Regulation by proteolysis: developmental switches
Susan Gottesman

The energy-dependent proteases originally defined in *Escherichia coli* have proven to have particularly important roles in bacterial developmental systems, including sporulation in *Bacillus subtilis* and cell cycle in *Caulobacter*. Degradation of key regulatory proteins participates, with regulation of synthesis and activity of the regulators, to one or two tight control and, where required, irreversible commitment of the cell to specific developmental pathways.

Addresses
Building 37, Room 2E18, Laboratory of Molecular Biology, National Cancer Institute, Bethesda, MD 20892-4255, USA;
e-mail: susang@helix.nih.gov

Current Opinion in Microbiology 1999, 2:142–147
http://biomednet.com/elecref/1369527400200142
© Elsevier Science Ltd ISSN 1369-5274

Introduction
Energy-dependent proteases help rid the cell of abnormal and misfolded proteins and play key roles in many regulatory cascades. In addition, the ATPase components of some of these proteases or their close relatives serve as chaperones. Our understanding of the mechanism of action of these proteases, based primarily but not exclusively on studies with the *Escherichia coli* prototypes, has continued to expand. At the same time, homologs of the *E. coli* enzymes in many other organisms have been identified and found to have critical regulatory roles, particularly in developmental pathways. This review briefly describes the energy-dependent proteases, approaches to studying their in vitro targets, new regulatory cascades in which they have been implicated.

The basic machinery: processive, selective proteases
In prokaryotes, on the order of six energy-dependent proteases must handle the degradation of a large variety of cellular proteins. Unlike eukaryotes, in which multiple ubiquitin tagging systems funnel multiple substrates into one cytoplasmic protease the 26S protease, substrate selection in prokaryotes must operate at the level of the protease itself. Energy-dependent proteases are generally large oligomeric assemblies. In those in which structure has been examined, the proteolytic sites are sequestered in a chamber, the entry to which is too small for most folded proteins [1,2]. Entry is believed to be mediated via the regulatory ATPase domains or subunits; it is these regulatory ATPases that also determine substrate specificity. The ATPase subunits from otherwise distinct proteases share sequence similarities that may imply similar mechanisms of action [3]. Because substrate recognition is by the regulatory ATPases, the substrate characteristics that lead to degradation are independent of any sequence requirements for peptide bond cleavage. Once a protein is recognized and bound properly by the regulatory ATPase domain, degradation of the whole polypeptide chain proceeds processively, with cleavage occurring every 5–10 amino acids.

Four ATP-dependent protease families have been recognized in *E. coli*.

Lon
*Lon* (also called *La*) is encoded by a single gene near the carboxyl terminus and an ATPase domain in the middle [4–10]. Recent studies of the protein from yeast mitochondria suggest that these domains can be expressed from separate polypeptides and still cooperate to restore full *Lon* function [11**].

**ClpAP and ClpXP**
*ClpP*, the proteolytic core of the ClpAP protease, has its serine active sites arrayed within a cavity formed from two seven-membered rings of ClpP [1]. The regulatory ATPase ring consists of six ClpA subunits [12]. ClpA can function as a chaperone in *vitro*, binding substrate in the presence of nucleotide and releasing it in an activated state when ATP is hydrolyzed; only a single round of binding and release is necessary for activation [13,14]. When ClpP is present but proteolytically inactive, substrates are translocated from the ATPase to the ClpP chamber following ATP hydrolysis [15**]. Parallel results observed in *vivo* for *Lon* protease (sequestration of substrates in the absence of active proteolysis) suggests that this may be a common feature of these two protease families (L van Melderen, MR Mauriz, S Gottesman, unpublished data).

ClpP consists of an ATPase ring, ClpX, associated with the ClpP rings; the substrate specificity of ClpXP is different from that of ClpAP [16–18]. ClpX can also act as a chaperone on its own [19–21].

**ClpYQ/HisUV**
This protease is a hybrid, in which a ClpX-like ATPase (known as ClpY or HisU) associates with a protease subunit, ClpQ or HisV. The protease subunit has a threonine active site and sequence similarities to the eukaryotic proteasome core subunits [2,22]. Recent studies suggest that this protease at least partially overlaps *Lon* in substrate specificity ([23,24]; W-F Wu, YN Zhou, S Gottesman, unpublished data), emphasizing that the substrate specificity is not a characteristic of a particular protease family.

**FtsH/HflB**
Another single subunit protease, FtsH (also known as HflB), is a zinc metalloprotease with two membrane spanning regions. The active site as well as the ATPase domain, a member of the AAA family of ATPases, reside...
in the cytoplasm (recently reviewed in [25]). FtsH is the only essential member of the energy-dependent proteases in *E. coli*, although in *Bacillus subtilis*, for instance, it is not essential [26,27]. FtsH has also been observed to form rings in EM studies [28].

**How are substrates recognized?**

It is not yet possible to examine a protein sequence and make even an educated guess as to whether it will be unstable and, if so, what protease will degrade it. For known unstable proteins, studies of the protease recognition signals suggest that the ends of proteins are frequently implicated in protease recognition. Whether the end *per se* is a recognition motif (implying that the NH2 or COOH are part of what is recognized) or whether recognition motifs have evolved to be near ends because the ends are more likely to be or to become accessible may vary with the particular case.

Proteins engineered to have certain abnormal amino-terminal amino acids in *vivo* are rapidly degraded (the N-end rule); degradation is dependent upon ClpAP [29-31]. It is not yet clear whether this is biologically relevant in bacteria. A carboxy-terminal tagging system leading to degradation of incomplete proteins is dependent upon a 'mobile messenger RNA' called 10Sa RNA or tmRNA, encoded by the *ssrA* gene of *E. coli*. This RNA serves as a template for the co-translational addition of 11 amino acids to the carboxyl terminus of proteins with truncated messenger RNAs [32]; the carboxy-terminal addition allows degradation by ClpAP, ClpXP, and, in some cases, FtsH [33*]. Thus ClpAP recognizes both the carboxyl terminus and the amino terminus of proteins under different circumstances. The degree of redundancy among these proteases varies with particular substrates, and presumably depends upon the structure of regions further into the body of the protein. A periplasmic protease, Tsp, also recognizes the same carboxy-terminal sequence [32]. In other cases, the carboxyl terminus is also the site of recognition, although the tagging sequence is not present. Recognition of the replication protein of bacteriophage Mu, A protein, by ClpX depends upon the last eight amino acids of MuA [20]. Specificity for the MuA sequence is at least in part encoded in the carboxy-terminal region of ClpX [35]. Degradation of a number of *Caulobacter* proteins, including a component of the flagellar motor, a chemoreceptor protein and CtrA, a regulatory protein, depends upon the carboxyl terminus of these proteins [36-38].

Degradation is also significantly modulated by the interaction of the substrate with other protein components, both to stimulate degradation, and, in other cases, to sequester substrates in protected complexes. As a result, sequences identified as being involved in substrate recognition may in fact be those modulating interaction with critical proteins. RpoS (σ2), a sigma factor degraded under certain conditions by ClpXP, must have at least some of the essential sequences for recognition in the middle of the protein, because fusions that delete the carboxyl terminus (and add on LacZ to make translational fusion proteins) are still degraded with the same pattern as the parent protein; deleting a bit further into the protein, however, leads to a stable protein [39,40]. RpoS degradation, however, requires an auxiliary protein, RssB/SprE [41,42]. Therefore, it is unclear if the sequences within RpoS are those responsible for recognition by RssB, recognition by the protease, or both. Similarly, degradation of the heat-shock sigma factor, RpoH or σ32, is dependent upon FtsH but is significantly affected by DnaJ, DnaK, and GrpE [43-46]. To complicate matters further, the activity of σ32 is also affected by DnaK and DnaJ, in the absence of the FtsH protease [47*]. Degradation of the wild-type Mu repressor is stimulated by the presence of a mutant repressor (Mu vir) [48].

**Regulatory roles for proteolysis**

In many cases, proteolysis in regulatory circuits is not itself regulated. The synthesis of the substrate is presumably changed with conditions, and constant rapid degradation ensures a tight coupling between synthesis and activity. In other cases, proteolysis changes with environmental or developmental cues. As noted above, regulation of the degradation of a given protein may depend on interactions with other proteins. Degradation of *UmuD′*, a component of the error-prone repair mutagenesis system in *E. coli*, depends on the availability of a second component of the system, *UmuC* and the *UmuD′* precursor, as well as the activity of the *UmuD* self-cleavage reaction (dependent upon RecA and DNA damage cues) [49]. MuA recognition by ClpX is affected by the interaction of MuA and MuB [20]. These protein–protein interactions are, not surprisingly, substrate specific. The requirement of RssB for RpoS degradation is also substrate specific [50*]; the regulatory step may be activation of RssB by phosphorylation [51]. It seems likely that most of the regulated proteolysis observed in both prokaryotes and eukaryotes will reflect substrate-specific protein interactions and/or modifications, rather than major changes in protease availability or activity. It is clear, however, that there are some exceptions to this. Bacteriophage T4 shuts off proteolysis by Lon and other proteases; *PinA*, a *T4* protein, acts as an inhibitor of the ATPase of Lon [52,53]. Bacteriophage lambda *RexB* protein inhibits ClpP, apparently helping to save lambda lysogens from killing their hosts after certain types of stresses [54*].

Protease involvement in regulatory pathways has been recognized either as a result of observable phenotypes of protease mutants or by the recognition of specific unstable proteins and the subsequent identification of the relevant protease. The sequencing of bacterial genomes coupled with the recognition of conserved families of proteases has simplified analysis via both of these pathways. Nonetheless, we still know the full story in very few cases.

**Phenotypes due to mutations in the protease**

Not all phenotypes associated with mutations in the proteases reflect regulatory proteolysis. Because some of the energy-dependent proteases are involved in dealing with
abnormal proteins, and some Clp ATPases are also acting as chaperones, it is not surprising that mutations in these proteases sometime yield strains that are sensitive to stress, in particular high temperature [55–58]. In addition, a mutation in a given protease frequently will lead to stabilization of more than one substrate, complicating analysis. The clpP mutation in B. subtilis is defective in high temperature growth, sporulation, competence development, and motility. A mutation in one likely target, MecA, suppresses some but not all phenotypes, and is not fully suppressed for those, suggesting that other substrates must exist [59**].

Most instructive in defining new roles for proteolysis are mutations isolated in a search for a given phenotype, which prove to be in a protease gene. For instance, mutations in clpP were found in a search for mutations that increased transcription of a cell-surface protein involved in Yersinia enterocolitica pathogenesis [60], suggesting the involvement of an unstable transcriptional regulator in this system. A number of other processes newly found to be modulated by protease mutants are reviewed below.

Swarming and biofilms
Mutations selected in Vibrio parahaemolyticus for their inability to down-regulate swarmer genes, normally expressed only when the bacteria are growing on solid surfaces, during liquid growth have proven to be in the lonS gene, encoding a Lon homolog [61]. In Pseudomonas fluorescens, mutants that are defective in biofilm formation include strains with insertions in clpP [62].

Restriction-modification systems
Introduction of new restriction systems into E. coli results in delayed expression of the restriction system, while modification is expressed immediately. Mutations in clpX or clpP lead to a significant lowering in the efficiency of acquiring the new system, suggesting that degradation of a component specific to restriction but not modification may be important [63]. The fact that a clpX mutation has a more significant phenotype than a clpP mutation might suggest that some ‘chaperone’ activity may participate as well.

mRNA stability
Kushner and colleagues [64*,65] looked for temperature sensitive mutants with changes in mRNA decay; one mutation that had a profound effect on mRNA decay, particularly in the context of mutations in other known nucleases, has proven to be an allele of ftsH. In addition, the classic ftsH29 allele of ftsH, which is not conditionally lethal, has the mRNA decay phenotypes at all temperatures. Therefore, defects in ftsH lead to defects in mRNA decay, and this is independent of the essential role of FtsH [64*,65]. The best current explanation for these observations would be that some (direct or indirect) regulator of mRNA decay is subject to FtsH-dependent degradation, and lack of degradation of the protein inhibits mRNA decay.

Unstable proteins in search of a protease
In some cases, proteins that are of interest have been shown to be unstable, such as lambda cII [66]. Because we now know there are a limited number of energy-dependent proteases, it is feasible to look at the degradation of a protein of interest in representative protease mutant hosts. For lambda Xis, recognized as an unstable protein in 1971 [67], the two overlapping proteases responsible have just recently been identified as FtsH and Lon [68]. Such a search of known proteases has also proven productive in studies in Caulobacter (see below).

Bacterial development and proteolysis
Developmental pathways, with their requirements for timed expression of proteins at one stage but not another, are prime cases where one might expect regulated proteolysis to be particularly important. Recent studies in B. subtilis and Caulobacter, summarized below, support this idea. Although the importance of degradation of specific substrates is apparent from these studies, it is more difficult to determine if proteolysis is the primary mechanism for throwing critical switches during development or serves as a locking mechanism to prevent reversal of switches thrown by other levels of regulation. B. subtilis
B. subtilis can grow vegetatively, develop competence under some starvation conditions, and, under more extreme conditions, sporulate. The energy-dependent proteases seem to operate in all of these pathways. Development of competence in B. subtilis requires a positive regulator of transcription of competence genes, ComK. ComK activates its own transcription as well as expression of downstream competence genes. If the cell is to avoid initiating the autoregulation of ComK, the level and/or activity of ComK need to be tightly regulated. Two proteins, MecA and ClpC (an ATPase with homology to ClpA), as well as ClpP, seem to participate in this process. In vitro, MecA and ClpC together bind ComK and may thus block its ability to autoregulate. ComS, the small protein that appears to mediate environmental signaling for competence development, overcomes this binding, possibly competing for MecA and ClpC and releasing active ComK [69]. Degradation of the bound ComK also contributes to down-regulating ComK activity; when ComS releases ComK, ComS itself becomes sensitive to degradation by the Clp protease [70*]. The probable multiple roles for ClpP and ClpC are demonstrated by the puzzling genetics of this system. ClpP mutants do not have the phenotype of mecA or clpC mutants (constitutive ComK synthesis); in fact, they fail to make ComK [59**]. High levels of MecA accumulating in the clpP mutant may be part of the explanation; clpP mecA double mutants partially suppress the clpP phenotypes [59**].

Null mutations in ftsH in B. subtilis, like clpP mutations, block sporulation at a very early stage [27*]. A second role
late in sporulation has been inferred from the isolation of mutations in ftsH as suppressors of an allele of spoVM. spoVM encodes a 26 amino acid polypeptide; null mutations in the gene are defective in spore formation [71]. Because the ftsH alleles do not suppress null mutations in spoVM, it seems possible the mutant allele (which encodes a fusion protein) is unstable and degraded by FtsH, rather than that FtsH plays a normal role at this stage of sporulation. It is intriguing, however, that, in *vitro*, the SpoVM peptide but not a mutant form is degraded by FtsH and can inhibit FtsH, not unlike the inhibition of FtsH by the small lambda cIII polypeptide, a part of the cII-dependent lysogenic switch mechanism [72]. If the only function of SpoVM, however, was to inhibit FtsH, the ftsH mutants would be expected to suppress a null spoVM mutation, which they did not [71].

**Caulobacter**

CtrA is a central regulator of development of cell type in *Caulobacter crescentus*, and acts as a timer for cell division as well. It is unstable at a particular stage of the cell cycle, and mutations in the carboxyl terminus of the protein, which stabilize the protein, perturb cell growth [37]. This observation suggests that the degradation of CtrA is critical for *Caulobacter* growth. Jenal and coworkers [73**] found that mutations in cpx or clpP were lethal for *Caulobacter* and depleting either protein stabilizes CtrA; whether this is the only target of these proteases is not yet known. Still to be examined is the basis for the regulated proteolysis of CtrA during only one part of the growth cycle; the proteases are known to be present at all times. Lon protease also is critical for cell cycle control in this organism, degrading CadM, an adenine DNA methyltransferase necessary for proper timing of DNA replication [74].

A number of other proteins are subject to periodic proteolysis in *Caulobacter*, including the cell division protein FtsZ. In this case, it seems likely that regulated degradation is secondary to regulated assembly of the protein into active complexes (only the free protein is degraded) [75**].

**Conclusions**

The energy-dependent Lon, Clp, and FtsH families of proteases were originally defined in *E. coli*, but it has now become evident that they are widespread throughout prokaryotes and also play critical roles in eukaryotic organelles. With functions both as proteases and in some cases, as chaperones, they have been found to act as modulators in many regulatory circuits. Mutations in Cip components are defective in virulence in bacterial pathogens [60], and a Clp-like protein in humans has been implicated in an inherited torsion dystonia [76]. In the past year, the involvement of ATP-dependent proteases as direct participants in regulatory circuits, particularly for microorganisms with developmental pathways, has provided new insight into ways in which proteolysis can act in regulation.

**References and recommended reading**

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

12. The ATPase and protease domains of *S. cerevisiae* mitochondrial Lon can be expressed from separate plasmids, and complement in vivo to restore function to a lon mutant (which has non-functional mitochondria). Thus the organization of this protease may be functional similar to the Clp proteases, in which ATPase and protease are in separate subunits.
17. Substrates for ClpP are normally bound by ClpA and then translocated to the proteolytic ClpP subunit for degradation. When complexes that are defective in ClpP proteolytic activity, via mutation or chemical inactivation of the active site serine, are formed, an intermediate in the degradation pathway can be isolated in which the substrate (nepa) has moved from the ATPase subunit into ClpP. This process requires ATP hydrolysis.


48. Although FtsH is essential for E. coli, null mutations are viable in the presence of a second mutation in shIC. This suppressor does not affect the requirement for FtsH for Omp32 degradation, so that proper degradation of Omp32 is not required for viability. Regulation of Omp32 activity, dependent upon DnaK and DnaJ, however, limits the deregulation of synthesis of heat shock proteins when Omp32 levels are abnormally high due to loss of degradation.

49. Welty DJ, Jones JM, Nakai H: Communication of Clp Protease protein with a cell-specific fashion; because it is not necessary for degradation of another Clp Protease substrate, plantain G protein.


52. RssR, a protein with homology to a response regulator, is required for the ClpXP-dependent degradation of RpoS. RssR is shown by act in a substrate-specific fashion, because it is not necessary for degradation of another Clp Protease substrate, plantain G protein.


Mutations in the mrsC gene lead to profound effects on mRNA decay when combined with mutations in multiple RNases. Surprisingly, mrsC is an allele of the ftsH gene encoding an ATP-dependent protease.


Regulation by proteolysis: developmental switches Gottesman 147