The Lon Protease from *Mycobacterium smegmatis*: Molecular Cloning, Sequence Analysis, Functional Expression, and Enzymatic Characterization‡

Stanislav G. Roudiak, Anjali Seth, Nancianne Knipfer, and Thomas E. Shrader*

Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Received July 15, 1997; Revised Manuscript Received October 15, 1997⊗

ABSTRACT: We have charterized a *Mycobacterium smegmatis* gene encoding a homolog of the ATP-dependent protease Lon (La). Our identification of a Lon homolog, in conjunction with our previous work, identifies *M. smegmatis* as the first known example of a eubacterium containing both Lon and a complete 20S proteasome (containing both α- and *β*-subunits). Despite the significant primary sequence divergence between *M. smegmatis* Lon (Ms-Lon) and *E. coli* Lon (Ec-Lon), expression of Ms-Lon was only moderately toxic to *E. coli* cells. The ability of *E. coli* cells to tolerate expression of Ms-Lon reveals that Ms-Lon does not recognize and degrade essential *E. coli* proteins. We conclude that discrimination against nonsubstrate proteins is broadly conserved between Ec-Lon and Ms-Lon. Additional conservation of substrate recognition was demonstrated by the ability of Ms-Lon to degrade efficiently RcsA, a natural substrate of Ec-Lon. Purified Ms-Lon displays chymotrypsin-like specificity in peptidase assays that are stimulated by unfolded protein and supported by nonhydrolyzed nucleotide analogs. Maximal peptidase activity requires ATP or dATP. Replacement of Ms-Lon's catalytic Ser with Ala (S675A), Thr (S675T), or Cys (S675C) reduced to background levels Ms-Lon's *in vitro* peptidase activity. However, by employing a sensitive *in vivo* assay, based on the degradation of RcsA, we demonstrated that the S675C variant retained specific protease activity. Finally, variants of Ms-Lon, with substututions at or near S675, reduce the enzyme's basal ATPase activity, suggesting a structural interaction between the peptidase and ATPase active sites of Ms-Lon.

Protein degradation mediates the turnover of damaged proteins and the conditional and constitutive turnover of naturally short-lived proteins (reviewed in *1* and *2*). Rapid degradation of key metabolic enzymes and regulatory proteins allows the cell to respond quickly to changing environmental conditions or developmental signals by tightly linking the protein's cellular concentration to its rate of synthesis. An additional role for protein degradation is manifested during conditions of extreme nutrient depravation or starvation. During their starvation-induced stationary phases, both eukaryotic (*3*) and prokaryotic (*4*) cells rely on enhanced degradation of nonessential proteins in order to provide amino acids for essential protein synthesis. The role of proteolysis in stationary phase survival mechanisms is particularly interesting within the mycobacterial group as the pathogen *Mycobacterium tuberculosis* may persist in stationary phase for decades and only "reawaken" to yield the disease state after diminution of the host immune response (*5*).

Protease Lon was the first ATP-dependent protease to be purified from *Escherichia coli* (called La or Lon, and differentiated as Ec-Lon here). The enzyme is a homotetramer of 87 kDa subunits, making it structurally less complex than other known ATP-dependent proteases (*6*). For example, the multicatalytic 26S proteasome contains ∼40 different subunits and has a native molecular mass of >2 000 000 daltons (reviewed in *7*) while the HslVU protease (*8*−10) and ClpAP proteases (*11*, *12*) of *E. coli* each contain >10 subunits. Despite its relative structural simplicity, Ec-Lon recognizes its substrates directly while the 26S proteasome delegates substrate recognition to the families of ubiquitin conjugating enzymes (*13*). These factors conspire to suggest Lon as a model enzyme to dissect the mechanistic details of the reaction cycle of ATP-dependent proteolysis and the complex process of substrate recognition by ATP-dependent proteases. Ec-Lon is responsible for the degradation of the majority of abnormal proteins, *in vivo* (*1*), suggesting that like the molecular chaperones (*14*) Ec-Lon recognizes a very broad range of polypeptide substrates. In addition, a growing number of normal cellular polypeptides have been identified as Ec-Lon substrates (*2*). The basis for substrate recognition by Ec-Lon, or any ATP-dependent protease, remains poorly understood. For example, it is not known if the specific substrates of Lon resemble unfolded or damaged proteins or if the enzyme has both broad modes (unfolded, damaged, and truncated protein) and specific modes (naturally short-lived protein) of substrate recognition. We report here that the highly divergent Ms-Lon protease degrades RcsA, a natural substrate of Ec-Lon (*15*), favoring models in which Lon enzymes recognize "unfolded" proteins, and naturally short-lived proteins are

‡ The nucleotide sequence reported here has been deposited in GenBank under Accession No. AF30688.

* To whom correspondence should be addressed. Phone: 718-430-2892. Fax: 718-430-8565. E-mail: shrader@aecom.yu.edu.

⊗ Abstract published in *Advance ACS Abstracts,* December 15, 1997.

S0006-2960(97)01732-7 CCC: $15.00
Published on Web 01/06/1998

designed to mimic, via unknown mechanisms, unfolded proteins.

Ec-Lon degrades certain proteins and small peptides with a chymotrypsin-like specificity during a reaction cycle that is nucleotide (NTP) dependent (*16*). Protein degradation requires NTP hydrolysis, but unstructured peptide substrates are hydrolyzed at near maximal rates in the presence of nonhydrolyzed analogs such as AMP-PNP (*17*, *18*). Both the basal peptidase and ATPase activities of Ec-Lon are stimulated by unfolded proteins, suggesting that the enzyme is stimulated by the very substrates that it degrades during conditions of cell stress (such as heat shock conditions) (*19*). We report here that these properties are conserved in Ms-Lon and are likely to be general properties of all Lon proteases. By contrast, both Ec-Lon and the Lon homolog from human mitochondria display a broad specificity for nucleotides and nucleotide analogs (*19, 20*), while we report that Ms-Lon greatly prefers ATP or dATP.

Previous data suggested that the stoichiometry of peptide bond hydrolysis to ATP hydrolysis by Ec-Lon was constant for several protein substrates (~2 ATP/peptide bond) (*6*). In addition, covalent modification of Ec-Lon's active site Ser residue inhibited the enzymes ATPase activity, suggesting that peptide and nucleotide hydrolysis might be linked reactions (*6*). However, the obligate coupling of these reactions was later ruled out via demonstrations that Ec-Lon's peptidase activity is supported by nonhyrolized nucleotide analogs (*17*) and removal of the active site Ser nucleophile abolished Ec-Lon's peptidase activity but not its ATPase activity (*21*). We report here that removal of the active site Ser nucleophile of Ms-Lon results in an enzyme that retains ATPase activity. However, we report that even conservative substitutions of the active site nucleophile of Ms-Lon have a measurable effect on its ATPase activity. This result, in part, forms the basis for our proposal that the peptidase and ATPase active sites of Ms-Lon interact structurally.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Media, Reagents and DNA Manipulations. *E. coli* strain MC1061 [*hsdR mcrB araD139* Δ*(araABC-leu)7679 galU galK rpsL thi*] was used for plasmid preparation and protein overexpression (*22*). Initial subclones of *M. smegmatis* DNA fragments from λgem-11 libraries were constructed in strain TS365 (MC1061 *pcnB80*). This strain underexpresses ColEI plasmids, increasing the stability of clones that are toxic to *E. coli* cells at high copy number (*23*, *24*). Strain Sgr-1 was constructed from OHP8 [MC4100 *cpsB10::lacZ(immλ)* leu::Tn10 sulB*] (*25*) by curing tetracycline resistance using fusaric acid plates (*26*). Strain Sgr-9 contains a *lon* mutation and was constructed from Sgr-1 via P1vir transduction and donor strain CAG9276 [C600 *lacZX90 lon146::Tn10/ F′ lacZX90*] (*27*). Where appropriate, *E. coli* cells were grown in Luria broth containing antibiotics at the following concentrations: Amp, 100 μg/mL; Kan, 20 μg/mL; Tet, 10 μg/mL. *Mycobacterium smegmatis* strain mc^2155 was treated as wild type in this study although this strain is known to contain mutations that result in an increased efficiency of transformation by plasmids (mc^2155 was derived from *M. smegmatis* strain ATCC 607) (*28*). *M. smegmatis* cells were grown in Middlebrook-ADC-TW media (*28*).

Identification of a Lon Protease Homolog Using the Polymerase Chain Reaction (PCR). Several sets of degenerate deoxyoligonucleotide pools were synthesized based on the conserved regions of the three Lon proteases whose sequences were known at that time (*29*−*31*). The most efficient deoxyoligonucleotide pools for the amplification of *lon* sequences were AMS1 = antisense of F(470)-S-D-H-Y-I/L = [5′-GCCGGATCCA(AGT)(AG)TA(AG)TG(AG)-TC(AG)CT(AG)AA-3′], AMS4 = D(405)-E-A-E-I-R = [5′-GCCGATCC(AGCT)GG(AGCT)GT(AGCT)GG(AGCT)-AA(AG)AC-3′], and AMS5 = P(376)-G-V-G-K-T = [5′-GCCGGATCCGA(CT)GA(AG)GC(AGCT)GA(AG)ATT-(ACT)(CA)G-3′] (numbers correspond to Ms-Lon, see Figure 2A). Deoxyoligonucleotide pools AMS1 and AMS5 were used to amplify regions of *M. smegmatis* and *E. coli* chromosomal DNA. Degenerate PCR often produces numerous "noise" DNA fragments along with the fragment of interest. By comparing PCR reactions from these two organisms, a DNA fragment amplified in both was identified as a putative *lon* segment. The PCR product generated from *M. smegmatis* DNA was purified and used as a template for further PCR amplification with deoxyoligonucleotide pools AMS4 and AMS5. This second, independent round of PCR amplification provided confidence that the amplified fragment was derived from a *lon* homolog. Sequence analysis of the PCR product from the second reaction revealed significant homology to the other known *lon* genes (Figure 1).

Isolation and Sequence Analysis of the M. smegmatis lon Gene. The PCR product described above was radiolabeled with ^{32}P and used as a probe to identify *lon*-containing clones from a λgem-11 library (Promega) containing 10−20 kb fragments of *M. smegmatis* chromosomal DNA isolated from strain mc^2-155 as described previously (*32*). Lambda phage plaque hybridizations were performed using standard techniques. Phage from one plaque, that showed strong hybridization with the probe, was chosen for further analysis. Southern hybridization to restriction digests of this phage, λlon-6, identified a 3.5 kb *Sac*I fragment that contained the PCR fragment (pAMS-90, see Figure 1). After initial sequence analysis of pAMS-90, a partially overlapping 4.3 kb *Bam*HI fragment was identified in λlon-6 (pAMS-110). These two plasmid clones together contain the entirety of the *M. smegmatis lon* gene and ~5 kb of surrounding chromosomal DNA (see Figure 1). Sequence determination employed dideoxynucleotide methods and was performed on both DNA strands for the regions indicated in Figure 1. This corresponds to the entire *lon* gene, its control region, and a small *Sal*I fragment that encodes the central portion of the 85-C antigen homolog. A small amount of flanking sequence information was determined on one DNA strand only as indicated in Figure 1. Sequence data are numbered from a *Sac*II site that is 337 bp upstream from the putative start codon of *lon*. In this numbering system, the *lon* stop codon starts at nucleotide 2670. The nucleotide sequence reported in this manuscript has been submitted to GenBank and is available under Accession Number AF30688.

Construction of a His_6-lon Allele and Characterization of an Affinity-Tagged Ms-Lon Protease. An N-terminally affinity-tagged Ms-Lon protease was constructed in order to facilitate enzyme overexpression and purification (*33*). PCR was used to generate a 173 bp *Bgl*II−*Kpn*I fragment

encoding the first 56 amino acids of the Ms-Lon protease proceeded by the affinity tag M-R-G-H_6-G-S. Using this fragment, the remainder of the His_6-*lon* gene was reconstructed from cloned fragments of the *lon* gene. The fragment generated by PCR was minimized to eliminate PCR-derived errors (this approach was also necessitated by the extreme difficulties encountered during attempts to use PCR to amplify longer regions of the G/C-rich *M. smegmatis* chromosome). Our final expression construct employs the vector pQE-8 (Qiagen) and contains *M. smegmatis* chromosomal DNA through the *Sph*I site (2805) that lies ∼130 nucleotides downstream of the *lon* termination codon. Overexpression of the His_6-Lon enzyme was carried out in *E. coli* strain MC1061 pREP4 (a kan^R, $lacI^Q$ containing vector; Qiagen). Overnight cultures of fresh transformants were diluted 1:100 into fresh LB medium (containing 100 μg/mL Amp and 25 μg/mL Kan) and grown at 37 °C until the A_{600} value for the culture reached ∼0.5, followed by growth in the presence of 1.0 mM IPTG for an additional 1−2 h. Cells were disrupted by sonication in lysis buffer (50 mM Tris, pH 7.9, 100 mM KCl, 1.0 mM imidazole). Nickel ion affinity chromatography involved binding cleared lysate from disrupted cells to Ni-NTA−agarose (Qiagen; 1.5 mL of Ni-NTA−agarose per liter of original culture volume), followed by washing with 100−200 volumes of lysis buffer containing 15 mM imidazole and elution with 5 volumes of lysis buffer containing 100 mM imidazole. Affinity-purified His_6-Lon was concentrated using a Centriprep 30 concentrator (Amicon) and further purified on a Sephacryl S-300 column (Pharmacia) equilibrated in lysis buffer. This material was concentrated to 5−10 mg/mL as judged by dye binding (BioRad) and analyzed for homogeneity using SDS−PAGE.

Mutation of the Active Site Region of Ms-Lon. All mutant Ms-Lon enzymes were expressed containing the aforementioned N-terminal His_6-peptide and were purified as described for the wild type enzyme. Mutations in the active site region of Ms-Lon were constructed by replacing the 31 nucleotide *Sty*I (2345)−*Bst*eII (2376) fragment of the *lon* gene, containing the catalytic residue S675, with double stranded deoxyoligonucleotides encoding the desired amino acid changes. To construct the indicated *lon* mutants, the following deoxyoligonucleotides were used (only the sense strand is shown; mutations are underlined): S675A = 5′-CAAGGACGGGCCGGCCCGCGGGCGTGACGATG-3′, S675T = 5′-CAAGGACGGGCCGACGGCGGGCGTGACGATG-3′, S675C = 5′-CAAGGACGGGCCGTGCGCGGGCGTGACGATG-3′, GPX = 5′-CAAGGACCCGGGGTCGGCGGGCGTGACGATG-3′, P674A = 5′-CAAGGACGGGGCCTCGGCGGGCGTGACGATG-3′, P-Ala = 5′-CAAGGACGCGGCCTCGGCCGCGGCCGCGATG-3′.

Peptidase and ATPase Assays. Peptidase assays contained 10 mM $MgCl_2$, 50 mM Tris (pH 7.9), 4.0 mM ATP, 0.3 mM peptide, and 5−10 μg of Ms-Lon in a total volume of 200 μL. Reactions were incubated for 30 or 60 min at 37 °C and stopped by the addition of 100 μL of 1% SDS and 1.2 mL of 0.1 M sodium borate (pH 9.2). Fluorescence was measured in a Perkin Elmer LS-5B spectrofluorometer with excitation at 335 or 380 nm, and emissions were monitored at 410 or 460 nm for peptides containing 4Mβ NA (4-methoxy-β-naphthylamide) or AMC (7-amido-4-methylcoumarin), respectively (*34*). The amount of free AMC or 4Mβ NA released during peptidase assays was calibrated using the free compound (Sigma). The peptides tested for their activity as substrates of Ms-Lon were Z-Gly-Gly-Arg-AMC, Z-Gly-Gly-Leu-AMC, Glt-Ala-Ala-Phe-4Mβ NA, and Suc-Leu-Tyr-AMC, Glt-Gly-Gly-Phe-AMC, Suc-Leu-Leu-Val-Tyr-AMC, and Suc-Phe-Leu-Phe-4Mβ NA. ATPase assays were performed using the method of Black and Jones for inorganic phosphate detection (*35*). Reaction mixtures contained 10 mM $MgCl_2$, 1.0 mM ATP, and 2−5 μg of Ms-Lon in a total volume of 200 μL. Reactions were stopped with the addition of 100 μL of 10% trichloroacetic acid. To this reaction was added a mixture of 150 μL of 14% ascorbic acid in 50% trichloroacetic acid and 100 μL of 2% ammonium molybdate, and the reactions were incubated for 10 min at room temperature. To this reaction was added 0.5 mL of 2% sodium acetate, 2% sodium arsenate, and 2% acetic acid and the reaction incubated for an additional 30 min at room temperature. The optical density of the final reaction was determined at 700 nm. Optical densities were converted to phosphate concentrations using KH_2PO_4 standards containing 1.0 mM ATP.

β-Galactosidase Assays of Sgr-9 Cells. The cellular levels of β-galactosidase in Sgr-9 cells containing Ms-Lon expressing plasmids were determined using the chromogenic β-galactosidase substrate 2-nitrophenyl β-D-galactopyranoside (ONPG). Transformants were grown overnight to saturation, without plasmid induction, and 1.0 mL of cells was pelleted in a microcentrifuge and resuspended in 50 μL of 25% sucrose, 0.25 M Tris (pH 8.0). To this mixture was added 10 μL of lysozyme (10 mg/mL) followed by incubation for 15 min on ice. The resulting protoplasts were lysed by the addition of 0.15 mL of 75 mM EDTA, 0.33 M NaCl and incubation on ice for 5 min, followed by the addition of 10 μL of 1% Triton X-100. Supernatant from the lysed cells was prepared by centrifugation at 12000*g* for 15 min at 4 °C. After centrifugation, the insoluble cell debris was removed from the supernatant fraction. ONPG assays were performed by incubating 50 μl of the above lysate in 650 μl of Z buffer [100 mM potassium phosphate (pH 7.0), 10 mM KCl, 1.0 mM $MgSO_4$, and 10 mM 2-mercaptoethanol] at room temperature. One hundred microliters of 10 mM ONPG was added to this reaction, and color was allowed to develop. The reaction was timed and stopped by the addition of 750 μL of 1 M Na_2CO_3 and the optical density of the final reaction determined at 420 nm. These values were normalized by dividing by the concentration of cellular protein as determined by the Bio-Rad protein detection reagent.

RESULTS AND DISCUSSION

Identification of the M. smegmatis lon Gene. Previous authors identified *lon* homologs by direct hybridization methods employing the *E. coli lon* gene as a probe (*30*). However, we expected that an *M. smegmatis lon* homolog, if present, would be highly divergent from the *E. coli lon* gene and also would have a significantly higher G/C content (the *M. smegmatis* genome is ∼65% G/C while the *E. coli* genome is ∼50% G/C). Therefore, we employed degenerate PCR in our search for homologs of the Lon protease in *M. smegmatis*. Using deoxyoligonucleotides corresponding to conserved regions of the Lon enzymes from *E. coli* (*29*), *Bacillus brevis* (*30*) and *Myxobacterium xanthus* (*31*), we

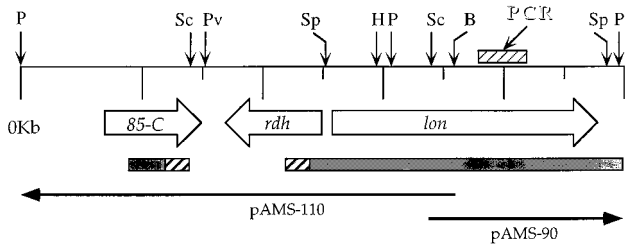

FIGURE 1: The *lon* gene of *Mycobacterium smegmatis*. A partial restriction map of the region of the *M. smegmatis* chromosome containing the *lon* gene. Shown are the positions of *lon*, a homolog of the ribitol dehydrogenase of *Klebsiella aerogens* (*rdh*), and a homolog of the common mycobacterial antigen (85-C). Also depicted is the portion of the *lon* gene originally isolated by degenerate PCR (lightly hatched box). The regions of the *M. smegmatis* chromosome that were sequenced on both strands are indicated by the stippled boxes while preliminary sequence determination has been obtained for the regions indicated by the heavily hatched boxes. The genomic sequences contained in plasmid clones pAMS-90 and pAMS-110 are indicated. Abbreviations: P = *Pst*I, Pv = *Pvu*II, K = *Kpn*I, Sc = *Sac*I, Sp = *Sph*I, H = *Hin*dIII, S = *Sal*I, B = *Bam*HI.

amplified a region of the *M. smegmatis* chromosome. Sequence analysis of this fragment revealed a clear homology to the Lon protease family (indicated in Figure 1). This internal region of the *M. smegmatis lon* gene was subsequently used to isolate and sequence large chromosomal DNA fragments encoding the entirety of the *M. smegmatis* Lon protease (Ms-Lon). In light of our previous demonstration that *M. smegmatis* contains a complete 20S proteasome (*32*), these sequence data identify *M. smegmatis* as the first known example of a eubacterium containing these two proteolytic systems. The region of the *M. smegmatis* chromosome containing the *lon* gene is schematized in Figure 1. Sequences upstream of *lon* encode a homolog from the common mycobacterial antigen family, which encode mycoyl-transferases (*36*) [most closely resembling antigen 85-C (*37*)] and a gene likely to encode a dehydrogenase [most closely resembling the ribitol dehydrogenase of *Klebsiella aerogenes* (*38*)]. Sequences downstream of *lon* neither encode significant open reading frames nor potential stem−loop structures, characteristic of factor-independent transcriptional terminators.

The deduced amino acid sequence of Ms-Lon is shown in Figure 2A. Also shown are the sequences of the Lon enzymes from *E. coli*, *B. brevis*, and *Bacillus subtilis*. As expected, based on the evolutionary distance between *E. coli* and *M. smegmatis*, the Lon proteases from these two organisms are highly divergent, and sequence alignments revealed only ∼40% identity. Surprisingly, Ms-Lon is also highly divergent from the Lon enzymes of the more closely related eubacteria *B. subtilis* and *B. brevis*, and sequence alignments again revealed ∼40% sequence identity. In fact, the Lon proteases from the Gram-positive eubacteria *B. subtilis* and *B. brevis* are more similar to the Lon protease from the Gram-negative eubacterium *E. coli* than they are to the Lon protease from the Gram-positive (acid fast) eubacterium *M. smegmatis* (Figure 2B). For example, in the 4-way sequence alignment (Figure 2A), there are 192 positions at which only 3 of the 4 sequences are identical, and in 112 of these positions it is the *M. smegmatis* sequence that does not match the consensus. In addition, the sequence alignment of Lon enzymes, depicted in Figure 2A, required the introduction of 4−5 gaps greater than a single amino acid to be introduced to optimize alignment. In each case, the gap was necessitated by the *M. smegmatis* sequence. The origin of the increased rate of sequence divergence between Ms-Lon and other Lon enzymes remains to be determined.

Several regions of Ms-Lon are conserved among all four of the Lon enzymes shown in Figure 2a. These include the ATPase domain [Walker A-box beginning at G367; Walker B-box beginning at V437 (*39*)] and the proteolytic active site region [catalytic nucleophile is S675 (*40*, *6*)]. In general, the N-terminal region of Lon homologs are not highly conserved. However, the reported sequence of the *B. subtilis lon* gene encodes a protein with N-terminal sequence M-A-E (*41*). This is identical to the putative N-terminus encoded by the *M. smegmatis lon* gene. In light of this identity and the strong ribosome binding site upstream of the putative initiation codon (GGAGG), this codon most likely encodes the first Met residue of the nascent Ms-Lon protease. We demonstrate below that Ms-Lon enzymes beginning at this codon are highly active.

Overexpression and Characterization of the M. smegmatis Lon Protease (Ms-Lon). Genes encoding homologs of the *E. coli* Lon protease have been identified in several bacteria and the mitochondria of eukaryotes. However, reports of successful overexpression of these enzymes have not appeared, and only the *lon* homolog from human mitochondria has been demonstrated to encode directly an ATP-dependent protease (*20*). We constructed an inducible overexpression system for Ms-Lon and expressed the enzyme as a fusion to an N-terminal hexa-His affinity domain (see Experimental Procedures) (*33*). Lysates from *E. coli* cells expressing and overexpressing His_6-Ms-Lon, and the resultant enzyme preparation, after initial and final purification, are shown in Figure 3. His_6-Ms-Lon represents 2−5% of the total cellular protein in uninduced cultures and >20% of the total cellular protein in cultures after induction (see Experimental Procedures). Final yields of purified Ms-Lon, and mutants (see below), were 10−20 mg of Ms-Lon/L of *E. coli* culture.

The Affinity Domain Does Not Alter the Essential Properties of Ms-Lon. We performed an important control experiment to ensure that the N-terminal modification of the *M. smegmatis* Lon protease did not alter the enzyme's essential properties. Briefly, we generated an alternate expression system that expresses the fusion protein His_6-ubiquitin−Ms-Lon (Figure 4). After expression, the His_6-ubiquitin domain was removed with the *S. cerevisiae* Ubp1 enzyme (*42*) to yield native Ms-Lon. Native Ms-Lon, His_6-Ms-Lon, and even unprocessed His_6-Ub−Ms-Lon all had identical peptidase and protease activity (data not shown and Figure 6B below). We conclude that small extensions at the N-terminus of Ms-Lon do not significantly alter its enzymatic activities and "Ms-Lon" is used to indicate the modified enzyme His_6-Ms-Lon.

Ms-Lon Has Chymotryptic Peptidase Specificity and both Its Peptidase and Its ATPase Activities Are Stimulated by Unfolded Protein. Previous authors employed small fluorogenic peptides to determine the cleavage site specificities of ATP-dependent proteases (*34*). Employing a similar series of substrates, we determined that Ms-Lon cleaves chymotrypsin-like substrates such as Suc-Phe-Leu-Phe-4MβNA and Glt-Ala-Ala-Phe-4MβNA but does not degrade trypsin-like substrates such as Z-Gly-Gly-Arg-AMC or other substrates

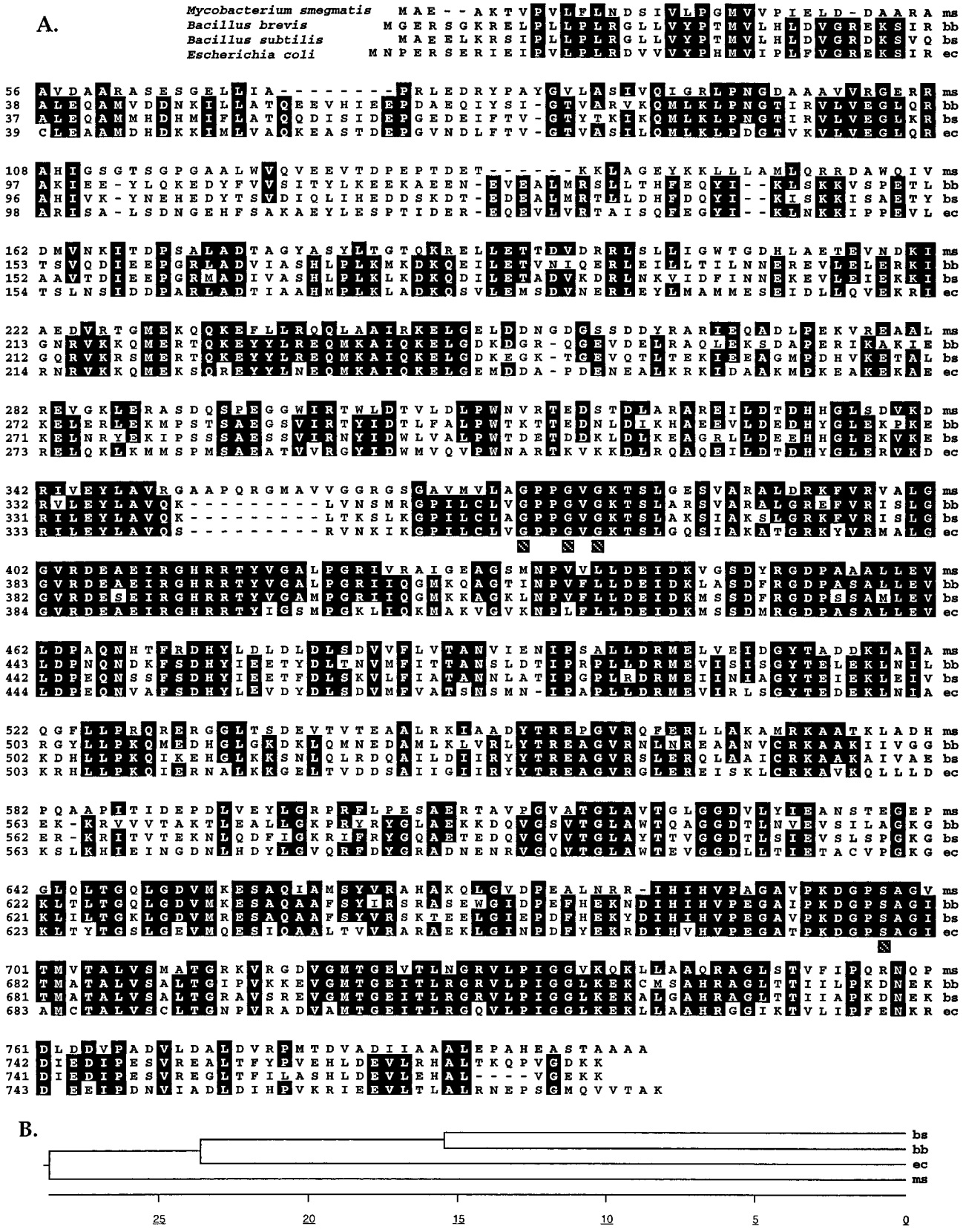

FIGURE 2: Sequence comparison of Ms-Lon with other Lon proteases. (A) The amino acid sequence deduced from the sequenced *M. smegmatis lon* gene is shown (ms) aligned with the amino acid sequences of the Lon proteases from *E. coli* (ec; *29*), *B. brevis* (bb; *30*), and *B. subtilis* (bs; *56*). The catalytic nucleophile S675 is indicated (40), as are three Gly residues that form part of the ATPase domain (39). The amino acid sequence is reported from the putative initiation codon of *lon* and may not represent the mature N-terminus of the enzyme. (B) Phylogenetic tree depicting the relationships of the Lon sequences shown in (A).

(see Experimental Procedures for a complete list of the peptides tested). Similar substrate specificities were reported for both Ec-Lon and human mitochondrial Lon (Hs-Lon) (*20*). Interestingly, the preferred substrate for the *M.*

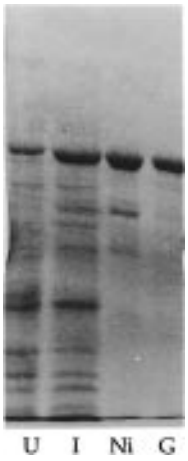

FIGURE 3: Expression and purification of Ms-Lon. An 8% SDS gel was used to fractionate the total cellular protein from *E. coli* cells overexpressing the *M. smegmatis* Lon protease. Lanes: U, culture before induction; I, culture after induction; Ni, Ni affinity-purified Ms-Lon protease; and G gel filtration-purified Ms-Lon protease.

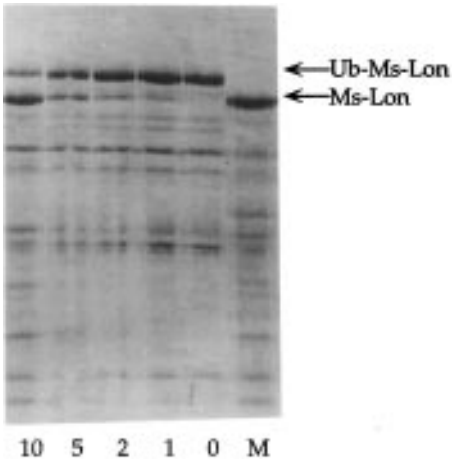

FIGURE 4: Expression of His_6-Ub−Ms-Lon fusion followed by removal of the His_6-Ub domain. Incubation of affinity-purified His_6-Ub−Ms-Lon with *E. coli* cell lysates from cells expressing the Ubp1 ubiquitin-fusion processing enzyme (42). Incubations are for 90 min with the indicated amount (in microliters) of *E. coli* lysate from cells expressing the Ubp1 enzyme. Lane M is independently purified Ms-Lon.

smegmatis 20S proteasome, Suc-Leu-Leu-Val-Tyr-AMC (*32*), was not cleaved by Ms-Lon, suggesting that these two large proteases are not highly redundant in their substrate recognition. For the preferred substrate, Glt-Ala-Ala-Phe-4MβNA, the cleavage rate (per minute) was ~2.0 mol of peptide/mol of Ms-Lon monomer (see Experimental Procedures). This rate of peptide cleavage was similar to the rate measured for overexpressed Ec-Lon, under similar conditions (*43*). As has been reported for Ec-Lon (*19*), cleavage of small peptide substrates by Ms-Lon was stimulated by the unfolded protein α-casein. For Ms-Lon, addition of 0.5 mg/mL α-casein, to peptidase reactions, increases the rate of degradation of Glt-Ala-Ala-Phe-4MβNA by 50%. Overall, our data reveal that the sequence divergence between Ms-Lon and Ec-Lon has left the essential peptidase activity relatively unchanged. In addition, stimulation of Ec-Lon's peptidase activity by unfolded protein, rather than the expected competition for Lon's peptidase active site, was conserved in Ms-Lon, suggesting that this remarkable property is broadly conserved among Lon homologs (*19*).

Table 1: Ms-Lon Is Activated to Near-Maximal Peptidase Activity Only by ATP and dATP[a]

nucleotide	4.0 mM	0.5 mM	0.05 mM	Ec-Lon 0.5 mM (% of ATP)
ATP	**100** (6)	58 (1)	11 (1)	100
dATP	109 (1)	89 (2)	nd	125
CTP	~0	~0	nd	68
GTP	~0	~0	nd	52
UTP	~0	~0	nd	56
ATPγS	6 (1)	nd	9 (1)	96
AMP-PNP	26 (1)	nd	~0	81
AMP-PCP	~0	nd	~0	52
ADP	~0	nd	nd	5
P-P-P	~0	nd	~0	72

[a] Degradation of the peptide Glt-Ala-Ala-Phe-4MβNA by Ms-Lon in the presence of the indicated nucleotide. Values are in percent relative to reactions containing 4.0 mM ATP (boldface text; acual rate is 2.0 molecules of Glt-Ala-Ala-Phe-4MβNA cleaved per Ms-Lon monomer per minute). Standard errors are shown in parentheses. Values for Ec-Lon are from reference *6*.

Ms-Lon Showed Increased Specificity for Nucleotide Substrates over Ec-Lon. Ms-Lon has basal ATPase activity (~25 molecules of ATP cleaved per Ms-Lon monomer per minute) which was stimulated 150% by the addition of 0.5 mg/mL α-casein. These values are very similar to those reported for Ec-Lon (*43*). However, Ec-Lon shows almost no discrimination between nucleotides, and even nonhydrolyzed nucleotide analogs and polyphosphate stimulated Ec-Lon to >50% of maximal levels during the degradation of peptide substrates (*6*). Similar promiscuity for nucleotides has been demonstrated for Hs-Lon which was activated, nearly equivalently, by ATP or CTP (*20*). By contrast, Table 1 reveals that Ms-Lon has a very strong preference for adenosine nucleotides. However, an intriguing aspect of Ec-Lon has been retained in Ms-Lon: Ms-Lon was activated by nonhydrolyzable nucleotide analogs (such as AMP-PNP and ATPγS). This activation demonstrates that only nucleotide binding by Ms-Lon is essential for activation of the proteolytic active site. The level of activation of Ms-Lon by AMP-PNP was less than that observed for Ec-Lon (*6*). However, given the increased nucleotide discrimination of Ms-Lon, this difference most likely results from a reduced tolerance of the slight structural differences between AMP-PNP and ATP, rather than a fundamental difference in the role of ATP in the reaction cycles of these two Lon enzymes.

Discrimination against Nonsubstrate Proteins Is Conserved between Ec-Lon and Ms-Lon. The mechanism of substrate recognition by Lon and other ATP-dependent proteases is poorly understood, and the assays described above confirm only that the essential properties of the peptidase active sites are conserved between Ms-Lon and Ec-Lon. Additional information concerning the overall conservation of substrate choice between these two enzymes can be inferred by determining the relative cellular toxicity of Ms-Lon to *E. coli* cells. Figure 5 depicts the growth of *E. coli* cultures containing Ms-Lon- and Ec-Lon-expressing plasmids (these expression plasmids are identical outside the *lon* coding regions). Contrary to our original expectations, *E. coli* cells tolerate expression of divergent Ms-Lon as well as they tolerate expression of indigenous Ec-Lon (*44*). In both cases, uninduced cultures grow slightly slower than cultures containing vector alone while induction slows, but

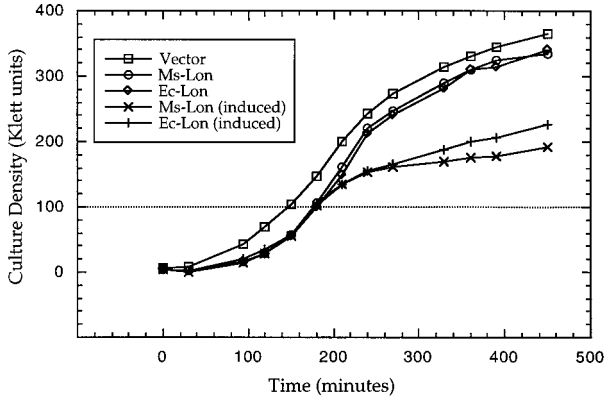

FIGURE 5: Growth curves of *E. coli* cells expressing either Ec-Lon or Ms-Lon. Cultures of *E. coli* strain MC1061 pREP4 containing the specified expression plasmids. Cultures were induced when the density of the culture reached 100 Klett units (dotted horizontal line) by the addition of IPTG to a final concentration of 1.0 mM.

does not stop immediately, cell growth (see Figure 5). This result is inconsistent with the rapid degradation of any essential *E. coli* protein by Ms-Lon. We conclude that both Ms-Lon and Ec-Lon are similar in their substrate recognition in that they both discrimate against nonsubstrate *E. coli* proteins.

In Vivo Complementation of an E. coli lon Mutant by Expression of Ms-Lon. To extend the results of the previous section, we determined the ability of Ms-Lon to degrade RcsA, a natural substrate of Ec-Lon. *E. coli* capsular polysaccharide synthesis is positively regulated by RcsA (*15*). Therefore, *E. coli lon* mutants contain increased steady-state levels of RcsA which results in excess polysaccharide synthesis and the slimy appearance of colonies of *E. coli lon* mutants (termed mucoid). To monitor the cellular level of RcsA, we utilized strain Sgr-9 which contains a lysogenized λ-bacteriophage containing *cpsB*−*lacZ* gene fusion (see Experimental Procedures) (*25*). Lysogens of the *cpsB*−*lacZ*-containing phage express β-galactosidase in response to cellular RcsA activity because RcsA is required for expression from the *cpsB* promoter and the *cpsB* promoter is responsible for expression of *lacZ* in this phage (see Figure 6A). Figure 6B indicates the cellular levels of β-galactosidase for cultures of Sgr-9 cells containing either Ec-Lon or Ms-Lon expression plasmids. Sgr-9 cells expressing Ec-Lon, Ms-Lon, or even unprocessed Ub-His_6−Ms-Lon display a >20-fold reduction in β-galactosidase activity relative to cells containing vector alone. We conclude that RcsA is degraded by Ms-Lon.

Overall, our results suggest that the divergent Ec-Lon and Ms-Lon enzymes recognize similar features in, at least, some of their substrates. However, the natural substrates of the Lon enzymes, from these two highly divergent organisms, are unlikely to be highly conserved in their primary sequence. For example, RcsA homologs have not been identified in any Gram-positive bacterium, and the *M. tuberculosis* genome, now 95% complete, has not been shown to contain a RcsA homolog. This seeming paradox is most easily explained if Lon enzymes recognize structural features of their substrates rather than conserved stretches of primary sequence. In view of the ability of Ec-Lon to degrade an enormous variety of damaged proteins (*6*), the recognized structural motif or feature must be very general. The most

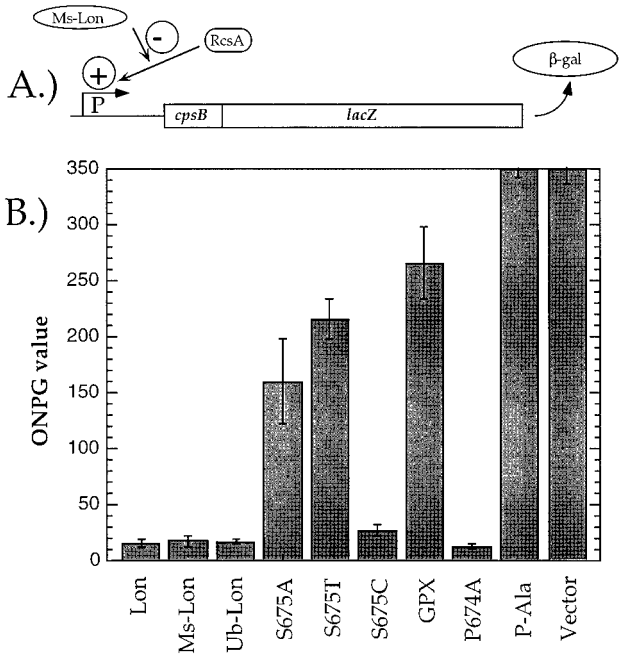

FIGURE 6: Ms-Lon degrades the Ec-Lon substrate RcsA. (A) Schematic of *cbsB*−*lacZ* gene fusion via which Sgr-9 cells express β-galactosidase in response to the cellular level of RcsA. (B) β-Galactosidase activity of cultures of Sgr-9 cells (*E. coli lon* mutant containing the *cpsB*−*lacZ* fusion) containing the indicted expression plasmids. Results are the average of two independent cultures, and standard errors are indicated.

likely feature shared by damaged and mutant proteins is a reduced compactness resulting from the inability to form a hydrophobic core and causing an increased exposure of hydrophobic residues to solvent. Therefore, it seems reasonable to conclude that Lon proteases recognize exposed hydrophobic residues in their substrates in analogy to molecular chaperones like GroESL (*14*). It is further suggested, by our data, that naturally short-lived eubacterial proteins will also lack the structural compactness characteristic of stable proteins and expose a unusual percentage of their hydrophobic residues to solvent. This general property of short-lived bacterial proteins would then account for their conserved recognition by members of the Lon protease family. This suggestion is consistent with previous demonstrations that RcsA is very insoluble and prone to aggregation when overexpressed, even in its native host, suggesting it is not a typical globular protein (*15*). The actual degree of the reduction in structural compactness of naturally short-lived proteins is unknown, and in light of the biological activity of short-lived proteins such as RcsA and SulA (*2*), almost certainly does not correspond to completely unstructured protein (random coil).

Active Site Mutants of Ms-Lon: Protease Activity. Based on a weak sequence homology between a region of Ec-Lon and other members of the chymotrypsin family of proteases, Amerik and colleagues identified the putative active site Ser nucleophile (*40*). This same residue was identified during Ec-Lon inactivation studies with peptidyl chloromethyl ketones (*6*). Indeed, replacement of this Ser with Ala eliminated both the protease and the peptidase activity of Ec-Lon (*21*, *40*). We found that this region of Ec-Lon was absolutely conserved in Ms-Lon (Figure 2A). To more fully characterize the active site region of Ms-Lon, we constructed variant Ms-Lons containing mutations in this region. The

active site Ser (S675) was changed to Ala (S675A), Thr (S675T), and Cys (S675C). In addition, the importance of small perturbations to the orientation of S675 was determined by changing the Pro residue immediately before S675 to Ala (P674A). To test the effect of more drastic structural changes to the active site region, the G673 and P674 residues were exchanged (GPX), and the entire region flanking S675 was converted to Ala (P-Ala = G673A + P674A + G677A + V678A + T679A; S675 and A676 were unchanged).

All of the active site mutations to Ms-Lon, except P674A, resulted in purified enzymes lacking detectable peptidase activity against Glt-Ala-Ala-Phe-4MβNA. The P674A enzyme retained 15% of wild-type Ms-Lon activity against Glt-Ala-Ala-Phe-4MβNA and the maximal measurable protease activity in more sensitive *in vivo* assays (Figure 6B, described above). The level of detection in this assay was ~2% of wild-type Ms-Lon activity (equivalent to background degradation in faux-purifications). We conclude that both the identity of the active site nucleophile and its precise orientation are critical determinants of Ms-Lon's proteolytic activity.

The S675C Mutant Retains in Vivo Protease Activity against RcsA. Our inability to detect peptidase activity in the S675C or S675T mutants was further examined by demonstrating that S675C retained <10% of wild-type activity against fluorescently labeled α-casein (data not shown). The large effect of replacing the catalytic nucleophile of Ms-Lon was to be expected as the same substitution to trypsin results in an enzyme with a K_{cat} that is $\sim 10^{-6}$ of the wild-type value (*45*). However, as a final test to detect residual protease activity, we employed the *in vivo* assay for the degradation of RcsA (described above). As demonstrated in Figure 6B, the S675C mutant was indeed active against RcsA. In addition, colonies of the *E. coli lon* mutant strain CAG9276 pREP4 appear nonmucoid when transformed with plasmids expressing Ec-Lon, Ms-Lon, S675C, and P674A but not when transformed with plasmids expressing S675A, S675T, GPX, or P-Ala (data not shown). We conclude that, like trypsin (*45*) and the *E. coli* leader peptidase (*46*), conservative replacement of the active site nucleophiles is tolerated. Although it is not possible to extract K_{cat} values from the data in Figure 6B, the observed near complete complementation of an *E. coli lon* mutation by S675C makes it likely that the K_{cat} of S675C was reduced by considerably less than the factor of 10^6 observed for the same substitution in trypsin (*45*).

Also shown in Figure 6b are data suggesting that expressing the S675A mutant reduces the cellular activity of RcsA. These results might mean that S675A retains some proteolytic activity, indicating that S675 is not the active site nucleophile. However, we favor an alternate interpretation in which the S675A mutant binds, unfolds, and renders RcsA a substrate for other cellular proteases or peptidases, or simply sequesters RcsA from its cellular targets. This interpretation is supported both by the demonstrated ability of molecular chaperones to unfold and present proteins to ATP-dependent proteases (*47*), and by data suggesting that both Lon and other ATP-dependent proteases have chaperone-like activity (*48*−*51*).

Active Site Mutants of Ms-Lon: ATPase Activity. In addition to determining the effect of active site mutations on the peptidase and protease activities of Ms-Lon, we determined the effect of these mutations on the enzyme's basal and stimulated ATPase activities (Figure 7). In contrast to peptidase activity, all of the mutants tested retain ATPase activity. However, substitution of the S675 nucleophile that removes a single hydroxyl group (S675A) results in a modest but measurable decrease in ATPase activity, while a more drastic structural alteration (P-Ala) has a larger effect. By contrast, active site substitutions that retain the overall size of the S675 nucleophile (S675C and S675T) do not reduce Ms-Lon's ATPase activity. The apparent increase in the ATPase activity of these Ms-Lon mutants, relative to wild-type Ms-Lon, probably reflects the slight self-proteolysis of wild-type Ms-Lon during assays, rather than a true increase (data not shown). In addition, the P674A mutation which retains 15% of peptidase activity does not alter Ms-Lon's ATPase activity. The trivial explanation of these results, that active site mutations result in the gross unfolding of Ms-Lon affecting all parts of the enzyme, was ruled out using circular dichroism (CD). Figure 8 reveals that Ms-Lon, S675A, and even P-Ala have superimposable CD spectra.

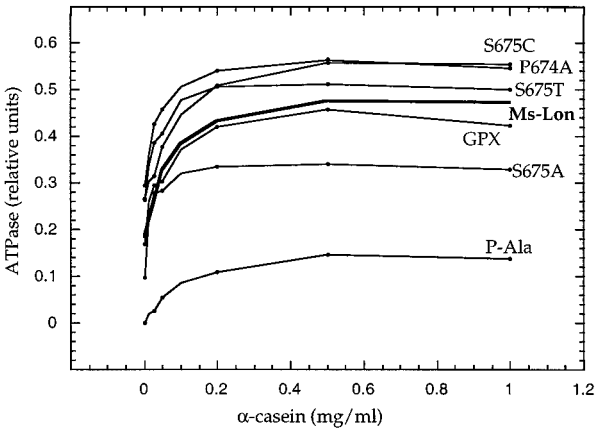

FIGURE 7: Basal and α-casein-stimulated ATPase activities of Ms-Lon active site variants. ATPase assays were performed on Ms-Lon and the indicated variants in the presence of increasing amounts of α-casein. For the α-casein concentrations (0, 0.05, 0.1, 0.2, 0.5. and 1.0 mg/mL), the standard errors in 10^{-3} relative units were Ms-Lon (12, 10, 0, 6, 4, 0), S675A (2, 3, 6, 9, 4, 13), S675T (8, 2, 17, 12, 0, 7), S675C (33, 6, 7, 4, 3, 7), P674A (16, 15, 5, 17, 5, 19), GPX (15, 20, 43, 42, 65, 49), and P-Ala (5, 1, 3, 3, 12, 6), respectively. The ratios of the ATPase activity in the plateau region to the basal (no α-casein) ATPase activities were Ms-Lon (2.55), S675A (3.41), S675T (1.87), S675C (1.92), P674A (2.0), GPX (2.61), and P-Ala (undefined as basal value is ~0). The plateau ATPase value is defined as (ATPase value at 0.5 mg/mL α-casein + ATPase value at 1.0 mg/mL α-casein)/2.

These data reveal that even small perturbations of the structure of Ms-Lon's peptidase active site result in changes in the ATPase active site. In other words, we conclude that these two sites interact structurally. This conclusion is consistent with the data of Goldberg and colleagues demonstrating that covalent modification of Ec-Lon's catalytic Ser residue inhibits the enzyme's ATPase activity (*6*). However, we suggest that the mechanism of the observed inhibition was via structural alteration of Ec-Lon's ATPase site, due to the modifying group in the peptidase active site.

In addition to changes in the basal ATPase rate of certain Ms-Lon mutants, we find significant differences in the magnitude of stimulation of their ATPase activity by α-casein. As seen in Figure 7, the degree of stimulation of Ms-Lon's ATPase activity, defined as the plateau ATPase

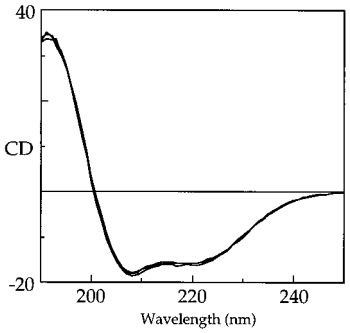

FIGURE 8: Circular dichroism spectrum of Ms-Lon, S675A, and P-Ala. Superposition of the circular dichroism spectra of Ms-Lon, S675A, and P-Ala in the region from 190 to 250 nm recorded on a Jasco J-720 instrument at 25 °C. Deconvoloution of these spectra by the method of Yang (57) yielded ~38% α-helix, 32% β-sheet, and 30% random coil; similar values were reported for Ec-Lon (21).

value over the basal ATPase value, is inversely correlated with the basal ATPase activity (numerical values are given in the Figure 7 legend). We interpret these data to suggest that unfolded protein binding, in part, optimizes the structure of the ATPase active site. The similar shapes of the substrate stimulation curves for the different mutants suggest that unfolded protein binding was not impaired in these mutants (resulting in a flatter curve), rather that the final form of the ATPase domain is suboptimal (resulting in a lower plateau). This interpretation is consistent with the data of Goldberg and colleagues revealing that the stimulation of Ec-Lon's ATPase activity by unfolded protein is mediated via the stimulation of ADP release (*6*), and further suggest that changes to Ms-Lon's peptidase active site result in impaired release of bound ADP from the enzyme.

Protein Degradation in the Mycobacteria. We demonstrated previously that *M. smegmatis* is one of a small number of eubacteria known to contain a complete homolog of the 20S proteasome originally isolated in eukaryotes (*32*). However, we were able to inactivate the *M. smegmatis* 20S proteasome without killing the cell (*32*). The dispensability of the eubacterial 20S proteasome contrasts the situation observed in eukaryotes where 20S proteasome subunits are encoded by essential genes (*7*) and the 20S proteasome is involved in the turnover of the vast majority of cytoplasmic protein (*52*). This result suggested either that ATP-dependent protein degradation is not particularly important to the mycobacterial cell or that these cells contain other ATP-dependent proteases. Here we have demonstrated that *M. smegmatis* contains the Lon protease, an enzyme that degrades 40−60% of the damaged protein in *E. coli* (*6*). This identifies *M. smegmatis* as the first known eubacterial cell to contain both Lon and a complete 20S proteasome (containing both α- and β-subunits). This would seem to be an excessive amount of proteolytic machinery for one cell; however, the list is unlikely to stop with these two enzymes. Genomic sequencing of *M. tuberculosis* reveals a homolog of a membrane-bound ATP-dependent protease (*51*). In addition, both the *M. tuberculosis* and *M. leprae* genomes (*53*) contain homologs of the ClpA subunit of the ClpAP protease (*11*, *12*).

Overall, we have demonstrated that even the 20S proteasome-containing eubacterium *M. smegmatis* contains multiple independent proteolytic systems. This same situation occurs in the distantly related eubacteria *E. coli* and *Haemophilus influenzae* (*54*) and the cyanobacterium *Synechocystis sp.* (*55*), suggesting it is a general property of all bacterial cells. Multiple ATP-dependent proteases are in contrast to the eukaryotic cell where the multicatalytic proteasome degrades essentially all cytoplasmic short-lived proteins (*52*). The physiological rationale for this disparity and also the overall importance of protein turnover to the cell remain central, but unresolved, questions.

ACKNOWLEDGMENT

We thank Dr. Ronald Gill for the *M. xanthus lon* sequence prior to publication, Dr. Georgi Abramochkin for help with Lon purification, and Drs. Alanna Schepartz Shrader and Thomas Leyh for comments on the manuscript.

REFERENCES

1. Goldberg, A. L. (1992) *Eur. J. Biochem. 203*, 9−23.
2. Gottesman, S., and Maurizi, M. R. (1992) *Microbiol. Rev. 56*, 592−621.
3. Backer, J. M., and Dice, J. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5830−5834.
4. Reeve, C. A., Bockman, A. T., and Matin, A. (1984) *J. Bacteriol. 157*, 758−763.
5. Wayne, L. G., and Hayes, L. G. (1996) *Infect. Immun. 64*, 2062−2069.
6. Goldberg, A. L., Moerschell, R. P., Chung, C. H., and Maurizi, M. R. (1994) *Methods Enzymol. 244*, 350−375.
7. Hochstrasser, M. (1995) *Curr. Opin. Cell. Biol. 7*, 215−223.
8. Rohrwild, M., Coux, O., Huang, H.-C., Moerschell, R. P., Yoo, S. J., Seol, J. H., Chung, C. H., and Goldberg, A. L. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5808−5813.
9. Missiakas, D., Schwager, F., Betton, J.-M., Georgopoulos, C., and Raina, S. (1996) *EMBO J. 15*, 6899−6909.
10. Rohrwild, M., Pfeifer, G., Santarius, U., Muller, S. A., Huang, H. C., Engel, A., Baumeister, W., and Goldberg, A. L. (1997) *Nature Struct. Biol. 4*, 133−139.
11. Hwang, B. J., Woo, K. M., Goldberg, A. L., and Chung, C. H. (1988) *J. Biol. Chem. 263*, 8727−8734.
12. Katayama, Y., Gottesman, S.; Pumphrey, J.; Rudikoff, S., Clark, W. P., and Maurizi, M. R. (1988) *J. Biol. Chem. 263*, 15226−15236.
13. Jentsch, S. (1992) *Ann. Rev. Genet. 26*, 179−207.
14. Fenton, W. A., and Horwich, A. L. (1997) *Protein Sci. 6*, 743−760.
15. Gottesman, S., and Stout, V. (1991) *Mol. Microbiol. 5*, 1599−1606.
16. Chung, C. H., and Goldberg, A. L. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4931−4935.
17. Goldberg, A. L., and Waxman, L. (1985) *J. Biol. Chem. 260*, 12029−34.
18. Van Melderen, L., Thi, M. H. D., Lecchi, P., Gottesman, S., Couturier, M., and Maurizi, M. R. (1996) *J. Biol. Chem. 271*, 27730−27738.
19. Waxman, L., and Goldberg, A. L. (1986) *Science 232*, 500−503.
20. Wang, N., Gottesman, S., Willingham, M. C., Gottesman, M. M., and Maurizi, M. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11247−11251.
21. Fisher, H., and Glockshuber, R. (1993) *J. Biol. Chem. 268*, 22502−22507.
22. Casadaban, M. J. and Cohen, S. N. (1980) *J. Mol. Biol. 138*, 179−207.
23. Lopilato, J., Bortner, S., and Beckwith, J. (1986) *Mol. Gen. Genet. 205*, 285−290.
24. Shrader, T. E., Tobias, J. W., and Varshavsky, A. (1993) *J. Bacteriol. 175*, 4364−4374.
25. Dervyn, E., Canceill, D., and Huisman, O. (1990) *J. Bacteriol. 172*, 7098−7103.
26. Bochner, B. R., Huang, H. C., Schieven, G. L., and Ames, B. N. (1980) *J. Bacteriol. 143*, 926−933.

27. Straus, D., Walter, W., and Gross, C. A. (1990) *Genes. Dev. 4*, 2202−2209.
28. Jacobs, W. R., Kalpana, G. V., Cirillo, J. D., Pascopella, L., Snapper, S. B., Udani, R. A., Jones, W., Barletta, R. G., and Bloom, B. R. (1991) *Methods Enzymol. 204*, 537−555.
29. Chin, D. T., Goff, S. A., Webster, T., Smith, T., and Goldberg, A. L. (1988) *J. Biol. Chem. 263*, 11718−11728.
30. Ito, K., Udaka, S., and Yamagata, H. (1992) *J. Bacteriol. 174*, 2281−2287.
31. Gill, R. E., Karlok, M., and Benton, D. (1993) *J. Bacteriol. 175*, 4538−4544.
32. Knipfer, N., and Shrader, T. E. (1997) *Mol. Microbiol. 25*, 375−383.
33. Smith, M. C., Furman, T. C., Ingnolia, T. D., and Pidgeon, C. (1988) *J. Biol. Chem. 263*, 7211−7215.
34. Waxman, L., and Goldberg, A. L. (1985) *J. Biol. Chem. 260*, 12022−12028.
35. Black, M. J., and Jones, M. E. (1983) *Anal. Biochem. 135*, 233−238.
36. Belisle, J. T., Vissa, V. D., Sievert, T., Takayama, K., Brennan, P. J., and Besra, G. S. (1997) *Science 276*, 1420−1422.
37. Wiker, H. G., and Harboe, M. (1992) *Microbiol. Rev. 56*, 648−661.
38. Azakami, H., Sugino, H., Iwata, N., Yokoro, N., Yamashita, M., and Murooka, Y. (1995) *Gene 164*, 89−94.
39. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) *EMBO J.* 1, 945−951.
40. Amerik, A. Y., Antonov, V. K., Gorbalenya, A. E., Kotova, S. A., Rotanova, T. V., and Shimbarevich, E. V. (1991) *FEBS Lett. 287*, 211−241.
41. Riethdorf, S., Volker, U., Gerth, U., Winkler, A., Engelmann, S., and Hecker, M. (1994) *J. Bacteriol. 176*, 6518−6527.
42. Tobias, J. W., Shrader, T. E., Rocap, G., and Varshavsky, A. (1991) *Science 254*, 1374−1377.
43. Thomas, C. D., Modha, J., Razzaq, T. M., Cullis, P. M., and Rivett, A. J. (1993) *Gene 136*, 237−242.
44. Goff, S. A., and Goldberg, A. L. (1987) *J. Biol. Chem. 262*, 4508−4515.
45. Higaki, J. N., Evnin, L. B., and Craik, C. S. (1989) *Biochemistry 28*, 9256−9263.
46. Tschantz, W. R., Sung, M., Delgado-Partin, V. M., and Dalbey, R. E. (1993) *J. Biol. Chem. 268*, 27349−27354.
47. Sherman, M. Y., and Goldberg, A. L. (1992) *EMBO J. 11*, 71−77.
48. Wickner, S., Gottesman, S., Skowyra, D., Hoskins, J., McKenney, K., and Maurizi, M. R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 12218−12222.
49. Levchenko, I., Luo, L., and Baker, T. A. (1995) *Genes Dev. 9*, 2399−2408.
50. Wawrzynow, A., Wojtkowiak, D., Marszalek, J., Banecki, B., Jonsen, M., Graves, B., Georgopoulos, C., and Zylicz, M. (1995) *EMBO J. 14*, 1867−1877.
51. Suzuki, C. K., Rep, M., Maarten van Dijl, J., Suda, K., Grivell, L. A., and Schatz, G. (1997) *Trends Biochem. Sci. 22*, 118−123.
52. Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A. L. (1994) *Cell 78*, 761−771.
53. Nath, I., and Laal, S. (1991) *Nucleic Acids Res. 18*, 4935.
54. Fleischmann, R. D., *et al.* (1995) *Science 269*, 496−512.
55. Kaneko, T., *et al.* (1996) *DNA Res. 3*, 109−136.
56. Wetzstein, M., Volker, U., Dedio, J., Lobau, S., Zuber, U., Schiesswohl, M., Herget, C., Hecker, M., and Schumann, W. (1992) *J. Bacteriol. 174*, 3300−3310.
57. Yang, J. T., Wu, C. S., and Martinez, H. M. (1986) *Methods Enzymol. 130*, 208−269.

BI971732F