THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 279, No. 9, Issue of February 27, pp. 8140–8148, 2004
Printed in U.S.A.

The Catalytic Domain of *Escherichia coli* Lon Protease Has a Unique Fold and a Ser-Lys Dyad in the Active Site*

Received for publication, November 7, 2003, and in revised form, November 28, 2003
Published, JBC Papers in Press, December 9, 2003, DOI 10.1074/jbc.M312243200

Istvan Botos‡, Edward E. Melnikov§, Scott Cherry‡, Joseph E. Tropea‡, Anna G. Khalatova§, Fatima Rasulova¶, Zbigniew Dauter‡, Michael R. Maurizi¶, Tatyana V. Rotanova§, Alexander Wlodawer‡, and Alla Gustchina‡‖

From the ‡Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, Maryland 21702-1201, the §Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow 117997, and the ¶Laboratory of Cell Biology, National Cancer Institute, Bethesda, Maryland 20892

ATP-dependent Lon protease degrades specific short-lived regulatory proteins as well as defective and abnormal proteins in the cell. The crystal structure of the proteolytic domain (P domain) of the *Escherichia coli* Lon has been solved by single-wavelength anomalous dispersion and refined at 1.75-Å resolution. The P domain was obtained by chymotrypsin digestion of the full-length, proteolytically inactive Lon mutant (S679A) or by expression of a recombinant construct encoding only this domain. The P domain has a unique fold and assembles into hexameric rings that likely mimic the oligomerization state of the holoenzyme. The hexamer is dome-shaped, with the six N termini oriented toward the narrower ring surface, which is thus identified as the interface with the ATPase domain in full-length Lon. The catalytic sites lie in a shallow concavity on the wider distal surface of the hexameric ring and are connected to the proximal surface by a narrow axial channel with a diameter of ~18 Å. Within the active site, the proximity of Lys^{722} to the side chain of the mutated Ala^{679} and the absence of other potential catalytic side chains establish that Lon employs a Ser^{679}-Lys^{722} dyad for catalysis. Alignment of the P domain catalytic pocket with those of several Ser-Lys dyad peptide hydrolases provides a model of substrate binding, suggesting that polypeptides are oriented in the Lon active site to allow nucleophilic attack by the serine hydroxyl on the *si*-face of the peptide bond.

Rapid proteolysis plays a major role in post-translational cellular control by the targeted degradation of short-lived regulatory proteins and also serves an important function in protein quality control by eliminating defective and potentially damaging proteins from the cell (1–4). In all cells, protein degradation is predominantly carried out by ATP-dependent proteases, which are complex enzymes containing both ATPase and proteolytic activities expressed as separate domains within a single polypeptide chain or as individual subunits in complex assemblies. Five ATP-dependent proteases, Lon, FtsH, ClpAP, ClpXP, and HslUV, have been discovered in *Escherichia coli*, and homologous proteases have been found in all eubacteria and in many eukaryotes. Alone among them, Lon protease has been found in virtually all living organisms, from Archaea to eubacteria to humans.

E. coli Lon protease was the first ATP-dependent protease to be identified (5–8). It is an oligomeric multidomain enzyme whose single polypeptide chain is composed of 784 amino acids (9). Comparison of the amino acid sequences of various members of the Lon family (9–11) suggested that Lon consists of three functional domains: a variable N-terminal domain, an ATPase domain, and a C-terminal proteolytic domain. The domain organization has been confirmed by expression of functional domains of the *E. coli* and yeast mitochondrial Lon (12, 13) and by limited proteolysis of the *E. coli* and *Mycobacterium* Lons (14–16). Despite extensive studies of this enzyme, many of its structural characteristics remain undetermined, although its function in selective energy-dependent proteolysis has been characterized in considerable detail (3, 17, 18).

As in other ATP-dependent proteases, the Lon ATPase domain belongs to the superfamily of AAA^+ proteins (ATPases associated with different cellular activities) (19). The characteristic AAA^+ domain consists of 220–250 amino acids that include hallmark Walker A and B motifs and several other regions of high sequence conservation (19). An AAA^+ module has two structural domains (20): a RecA-like α/β domain and an α domain that may interact with protein substrates. In response to ATP hydrolysis, these domains undergo changes in conformation and orientation. Transduction of the mechanical motions within the AAA^+ module to bound substrates provides the driving force for various functions, including binding and unfolding of target proteins, translocation of proteins to an associated functional domain (in this case, the protease), and coordinated activation of the functional domain (1, 20–23).

The heterooligomeric ATP-dependent proteases (ClpAP, ClpXP, and HslUV), as well as 26 S proteasomes, display similar overall architecture despite a lack of similarity in sequence and three-dimensional folds of their proteolytic subunits (22). The AAA^+ modules are contained in separate subunits, which assemble into six- or seven-membered rings. The proteolytic subunits also assemble into six- or seven-membered rings that stack upon each other to form barrel-shaped complexes with a central cavity containing the proteolytic active sites accessible by narrow axial channels in the rings. The ATPase rings interact at both ends of the protease barrels.

Despite considerable efforts, the spatial arrangement and

* This work was supported in part by Russian Foundation for Basic Research Grant 02-04-48481 (to T. V. R.) and by United States Civilian Research and Development Foundation Grant RB1–2505-MO-03 (to T. V. R. and A. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 1rr9 and 1rre) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‖ To whom correspondence should be addressed: National Cancer Institute, MCL, Bldg. 539, Rm. 143, Frederick, MD 21702-1201. Tel.: 301-846-5338; Fax: 301-846-6128, E-mail: alla@ncifcrf.gov.

oligomeric state of intact, homooligomeric Lon and FtsH proteases are still not known with certainty. Analysis by proteolysis and expression of functional domains of Lon proteases from several sources points to self-association of the isologous domains into hexameric or heptameric rings (13, 24).[1] High resolution structures of individual domains of FtsH and Lon have now been reported. The crystal structure of the AAA^+ module of FtsH was used to construct a model of a hexameric ring similar to those formed by other AAA^+ modules (25, 26). The α domain of the *E. coli* Lon AAA^+ module, the only structural fragment of Lon to be crystallized previously, displays a fold typically found in AAA^+ proteins (27). The quality of the structure of the isolated α domain confirmed the unique stability of these α domains, which are thought to undergo rigid body motions during the catalytic cycle of AAA^+ proteins. However, the oligomerization state of the intact protein could not be deduced from the α domain structure (27).

The proteolytic components of ATP-dependent proteases contain several different active site types. ClpP has a classic serine protease triad (28); HslV, like the proteasome, has a catalytic N-terminal threonine residue (29); and FtsH is a zinc-dependent metalloprotease (30). The catalytic residues in Lon have remained uncertain until quite recently. Mutational studies suggested that Lon had a catalytic serine (31); however, other candidate catalytic residues could not be definitively identified (32). Extensive sequence comparisons suggested that Lon contains a catalytic Ser-Lys dyad (32–34), and this model was experimentally supported by site-directed mutagenesis of candidate residues in *E. coli* Lon (34). Here, we report the high resolution crystal structure of the proteolytic domain of *E. coli* Lon. The structure confirms the presence of a Ser-Lys catalytic dyad in the active site and reveals a unique structural fold distinct from both the classical serine proteases containing active site catalytic triads and from other hydrolytic enzymes that are utilizing Ser-Lys catalytic dyads. The catalytic domain of Lon in the crystals assembles into hexameric rings, which provides strong support for a hexameric structure of the holoenzyme.

EXPERIMENTAL PROCEDURES

Expression and Purification of the Full-length Lon-S679A—The intact, proteolytically inactive mutant, Lon-S679A, was expressed and purified as described previously (27, 35). This procedure yielded ~100 mg of Lon-S679A from 40 g of cell paste, with ~90% purity. All of the purification procedures were performed at 4 °C and monitored by SDS-PAGE on 4–12% NuPAGE gels (Invitrogen). Protein concentration was estimated with a Bio-Rad protein assay using bovine serum albumin as the standard. All of the chromatography columns were from Amersham Biosciences, and the filter membranes were from Millipore (Bedford, MA).

Limited Proteolysis of Intact Lon and Purification of the P Domain—Purified Lon-S679A (100 mg) was cleaved with 250 μg of α-chymotrypsin (Sigma) in 50 ml of 20 mM potassium phosphate, pH 8, containing 0.3 M NaCl. After 2 h of incubation at 30 °C, the reaction was stopped by adding phenylmethylsulfonyl fluoride to 1 mM. The reaction solution was cooled to 4 °C, diluted 6-fold with 20 mM potassium phosphate, pH 8.0, containing 1 mM DTT,[2] filtered through a 0.45-μm polysulfone membrane and loaded onto a 5-ml HiTrap Heparin HP column equilibrated with the dilution buffer. The target protein was collected in the flow-through, concentrated using a YM-10 membrane, diluted 10-fold with buffer A (20 mM HEPES, pH 7.0, 1 mM DTT), and applied to a 5-ml HiTrap Q-Sepharose HP column equilibrated in buffer A. The column was washed and the protein was eluted with a 100-ml linear gradient of NaCl from 0 to 1 M in buffer A. Fractions eluted within 0.1–0.2 M NaCl were pooled, concentrated on Centriprep-10, and applied to a HiLoad 26/60 Superdex 75 column equilibrated in 20 mM Tris-HCl, pH 8, 0.15 M NaCl, 1 mM DTT buffer. The purity and homogeneity of the target fragment (residues 585–784) were verified by N-terminal sequencing and electrospray ionization mass spectroscopy (Agilent 1100 series).

Cloning, Expression, and Purification of the P Domain—Coding regions for the Lon P domain were amplified from a plasmid carrying the gene for Lon-S679A (31), using PCR with the following oligonucleotide primers: 5′-GAG AAC CTG TAC TTC CAG GAC TAT GGT CGC GCT GAT AAC GAA AAC-3′ and 5′-GGG GAC CAC TTT GTA CAA GAA AGC TGG GTT ATT ACT ATT TTG CAG TCA CAA CCT GCA TG-3′ (primer Lon R). The PCR amplicon was subsequently used as template for a second PCR with the following primers: 5′-GGG GAC AAG TTT GTA CAA AAA AGC AGG CTC GGA GAA CCT GTA CTT CCA G-3′ and primer Lon R. The amplicon from the second PCR was inserted by recombinational cloning into the entry vector pDONR201 (Invitrogen), and the nucleotide sequence was confirmed by DNA sequencing. The open reading frame encoding the P domain of Lon (Asp^{585}-Lys^{784}) containing an N-terminal recognition site (ENLYF(Q/D)) for tobacco etch virus protease was cloned into the destination vector pDEST-His-maltose-binding protein to produce pELP2. pELP2 directs the expression of the inactive, S679A P domain mutant of *E. coli* Lon as a fusion protein with *E. coli* maltose-binding protein with an intervening tobacco etch virus protease recognition site. The maltose-binding protein contains an N-terminal His_6 tag for affinity purification by immobilized metal affinity chromatography. The fusion protein was expressed in the *lon*-deficient *E. coli* strain BL21 (DE3) (Novagen, Madison, WI). The cells were grown to mid-log phase (A_{600} = ~0.5) at 37 °C in Luria broth containing 100 μg/ml ampicillin and 0.2% glucose. Overproduction of fusion protein was induced with isopropyl-α-D-thiogalactopyranoside at a final concentration of 1 mM for 4 h at 30 °C. The cells were pelleted by centrifugation and stored at −80 °C. The selenomethionine-substituted P domain of Lon was produced using the saturation of the methionine biosynthetic pathway protocol (36).

Cell paste was suspended in 25 mM HEPES, pH 8, containing 100 mM NaCl, 25 mM imidazole (buffer B), and Complete EDTA-free protease inhibitor mixture (Roche Applied Science). The cells were disrupted with an APV Gaulin Model G1000 homogenizer at 10,000 p.s.i. The homogenate was centrifuged at 20,000 × *g* for 30 min, and the supernatant was applied to a 10-ml nickel-nitrilotriacetic acid Superflow column (Qiagen) equilibrated in buffer B. The column was washed and eluted with a 100-ml linear gradient of imidazole from 25 to 200 mM. Fractions containing recombinant fusion protein were pooled and incubated overnight at 4 °C with 4–5 mg of His_6-tagged tobacco etch virus protease. The digest was diluted 4-fold with 25 mM HEPES, pH 8, containing 100 mM NaCl and applied to a 25-ml nickel-nitrilotriacetic acid Superflow column equilibrated in buffer B. The column effluent containing the Lon P domain was collected, mixed with an equal volume of 25 mM HEPES, pH 8, 1 mM DTT (buffer C), and applied to a 16/5 Mono Q HR column equilibrated in buffer C. The column was washed, and protein was eluted with a 300-ml linear gradient of NaCl from 0 to 1 M. Desired fractions were pooled, concentrated using YM10 membrane, and applied to a HiPrep 26/60 Sephacryl S-100 HR column equilibrated in 20 mM Tris-HCl, pH 7.5, containing 150 mM NaCl and 1 mM DTT. Fractions with the P domain were pooled, concentrated as above, flash-frozen in liquid nitrogen, and stored at −80 °C until use.

In all cases the final product was better than 95% pure on the basis of silver staining after SDS-PAGE. Protein concentration was estimated spectrophotometrically using a calculated molar extinction coefficient (37) of 9650 M^{-1} cm^{-1} at 280 nm. The molecular weights of the recombinant P domains of Lon were confirmed by electrospray mass spectrometry; selenomethionine substitution exceeded 99%.

Protein Crystallization—Only the inactive Lon-S679A P domain was successfully crystallized, either in its native form or as a Se-Met derivative. For crystallization, the native form was concentrated to 18 mg/ml, and the Se-Met form was concentrated to 23.7 mg/ml. Initial screening of crystallization conditions (38) was carried out by the hanging drop, vapor diffusion method (39), using the Hampton (Hampton Research, Laguna Niguel, CA) and Wizard (Emerald Biostructures, Bainbridge Island, WA) screens. Native P domain, whether proteolytically obtained or recombinant, crystallized under very similar conditions (2 M ammonium sulfate, 0.1 M PIPES, pH 6.5). Typical crystals grew to 0.2 × 0.1 × 0.1 mm in 5–10 days at room temperature. The crystals of the Se-Met derivative grew in 2 M ammonium sulfate, 0.1 M MES, pH 6.0, with the largest crystal growing to the size of 0.5 × 0.45 × 0.2 mm in 14 days at room temperature. Before flash freezing, the crystals were transferred into a cryoprotectant solution, consisting of 80% mother liquor and 20% ethylene glycol for the native and 100% paratone-N for the Se-Met derivative.

[1] M. R. Maurizi, F. Rasulova, G. Leffers, R. Leapman, and A. C. Steven, unpublished results.

[2] The abbreviations used are: DTT, dithiothreitol; Se-Met, selenomethionine; PIPES, 1,4-piperazinediethanesulfonic acid; MES, 4-morpholineethanesulfonic acid; r.m.s., root mean square; SPase, signal peptidase.

TABLE I
Statistics of data collection and structure refinement
FOM, figure of merit; FC, calculated structure factors; FP, experimental structure factors; CC, correlation coefficient.

	Native	Se-Met
Data collection		
Space group	$P3_1$	
Molecules/a.u.	6	
Wavelength (Å)	1.54	0.97915
Unit cell parameters (Å)	$a = b = 86.21$, $c = 122.68$	$a = b = 86.37$, $c = 124.16$
Resolution (Å)	30–2.1	20–1.75
Total reflections	127,371	460,655
Unique reflections	58,314	105,429
Completeness (%)[a]	98.0 (99.5)	100.0 (100.0)
Average I/σ	14.1 (1.1)	19.2 (2.1)
R_{merge} (%)[b]	5.6 (62.4)	7.5 (50.5)
Phasing statistics (20-1.75 Å)		
Number of Se sites		19
Mean FOM of phasing (SOLVE)		0.26
Overall FOM of phasing (RESOLVE)		0.47
R for FC *vs.* FP (%)		29.7
CC of recovered map with RESOLVE map		0.68
Correlation of noncrystallographic symmetry regions (%)		93
Refinement statistics		
R (%)[c]	24.3	19.7
R_{free} (%)[d]	29.6	26.2
Root mean square deviation bond lengths (Å)	0.006	0.021
Root mean square deviation angles (degrees)	1.3	3.4
Temperature factor (protein, Å^3)	47.8	19.5
Temperature factor (solvent, Å^3)	41.4	41.9
Number of protein atoms	8461	8212
Number of solvent molecules	259	792

[a] The values in parentheses relate to the highest resolution shell.
[b] $R_{merge} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the observed intensity, and $\langle I\rangle$ is the average intensity obtained from multiple observations of symmetry-related reflections after rejections.
[c] $R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.
[d] R_{free} defined in Ref. 47.

Crystallographic Procedures—X-ray data for the native Lon P domain were collected on a Mar345 detector, using a Rigaku rotating anode x-ray source with CuKα radiation focused by an MSC/Osmic mirror system. A series of putative heavy atom derivatives produced by soaking were tested without success. Se-Met derivative data were collected on beamline X9B at Brookhaven National Laboratory, National Synchrotron Light Source, on a Quantum4 ADSC detector. Native data were processed using program DENZO and scaled using program SCALEPACK, whereas derivative data collected at the selenium peak wavelength was processed using the HKL2000 package (40) (Table I). Single-wavelength anomalous dispersion phasing was carried out using programs SOLVE and RESOLVE (41). RESOLVE used density averaging based on the 6-fold noncrystallographic symmetry and yielded an excellent map and a partial model with 750 residues. The remainder of the model was built manually into the initial map using program O (42). Initial rigid body refinement with CNS (43), using a maximum-likelihood target, was followed by simulated annealing (44) with Engh and Huber parameters (45). The model was rebuilt into density using program O. Subsequent refinement of the Se-Met structure was carried out with SHELXL (46) modeling in water and alternate conformations for side chains. TLS refinement was carried out with REFMAC5 (CCP4 suite) to an R factor of 19.7% and R_{free} (47) of 26.2%. The TLS parameters describe the thermal motion of a rigid body in terms of translation and libration of the group. In simple terms, the T tensor describes the mean square translation, the L tensor describes the mean square libration, and S describes the cross-correlation, respectively (48). The refinement of the native structure was completed with CNS to an R factor of 24.3% and R_{free} of 29.6%. The coordinates and structure factors have been submitted to the Protein Data Bank with the accession codes 1rr9 for the native and 1rre for the Se-Met derivative, respectively.

Fold Assignment—The preliminary coordinates of molecule A of the catalytic domain of Lon were submitted to the DALI server (49) to search for other proteins with a similar fold. The highest Z score was 5.3 with r.m.s. deviation of 4.3 Å for 107 Cα atoms of phosphomevalonate kinase (Protein Data Bank code 1k47). This level of similarity does not support any evolutionary relationship, especially with only 14% sequence identity. The similarity to any proteolytic enzymes is even lower, with the Z score for fibroblast collagenase of 2.6 with 2.9 Å r.m.s. deviation for only 76 Cα atoms. Thus, we can conclude that the fold of the catalytic domain of Lon is unique and that this enzyme is the first structurally characterized member of a novel family of proteases.

RESULTS AND DISCUSSION

Structure Solution and Refinement—The crystal structure of the Lon-S679A P domain was solved using single-wavelength anomalous data for a Se-Met variant. The experimental map based on the anomalous signal, enhanced by solvent flattening and noncrystallographic symmetry averaging, was quite clear and enabled unambiguous tracing of all but a few terminal residues in each of six molecules in the asymmetric unit. This structure was refined at 1.75 Å resolution in parallel with the native protein data extending to 2.1 Å. The two structures were very similar (r.m.s. deviation 0.43 Å for Cα atoms), and only the former one will be discussed here in detail. Multiple conformations were evident for several side chains even in the original experimental map, in particular for methionines and cysteines. In the course of the refinement, it was crucial to model the alternate conformations for the heavy scatterers (selenium and sulfur atoms) to substantially reduce the phasing errors. The absence of density for 18 terminal residues/molecule (9 from either end) accounts for the ~10% of the structure missing in our final model and contributes to the slightly elevated R factors. To further alleviate this problem, the 1.75 Å structure was refined with the TLS routine in program REFMAC5, reducing the free R factor by a further ~2%.

Earlier efforts to solve the crystal structure were hampered by twinning of the vast majority of the crystals. Most diffraction data, including, for example, data obtained with a mercury derivative, could be scaled equally well in both a trigonal ($P3_1$) and hexagonal ($P6_2$) space group, with a clear indication that the latter was due to twinning. The UCLA twinning server (50) identified a 50:50 twinning ratio of the twin components, also supported by the analysis of statistics in the CCP4 program TRUNCATE (51). The same phenomenon was later observed with data collected from several crystals of the Se-Met variant, and phasing efforts were successful only after we identified the one nontwinned crystal of the Se-Met derivative.

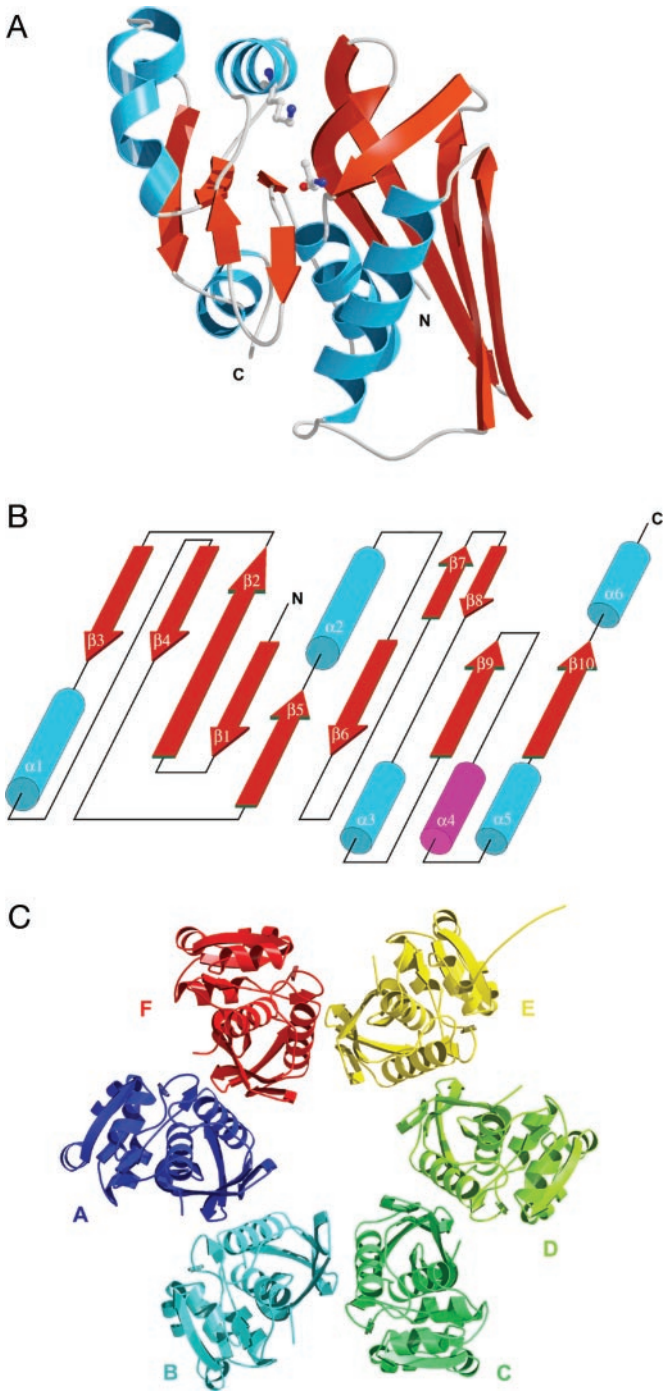

FIG. 1. **Structure of *E. coli* Lon P domain.** *A*, ribbon diagram with the catalytic dyad residues highlighted in ball and stick. *B*, topology of the P domain of Lon. The 3_{10}-helix is marked in *light purple*. *C*, the monomers A–F form a dome-shaped hexamer with a pore through the middle. Molecule E has all of the C-terminal residues. The figures were generated with Bobscript (67) and Raster3D (68).

Description of the Structure—The structure of the P domain contains a novel fold, consisting of six α helices and ten β strands (Fig. 1*A*). Six such domains (molecules A–F) form a ring-shaped hexamer, but the approximate 6-fold symmetry is not crystallographic. Although the P domain obtained by either limited proteolysis or by expression of the recombinant construct contained residues 585–784 of the full-length Lon, the first nine residues on the N terminus were not visible in the electron density maps. The traced structures of five of the monomers are composed of residues 594–775, whereas the sixth monomer contains residues 594–784. We will refer to the ring surface closest to the N terminus as the proximal surface and the opposite surface as the distal surface.

The N-terminal β strand 1 (residues 594–605) and an antiparallel β strand 2 (607–619) form a long β-hairpin loop. This loop and parallel β strands 3 (621–628) and 4 (661–670), which are separated by helix 1 (632–649), form the first large β sheet (Fig. 1*B*), which lies in a plane aligned with the 6-fold axis. Strand 5 (673–678) is perpendicular to the plane of this β sheet and is connected to helix 2 (681–693) by a loop that contains the catalytic Ser^{679} (here replaced by alanine). Helices 1 and 2 interact with the first β sheet, forming a relatively compact subdomain. On the distal surface of the domain, strand 5 forms a second, small β sheet, forming two hydrogen bonds with the tip of the β-hairpin loop (residue 603) and another two hydrogen bonds with the segment connecting strand 3 and helix 1 (residues 629–631). This small β sheet is on the surface of a shallow concavity in the center of the distal ring surface. A disulfide bridge between Cys^{617} and Cys^{691} connects the end of helix 2 to the end of β strand 2, stabilizing the subdomain. This unusual surface-exposed disulfide bond is unique to Lon proteases from closely related enteric bacteria.

Following helix 2, a random coil (695–699) forms a bridge to the second subdomain. A short β strand 6 (700–705) leads into another β loop (706–714) formed by antiparallel strands 7 and 8, followed by helix 3 (719–730). Helix 3 runs nearly parallel along the distal surface and carries the second catalytic residue, Lys^{722}. Strand 9 (732–737) returns along helix 3, followed by a short 3_{10} helix 4 (738–747), α helix 5 (748–755), and then parallel strand 10 (756–761). Strands 6, 9, and 10 form a third small β sheet, sandwiched by helix 3 and C-terminal helix 6. Most molecules in the structure terminate after helix 6 (762–772), which runs along the outside of the ring about one-third below and parallel to the proximal surface. The last nine residues are visible in only one molecule, designated molecule E. These residues adopt an extended conformation, reaching to the active site of molecule A′ of a symmetry-related hexamer (because of crystallographic symmetry designations E-A′ and E′-A are equivalent).

Oligomeric Structure of the Lon P Domain—The six molecules in the crystallographic asymmetric unit are related by a 6-fold noncrystallographic symmetry axis (Fig. 1*C*). Individual molecules relate to each other by the same noncrystallographic symmetry matrix, but superposition of their Cα traces yields a range of r.m.s. deviations between 0.42 and 0.82 Å. Thus, the crystal structure reveals a hexamer in which individual monomers form a ring (Fig. 2*A*). Viewed from the side, the ring is dome-shaped, with a diameter of ~100 Å at the base and ~50 Å at the top. The total buried surface area of the six monomers upon hexamerization is 9,169 $Å^2$. The subunit interface is characterized by mostly hydrophilic interactions, in which a number of structural water molecules are also recruited. Helix 1 packs against β sheet 1 of the adjacent subunit forming a major interface with a number of hydrogen bonds. In addition, Glu^{660} in the loop following helix 1 makes a salt bridge with the adjacent Lys^{621}, located in the loop between strands 2 and 3 of the β sheet. The proximal surface of the hexamer is slightly positively charged, and the distal surface, particularly within the concave depression, is negatively charged (Fig. 2*A*).

A solvent-accessible central pore ~32 Å long runs through the hexamer. It has a diameter of ~18 Å at the entrance from the proximal surface and widens slightly toward the distal end. The pore entrance from the proximal surface is similar in size to the entry channel in *E. coli* ClpP (28) and can be expected to accommodate a folded α-helix or β strand. Side chains lining the proximal entrance of the pore come from Lys^{621}, Lys^{623}, Glu^{660}, and Tyr^{659}. A number of negatively charged and polar

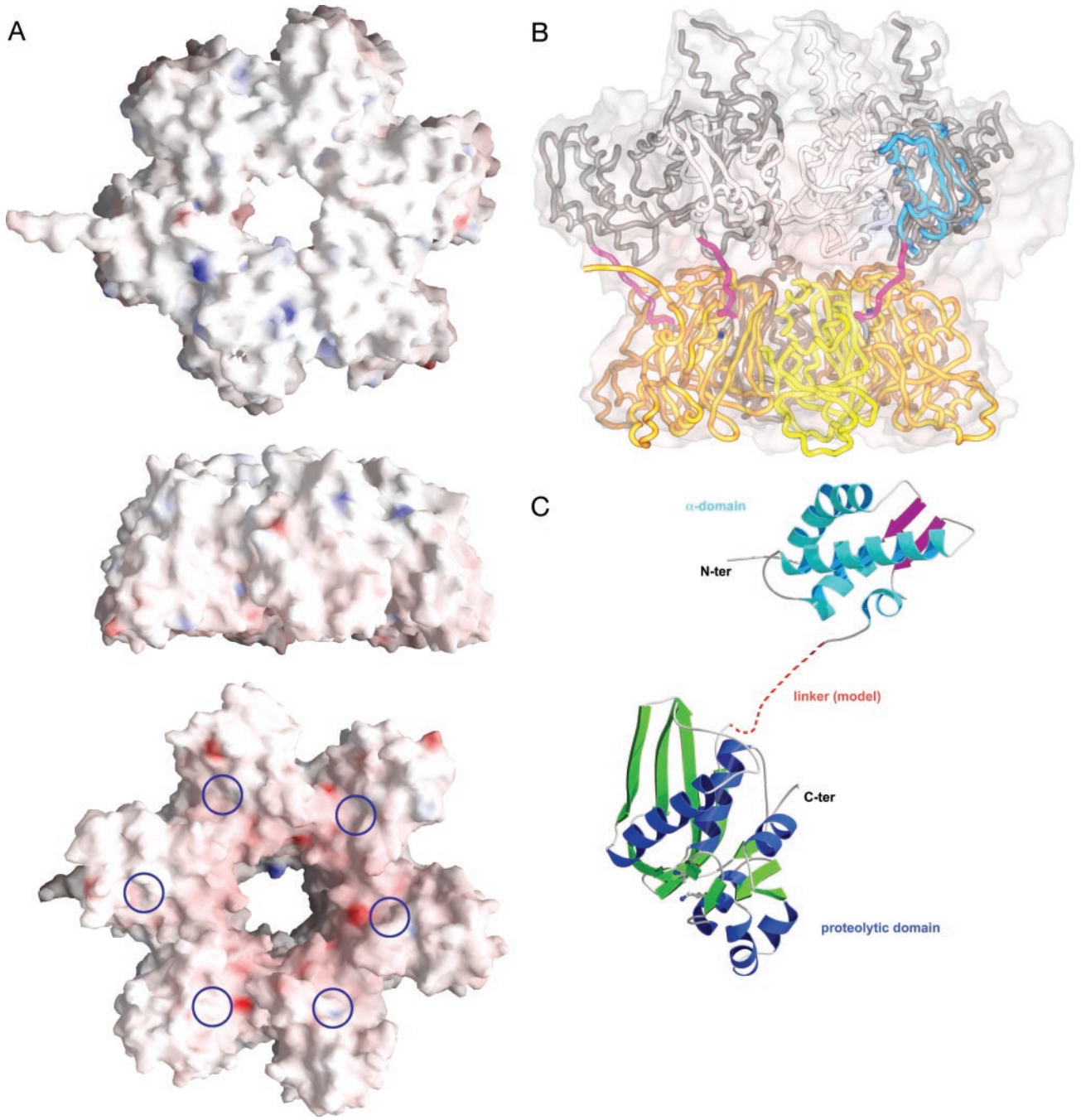

FIG. 2. **Molecular surface of the Lon P domain.** *A*, the hexamer has been rendered to show the electrostatic surface potential of the proximal surface expected to form the interface with the ATPase domain (*top panel*), the lateral edge of the ring (*middle panel*), and the distal surface with the positions of the active sites circled (*bottom panel*). *B*, *E. coli* Lon α domain from the ATPase functional domain (27) (*cyan*) and P domain (shades of *yellow*) superimposed on the corresponding domains in *Hemophilus influenzae* HslUV (shades of *gray*; Protein Data Bank code 1kyi). The transparent molecular surface is of the HslUV complex. The α domain superposition was structure-based, whereas the hexameric ring of the P domain was fitted into the molecular surface of the HslV. The rotation of the P domain about the 6-fold axis is arbitrary. The N-terminal residues of the P domain structure (Arg^{594}) are marked in *dark blue*. The HslU C-terminal region mimics the connecting 10 residues from the *E. coli* Lon C terminus of the α domain to the N terminus of the P domain (*magenta*). *C*, model of the linker region connecting the α domain to the P domain. The figures were generated with SPOCK (69) and GRASP (70).

residues (Glu^{632}, Glu^{636}, Gln^{635}, and Tyr^{626}) are located toward the distal end of the pore, making the distal end rather negatively charged, whereas the proximal half of the pore has several positively charged residues, which might serve a gating function for substrate entry.

The 6-fold axis of each hexameric ring, running through the central pore, is tilted 7° relative to the crystallographic 3_1 axis, with the result that the hexagonal rings within the lattice are not parallel. The tilt appears to be caused by an interaction between molecule A and the ordered C-terminal peptide in molecule E′ (residues 774–784) of a symmetry-related hexamer (Fig. 3*A*). The electron density for the extended C-terminal peptide of molecule E′ is very well defined, including the side chain of the C-terminal Lys^{784}, which is involved in a strong ion-pair interaction with the catalytic residue Lys^{722} in the active site of molecule A. The active sites of the other five molecules contain tightly bound sulfate ions in the corresponding locations, with an oxygen atom participating in similar interactions with Lys^{722} (Fig. 3*A*). This ion is possibly derived from the ammonium sulfate included in the crystallization

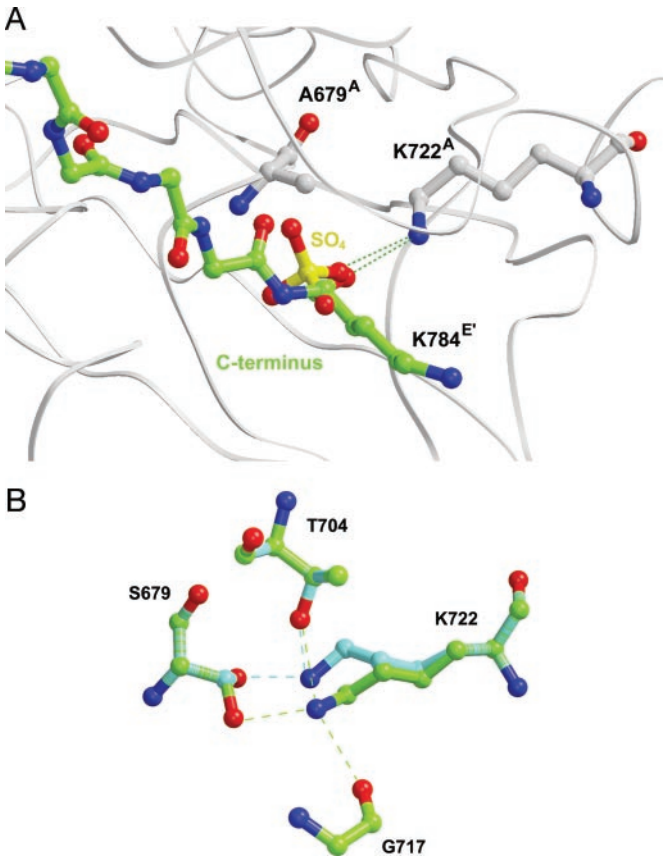

FIG. 3. **Active site of *E. coli* Lon protease.** *A*, the C terminus of a symmetry-related molecule E′ (*green*) points into the active site of molecule A. The carboxyl oxygen is hydrogen-bonded to Lys^{722} Nζ. A sulfate molecule (*yellow*) bound in the active sites of molecules B-F also forms a hydrogen bond with Lys^{722} Nζ. *B*, alternative models of the active enzyme. In the first model (*green*) Lys^{722} hydrogen bonds to Ser^{679}, Thr^{704}, and Gly^{717}, while in the second model (*cyan*) Lys^{722} hydrogen bonds to Ser^{679} and Thr^{704}. See text for further discussion.

solution or might represent part of a MES buffer molecule, because there are patches of weak electron density extending from it. These interactions might affect the position of the Nζ atom of Lys^{722}, but it is unlikely that the orientation of the hydrophobic part of this side chain could be changed.

Modeling of Connectivity between the P and α Domains of Lon—In other ATP-dependent proteases like HslUV, the hexameric ring of the ATPase domains is positioned on the proximal surface of the hexameric ring of the proteolytic domains (Fig. 2*B*). The N-terminal residue of the P domain used for crystallization is located within the sequence of intact Lon immediately following the C-terminal residue of the small α domain of the AAA^+ module. We could therefore predict the orientation of the P domain relative to the ATPase domain in Lon from the positions of the N-terminal residue in the P domain structure determined here and the C-terminal residue in the α domain structure solved recently (27). The α domains can be placed in a hexameric arrangement by analogy with the structure of HslUV (Fig. 2*B*), which in turn conforms to the topological arrangement observed in all AAA and AAA^+ modules crystallized to date. Without major structural rearrangement, the proximal surface of the P domain hexamer must be oriented toward the surface of the ring formed by the ATPase domain. Although nine N-terminal residues of the P domain were not visible in the structure presented here, they are sufficient to fill completely the missing part of the first β strand, which then terminates at the proximal surface, leaving only a short loop required to connect to the C-terminal helix of the α domain (Fig. 2*C*). This arrangement puts the proteolytic active site toward the distal surface at the end of the axial pore, as is seen in both ClpP and the proteolytic subunits of the proteasome. The linker region joins the P domain at β sheet 1, which plays a critical role in forming the subunit interface. Interestingly, the C-terminal peptide of the α domain of HslU also penetrates into the intersubunit space in HslV in the assembled holoenzyme. This similarity in the sites where the ATPase is linked to the protease suggests that conformational changes in the AAA^+ module may have an effect on subunit interfaces within the P domain. It has been shown that the C-terminal peptides of HslU alone are sufficient for allosteric activation of HslV (52, 53).

Binding of Polypeptide Substrates—The catalytic residues of individual molecules within the Lon P domain hexamer are spaced ~31 Å apart, similar to the distances between catalytic centers in hexameric ClpP (28). However, unlike in ClpP, there is no continuous hydrophobic groove on the surface of Lon that would form a continuous substrate-binding surface connecting the active sites. Instead, on the distal surface of the ring, there is a shallow hydrophobic groove lined by four negatively charged residues leading up to the catalytic serine. Beyond the catalytic center, the groove bends 90° and forms a cleft that extends to the edge of the ring. An unfolded substrate emerging from the pore could be threaded through this groove, presented for cleavage in the active site, and released through the cleft. It cannot be ruled out that peptides also bind within the outer cleft on the way into the active site, a substrate binding mode that could be used by Lon to cleave peptides in the absence of nucleotide.

Modeling of the Active Site of the Wild Type Enzyme—Although the catalytic Ser^{679} was replaced by an alanine in the structure reported here, structural analysis of other serine proteases indicates that this substitution is unlikely to alter the conformation of this residue in a significant way. In the experimental structure, Lys^{722} makes a hydrogen bond to the carbonyl oxygen of Gly^{717} and also interacts with either a bound sulfate or the C terminus of another monomer. The position of the hydroxyl group of Ser^{679} can only be modeled by postulating its most likely interactions. Several models that create a catalytic dyad are possible. Assuming that Lys^{722} interacts not only with Ser^{679} but also with other adjacent polar groups in the enzyme, the only likely candidates are Oγ1 of Thr^{704} and the carbonyl O of Gly^{717}. One model (Fig. 3*B*) can be obtained by converting Ala^{679} to a serine and setting the χ1 torsion angle to about −60°. In this model the side chain of Lys^{722} needs only a minor adjustment of the χ4 torsion angle to 148°, which brings the Nζ atom of Lys^{722} to within ~3 Å of Oγ1 of Thr^{704} and O of Gly^{717}. The resulting distance between Ser^{679} Oγ and Lys^{722} Nζ is ~2.8 Å. In a second model (Fig. 3*B*), the χ1 torsion angle of Ser^{679} is set to about 70°, and more extensive conformational rearrangement of the side chain of Lys^{722} allows this residue to form hydrogen bonds with Oγ of Ser^{679} and Oγ1 of Thr^{704}. In this orientation Lys^{722} no longer interacts with O of Gly^{717}. These models can be utilized for the analysis of the stereochemistry of the catalytic reaction, as discussed below.

Comparison of the Lon P Domain to Other Proteases That Utilize a Ser-Lys Dyad—The structure of the Lon P domain was compared with the structures of four other proteins that utilize a serine-lysine catalytic dyad mechanism. These enzymes are: *E. coli* type 1 signal peptidase (SPase) (Protein Data Bank codes 1kn9 and 1b12) (54), LexA (Protein Data Bank codes 1jhc and 1jhe) (55), UmuD′ (Protein Data Bank code 1umu) (56), and λ cI protein (Protein Data Bank code 1f39) (57). Structural comparison of these proteins revealed the overall similarity of

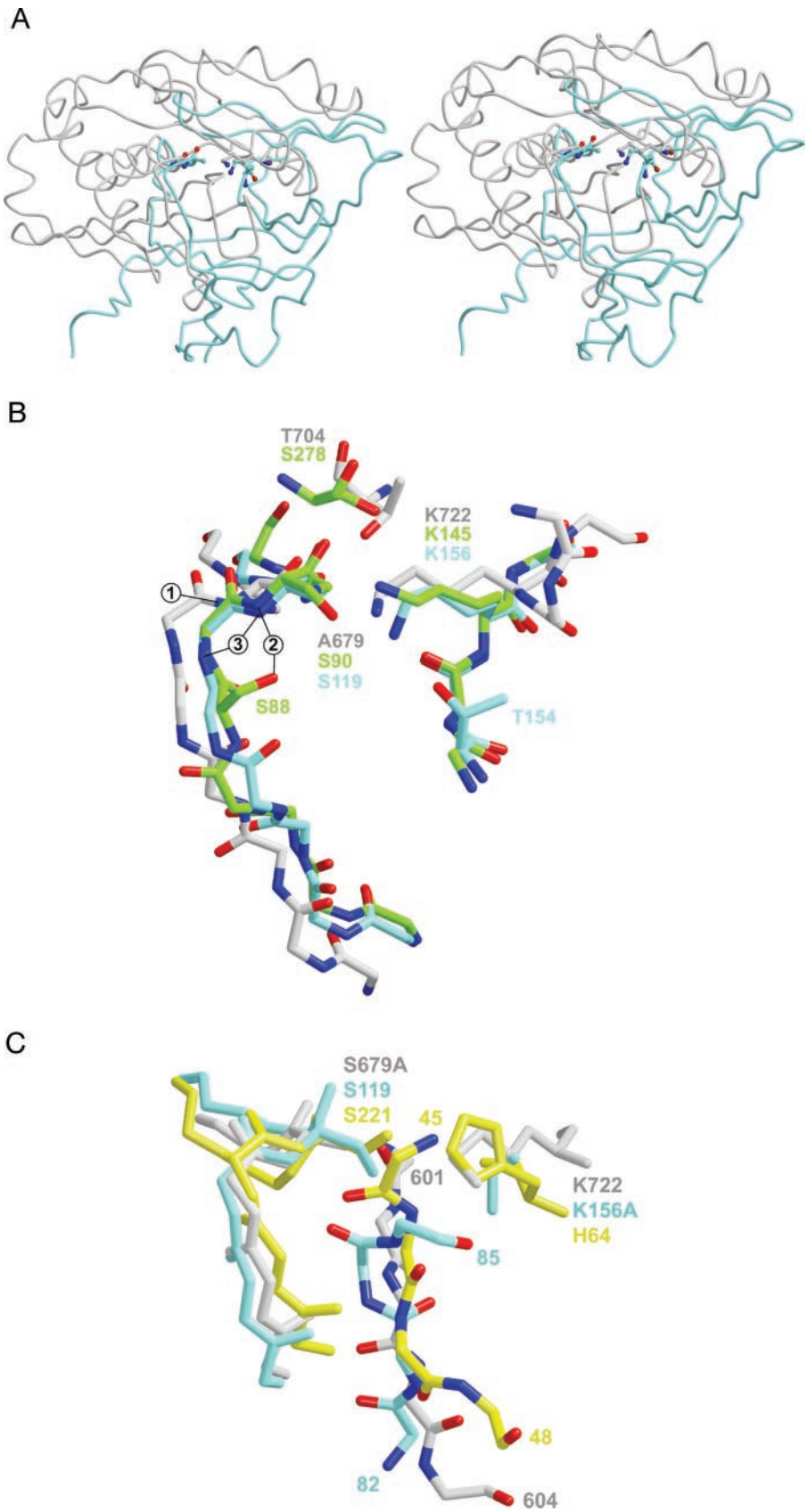

FIG. 4. **The active site region of Lon compared with that of other serine proteases.** *A*, stereo diagram of Lon P domain (*gray*) superimposed on LexA (*cyan*; Protein Data Bank code 1jhc). The superposition is based on the active sites, and the active site Ser and Lys are marked in ball-and-stick. *B*, comparison of active site residues from Lon P domain (*gray*), bacterial SPase (*green*; Protein Data Bank code 1kn9), and LexA (*cyan*; Protein Data Bank code 1jhc). The side chains of the catalytic residues are shown. The *circled numbers* mark the possible atoms of the oxyanion hole of Lon (*circled 1*), and the corresponding atoms in bacterial SPase (*circled 2*) and LexA (*circled 3*). *C*, active site superposition of subtilisin complexed with eglin (*yellow*; Protein Data Bank code 1cse), LexA (*cyan*; Protein Data Bank code 1jhe), and Lon (*gray*). The two central superimposed strands indicate the opposite direction of the main chain in eglin and the strand of the self-cleavable loop in LexA. This correlates with the *re*- or *si*-face attack of the amide bond by the catalytic serine. The third central strand belongs to the N-terminal loop of the Lon P domain.

their folds within the catalytic cores. The catalytic domain of SPase was superimposed on the other three structures with an r.m.s. deviation ~1.5 Å (54). Despite the fact that the P domain of Lon has a completely different overall fold (Fig. 4*A*), the stretch of residues that includes strand 5 with the catalytic Ser^{679} and the beginning of helix 2 could be superimposed quite well onto the corresponding segments in the four structures listed above. This alignment also brought the catalytic lysines into excellent superposition (Fig. 4*B*).

Comparative analysis of the active sites of these enzymes

identified a third residue that might assist the Ser-Lys dyad during catalysis. This residue is either a serine or a threonine with the side chain Oγ within hydrogen bonding distance of the catalytic lysine. The role of such a residue might be similar to the aspartate present in the classic catalytic triad of serine proteases (58). In the structures of LexA, UmuD′, and cI protein, this threonine is conserved and located at position −2 relative to the catalytic lysine (for example, Thr^{154} and Lys^{156} in LexA). In SPase the same role is assigned to Ser^{278}, which is located in a distinctly different position (59). We propose a similar role for Thr^{704} of Lon, which is highly conserved in the Lon family (33). The side chain of Thr^{704} is close to Lys^{722} (Fig. 3*B*) and in the vicinity of Ser^{278} in SPase (Fig. 4*B*). Although the structure of the mutant P domain does not show a hydrogen bond between Oγ of Thr^{704} and Nζ of Lys^{722}, the distance between these two atoms is ~3.5 Å, and we hypothesize that the presence of Ser^{679} in the active site of the wild type enzyme allows the postulated hydrogen bond network to be created (Fig. 3*B*).

Another critical structural feature, the "oxyanion hole," stabilizes the formation of a tetrahedral intermediate during catalysis in serine proteases (60). In the structure of a LexA mutant (Protein Data Bank code 1jhe), the oxyanion hole was identified on the basis of the hydrogen bonded interactions of the carbonyl oxygen of Ala^{84} (the scissile bond Ala^{84}-Gly^{85} is one of the two sites that are involved in self-cleavage in LexA) (55). It was shown to be formed by two main chain amide nitrogens: one originating from the catalytic serine and the other from the preceding residue (Fig. 4*B*). In three other structures of LexA in which the loop containing the scissile bond Ala^{84}-Gly^{85} is not in the conformation that would allow self-cleavage, the position of Ala^{84} O is occupied by a water molecule. An equivalent water is found in the structures of the Lon P domain, the λ cI protein, and UmuD′. In these structures, the water makes an additional hydrogen bond to the carbonyl oxygen of the residue at position −3 relative to the catalytic serine (Asp^{676} in the case of Lon). SPase lacks the conserved water molecule, and its oxyanion hole has a slightly different arrangement. It is formed by the backbone amide nitrogen of the catalytic serine and the side chain hydroxyl of Ser^{88} (Fig. 4*B*). Considering the similarity of the active site configurations of the P domain and the other enzymes, it seems logical to expect that the oxyanion hole in Lon is also formed by the amide nitrogens of Ser^{679} and the preceding residue, Pro^{678}. However, because proline has not been shown to participate in the formation of oxyanion holes in other enzymes, its contribution is rather questionable. The only other candidate would be the amide nitrogen of Trp^{603}, which is present in this vicinity in the Lon P domain.

Substrate Binding Differences between Lon and Serine Proteases with the Ser-His(Glu)-Asp Triad—The structure of the Lon P domain was compared with those of subtilisin (Protein Data Bank code 1cse) (61), chymotrypsin (Protein Data Bank code 1acb) (62), and sedolisin (Protein Data Bank code 1ga4) (63, 64), representative serine proteases that utilize catalytic triads. As in Lon, the catalytic serine in subtilisin and sedolisin is located between the end of a strand and the beginning of a helix, and the polypeptide region adjacent to the serine superimposes well onto the corresponding segment in the Ser-Lys dyad family. The catalytic His^{64} in subtilisin and its functionally equivalent Glu^{80} in sedolisin superimpose on the catalytic Lys^{722} in Lon (Fig. 4*C*). In chymotrypsin, the main chain in the vicinity of catalytic Ser^{195} has a different conformation. Ser^{195} is located within a helical turn, and therefore the preceding strand, which was superimposed in the structures compared above, is not present. Nevertheless, active site residues 193–195, which include the catalytic Ser^{195} and amide nitrogens 193 and 195, forming the oxyanion hole in this enzyme, can be superimposed quite well on the corresponding residues in Lon P domain (677–679), placing the catalytic His^{57} and Lys^{722} in reasonably close proximity (not shown). Substrate binding to subtilisin and chymotrypsin has been modeled based on crystal structures of these enzymes with bound inhibitor, eglin. Residues 43–48 in eglin map the S3-S3′ substrate-binding sites in these enzymes (Fig. 4*C*). In contrast, the structure of LexA revealed that the strand containing the scissile bond 84–85 runs in the opposite direction to the corresponding strand of eglin in the subtilisin complex (Fig. 4*C*), and thus it is clear that the main chain direction of the bound substrates would also be opposite. The hydroxyl moieties of the catalytic serines in these two enzymes also point in opposite directions.

These differences between the two families of serine proteases are probably connected to the preferences for the *re*-face *versus si*-face types of nucleophilic attack on the amide bonds of their substrates. Whereas the proteases utilizing the catalytic triad mechanism were shown to attack their substrates from the *re*-face of the amide bond (65), it seems that the enzymes with the catalytic dyad in the active site prefer a *si*-face attack. That was indeed shown to be the case for bacterial SPase in complex with a β-lactam inhibitor (66). The orientation of the hydroxyl of the catalytic serine in SPase is the same as in LexA, shown in Fig. 4 (*B* and *C*), and the direction of the modeled substrate corresponds to that of the strand containing the scissile bond 84–85 of LexA.

As described above, two likely conformations of Ser^{679}-Lys^{722} dyad can be modeled (Fig. 3*B*). At this stage we cannot rule out the possibility of the *re*-face attack because the modeling also allowed the interactions for the catalytic dyad to be maintained with the negative χ1 value for Ser^{679} (Fig. 3*B*). However, by analogy to SPase and LexA, we predict that Lon will attack its substrate from the *si*-face of the amide bond (corresponding to the conformation of the side chain of Ser^{679} with a positive χ1 value) and that the direction of the main chain of the substrate will be most likely the same as in the strand 82–85 in LexA. Strand 1 of the N-terminal loop in Lon P domain (Fig. 1*B*) runs in a parallel fashion to the strand that contains residues 45–48 in eglin and in an opposite direction to the corresponding strand in LexA (Fig. 4*C*). We suggest that this strand might participate in substrate binding in Lon, forming an antiparallel β sheet with the N-terminal half of the substrate. Some movement of the N-terminal loop in Lon will be required to accommodate the substrate in this orientation, and we note that this loop is directly connected to the α domain of the AAA^+ module, which is expected to undergo some conformational rearrangement in response to nucleotide binding and hydrolysis. An experimental verification of these models will require a structure of the wild type enzyme complexed with an appropriate peptide ligand.

Acknowledgments—We thank the Structural Biophysics Laboratory Biophysics Resource (National Cancer Institute at Frederick) for the access to the electrospray mass spectrometer.

REFERENCES

1. Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) *Science* **286,** 1888–1893
2. Gottesman, S., Wickner, S., and Maurizi, M. R. (1997) *Genes Dev.* **11,** 815–823
3. Goldberg, A. L. (1992) *Eur. J. Biochem.* **203,** 9–23
4. Gottesman, S., and Maurizi, M. R. (1992) *Microbiol. Rev.* **56,** 592–621
5. Swamy, K. H., and Goldberg, A. L. (1981) *Nature* **292,** 652–654
6. Goldberg, A. L., Swamy, K. H., Chung, C. H., and Larimore, F. S. (1981) *Methods Enzymol.* **80,** 680–702
7. Charette, M. F., Henderson, G. W., and Markovitz, A. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 4728–4732
8. Goldberg, A. L., Moerschell, R. P., Chung, C. H., and Maurizi, M. R. (1994) *Methods Enzymol.* **244,** 350–375
9. Amerik, A. Y., Antonov, V. K., Ostroumova, N. I., Rotanova, T. V., and Chistyakova, L. G. (1990) *Bioorgan. Khim.* **16,** 869–880
10. Gottesman, S., Wickner, S., Jubete, Y., Singh, S. K., Kessel, M., and Maurizi,

M. (1995) *Cold Spring Harb. Symp. Quant. Biol.* **60,** 533–548
11. Rotanova, T. V. (1999) *Bioorgan. Khim.* **25,** 883–891
12. Rasulova, F. S., Dergousova, N. I., Starkova, N. N., Melnikov, E. E., Rumsh, L. D., Ginodman, L. M., and Rotanova, T. V. (1998) *FEBS Lett.* **432,** 179–181
13. van Dijl, J. M., Kutejova, E., Suda, K., Perecko, D., Schatz, G., and Suzuki, C. K. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 10584–10589
14. Ovchinnikova, T. V., Martynova, N. Yu., Potapenko, N. A., Vasilyeva, O. V., Golikova, N. Yu., Ginodman, L. M., and Rotanova, T. V. (1998) *Biomed. Tekhnol.* **8,** 62–65
15. Vasilyeva, O. V., Kolygo, K. B., Leonova, Y. F., Potapenko, N. A., and Ovchinnikova, T. V. (2002) *FEBS Lett.* **526,** 66–70
16. Roudiak, S. G., and Shrader, T. E. (1998) *Biochemistry* **37,** 11255–11263
17. Gottesman, S. (1996) *Annu. Rev. Genet.* **30,** 465–506
18. Frank, E. G., Ennis, D. G., Gonzalez, M., Levine, A. S., and Woodgate, R. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 10291–10296
19. Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. (1999) *Genome Res.* **9,** 27–43
20. Lupas, A. N., and Martin, J. (2002) *Curr. Opin. Struct. Biol.* **12,** 746–753
21. Maurizi, M. R., and Li, C. C. (2001) *EMBO Rep.* **2,** 980–985
22. Lupas, A., Flanagan, J. M., Tamura, T., and Baumeister, W. (1997) *Trends Biochem. Sci.* **22,** 399–404
23. Zwickl, P., Baumeister, W., and Steven, A. (2000) *Curr. Opin. Struct. Biol.* **10,** 242–250
24. Stahlberg, H., Kutejova, E., Suda, K., Wolpensinger, B., Lustig, A., Schatz, G., Engel, A., and Suzuki, C. K. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 6787–6790
25. Krzywda, S., Brzozowski, A. M., Verma, C., Karata, K., Ogura, T., and Wilkinson, A. J. (2002) *Structure* **10,** 1073–1083
26. Niwa, H., Tsuchiya, D., Makyio, H., Yoshida, M., and Morikawa, K. (2002) *Structure* **10,** 1415–1423
27. Botos, I., Melnikov, E. E., Cherry, S., Khalatova, A. G., Rasulova, F., Tropea, J., Maurizi, M. R., Rotanova, T. V., Gustchina, A., and Wlodawer, A. (2003) *J. Struct. Biol.*, in press
28. Wang, J., Hartling, J. A., and Flanagan, J. M. (1997) *Cell* **91,** 447–456
29. Yoo, S. J., Shim, Y. K., Seong, I. S., Seol, J. H., Kang, M. S., and Chung, C. H. (1997) *FEBS Lett.* **412,** 57–60
30. Tomoyasu, T., Gamer, J., Bukau, B., Kanemori, M., Mori, H., Rutman, A. J., Oppenheim, A. B., Yura, T., Yamanaka, K., Niki, H., and. (1995) *EMBO J.* **14,** 2551–2560
31. Amerik, A. Y., Antonov, V. K., Gorbalenya, A. E., Kotova, S. A., Rotanova, T. V., and Shimbarevich, E. V. (1991) *FEBS Lett.* **287,** 211–214
32. Starkova, N. N., Koroleva, E. P., Rumsh, L. D., Ginodman, L. M., and Rotanova, T. V. (1998) *FEBS Lett.* **422,** 218–220
33. Birghan, C., Mundt, E., and Gorbalenya, A. E. (2000) *EMBO J.* **19,** 114–123
34. Rotanova, T. V., Melnikov, E. E., and Tsirulnikov, K. B. (2003) *Bioorgan. Khim.* **29,** 97–99
35. Rotanova, T. V., Kotova, S. A., Amerik, A. Yu., Lykov, I. P., Ginodman, L. M., and Antonov, V. K. (1994) *Bioorgan. Khim.* **20,** 114–125
36. Doublie, S. (1997) *Methods Enzymol.* **276,** 523–530
37. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem.* **182,** 319–326
38. Jancarik, J., and Kim, S. H. (1991) *J. Appl. Crystallogr.* **21,** 916–924
39. McPherson, A. (1982) *Preparation and Analysis of Protein Crystals*, John Wiley & Sons, Inc., New York
40. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
41. Terwilliger, T. C. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57,** 1763–1775
42. Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol.* **277,** 173–208
43. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
44. Brünger, A. T., Krukowski, A., and Erickson, J. W. (1990) *Acta Crystallogr. Sect. A* **46,** 585–593
45. Engh, R., and Huber, R. (1991) *Acta Crystallogr. Sect. A* **47,** 392–400
46. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol.* **277,** 319–343
47. Brünger, A. T. (1992) *Nature* **355,** 472–474
48. Schomaker V., and Trueblood, K. N. (1968) *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* **24,** 63–76
49. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
50. Yeates, T. O. (1997) *Methods Enzymol.* **276,** 344–358
51. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
52. Ramachandran, R., Hartmann, C., Song, H. K., Huber, R., and Bochtler, M. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7396–7401
53. Kwon, A. R., Kessler, B. M., Overkleeft, H. S., and McKay, D. B. (2003) *J. Mol. Biol.* **330,** 185–195
54. Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (2002) *J. Biol. Chem.* **277,** 9512–9519
55. Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W., and Strynadka, N. C. (2001) *Cell* **106,** 585–594
56. Peat, T. S., Frank, E. G., McDonald, J. P., Levine, A. S., Woodgate, R., and Hendrickson, W. A. (1996) *Nature* **380,** 727–730
57. Bell, C. E., Frescura, P., Hochschild, A., and Lewis, M. (2000) *Cell* **101,** 801–811
58. Dodson, G., and Wlodawer, A. (1998) *Trends Biochem. Sci.* **23,** 347–352
59. Klenotic, P. A., Carlos, J. L., Samuelson, J. C., Schuenemann, T. A., Tschantz, W. R., Paetzel, M., Strynadka, N. C., and Dalbey, R. E. (2000) *J. Biol. Chem.* **275,** 6490–6498
60. Kraut, J. (1977) *Annu. Rev. Biochem.* **46,** 331–358
61. Bode, W., Papamokos, E., and Musil, D. (1987) *Eur. J. Biochem.* **166,** 673–692
62. Frigerio, F., Coda, A., Pugliese, L., Lionetti, C., Menegatti, E., Amiconi, G., Schnebli, H. P., Ascenzi, P., and Bolognesi, M. (1992) *J. Mol. Biol.* **225,** 107–123
63. Wlodawer, A., Li, M., Dauter, Z., Gustchina, A., Uchida, K., Oyama, H., Dunn, B. M., and Oda, K. (2001) *Nat. Struct. Biol.* **8,** 442–446
64. Wlodawer, A., Li, M., Gustchina, A., Oyama, H., Dunn, B. M., and Oda, K. (2003) *Acta Biochim. Polon.* **50,** 81–102
65. James, M. N. G. (1994) in *Proteolysis and Protein Turnover* (Bond, J. S., and Barrett, A. J., eds) pp. 1–8, Portland Press, Brookfield, VT
66. Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (1998) *Nature* **396,** 186–190
67. Esnouf, R. M. (1997) *J. Mol. Graph. Model.* **15,** 132–134
68. Meritt, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 869–873
69. Christopher, J. A. (1998) *SPOCK: The Structural Properties Observation and Calculation Kit*, Center for Macromolecular Design, Texas A&M University, College Station, TX
70. Nicholls, A. (1992) *GRASP: Graphical Representation and Analysis of Surface Properties*, Columbia University, New York, NY