lon *in the absence or the presence of ATP* [177]. Even at relatively high concentrations of proteins, peptidase activity is activated and there is little or no competition between the

protein and the peptide for the active site. Thus, protein binding at the allosteric site also induces heterogeneity in subunit conformations such that one subunit active site is always open. The cleavable sites on the protein itself must not have access to the active site unless ATP is also present.

The above observations led Goldberg and colleagues[103, 177] to propose a model for Lon-dependent protein degradation with 3 critical features:

(a) ATP and ADP bind at the ATPase active site on Lon and allosterically affect the conformation of the proteolytic active site.
(b) Protein substrates bind at an allosteric site on the enzyme and affect the conformation at the proteolytic active site.
(c) Proteins binding at the allosteric site affect the binding of nucleotides, and, specifically 'good' protein substrates promote the exchange of ATP for tightly bound ADP.

In figure 1, I have presented a scheme for the catalytic cycle of Lon protease, which is largely based on the Goldberg model, but which includes some speculation as well. It should be emphasized that detailed kinetic studies

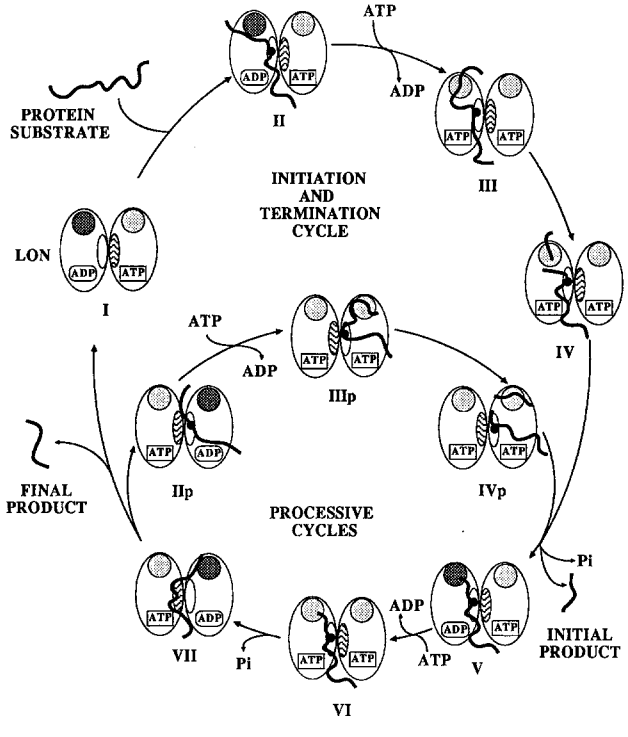

Figure 1. Model for the catalytic cycle of Lon protease. Only 2 of the 4 subunits of Lon are shown. It is assumed that the subunits function in pairs, one with high affinity and one with lower affinity for ATP and that the subunit conformations can alternate between the two. The proteolytic sites are shown as circles and the allosteric sites for proteins are shown as ovals. Proteolytic sites and allosteric sites are open (light figures) or closed (dark figures), depending on whether ATP or ADP is bound to the nucleotide site. Because degradation is shown as occurring on different subunits in each cycle, it is necessary to invert the figure in proceeding from IVp to V.

of binding order, hydrolysis steps, and product release have not been reported. For simplicity, the catalytic cycle is shown for only two subunits, although all four subunits might take part in degrading a single large protein. Since there is some suggestion that Lon subunits cooperate in protein degradation, the model assumes that the subunits act in pairs and that the peptide bond cleavage reaction flip-flops between the two active sites. The model also assumes that the subunits of Lon are identical and are assembled in a symmetrical manner, and that any heterogeneity in the enzyme during catalysis is due to interactions between subunits subsequent to ligand binding.

In the absence of protein substrates, Lon is in a state with both ATP and ADP bound (I). Binding of protein substrates can be visualized as occurring at the allosteric site on the subunit with ADP bound (II), promoting exchange between ATP and ADP (III). The binding of ATP to the subunit with the protein bound opens that active site and allows a portion of the protein to enter (III), and if an appropriate cleavage site is available on the protein, the protein is cleaved (IV). Although hydrolysis of ATP would not be necessary in cases where the products readily dissociate from the active site, ATP hydrolysis would produce the ADP-induced closed state of the active site, favoring dissociation of the cleavage product (V). The continued binding of the protein in the allosteric site could promote a second ATP/ADP exchange giving rise to (VI), which is similar to state (III) except that the substrate protein is not positioned in the active site. Since binding of the protein at the allosteric site might now be inhibitory, a mechanism must exist to either expel the protein from the site or allow it to move in order to position new sites in the protein for cleavage. One possibility is that ATP antagonizes binding of the protein to the allosteric site on the same subunit. This idea is also consistent with the proposal above that the protein initially binds to the subunit with ADP rather than the one with ATP. Hydrolysis of the ATP on the adjacent subunit (VII) would open the allosteric site on that subunit for binding the protein once it is released from the other subunit. At this point the decision must be made whether the protein will be further degraded or released. If the protein completely dissociates from state (VII), then the original state of the Lon protease (I) is regenerated. Alternatively, if the protein still has suitable sites for interaction with Lon, it can bind again at the allosteric site on the adjacent subunit (IIp), favoring ATP/ADP exchange (IIIp), cleavage of the next site will occur (IVp), and the cycle can continue 'processively' for as long as the protein has appropriate sites for recognition at the allosteric site and for cleavage.

There are two intriguing aspects of this scheme, which have to do with in vivo functions of Lon and other ATP-dependent proteases. First, the proposal that ATP binding and hydrolysis disrupt the binding of the protein at the allosteric site is analogous to the proposed role of

ATP in the activity of heat shock proteins or molecular chaperones; for example, complexes between DnaK and proteins are dissociated upon addition of ATP[76, 152]. Lon is, in fact, a heat shock protein in *E. coli*, and it will be interesting to see if homologous heat shock proteins are found in eukaryotic cells. Second, the requirement that protein substrates have two sites that can interact specifically with the protease provides an enormous increase in selectivity, especially since one of the interactions is accompanied by ATP hydrolysis. This mechanism is analogous to the 'kinetic proofreading' mechanism proposed by Hopfield[66] to explain how enzymes can show increased discrimination between substrates whose intrinsic affinities are not sufficiently different. The discriminatory ability increases exponentially with the number of interactions separated by thermodynamically favorable reactions. Thus, in vivo, the dependence on ATP provides a mechanism for increasing the specificity of protein degradation.

Specificity of Lon protease. Lon protease will degrade a large number of denatured proteins in solution. Most proteins are cleaved at a multiple sites yielding small peptides without the appearance of partially degraded intermediates[38, 94b]. The lack of accumulation of intermediates reflects a processive mechanism for degradation by Lon, because the same results are obtained in the presence of a large excess of the undegraded protein. Surprisingly, degradation appears processive even for very short proteins with relatively low affinity for the enzyme. Insulin B chain ($K_m \sim 160$ mM) is consistently cut at two sites with minimal release of intermediates. Casein is cleaved into about 15 peptides[104] and the rate of appearance of all of the peptides is essentially the same. Since the turnover numbers for degradation are only 8–16 min^{-1} (per Lon tetramer), Lon must retain the uncut portions of casein for at least 1–2 min, and possibly longer. The mechanism by which Lon retains proteins for the extended times required for processive cleavage is not known.

Lon cleaves proteins at peptide bonds between a large number of different pairs of amino acids, indicating that the binding pocket for substrates can accommodate a variety of side chains[94b]. Of the 14 bonds cleaved in 3 proteins, 12 bonds had a hydrophobic amino acid in the P_1 position, and in most of the cases there was a basic residue in the P_{-1}–P_{-4} positions[94b]. Cleavage of short peptidyl naphthylamides occurred preferentially with hydrophobic amino acids in positions P_1–P_3 and an acidic blocking group on the amino terminus[176]. It does not seem likely that these data are indicative of how Lon selects substrates in vivo. Since protein substrates do not inhibit cleavage of peptidyl amide bonds, the primary interaction between potential protein substrates and Lon should be at the allosteric site. So far, the nature of the interactions at that site are not known. The amino acid sequences and the various predicted physical or structural properties of four proteins known to be degraded by Lon in vivo show minor similarities around sites known to be cleaved in λ N protein[96], but these similarities probably reflect features tolerated by the proteolytic active site rather than determinants of recognition. Since either fusions to the carboxy terminus (V. Stout, unpublished) or deletions of the carboxy terminal regions of natural substrates[33] affect either interaction with or degradation by Lon, it seems likely that the carboxy end of proteins may be involved in recognition by Lon.

In vivo activities of Lon. A thorough review of the genetics and physiology of Lon was published recently by Gottesman, and readers should refer to that review for details of Lon function in vivo[50]. Studies with *lon* mutants indicate that, in vivo, Lon participates in the energy-dependent degradation of canavanyl proteins, puromycyl peptides, various missense mutant proteins, and nonsense fragments of proteins[12, 59, 61], and, in addition, carries out the rate-limiting step in the degradation of several naturally unstable proteins (table 2). All of the proteins degraded by Lon are also degraded by other proteases, but the relative contributions of Lon and other proteases to the degradation of different proteins varies. For example, the 50% defect in the degradation of canavanyl proteins in *lon* mutants suggests that Lon degrades these abnormal proteins at about the same rate as all other proteases combined, whereas Lon degrades SulA, RcsA, and N protein at least 10 times faster than do the other proteases. Thus, the absolute Lon-dependent degradation rate of naturally unstable targets is higher than that of abnormal proteins.

One important difference between abnormal and specific protein degradation in vivo is that heat shock proteins are involved in the former[162] and may not be involved in the latter. Lon itself is a heat shock protein, and its synthesis rate increases 2–4-fold upon temperature shift and upon causing the production of abnormal proteins in cells[48, 133]. However, mutations in the heat shock regulator or in other heat shock genes do not produce specific *lon* phenotypes such as UV sensitivity or mucoidy[50]. Nevertheless, *dnaJ* mutants are as defective in degradation of canavanyl proteins as are *dnaJ lon* double mutants, which implies that DnaJ is needed for all Lon dependent degradation of abnormal proteins. As pointed out by Gottesman[50], conditions that lead to production of abnormal proteins induce the heat shock response, so Lon-dependent degradation of abnormal proteins usually takes place in the presence of excess heat shock proteins. Heat shock proteins might form complexes with Lon to activate it for cleavage of abnormal proteins. Another alternative is that heat shock proteins are required to keep the abnormal proteins soluble and accessible to Lon; the abnormal proteins complexes to heat shock proteins might be recognized better by Lon. The issue of accessibility applies to the specific targets of Lon also. Lon can degrade extremely small amounts (< 10 molecules per cell) of some proteins[158]. Either Lon has a very high affinity for these proteins, or there is a mech-

anism for bringing Lon together with its substrates. In this regard, the DNA binding activity of Lon[24,181] is intriguing. Two of the substrates for Lon (RcsA and transposase) are probably DNA binding proteins[32,158], N protein binds RNA (A. Das, personal communication), and SulA may interact with the septation apparatus[69]. One explanation for the large size of Lon may be that it possesses binding sites for different cellular components and that at least a portion of it is localized in cells at functional foci where it would easily come into contact with its specific targets.

The two-component Clp protease

Cells lacking Lon protease still carry out energy-dependent protein degradation, and extracts of *lon* cells have casein degrading activity that is dependent on ATP[100]. We attempted to characterize the ATP-dependent casein-degrading activity and succeeded in purifying Clp, a protease composed of two components, ClpA and ClpP, that degrades casein and other proteins only in the presence of ATP[79]. The two components of Clp protease were readily separated from each other during purification, and in the absence of the other component, neither had proteolytic activity against casein or other proteins. Activity was restored upon combining the two components in the presence of ATP. Subsequently, others confirmed the two-component nature of Clp which they called Protease Ti[70].

Structure and properties of Clp protease. The two components of Clp are functionally distinct proteins and the products of separate genes. ClpA has a subunit M_r of 87,000 and has an intrinsic ATPase activity that is increased in the presence of ClpP and substrates[71,78]. The amino acid sequence of ClpA, derived from the DNA sequence, contains two consensus sequences for ATP binding sites[56,58]. The two sites are in separate regions of the protein and probably correspond to separate structural or functional domains. ClpP has a subunit M_r of 21,500 and is the proteolytic component of Clp[71,98,179]. ClpP is inhibited by diisopropylfluoro phosphate[70], and Ser111 has been identified as the site of modification[98]. Site-directed mutagenesis of Ser111 and of His135 identified these residues as elements of the catalytic triad expected for a classical serine protease active site[98].

ClpA purified in the absence of metal ions and nucleotides has an apparent M_r of 110–140,000[71,79]. Addition of Mg^{2+} and ATP to ClpA converts it to an associated form with a M_r of 450–500,000[95]. Thus, the structure of ClpA in vivo is most probably an ATP-bound hexamer. ClpP is purified as a 220–240,000 M_r species in the presence of $\geq$ 100 mM KCl but at low salt forms a species with M_r 460–500,000, an apparent dimer of the high salt form[71,97]. Electron micrographs of negatively stained ClpP[97] reveal that ClpP subunits are arranged in two rings of six subunits each which are superimposed to form a dodecamer (fig. 2). Slight asymmetries are apparent in many of the particles seen in the electron micrographs, which raises the possibility that the subunit associations might not all be identical or that there is some heterogeneity in the subunits themselves. The geometry of ClpP is similar to the basic unit of the proteasome, which is composed of superimposed rings of heterologous subunits.

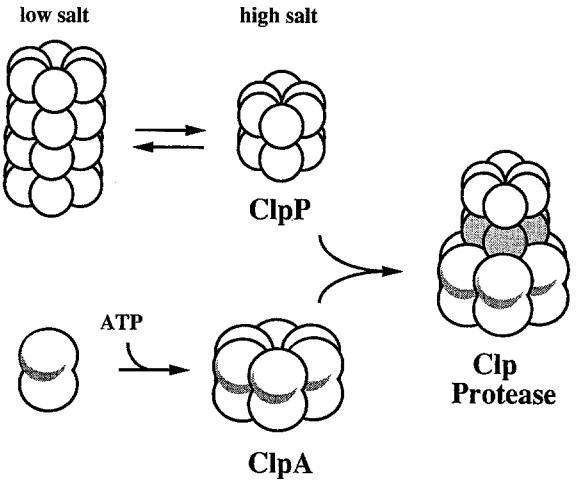

Figure 2. Structure of Clp protease. The structure of ClpP was determined from electron micrographs. The size of ClpA and the size of the ClpA/ClpP complex were determined by gel filtration and sedimentation velocity ultracentrifugation in the presence of ATP.

ClpA and ClpP do not form a complex in the absence of ATP, but addition of ATP to a mixture of ClpA and ClpP results in a rapid and specific association between the two proteins. The M_r of the associated species is $\sim$ 750,000[95] (MRM, unpublished), which would correspond to an association of 1 hexamer of ClpA with 1 dodecamer of ClpP (fig. 2). It is possible that, under the conditions for forming the active complex in vitro, the enzyme is partially dissociated. Activity titrations indicate that saturation of the active complex is achieved with a 1:2 subunit molar ratio of ClpA and ClpP, and thus one hexamer of ClpA would be expected to interact with a dodecamer of ClpP. ClpP inactivated with diisopropylfluoro phosphate (DIP-ClpP) retains the same native size as active ClpP and can interact with ClpA in an ATP-dependent manner[98]. The rate at which DIP-ClpP displaced active ClpP from its complex with ClpA was used to calculate the off rate of ClpP from the active Clp protease under assay conditions[95]. k_{off} for ClpP was $\sim 2 \times 10^{-3}\,s^{-1}$, which assuming an association rate constant of $10^6\,M^{-1}s^{-1}$ would give a dissociation constant of $\sim 2 \times 10^{-9}\,M^{-1}$ for the complex.

ATP-Dependence and other catalytic properties of Clp. Clp protease hydrolyzes a number of proteins endoproteolytically generating short acid-soluble peptides. As with Lon, Clp degrades protein substrates to small peptides even the presence of excess protein substrate, indicating a processive mechanism in which multiple peptide bonds are cleaved in the same substrate without the release of large intermediates. The k_{cat} for peptide bond cleavage of

[^{3}H]α-casein is 9–15 min^{-1} (table 3), somewhat faster than Lon protease. Clp shows no measurable activity on casein or other large proteins without ATP. dATP will also activate Clp proteolytic activity [70], but no other nucleotide triphosphates or any non-hydrolyzable analogs of ATP can promote proteolytic activity [79]. Thus, Clp has a stricter requirement for ATP than does Lon protease. ClpA has an ATPase activity in the absence of ClpP and other proteins, and proteolysis is accompanied by an increase of 80–100% in the ATPase activity. The coupling between peptide bond cleavage and ATP hydrolysis has not been studied in detail but the k_{cat} for ATPase in the presence of ClpP and saturating protein substrate is 80–100 min^{-1}, which would correspond to an increase of 3–4 and a total of 6–8 ATP molecules hydrolyzed per peptide bond cleaved (MRM, unpublished).

The association of ClpA and ClpP does not lead to an increase in ATPase activity. Also, the non-hydrolyzable analog AMPPNP, weakly promotes self-association of ClpA and formation of the ClpA/ClpP complex but does not activate proteolysis [95]. The requirement for ATP binding in the assembly of the active Clp protease illustrates the allosteric role that ATP can play in activating proteolysis by ATP-dependent proteases. Since ATP hydrolysis is required for the degradation of proteins, it is clear that ATP functions both in the activation of Clp and in the catalytic cycle. These two roles for ATP are somewhat analogous to the two roles of ATP with Lon protease: the first ATP is primarily an allosteric activator, opening the active site of Lon or allowing association of ClpA and ClpP to open the active site of ClpP; the second ATP is involved at a poorly-understood second step that is necessary for cleavage of proteins. The latter step could be involved in unfolding or translocation of the protein, as was suggested for Lon.

The proteolytic active site lies entirely within ClpP. ClpP alone will hydrolyze the amide bond of short fluorogenic peptides [179], and ClpP will also endoproteolytically degrade small polypeptides in the absence of ClpA and ATP, but only about 1% of the rate of cleavage observed in the presence of ClpA and ATP (MRM, unpublished). The degradation of peptides by ClpP is not activated by casein or other protein substrates as occurs with Lon. At the same time, casein is not an inhibitor of peptide cleavage by ClpP in the absence of ClpA and ATP but is an inhibitor when ClpA and ATP are both present (MRM, unpublished). It thus appears that the active site of ClpP is accessible to peptides without activation but can bind protein substrates only as part of the active complex with ClpA. With insulin B chain as a substrate, cleavage by ClpP alone occurs at the same peptide bonds as with the ClpA/ClpP complex, thus the specificity of ClpP is not changed by its association with ClpA (MRM, unpublished). ClpP alone makes several cuts in insulin B chains in the presence of excess uncut substrate, displaying a similar processivity as the ATP-dependent degradation reaction. If ATP hydrolysis is not required for processive cleavage of proteins, it is likely that the multiple-site cleavage of protein substrates is a function of the multisubunit structure of the proteases rather than any energy-consuming protein translocation event.

Clp degrades a broad range of peptide bonds (MRM, unpublished). Clp did not show the same specificity towards fluorogenic peptides shown by Lon, although the peptide cleaved, succinyl-Leu-Tyr-amidomethylcoumarin [176], had a similar hydrophobic peptide with an acidic blocking group at the amino terminus. No clear specificity emerged when the peptide bonds cut in glucagon or insulin B chain were examined; a number of peptide bonds with hydrophobic residues in the P_1 position were cleaved but sites containing non-hydrophobic residues were also cut (MRM, unpublished). As with Lon, it is likely that the specificity for protein degradation depends on interactions other than at the proteolytic active site.

In vivo activities of Clp protease. Mutations in *clpA* or *clpP* lead to a partial defect in the degradation of abnormal proteins, which is more easily seen in cells lacking Lon protease [78]. About 15% of canavanyl protein turnover is attributable to Clp, and the additional effect of a *clp* mutation on the degradation of individual canavanyl proteins degraded by Lon is evident on SDS gels [96]. Clp can also specifically degrade certain LacZ-fusion proteins, in particular, a highly unstable protein that contains the first 40 amino acids of ClpA itself fused to LacZ [56]. Natural substrates for Clp have not been identified, and mutants lacking Clp protease have no obvious growth defects [56].

Both ClpA and ClpP are present in cells grown under a variety of conditions; ClpA and ClpP levels increase toward the end of exponential growth. ClpP synthesis increases about two-fold upon heat shock, and the *clpP* gene is at least partly under control of the heat shock sigma factor [82]. ClpA levels are slightly higher in cells grown at elevated temperatures but synthesis of ClpA at 42 °C is not dependent on $\sigma 32$ [78]. The separate regulation of *clpA* and *clpP* at high temperatures suggests that, at least under some physiological conditions, ClpP might have activity in vivo independent of ClpA. In this light, the discovery of *E. coli* ClpB, which is highly homologous to ClpA [58] raises the intriguing possibility that ClpP may have more than one activator in vivo.

One activity of ClpP that is expressed independently of ClpA is an unusual self-processing reaction [98]. ClpP is synthesized with a 14-amino acid extension at the amino terminus which is removed very rapidly ($t_{1/2} \leq 2$ min) to give the mature protein which constitutes > 90% of the ClpP in the cell [97, 98]. Processing of ClpP requires an active ClpP but not an active ClpA in the cell. ClpP made from a single copy of *clpP* in the chromosome is sufficient to allow processing of inactive mutants of ClpP cloned in multi-copy. Thus, ClpP auto-processing can occur intermolecularly, although whether the ClpP subunits must be in the same or separate dodecamers has not been established. The

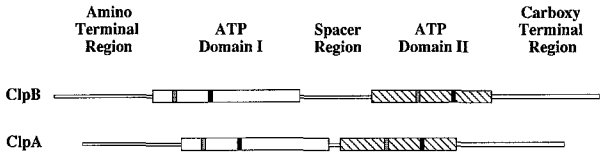

Figure 3. Sequence domains in ClpA and ClpB subsets of ClpA-like proteins. The ClpA subset is found in *E. coli* and *R. blastica*. The ClpB subset is found in *E. coli, B. nodosus*, and *T. brucei*. Tomato has a spacer region about one-half the size of the spacer in ClpB.

function of this auto-processing reaction for ClpP is not known.

The Clp family of ATP-dependent proteases. ClpA and ClpP are highly conserved in many if not all organisms[154b]. Sequence comparisons between *clpA* and *clpP* and a number of genes of unknown function gave the first indications of the widespread occurrence of these genes[58,98]. The high degree of conservation of the translated amino acid sequences of these genes strongly suggests that the proteins have conserved the function of ATP-dependent proteolysis as well. A schematic diagram of the ClpA family of proteins is shown in figure 3. The central features of the protein sequences are the two highly conserved domains surrounding the ATP binding site consensus sequences. Conservation within the two regions is very high: 55–80% identical, 20–35% similar in the 230 amino acids in ATP domain 1 and 50–80% identical, 20–40% similar in the 190 amino acids in ATP domain 2. The two domains on the other hand show almost no homology with each other, indicating that the two regions arose from evolutionarily distant ancestors. Conservation of the ClpP amino acid sequences in homologous plant chloroplast proteins are also quite extensive: 36–46% identical and 24–34% similar amino acids throughout the entire lengths of the proteins. The active site serine and histidine lie in very well conserved regions lending support to the idea that the ClpP homologs are also proteolytic enzymes. Although ClpP sequences have only been observed in *E. coli* and plant chloroplasts, immunochemical screening indicates that proteins with conserved structures similar to ClpP are found in a number of prokaryotic and eukaryotic organisms[98].

E. coli contains two members of the ClpA family, ClpA and ClpB[58]. ClpB contains a 120-amino acid spacer region between the ATP-binding domains. The central spacer is moderately conserved in some members of the family but is not present in all members, and thus may serve to further divide the family into ClpA-like and ClpB-like subfamilies[154b]. It is possible that these subfamilies have evolved different catalytic functions which are reflected in the central regions or in the less well-conserved amino and carboxy terminal portions of the proteins. It seems more likely that the differences reflect rather the specificities of the ATPase subunits for their corresponding proteases or for the types of protein substrates with which they interact. ClpB may be more specific for directing proteolysis at the types of abnormal proteins that arise under heat shock conditions or ClpB may have specific regulatory targets under conditions of stress. Comparison of the functions of ClpA and ClpB should provide valuable insights into the specificity of action and function of intracellular proteases.

RecA-mediated protein degradation

Damage to *E. coli* DNA results in the proteolytic inactivation of the LexA repressor and the consequent induction of a number of DNA repair functions[92,174]. In lysogens of bacteriophage λ, the CI repressor is also cleaved following DNA damage, resulting in induction of λ lytic functions[142]. Both repressors are stable in vivo and degradation is initiated only upon activation of the RecA protein, a multifunctional protein also involved in DNA recombination[174]. The mechanism of activation of RecA in vivo is not completely understood but it may be mediated by binding of single-stranded DNA fragments resulting from the damage to DNA[92]. Cleavage of both repressors in vivo occurs at an Ala-Gly bond, which splits the protein into two large fragments[142]. Subsequent studies have shown that another protein, UmuD, is activated by cleavage in a RecA-dependent manner in vivo[13,15,116,148].

The mechanism of repressor cleavage in the presence of RecA has been studied by Little and co-workers[90,91,143,154], who have been able to obtain cleavage of LexA and CI at neutral pH in vitro in the presence of RecA, ATP, and single stranded DNA. ATP apparently acts as an allosteric effector for RecA, since non-hydrolyzable analogs activate as well[29,91]. Remarkably, at high pH in the presence of divalent metal ions, LexA and CI are spontaneously cleaved in the absence of RecA and ATP[90]. Both the RecA-mediated and the spontaneous cleavage of the repressors occur at the same Ala-Gly bond cleaved in vivo[90,154]. These observations led Little to propose that degradation of LexA and other 'substrates' for RecA occurs via a self-cleaving reaction following conformational changes produced either by the interaction with RecA at neutral pH or by the combination of high pH and metal ions for the spontaneous reaction. Kinetic studies suggest that the auto-cleavage reaction is intramolecular, indicating that the capacity for degradative activity lies within a single LexA molecule[90]. Mutational analysis of LexA showed that Ser119 and Lys156 in the carboxy terminal half of the protein, far from the cleavage site, are required for cleavage of the repressor both in vivo and in vitro[154]. Both residues are conserved in lambda repressor and UmuD[131,154]. A catalytic model has been proposed in which Ser119 serves as the nucleophilic group attacking the Ala-Gly peptide bond, similar to its role in classical serine proteases, and Lys156 functions by partially deprotonating the serine hydroxyl[154]. Major support for this model has come from the demonstration that diisopropylfluoro phosphate reacts with Ser119 of LexA and inactivates the protein for auto-degradation[143].

The details regarding the mechanism of auto-degradation of LexA reveal a number of instructive parallels with the properties of intracellular protein degradation. LexA degradation is regulated in that it is initiated only in response to a specific metabolic signal. The degradation is highly selective; there is no evidence that LexA can cleave any protein other than itself, although a carboxy-terminal fragment lacking the cleavage site is reactive with DFP[143]. The combination of LexA and RecA acts as a two-component protease, in which the regulatory subunit (RecA) activates the protease (LexA). Interaction with RecA could result in positioning the 'substrate' portion of LexA in the active site or might be necessary to form the catalytically competent active site. Degradation is ATP-dependent, and ATP acts as an allosteric effector for the regulatory subunit to promote a conformation favoring interaction with the protease. The emergence of common principles in degradation of proteins, extending even to the specialized instances of auto-degradation, is encouraging and promises that we are, in fact, on the road to understanding the order that necessarily underlies this potentially chaotic process.

Other possible energy-dependent proteases: Alp, ClpB

E. coli possesses energy-dependent proteases other than Lon, Clp, and RecA. As indicated earlier, *lon clp* double mutants retain 20–30% of the wild-type energy-dependent degradative capacity toward canavanyl proteins and essentially all of the degradative capacity during starvation. Two candidates for energy-dependent proteases are ClpB[58] and the protease associated with the *alp* function[169]. ClpB is a close homolog of ClpA, but its involvement in proteolysis is only speculative at this point. *clpB* mutants are not defective in abnormal protein degradation and do not affect the expression of Lon^- or Clp^- phenotypes (Gottesman, Maurizi, unpublished). Nevertheless, it seems likely that ClpB alone or in combination with another protein (ClpQ?) will have possibly quite specific ATP-dependent proteolytic activity.

The *alp* function was identified by Trempy and Gottesman[169] as an activity that suppressed *lon*$^-$ phenotypes when expressed on a multi-copy plasmid. The degradation of SulA, which is blocked in a *lon* mutant, is restored by multi-copy *alp*, and the *alp*-dependent degradation in vivo is inhibited by cyanide. *alpA* encodes a protein of 75 amino acids. The AlpA protein is probably not itself a protease or part of a protease, but appears to be required for a genetic rearrangement which gives rise to an energy-dependent protease capable of cleaving Lon substrates (J. Kirby, S. Gottesman, unpublished). *alp*-dependent degradation does not involve RecA, ClpA, ClpB, ClpP, HflA, or heat shock proteins. The natural substrates for the protease induced by *alp* are not known.

Non-ATP-dependent proteases

E. coli possesses a large number of proteases that do not require ATP or other forms of metabolic energy for activity. Several of these proteins have specialized functions (e.g., signal peptidases and outer membrane protease, OmpT) or are found in locations that suggest that they have access to a limited range of possible substrates (e.g., Protease Pi and other periplasmic proteases). For a majority of the proteases, the genes have not been identified and no mutations are available. Some properties of these proteases are listed in table 1. A few proteases that have unusual properties or specificity will be described further.

Signal peptidases and signal peptide peptidase

Protein secretion and the processing of secreted proteins have been reviewed previously[105,119]. The proteases that remove the signal peptide from secreted proteins do not appear to have general protein degrading activity but catalyze a highly specific limited proteolysis. *E. coli* has at least two such signal peptidases, which act on different groups of proteins. Signal peptidase I (SPase I) was purified on the basis of its cleaving the signal peptide from M13 coat protein, and has now been shown to remove the signal peptide from a large number of secreted proteins[182,183]. Signal peptidase II removes the signal peptide from *E. coli* lipoproteins[74,170,171].

Precursors of secreted proteins are synthesized and extruded through the membrane co-translationally. The secreted proteins are anchored in the membrane by the amino terminal signal peptide, which has 1–3 basic amino acids near the terminus followed by a membrane-spanning non-polar stretch of amino acids. The processing site, where the precursor is cleaved to yield the mature protein, is usually within 15–30 amino acids of the amino terminus[83,105]. Processing of membrane proteins proceeds in a similar manner, except that the protein is extruded in several portions separated by membrane-spanning domains rather than in its entirety and the processing site lies within the first periplasmic loop. In vivo, cleavage of signal peptides is very efficient ($t_{1/2} < 1$ min), but numerous pulse-chase experiments indicate that the entire protein is usually synthesized before the signal is removed.

SPase I removes the signal peptide from all secreted proteins except lipoproteins[183]. Cleavage of precursors ususally occurs at sites with Ala in the P_1 position, and in > 50% of the proteins the cut follows the sequence Ala-X-Ala[105]. The minimum requirement for processing is the presence of amino acids with small, uncharged side chains in the P_3 position and helix-breaking amino acids in positions P_4–P_6. Purified SPase I will remove signal peptides from proteins in vitro in the presence of detergents[178,182], indicating that the protease can recognize and interact with processing sites directly. In vivo, however, the precursor is anchored to the membrane by the signal peptide, which must alter the conformation or the availability of the cleavage site. Mutational studies indicate that in addition to the sequence around the cleavage site, structural or sequences elements in the signal peptide and elsewhere in the protein affect processing[83]. Thus,

the efficiency and the specificity of proteolysis by signal peptidases appears to depend on the principle of bipartate recognition as suggested above for energy-dependent proteases.

SPase II is an outer membrane protease that removes the signal peptide from *E. coli* lipoproteins [170–172]. SPase II cleaves after the glycine residue in the sequence Leu-Ala-Gly-Cys(glyceride) and specificity appears to rely primarily on the recognition of this unusual modified amino acid residue. Precursors without the glyceride modification are not cleaved by either signal peptidase [172].

The signal peptides removed from the precursors are rapidly degraded in vivo by signal peptide peptidases; several proteases have been found that will degrade signal peptides in vitro [72, 117]. Protease IV, a cytoplasmic membrane protease, was shown to be identical to a protease purified on the basis of its ability to degrade the signal peptide from *E. coli* lipoprotein [72, 73]. Protease IV cleaves the signal peptide in solution but does not degrade the precursor protein with the signal peptide attached. The protease is inhibited by leupeptin and chymostatin and is thought to be a serine protease based on its sensitivity to diisopropylfluoro phosphate [117]. Mutants lacking Protease IV accumulate large amounts of lipoprotein signal peptide [73].

OmpT, DegP, Protease III

DegP (HtrA) is a 48,000 M_r serine protease found in the periplasm of *E. coli* [89, 160]. The *degP* function was originally identified in vivo by mutational analysis because *degP* mutants could not degrade alkaline phosphatase fusion proteins [160]. DegP is a heat-shock protein and was identified independently as a function needed for growth of *E. coli* at high temperatures [88]. Synthesis of DegP (HtrA) is not under control of $\sigma 32$ but appears to be part of another heat shock regulon controlled by another sigma factor, σE [88]. DegP may normally function to destroy damaged membrane or periplasmic proteins or proteins that are not properly processed. For example, DegP degrades the precursor of colicin A lysis protein that accumulates in cells treated with globomycin or in over-producing cells [17]. DegP does not appear to have an essential activity in cells grown at low temperatures, since *degP* mutants are viable at low temperatures. The temperature sensitivity of *degP* mutants may reflect an accumulation at high temperatures of damaged proteins that interferes with essential periplasmic functions. No regulatory targets for DegP have been identified.

The outer membrane protease, OmpT, is interesting because of its specificity in cleaving proteins between basic residues [60, 165]. OmpT, which is the same as Protease VII [165], is a serine protease with subunit M_r of 36,000 and a native M_r of 180,000. Although OmpT has been implicated in the degradation of a number of proteins in crude cell extracts, no in vivo functions have been associated with any of these reactions. A physiological function for OmpT has been suggested by the observation that colicin entry into and exit from *E. coli* cells is accompanied by OmpT-dependent cleavage [16]. OmpT has also been implicated in the degradation of certain secreted fusion proteins [6].

Protease III (Protease Pi) is a 110,000 M_r periplasmic protease that requires Zn^{2+} for activity [21]. The protease cleaves a number of low molecular weight proteins endoproteolytically but has low activity against high molecular weight proteins. Protease III cuts insulin B chain between a Tyr-Leu bond and appears to have a chymotrypsin-like specificity. Mutants lacking Protease III have no obvious growth defects and are not deficient in the degradation of abnormal proteins [22]. Protease III is a member of a family of metalloproteases that have homology to human insulinase [28, 37, 42, 138]. This family includes a number of proteases thought to be involved in the processing of mitochondrial precursor proteins. Members of the insulinase superfamily bind Zn^{2+} but do not have the Zn^{2+} binding motif found in the majority of Zn-metalloproteases [138].

HFLA and the degradation of the λ CII protein

The λ CII protein is one of the most rapidly degraded proteins found in *E. coli* cells ($t_{1/2} < 2$ min) [57]. The regulatory effects of CII and control of its degradation were recently reviewed by Gottesman [50]. The rate of degradation of CII in vivo is reduced about 50% by mutations in either of two loci, *hflA* or *hflB*. The *hflA* locus encodes two proteins, HflK and HflC, and purified preparations of these proteins degraded purified CII protein very slowly in vitro [4]. Complete degradation of CII required incubations of ~ 1 h and relatively large amounts of the protease. Degradation was inhibited by serine protease inhibitors as well as by heavy metal ions, but not by sulfhydryl reagents. CIII protein, which partially protects CII in vivo, inhibited degradation only at very high concentrations. The slowness of the in vitro degradation reaction was unexpected, given the rapid turnover of CII in vivo. Possibly, the product of *hflB* is required along with the *hflA* products for efficient degradative activity.

Other cytoplasmic proteases

Proteases Do, Re, and So [26, 51, 126, 166] are soluble, cytoplasmic proteases whose in vivo functions are not known. All three proteases have been purified and shown to be serine proteases that can degrade casein endoproteolytically. A number of *E. coli* proteins have been tested as substrates for Do, Re, and So, and some attempts have been made to assign a function for the proteases in vivo based on in vitro activities. Protease Re appears similar to a protease reported to degrade oxidized glutamine synthetase in vitro [126, 144]. Protease So also degrades oxidized glutamine synthetase in vitro, at about five times the rate as Protease Re [86]. While it is tempting to suggest that these proteases may be responsible for the degradation of oxidized proteins, which are known to be degraded at an accelerated rate in vivo [30], there are no specific data linking Proteases Re and So to control of oxidative

damage in cells. Proteases Do and So have been shown to degrade the Ada protein in vitro [85]. Ada takes part in repair of methylated DNA [87] by carrying out an unusual suicide reaction in which Cys321 of the protein accepts a methyl group from O^6-methylguanine in DNA. The methylated Ada is no longer functional in demethylation and has a short half-life in vivo. Several other proteases will degrade Ada in vitro, although one of them is an outer membrane protease and is not likely to be the responsible protease in vivo [147]. Protease So has also been shown to degrade signal peptides after release from precursor proteins in vitro [117]. While all of the above activities show some degree of specificity in that only a limited number of proteases have been shown to have a particular degradative activity, there are no biochemical data on the degradative reactions in vivo to correlate with the in vitro findings.

Undiscovered functions, undiscovered proteases

Of the 25 proteases listed in table 1, only six have known physiological functions and target specific proteins or classes of proteins for degradation. Three others have been shown to affect the degradation of abnormal proteins in vivo. Identification of the physiological functions of the remaining proteases will require mutants lacking each of the proteases. One may question whether all proteases found in cells will have unique targets, but the data obtained with Lon and Clp suggest that proteases have activity specifically directed at a few physiological substrates as well as general degradative activity toward abnormal proteins. Since most abnormal protein degradation in the cytoplasm is energy-dependent, it is unlikely that non-energy-dependent cytoplasmic proteases are involved in the initial degradative steps. These proteases could be responsible for the secondary steps in the turnover of proteins that have been partially degraded, or they might act in concert with energy-dependent enzymes, such as heat shock proteins. The proteases might overlap in their specificities for different kinds of abnormal proteins, as Lon and Clp overlap in their activities toward canavanyl proteins.

There are a number of proteolytic functions in vivo for which the proteases have not yet been identified. The instances of rapid degradation are particularly intriguing, e.g., σ32; λ O protein, CIII, and Xis (table 2), since these reactions probably require highly selective proteases. The proteases involved in degradation during starvation, which is energy-dependent, have not been identified, nor have the energy-dependent proteases that degrade abnormal proteins in the absence of Lon and Clp. While it is possible that some of the known proteases may be responsible for some of the above functions, it is probable that still more proteases remain to be discovered in *E. coli*, at least one of which should be energy- or ATP-dependent.

Perspective

Regulatory significance of protein degradation

Proteolytic enzymes serve useful and sometimes essential functions for living cells. The importance of limited proteolysis in regulating enzyme activities or in the assembly of macromolecular complexes in prokaryotic and eukaryotic cells has been understood for some time. Proteolytic turnover of enzymes has been recognized as a major element of the response to developmental changes and to shifting metabolic conditions. In animal cells degradation of enzymes contributes to the maintenance of metabolic balance, and in prokaryotes and eukaryotes degradation of subunits is a mechanism for adjusting the stoichiometry of the components of multimeric assemblies. In recent years, two new aspects of intracellular proteolysis have become apparent: (a) a significant number of regulatory proteins are very rapidly degraded in cells, and this degradation is a major post-translational mechanism for regulating their intracellular levels and activities, and (b) mutations in specific proteases can impair growth or be conditionally lethal.

In *E. coli*, most of the highly unstable proteins are from bacteriophage lambda (table 2). The constitutively unstable λ proteins are positive regulators and degradation would serve to modulate their activity. The conditionally unstable protein, CI repressor, is a negative regulator that is inactivated in response to environmental conditions. Two of the *E. coli* proteins, σ32 and RcsA, are also positive regulators and blocking degradation leads to an increase in the functional level. The third rapidly degraded *E. coli* protein, SulA, represents a different class of unstable protein – a protein made during an emergency response whose activity could be detrimental to the cell after the emergency is over [54, 55]. In the case of SulA, which is made in response to DNA damage and inhibits septation, failure to degrade the protein leads to irreversible filamentation and cell death [69]. In eukaryotic cells a number of rapidly degraded regulatory proteins are also positive regulators. Cyclin is required to activate cdc2 kinase and degradation of cyclin during mitosis is needed for proper timing of the cell cycle in developing cells [111]. Several viral positive transcriptional regulators, including adenovirus E1A [153], are also unstable. Rapid degradation of regulatory proteins, particularly positive regulatory proteins, appears to be quite common in both eukaryotic and prokaryotic cells and further examples are sure to be demonstrated.

In *E. coli*, mutations in any single protease (except signal peptidases) or in combinations of proteases are not lethal under normal conditions. However, as indicated above, *lon* mutations are conditionally lethal under conditions that induce SulA. In yeast, temperature-sensitive mutations in the ubiquitin conjugating system, which marks proteins for degradation, produce growth defects at high temperature [41], implying that some ubiquitin-dependent degradative activity is required for normal cell division.

More recently, it has been shown that mutations in components of the multicatalytic protease (proteasome) in yeast also cause loss of viability[43,63a]. Identification of the proteins targeted by the ubiquitin system or by the proteasomes and learning how blocking the degradation of these proteins affects cell growth promises to provide additional insights into the complex regulatory roles played by proteases in living cells.

Similarities in the energy-dependent proteolytic systems of E. coli and eukaryotic cells

Proteases with enzymatic properties similar to both Clp and Lon have been found in prokaryotic and eukaryotic cells. The *lon* function is found in other bacteria[36,145], but the enzymes have not been studied. A Lon-like protease was found in mitochondria, but this protease has not been characterized thoroughly[34]. On the other hand, the major cytoplasmic proteases in eukaryotic cells, the proteasome[120,141] and the 26S ATP-dependent protease[67], have remarkable similarities to Clp protease. The arrangement of subunits in proteasomes and in ClpP are quite similar. ClpP has superimposed hexagonal rings of subunits, which resemble proteasome particles on face in the electron microscope[40]. The four-layered cylindrical structure of the proteasome has not been observed for ClpP; however, in low salt, ClpP does form a higher molecular weight structure that should have 24 subunits (fig. 2). Although proteasome preparations are known to contain 12–15 subunits and at least 5 different genes for proteasome subunits have now been identified, it is not known whether each proteasome is composed of a full complement of subunits or whether different particles have different compositions. Proteasomes contain at least three different proteolytic activities. ClpP has identical subunits but it appears that subunit interactions give rise to differential expression of protease and peptidase activities. ClpA and ClpP interact in the presence of ATP to form the active Clp protease. This reaction is similar to that observed in formation of the 26S protease from the proteasome and two soluble proteins in the presence of ATP[39,44]. The complex of the proteasome and the other proteins converts the enzyme into an ATP-dependent protease. Thus, both Clp and the 26S protease employ a similar enzymatic strategy in assembling multiple proteolytic active sites in a complex with an ATP-dependent regulatory subunit.

One difference in the two systems is that, whereas ClpA undergoes a self-association in the presence of ATP and the associated ClpA interacts with ClpP, the regulatory components of the 26S protease do not self-associate nor do they associate with each other in the presence of ATP. Also, none of the components of the eukaryotic system has ATPase activity in the absence of the other components. The 26S protease can degrade ubiquitin-conjugated proteins and thus has no need to recognize protein substrates directly but should have a binding site for ubiquitin and a mechanism to present the substrate portion of the conjugates to the proteolytic active sites. Clp can degrade proteins without modifications, and therefore must have a site for interaction with protein substrates directly. The substrate recognition portion of ClpA might function analogously to the E3 protein of the ubiquitin conjugating system[184]. Instead of presenting the protein for conjugation, however, ClpA would present the substrate directly in the proteolytic active site. Given the major differences between Clp and the 26S protease, it seems unlikely that the proteases are closely related or homologous. At the same time, the similarity

Table 4. Comparison of Lon and Clp proteases

Properties	Clp	Lon
Subunit size	A: 83,000 P: 21,500	87,000
Native size	A plus ATP: > 450,000 P: 240,000 $P_{12}A_6$: ≥ 700,000 (multimeric state requires ATP)	450,000 (stable tetramer)
Protease active site	Serine protease	Probably serine protease
ATP binding sites	2 per A subunit (similar affinities)	1 per subunit (2 strong, 2 weak per tetramer)
ATPase activity	Substrate activated	Substrate activated
ATP/peptide bond	6–10	2–4
Peptidase activity	P only; no nucleotide required	Requires nucleotide, but not ATP hydrolysis; activated by proteins
Protease activity	Activated by ATP	Activated by ATP for protein hydrolysis; AMPPNP stimulates cleavage of N protein
In vivo substrates	Some fusion proteins; few abnormal proteins	Sul A; Rcs A; N protein; many abnormal proteins
Regulation	A: high cell densities P: heat shock (σ32)	Heat shock (σ32) and other stresses

in basic biochemical properties between the two proteases promises that information gained from either system will be instructive in understanding the other.

Clp protease is highly conserved in all organisms, but no Lon homologs have yet been found outside of enteric bacteria. Catalytic properties of Clp and Lon in vitro are somewhat similar (table 4) but the structures of the proteins are quite different. The sequence similarities between ClpA and Lon are restricted to the ATP-binding consensus sequences, with domain 2 of ClpA having slightly more similarity in predicted structure for that region. There does not appear to be much sequence similarity between ClpP and Lon, and in particular the sequences around the active site serine and histidine in ClpP are not found in Lon.

Other bacteria have ClpA homologs with the same short spacer region between ATP domains (fig. 3) and are surely components of ATP-dependent proteases. The plant homologs have a longer spacer region that is not the same as that found in ClpB. Given the chloroplast transit sequence found at the amino terminal part of the tomato ClpA and the presence of ClpP in the chloroplast, it is highly probably that the Clp homologs constitute a functional protease in plants. The other eukaryotic homologs of ClpA are more similar to ClpB, since they have the longer spacer region and show marginally greater conservation of amino acids[58]. ClpB has recently been reported to be a heat shock protein in *E. coli*[80a, 154a] and to have homology with the yeast heat shock protein, HSP104[126a]. *E. coli* ClpB has not yet been shown to have proteolytic activity alone or in combination with another protein, but there should be little doubt that ClpB and the related eukaryotic proteins are parts of ATP-dependent proteases. It may be that ClpA and ClpB direct degradation of different proteins, and it will be interesting to see if more classes of ClpA-like proteins will be found in *E. coli* and in other cells.

Acknowledgment. I am indebted to Susan Gottesman for the many ideas she has generously and enthusiastically shared over the years.

1 Amerik, A. Y., Antonov, V. K., Gorbalenya, A. E., Kotova, S. A., Rotanova, T. V., and Shimbarevich, E. V., Site-directed mutagenesis of La protease. FEBS Lett. *287* (1991) 211–214.
2 Bachmair, A., Finley, D., and Varshavsky, A., In vivo half-life of a protein is a function of its amino-terminal residue. Science *234* (1986) 179–186.
3 Bachmair, A., and Varshavsky, A., The degradation signal in a short-lived protein. Cell *56* (1989) 1019–1032.
4 Bahl, H., Echols, H., Straus, D. B., Court, D., Crowl, R., and Georgopoulos, C. P., Induction of the heat shock response of *E. coli* through stabilization of sigma 32 by the phage lambda cIII protein. Genes Development *1* (1987) 57–64.
5 Baker, T. A., Grossman, A. D., and Gross, C. A., A gene regulating the heat shock response in *Escherichia coli* also affects proteolysis. Proc. natl Acad. Sci. USA *81* (1984) 6779–6783.
6 Baneyx, F., and Georgiou, G., In vivo degradation of secreted fusion proteins by the *Escherichia coli* outer membrane protease OmpT. J. Bact. *172* (1990) 491–494.
7 Banuett, F., Hoyt, M. A., McFarlane, L., Echols, H., and Herskowitz, I., *hflB*, a new *Escherichia coli* locus regulating lysogeny and the level of bacteriophage lambda cII protein. J. molec. Biol. *187* (1986) 213–224.
8 Ben-Bassat, A., Bauer, K., Chang, S. Y., Myambo, K., Boosman, A., and Chang, S., Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine peptidase and its structure. J. Bact. *169* (1987) 751–757.
9 Bond, J. S., and Butler, P. E., Intracellular proteases. A. Rev. Biochem. *56* (1987) 333–364.
10 Bonnefoy, E., Almeida, A., and Rouviere-Yaniv, J., Lon-dependent regulation of the DNA-binding protein HU in *Escherichia coli*. Proc. natl Acad. Sci. USA *86* (1990) 7691–7695.
11 Bowie, J. U., and Sauer, R. T., Identification of C-terminal extensions that protect proteins from intracellular proteolysis. J. biol. Chem. *264* (1989) 7596–7602.
12 Bukhari, A. I., and Zipser, D., Mutants of *Escherichia coli* with a defect in the degradation of nonsense fragments. Nature *243* (1973) 238–241.
13 Burckhardt, S. E., Woodgate, R., Scheuermann, R. H., and Echols, H., UmuD mutagenesis protein of *Escherichia coli*: Overproduction, purification, and cleavage by RecA. Proc. natl Acad. Sci. USA *85* (1988) 1811–1815.
14 Canceill, D., Dervyn, E., and Huisman, O., Proteolysis and modulation of the activity of the cell division inhibitor SulA in *Escherichia coli lon* mutants. J. Bact. *172* (1990) 7297–7300.
15 Caron, P. R., and Grossman, L., Potential role of proteolysis in the control of UvrABC incision. Nucl. Acids Res. *16* (1988) 10903–10912.
16 Cavard, D., and Lazdunski, C., Colicin cleavage by OmpT protease during both entry into and release from *Escherichia coli* cells. J. Bact. *172* (1990) 648–652.
17 Cavard, D., Lazdunski, C., and Howard, S. P., The acylated precursor form of the Colicin A lysis protein is a natural substrate of the DegP protease. J. Bact. *171* (1989) 6316–6322.
18 Charette, M., Henderson, G. W., and Markovitz, A., ATP hydrolysis-dependent activity of the *lon*(*capR*) protein of *E. coli* K12. Proc. natl Acad. Sci. USA *78* (1981) 4728–4732.
19 Charette, M. F., Henderson, G. W., Doane, L. L., and Markovitz, A., DNA Stimulated ATPase Activity of the Lon (CapR) Protein. J. Bact. *158* (1984) 195–201.
20 Cheng, H. H., Muhlrad, P. J., Hoyt, A., and Echols, H., Cleavage of the cII protein of phage lambda by purified HflA protease: control of the switch between lysis and lysogeny. Proc. natl Acad. Sci. USA *85* (1988) 7882–7886.
21 Cheng, Y. S., and Zipser, D., Purification and characterization of protease III from *Escherichia coli*. J. biol. Chem. *254* (1979) 4698–4706.
22 Cheng, Y.-S. E., Zipser, D., Cheng, C.-Y., and Roiseth, S. J., Isolation and characterization of mutations in the structural gene for protease III (*ptr*). J. Bact. *140* (1979) 125–130.
23 Chin, D. T., Goff, S. A., Webster, T., Smith, T., and Goldberg, A. L., Sequence of the *lon* gene in *Escherichia coli*: A heat-shock gene which encodes the ATP-dependent protease La. J. biol. Chem. *263* (1988) 11718–11728.
24 Chung, C. H., and Goldberg, A. L., DNA stimulates ATP-dependent proteolysis and protein-dependent ATPase activity of protease La from *Escherichia coli*. Proc. natl Acad. Sci. USA *79* (1982) 795–799.
25 Chung, C. H., and Goldberg, A. L., The product of the *lon*(*capR*) gene in *Escherichia coli* is the ATP-dependent protease, protease La. Proc. natl Acad. Sci. USA *78* (1981) 4931–4935.
26 Chung, C. H., and Goldberg, A. L., Purification and characterization of protease So, a cytoplasmic serine protease in *Escherichia coli*. J. Bact. *154* (1983) 231–238.
27 Chung, C. H., Ives, H. E., Almeda, S., and Goldberg, A. L., Purification from *Escherichia coli* of a periplasmic protein that is a potent inhibitor of pancreatic proteases. J. biol. Chem. *258* (1983) 11032–11038.
28 Claverie-Martin, F., Diaz-Torres, M. R., Kushner, S. R., Analysis of the regulatory region of the protease III (*ptr*) gene of *Escherichia coli* K12. Gene *54* (1987) 185–195.
29 Craig, N. L., and Roberts, J. W., Function of nucleoside triphosphate and polynucleotide in *Escherichia coli recA* protein directed cleavage of phage lambda repressor. J. biol. Chem. *256* (1981) 8039–8044.
30 Davies, K. J. A., and Lin, S. W., Degradation of oxidatively denatured protein in *Escherichia coli*. Free Radic. Biol. Med. *5* (1988) 215–223.
31 Dennis, P. P., Synthesis and stability of individual ribosomal proteins in the presence of rifampicin. Mol. gen. Genet. *134* (1974) 39–47.

32 Derbyshire, C., Kramer, M., and Grindley, N. D. F., Role of instability in the cis action of the insertion sequence IS903 transposase. Proc. natl Acad. Sci. USA *87* (1990) 4048–4052.

33 Dervyn, E., Canceill, D., and Huisman, O., Saturation and specificity of the Lon protease of *Escherichia coli*. J. Bact. *172* (1990) 7098–7103.

34 Desautels, M., and Goldberg, A. L., Liver mitochondria contain an ATP-dependent, vanadate-sensitive pathway for the degradation of proteins. Proc. natl Acad. Sci. USA *79* (1982) 1869–1873.

35 Donch, J., and Greenberg, J., Genetic analysis of *lon* mutants of strain K-12 of *Escherichia coli*. Mol. gen. Genet. *103* (1968) 105–115.

36 Downs, D., Waxman, L., Goldberg, A. L., and Roth, J., Isolation and characterization of *lon* mutants in *Salmonella typhimurium*. J. Bact. *165* (1986) 193–197.40.

37 Dykstra, C. C., and Kushner, S. R., Physical Characterization of the cloned Protease III gene from *Escherichia coli* K-12. J. Bact. *163* (1985) 1055–1059.

38 Edmunds, T., and Goldberg, A. L., Role of ATP hydrolysis in the degradation of proteins by protease La from *Escherichia coli*. J. cell. Biochem. *32* (1986) 187–191.

39 Eytan, E., Ganoth, D., Armon, T., and Hershko, A., ATP-dependent incorporation of 20S protease into the 26S complex that degrades proteins conjugated to ubiquitin. Proc. natl Acad. Sci. USA *86* (1989) 7751–7755.

40 Falkenburg, P. E., Haass, C., Kloetzel, P. M., Niedel, B., Kopp, F., Kuehn, L., and Dahlmann, B., Drosophila small cytoplasmic 19S ribonucleoprotein is homologous to the rat multicatalytic proteinase. Nature *331* (1988) 190–192.

41 Finley, D., Ciechanover, A., and Varshavsky, A., Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant *ts85*. Cell *37* (1984) 43–55.

42 Finch, P. W., Wilson, R. R., Brown, K., Hickson, I. D., and Emmerson, P. T., Complete nucleotide sequence of the *Escherichia coli ptr* gene encoding Protease III. Nucl. Acids Res. *14* (1986) 7695–7703.

43 Fujiwara, T., Tanaka, K., Orino, E., Yoshimura, T., Kumatori, A., Tamura, T., Chung, C. H., Nakai, T., Yamaguchi, K., Shin, S., Kakizuka, A., Nakanishi, S., and Ichihara, A., Proteasomes are essential for yeast proliferation. J. biol. Chem. *265* (1990) 16604–16613.

44 Ganoth, D., Leshinsky, E., Eytan, E., and Hershko, A., A multicomponent system that degrades proteins conjugated to ubiquitin. J. biol. Chem. *263* (1988) 12412–12419.

45 Georgopoulos, C., Ang, D., Libeleric, K., and Zylicz, M., Properties of the *Escherichia coli* heat shock proteins and their role in bacteriophage λ growth in: Stress Proteins in Biology and Medicine, pp. 191–221. Eds R. Morimoto, A. Tissieres and C. Georgopoulos. Cold Spring Harbor Press 1990.

46 Goff, S. A., and Goldberg, A. L., An increased content of protease La, the *lon* gene product, increases protein degradation and blocks growth in *Escherichia coli*. J. biol. Chem. *262* (1987) 4508–4515.

47 Goff, S. A., Casson, L. P., and Goldberg, A. L., Heat shock regulatory gene *htpR* influences rates of protein degradation and expression of the *lon* gene in *Escherichia coli*. Proc. natl Acad. Sci. USA *81* (1984) 6647–6651.

48 Goff, S. A., and Goldberg, A. L., Production of abnormal proteins in *E. coli* stimulates transcription of *lon* and other heat shock genes. Cell *41* (1985) 587–595.

49 Goldberg, A. L., Degradation of abnormal proteins in *Escherichia coli*. Proc. natl Acad. Sci. USA *69* (1972) 422–426.

50 Goldberg, A. L., and St. John, A. C., Intracellular protein degradation in mammalian and bacterial cells: part 2. A. Rev. Biochem. *45* (1976) 747–803.

51 Goldberg, A. L., Sreedhara Swamy, K. H., Chung, C. H., and Larimore, F. S., Proteases of *Escherichia coli*. Meth. Enzym. *80* (1983) 680–702.

52 Goldberg, A. L., and Waxman, L., The role of ATP hydrolysis in the breakdown of proteins and peptides by protease La from *Escherichia coli*. J. biol. Chem. *260* (1985) 12029–12034.

53 Goldschmidt, R., In vivo degradation of nonsense fragments in *E. coli*. Nature (London) *228* (1970) 1151–1154.

54 Gottesman, S., Regulation by proteolysis, in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, pp. 1308–1312. Eds F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaecter and H. E. Umbarger. American Society for Microbiology, Washington, D.C. 1987.

55 Gottesman, S., Genetics of proteolysis in *Escherichia coli*. A. Rev. Genet. *23* (1989) 163–198.

56 Gottesman, S., Clark, W. P., and Maurizi, M. R., The ATP-dependent Clp protease of *Escherichia coli* sequence of *clpA* and identification of a Clp-specific substrate. J. Biol. Chem. *265* (1990) 7886–7893.

57 Gottesman, S., Gottesman, M., Shaw, J., and Pearson, M. L., Protein degradation in *E. coli*: The *lon* mutation and bacteriophage lambda N and cII protein stability. Cell *24* (1981) 225–233.

58 Gottesman, S., Squires, C., Pichersky, E., Carrington, M., Hobbs, M., Mattick, J. S., Dalrymple, B., Kuramitsu, H., Shiroza, T., Foster, T., Clark, W. P., Ross, B., Squires, C., and Maurizi, M. R., Conservation of the regulatory subunit for the Clp ATP-dependent protease in prokaryotes and eukaryotes. Proc. natl Acad. Sci. USA *87* (1990) 3513–3517.

59 Gottesman, S., and Zipser, D., The Deg phenotype of *Escherichia coli lon* mutants. J. Bact. *133* (1978) 844–851.

60 Grodberg, J., and Dunn, J. J., *ompT* Encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. J. Bact. *170* (1988) 1245–1253.

61 Grossman, A. D., Burgess, R., Walter, W., and Gross, C., Mutations in the *lon* gene of *E. coli* K12 phenotypically suppress a mutation in the sigma subunit of RNA polymerase. Cell *32* (1983) 151–159.

62 Grossman, A. D., Erickson, J. W., and Gross, C. A., The *htpR* gene product of *E. coli* is a sigma factor for heat shock promoters. Cell *38* (1984) 383–390.

63 Grossman, A. D., Straus, D. B., Walter, W. A., and Gross, C. A., Sigma 32 synthesis can regulate the synthesis of heat shock proteins in *Escherichia coli*. Genes Dev. *1* (1987) 179–184.

63a Heinemeyer, W., Kleinschmidt, J. A., Saidowsky, J., Escher, C., and Wolf, D. H., Proteinase YscE, the yeast proteasome/multicatalytic-multifunctional proteinase: mutants unravel its function in stress induced proteolysis and uncover its necessity for cell survival. EMBO J. *10* (1991) 555–562.

64 Hershko, A., Ubiquitin-mediated protein degradation, J. biol. Chem. *263* (1990) 15237–15240.

65 Holck, A., and Kleppe, K., Cloning and sequence of the gene for the DNA-binding 17K protein of *Escherichia coli*. Gene *67* (1988) 117–124.

66 Hopfield, J. J., Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc. natl Acad. Sci. USA *71* (1974) 4135–4139.

67 Hough, R., Pratt, G., and Rechensteiner, M., Purification of two high molecular weight proteases from rabbit reticulocyte lysate. J. biol. Chem. *262* (1987) 8303–8313.

68 Hoyt, M. A., Knight, D. M., Das, A., Miller, H. I., and Echols, H., Control of phage lambda development by stability and synthesis of cII protein: Role of the viral cIII and host *hflA*, *himA* and *himD* genes. Cell *31* (1982) 565–573.

69 Huisman, O., D'Ari, R., and Gottesman, S., Cell division control in *Escherichia coli*: specific induction of the SOS SfiA protein is sufficient to block septation. Proc. natl Acad. Sci. USA *81* (1984) 4490–4494.

70 Hwang, B. J., Park, W. J., Chung, C. H., and Goldberg, A. L., *Escherichia coli* contains a soluble ATP-dependent protease (Ti) distinct from protease La. Proc. natl Acad. Sci. USA *84* (1987) 5550–5554.

71 Hwang, B. J., Woo, K. M., Goldberg, A. L., and Chung, C. H., Protease Ti, a new ATP-dependent protease in *Escherichia coli* contains protein-activated ATPase and proteolytic functions in distinct subunits. J. biol. Chem. *263* (1988) 8727–8734.

72 Ichihara, S., Beppu, N., and Mizushima, S., Protease IV, a cytoplasmic membrane protein of *Escherichia coli*, has signal peptide peptidase activity. J. biol. Chem. *259* (1984) 9853–9857.

73 Ichihara, S., Suzuki, T., Suzuki, M., and Mizushima, S., Molecular cloning and sequencing of the *sppA* gene and characterization of the encoded protease IV, a signal peptide peptidase, of *Escherichia coli*. J. biol. Chem. *261* (1986) 9405–9411.

74 Innis, M. A., Tokunaga, M., Williams, M. E., Loranger, J. M., Chang, S. Y., Chang, S., and Wu, H. C., Nucleotide sequence of the *Escherichia coli* prolipoprotein signal peptidase (*lsp*) gene. Proc. natl Acad. Sci. USA *81* (1984) 3708–3712.

75 Ishihama, A., Fujita, N., and Glass, R. E., Subunit assembly and metabolic stability of *E. coli* RNA polymerase. Prot. Struct. Funct. Gen. *2* (1987) 42–53.

76 Johnson, C., Chandrasekhar, G. N., and Georgopoulos, C., *Escherichia coli* DnaK and GrpE heat shock proteins interact both in vivo and in vitro. J. Bact. *171* (1989) 1590–1596.

77 Jones, C. A., and Holland, I. B., Role of the SfiB (FtsZ) protein in division inhibition during the SOS response in *E. coli*: FtsZ stabilizes the inhibitor SfiA in maxicells. Proc. natl Acad. Sci. USA *82* (1985) 6045–6049.

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

79 Katayama-Fujimura, Y., Gottesman, S., and Maurizi, M. R., A multiple-component, ATP-dependent protease from *Escherichia coli*. J. biol. Chem. *262* (1987) 4477–4485.

80 Keller, J. A., and Simon, L. D., Divergent effects of a *dnaK* mutation on abnormal protein degradation in *Escherichia coli*. Molec. Microbiol. *2* (1988) 31–41.

80a Kitagawa, M., Wada, C., Yoshioka, S., and Yura, T., Expression of ClpB, an analog of the ATP-dependent protease-regulatory subunit in *Escherichia coli* is controlled by heat shock σ factor (σ^{32}) J. Bact. *173* (1991) 4247–4253.

81 Kornitzer, D., Altuvia, S., and Oppenheim, A. B., The activity of the CIII regulator of lamboid bacteriophages resides within a 24-amino acid protein domain. Proc. natl Acad. Sci. USA *88* (1991).

82 Kroh, H. E., and Simon, L. E., The ClpP component of Clp protease is the σ32-dependent heat shock protein F21.5. J. Bact. *172* (1990) 6026–6034.

83 Kuhn, A., and Wickner, W., Conserved residues of the leader peptide are essential for cleavage by leader peptidase. J. biol. Chem. *260* (1985) 15914–15918.

84 Lazarides, E., and Moon, R. T., Assembly and topogenesis of the spectrin-based membrane skeleton in erythroid development. Cell *37* (1984) 354–356.

85 Lee, C. S., Hahm, J. K., Hwang, B. J., Park, K. C., Ha, D. B., Park, S. D., and Chung, C. H., Processing of Ada protein by two serine endoproteases Do and So from *Escherichia coli*. FEBS Lett. *262* (1990) 310–312.

86 Lee, Y. S., Park, S. C., Goldberg, A. L., and Chung, C. H., Protease So from *Escherichia coli* preferentially degrades oxidatively damaged glutamine synthetase. J. biol. Chem. *263* (1988) 6643–6646.

87 Lindahl, T., Sedgwick, B., Sekiguchi, M., and Nakabeppu, Y., Regulation and expression of the adaptive response to alkylating agents. A. Rev. Biochem. *57* (1988) 133–157.

88 Lipinska, B., Fayet, O., Baird, L., and Georgopoulos, C., Identification, characterization, and mapping of the *Escherichia coli htrA* gene, whose product is essential for bacterial growth only at elevated temperatures. J. Bact. *171* (1989) 1574–1584.

89 Lipinska, B., Zylicz, M., and Georgopoulos, C., The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase. J. Bact. *172* (1990) 1791–1797.

90 Little, J. W., Autodigestion of LexA and phage lambda repressors. Proc. natl Acad. Sci. USA *81* (1984) 1375–1379.

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

92 Little, J. W., and Mount, D. W., The SOS regulatory system of *Escherichia coli*. Cell *29* (1982) 11–22.

93 Mandelstam, J., Turnover of protein in growing and nongrowing population of *Escherichia coli*. Biochem. J. *169* (1958) 110–119.

94 Maurizi, M. R., Degradation in vitro of bacteriophage lambda N protein by Lon protease from *Escherichia coli*. J. biol. Chem. *262* (1987) 2696–2703.

95 Maurizi, M. R., ATP-promoted interaction between ClpA and ClpP in activation of Clp protease from *Escherichia coli*. Biochem. Soc. Trans. (1991) in press.

96 Maurizi, M. R., Katayama, Y., and Gottesman, S., Selective ATP-dependent degradation of proteins in *Escherichia coli*, in: The Ubiquitin System. Current Communications in Molecular Biology, pp. 147–154. Ed. M. J. Schlesinger. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 1988.

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

98 Maurizi, M. R., Clark, W. P., Kim, S. H., and Gottesman, S. J., ClpP represents a unique family of serine proteases. Biol. Chem. *265* (1990) 12546–12552.

99 Maurizi, M. R., and Switzer, R. L., Proteolysis in bacterial sporulation. Curr. Top. Cell Regul. *16* (1979) 163–224.

100 Maurizi, M. R., Trisler, P., and Gottesman, S., Insertional mutagenesis of the *lon* gene in *Escherichia coli*: *lon* is dispensable. J. Bact. *164* (1985) 1124–1135.

101 McGrath, M. E., Hines, W. M., Sakanari, J. A., Fletterick, R. J., and Craik, C. S., The sequence and reactive site of Ecotin. J. biol. Chem. *266* (1991) 6620–6625.

102 Menon, A. S., and Goldberg, A. L., Binding of nucleotides to the ATP-dependent protease La from *Escherichia coli*. J. biol. Chem. *262* (1987) 14921–14928.

103 Menon, A. S., and Goldberg, A. L., Protein substrates activate the ATP-dependent protease La by promoting nucleotide binding and release of bound ADP. J. biol. Chem. *262* (1987) 14929–14934.

104 Menon, A. S., Waxman, L., and Goldberg, A. L., The energy utilized in protein breakdown by the ATP-dependent protease La from *Escherichia coli*. J. biol. Chem. *262* (1987) 722–726.

105 Michaelis, S., and Beckwith, J., Mechanism of incorporation of cell envelop proteins in *Escherichia coli*. A. Rev. Microbiol. *36* (1982) 435–465.

106 Miller, C. G., Protein degradation and proteolytic modification, in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, pp. 680–691. Eds F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter and H. E. Umbarger. American Society for Microbiology, Washington, D.C. 1987.

107 Miller, C. G., Genetics and physiological roles of *Salmonella typhimurium* peptidases, in: Microbiology 1985, pp. 346–349. Ed. L. Leive. American Society for Microbiology, Washington, D.C. 1985.

108 Miller, C. G., and Schwartz, G., Peptidase-deficient mutants of *Escherichia coli*. J. Bact. *135* (1978) 603–611.

109 Mizusawa, S., and Gottesman, S., Protein degradation in *Escherichia coli*: The *lon* gene controls the stability of the SulA protein. Proc. natl Acad. Sci. USA *80* (1983) 358–362.

110a Mosteller, R. D., Goldstein, R. V., and Nishimoto, K. R., Metabolism of individual proteins in exponentially growing *Escherichia coli*. J. biol. Chem. *255* (1980) 2524–2532.

110b Moesteller, R. D., Nishimoto, K. R., and Goldstein, R. V., Inactivation and partial degradation of phosphoribosylanthranilate isomerase-indoleglycerol phosphate synthetase in nongrowing cultures of *Escherichia coli*. J. Bact. *131* (1977) 153–162.

111 Murray, A. W., Solomon, M. J., and Kirschner, M. W., The role of cyclin synthesis and degradation in the control of maturation promoting factor activity. Nature *339* (1989) 280–286.

112 Nash, H. A., Robertson, C. A., Flamm, E., Weisberg, R. A., and Miller, H. I., Overproduction of *Escherichia coli*: integration host factor, a protein with nonidentical subunits. J. Bact. *169* (1987) 4124–4127.

113 Neidhardt, F. C., VanBogelen, R. A., and Vaughn, V., The genetics and regulation of the heat shock proteins. A. Rev. Genet. *18* (1984) 295–329.

114 Neurath, H., Evolution of proteolytic enzymes. Science *224* (1984) 350–357.

115 Nishi, K., and Schnier, J., The phenotypic suppression of a mutation in the gene *rplX* for ribosomal protein L24 by mutations affecting the lon gene product for protease La in *Escherichia coli* K12. Molec. gen. Genet. *212* (1988) 177–181.

116 Nohmi, T., Battista, J. R., Dodson, L. A., and Walker, G. C., RecA-mediated cleavage activates UmuD for mutagenesis: Mechanistic relationship between transcriptional derepression and posttranslational activation. Proc. natl Acad. Sci. USA *85* (1988) 1816–1820.

117 Novak, P., Ray, P. H., and Dev, I. K., Localization and purification of two enzymes from *Escherichia coli* capable of hydrolyzing a signal peptide. J. biol. Chem. *261* (1986) 420–427.

118 Olden, K., and Goldberg, A. L., Studies on the energy requirement for intracellular protein degradation in *Escherichia coli*. Biochim. biophys. Acta *542* (1978) 385–598.

119 Oliver, D., Protein secretion in *Escherichia coli*. A. Rev. Microbiol. *39* (1985) 615–648.

120 Orlowski, M., The multicatalytic proteinase complex, a major extralysosomal proteolytic system. Biochemistry *29* (1990) 10289–10297.

121 Pacaud, M., Sibilli, L., and Le Bras, G., Protease I from *Escherichia coli*. Eur. J. Biochem. *69* (1976) 141–151.

122 Pacaud, M., Protease II from *Escherichia coli*: substrate specificity and kinetic properties. Eur. J. Biochem. *82* (1978) 439–451.

123 Pacaud, M., Purification and characterization of two novel proteolytic enzymes in membranes of *Escherichia coli*. J. biol. Chem. *257* (1982) 4333–4339.

124 Pakula, A. A., Young, V. B., and Sauer, R. T., Bacteriophage λ Cro mutations: effects on activity and intracellular degradation. Proc. natl Acad. Sci. USA *83* (1986) 8829–8833.

125 Palmer, S. M., and St. John, A. C., Characterization of a membrane-associated serine protease in *Escherichia coli*. J. Bact. *169* (1987) 1474–1479.

126 Park, J. H., Lee, Y. S., Chung, C. H., and Goldberg, A. L., Purification and characterization of protease Re, a cytoplasmic endoprotease in *Escherichia coli*. J. Bact. *170* (1988) 921–926.

126a Parsell, D. A., Sanchez, Y., Stitzel, J. D., and Lindquist, S., Hsp 104 is a highly conserved protein with two essential nucleotide-binding sites. Nature (London) *353* (1991) 270–273.

127 Parsell, D. A., and Sauer, R. T., The structural stability of a protein is an important determinant of its proteolytic susceptibility in *E. coli*. J. biol. Chem. *264* (1989) 7590–7595.

128 Parsell, D. A., Silber, K. R., and Sauer, R. T., Carboxy-terminal determinants of intracellular protein degradation. Genes Devl. *4* (1990) 277–286.

129 Pato, M. L., and Reich, C., Instability of transposase activity: evidence from bacteriophage Mu DNA replication. Cell *29* (1982) 219–225.

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

131 Perry, K. L., Elledge, S. J., Mitchell, B. B., Marsh, L., and Walker, G. C., *umuDC* and *mucAB* operons whose products are required for UV light- and chemical-induced mutagenesis: UmuD, MucA, and LexA proteins share homology. Proc. natl Acad. Sci. USA *82* (1985) 4331–4335.

132 Peterson, K. R., Wertman, K. F., Mount, D. W., and Marinus, M. G., Viability of *Escherichia coli* K-12 DNA adenine methylase (S) mutants requires increased expression of specific genes in the SOS regulon. Molec. gen. Genet. *201* (1985) 14–19.

133 Phillips, T. A., VanBogelen, R. A., and Neidhardt, F. C., *lon* gene product of *Escherichia coli* is a heat shock protein. J. Bact. *159* (1984) 283–287.

134 Pine, M. J., Response of intracellular proteolysis to alteration of bacterial protein and the implications in metabolic regulation. J. Bact. *93* (1967) 1527–1533.

135 Pine, M. J., Steady-state measurements of the turnover of amino acid in the cellular protein of growing *Escherichia coli*: existence of two kinetically distinct reactions. J. Bact. *103* (1970) 207–215.

136 Pine, M. J., Regulation of intracellular proteolysis in *Escherichia coli*. J. Bact. *115* (1973) 1097–1116.

137 Platt, T., Miller, J. H., and Weber, K., In vivo degradation of mutant *lac* repressor. Nature (London) *228* (1970) 1154–1156.

138 Rawlings, N. D., and Barrett, A. J., Homologues of insulinase, a new superfamily of metallopeptidases. Biochem. J. *274* (1991) in press.

139 Rechsteiner, M., Ubiquitin-mediated pathways for intracellular proteolysis. A. Rev. Cell Biol. *3* (1987) 1–30.

139a Regnier, P., The purification of protease IV and the demonstration that it is a proteolytic enzyme. Biochem. biophys. Res. Commun. *99* (1981) 1369–1376.

140 Reiss, Y., Kaim, D., and Hershko, A., Specificity of binding of NH_2-terminal residue of proteins to Ubiquitin-protein ligase. J. biol. Chem. *263* (1988) 2693–2698.

141 Rivett, A. J., The multicatalytic proteinase of mammalian cells. Archs Biochem. Biophys. *268* (1989) 1–8.

142 Roberts, J. W., and Roberts, C. W., Proteolytic cleavage of bacteriophage lambda repressor in induction. Proc. natl Acad. Sci. USA *72* (1975) 147–151.

143 Roland, K., and Little, J. W., Reaction of LexA repressor with diisopropylfluoro phosphate: a test of the serine protease model. J. biol. Chem. *265* (1990) 12828–12835.

144 Roseman, J. E., and Levine, R. L., Purification of a protease from *Escherichia coli* with specificity for oxidized glutamine synthetase. J. biol. Chem. *262* (1987) 2101–2110.

145 Rupprecht, K. R., and Markovitz, A., Conservation of *capR* (*lon*) DNA of *Escherichia coli* K-12 between distantly related species. J. Bact. *155* (1983) 910–914.

146 Schroer, D. W., and St. John, A. C., Relative stability of membrane proteins in *Escherichia coli*. J. Bact. *146* (1981) 476–483.

147 Sedgwick, B., in vitro proteolytic cleavage of the *Escherichia coli* Ada protein by the *ompT* gene product. J. Bact. *171* (1989) 2249–2251.

148 Shinagawa, H., Iwasaki, H., Kato, T., and Nakata, A., RecA protein-dependent cleavage of UmuD protein and SOS mutagenesis. Proc. natl Acad. Sci. USA *85* (1988) 1806–1810.

149 Shineberg, B., and Zipser, D., The *lon* gene and degradation of β-galactosidase nonsense fragments. J. Bact. *116* (1973) 1469–1471.

150 Simon, L. D., Tomczak, K., and St. John, A. C., Bacteriophages inhibit degradation of abnormal proteins in *E. coli*. Nature *275* (1978) 424–428.

151 Skorupski, K., Tomaschewski, J., Ruger, W., and Simon, L. D., A bacteriophage T4 gene which functions to inhibit *Escherichia coli* Lon protease. J. Bact. *170* (1988) 3016–3024.

152 Skowyra, D., Georgopoulos, C., and Zylicz, M., The *E. coli dnaK* gene product, the hsp70 homologg, can reactivate heat-inactivated RNA polymerase in an ATP hydrolysis-dependent manner. Cell *62* (1990) 939–944.

153 Slavicek, J. M., Jones, N. C., and Richter, J. D., Rapid turnover of adenovirus E1A is determined through a co-translational mechanism that requires an aminoterminal domain. EMBO J. *7* (1988) 171–180.

154 Slilaty, S. N., and Little, J. W., Lysine-156 and serine-119 are required for LexA repressor cleavage: A possible mechanism. Proc. natl Acad. Sci. USA *84* (1987) 3987–3991.

154a Squires, C. L., Pedersen, S., Ross, B. M., and Squires, C., ClpB is the *Escherichia coli* heat shock protein F84.1. J. Bact. *173* (1991) 4254–4262.

154b Squires, C. L., and Squires, C., The Clp proteins – proteolysis regulators or molecular chaperones? J. Bact. *174* (1992) in press.

155 St. John, A. C., and Goldberg, A. L., Effects of reduced energy production on protein degradation, guanosine tetraphosphate, and RNA synthesis in *Escherichia coli*. J. biol. Chem. *253* (1978) 2705–2711.

156 St. John, A. C., and Goldberg, A. L., Effects of starvation for potassium and other inorganic ions on protein degradation and ribonucleic acid synthesis in *Escherichia coli*. J. Bact. *143* (1978) 1223–1233.

157 St. John, A. C., Jakubas, K., and Beim, D., Degradation of proteins in steady-state cultures of *Escherichia coli*. Biochim. biophys. Acta *586* (1979) 537–544.

158 Stout, V., Torres-Cabassa, A., Maurizi, M. R., Gutnick, D., and Gottesman, S., RcsA, an unstable regulator of capsular polysaccharide synthesis. J. Bact. *173* (1991) 1738–1747.

159 Strauch, K., Johnson, K., and Beckwith, J., Characterization of *degP*, a gene required for proteolysis in the cell envelope and essential for growth of *Escherichia coli* at high temperature. J. Bact. *171* (1989) 2689–2696.

160 Strauch, K. L., and Beckwith, J., An *Escherichia coli* mutation preventing degradation of abnormal periplasmic proteins. Proc. natl Acad. Sci. USA *85* (1988) 1576–1580.

161 Straus, D. B., Walter, W. A., and Gross, C. A., The heat shock response of *E. coli* is regulated by changes in the concentration of sigma 32. Nature (London) *329* (1987) 348–391.

162 Straus, D. B., Walter, W. A., and Gross, C. A., *Escherichia coli* heat shock gene mutants are defective in proteolysis. Genes Devl. *2* (1988) 1851–1858.

163 Straus, D. B., Walter, W., and Gross, C. A., DnaK, DnaJ, and GrpE heat shock proteins negatively regulate heat shock gene expression by controlling the synthesis and stability of σ32. Genes Devl. *4* (1990) 2202–2209.

164 Strongin, A. Y., Gorodetsky, D. I., and Stepanov, V. M., The study of *Escherichia coli* proteases. Intracellular serine protease of *E. coli* – an analog of *Bacillus* proteases. J. gen. Microbiol. *110* (1979) 443–451.

165 Sugimura, K., and Nishihara, T., Purification, characterization, and primary structure of *Escherichia coli* protease VII with specificity for paired basic residues: identity of protease VII and OmpT. J. Bact. *170* (1988) 5625–5632.

166 Swamy, K. H. S., Chung, C. H., and Goldberg, A. L., Isolation and characterization of protease Do from *Escherichia coli*, a large serine protease containing multiple subunits. Archs Biochem. Biophys. *224* (1983) 543–554.

167 Tilly, K., Spence, J., and Georgopoulos, C., Modulation of stability of the *Escherichia coli* Heat Shock Regulatory Factor sigma 32. J. Bact. *171* (1989) 1585–1589.

168 Torres-Cabassa, A. S., and Gottesman, S., Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. J. Bact. *169* (1987) 981–989.

169 Trempy, J. E., and Gottesman, S., Alp: A suppressor of Lon protease mutants in *Escherichia coli*. J. Bact. *171* (1989) 3348–3353.

170 Tokunaga, M., Loranger, J. M., Wolfe, P. B., and Wu, H. C., Prolipoprotein signal peptidase in *Escherichia coli* is distinct from the M13 precoat protein signal peptidase. J. biol. Chem. *257* (1982) 9922–9925.

171 Tokunaga, M., Loranger, J. M., Chang, S. Y., Regue, M., Chang, S., and Wu, H. C., Identification of prolipoprotein signal peptidase and genomic organization of the Isp gene in *Escherichia coli*. J. biol. Chem. *260* (1985) 5610–5616.
172 Tokunaga, M., Tokunaga, H., and Wu, H. C., Post-translational modification and processing of *Escherichia coli* prolipoprotein in vitro. Proc. natl Acad. Sci. USA *79* (1982) 2255–2259.
173 Vaithilingam, I., and Cook, R. A., High-molecular-mass proteases (possibly proteasomes) in *Escherichia coli* K12. Biochem. Int. *19* (1989) 1297–1307.
174 Walker, G. C., The SOS Response of *Escherichia coli*, in: *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, pp.41346–1357. Eds F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger. American Society for Microbiology, Washington, D.C. 1987.
175 Waxman, L., and Goldberg, A. L., Protease La from *Escherichia coli* hydrolyzes ATP and proteins in a linked fashion. Proc. natl Acad. Sci. USA *79* (1982) 4883–4887.
176 Waxman, L., and Goldberg, A. L., Protease La, the *lon* gene product, cleaves specific fluorogenic peptides in an ATP-dependent reaction. J. biol. Chem. *260* (1985) 12022–12028.
177 Waxman, L., and Goldberg, A. L., Selectivity of intracellular proteolysis: protein substrates activate the ATP-dependent protease (La). Science *232* (1986) 500–503.
178 Wolfe, P. B., Silver, P., and Wickner, W., The isolation of homogeneous leader peptidase from a strain of *Escherichia coli* which overproduces the enzyme. J. biol. Chem. *257* (1982) 7898–7902.
179 Woo, K. M., Chung, W. J., Ha, D. B., Goldberg, A. L., and Chung, C. H., Protease Ti from *Escherichia coli* requires ATP hydrolysis for protein breakdown but not for hydrolysis of small peptides. J. biol. Chem. *264* (1989) 2088–2091.
180 Yen, C., Green, L., and Miller, C. G., Degradation of intracellular protein in *Salmonella typhimurium* peptidase mutants. J. molec. Biol. *143* (1980) 21–33.
181 Zehnbauer, B. A., Foley, E. C., Henderson, G. W., and Markovitz, A., Identification and purification of the *lon*$^+$ (*capR*$^+$) gene product, a DNA-binding protein. Proc. natl Acad. Sci. USA *78* (1981) 2043–2047.
182 Zwizinski, C., and Wickner, W., Purification and characterization of leader (signal) peptidase from *Escherichia coli*. J. biol. Chem. *255* (1980) 7973–7977.
183 Zwizinski, C., Date, T., and Wickner, W., Leader peptidase is found in both the inner and outer membranes of *Escherichia coli*. J. biol. Chem. *256* (1981) 3593–3597.
184 **Note in proof:** Recently, Varshavsky and colleagues have shown that a modified N-end rule applies to degradation of aminoterminally mutated β-galactosidase in *Escherichia coli* also. Proteins unstable in wild-type cells were stabilized by *clpA* mutations, further suggesting an analogous function for ClpA and the E3 protein of the ubiquitin system, which may be a mediator of the N-end rule in eukaryotic cells. Tobias, J. W., Shrader, T. E., Rocap, G., and Varshavsky, A., The N-end rule in bacteria. Science *254* (1991) 1374–1377.

0014-4754/92/020178-24$1.50 + 0.20/0

The role of proteolytic processing in the morphogenesis of virus particles

C. U. T. Hellen and E. Wimmer

Department of Microbiology, State University of New York at Stony Brook, Stony Brook (New York 11794-8631, USA)

Abstract. Proteinases are encoded by many RNA viruses, all retroviruses and several DNA viruses. They play essential roles at various stages in viral replication, including the coordinated assembly and maturation of virions. Most of these enzymes belong to one of three (Ser, Cys or Asp) of the four major classes of proteinases, and have highly substrate-selective and cleavage specific activities. They can be thought of as playing one of two general roles in viral morphogenesis. Structural proteins are encoded by retroviruses and many RNA viruses as part of large polyproteins. Their proteolytic release is a prerequisite to particle assembly; consequent structural rearrangement of the capsid domains serves to regulate and direct association and assembly of capsid subunits. The second general role of proteolysis is in assembly-dependent maturation of virus particles, which is accompanied by the acquisition of infectivity.
Key words. Virus polyprotein; virus assembly; virus maturation; retrovirus; picornavirus; cysteine proteinase; aspartic proteinase.

Introduction

The replication of many viruses is entirely dependent on proteolytic processing. Virus-encoded proteinases play essential roles at various stages in viral replication, such as the separation of structural and non-structural proteins, the activation of specific enzymes, and the coordinated assembly and maturation of virions [52, 67, 68]. The high substrate selectivity and the exquisite specificity of cleavage of these enzymes enable viruses to regulate successive stages in the replication and encapsidation of their genomes. Virus assembly is the penultimate step in replication, immediately preceding viral release from infected cells, and is a complex process that may be catalyzed by a combination of host- and virus-encoded proteinases. Host-encoded proteinases cleave viral glycoprotein precursors in a conventional manner after the precursor polypeptides have been transported to vesicular compartments. In contrast, cleavages catalyzed by virus-encoded proteinases are more complex. They occur