

# Cleavage Site Selection within a Folded Substrate by the ATP-dependent Lon Protease\*<sup>§</sup>

Received for publication, March 14, 2005, and in revised form, May 2, 2005  
Published, JBC Papers in Press, May 3, 2005, DOI 10.1074/jbc.M502796200

Gabriela Ondrovičová<sup>‡§</sup>, Tong Liu<sup>§¶</sup>, Kamalendra Singh<sup>¶</sup>, Bin Tian<sup>¶</sup>, Hong Li<sup>¶</sup>,  
Oleksandr Gakh<sup>||\*\*</sup>, Dušan Perečko<sup>‡</sup>, Jiří Janata<sup>||</sup>, Zvi Granot<sup>‡‡</sup>, Joseph Orly<sup>‡‡</sup>, Eva Kutejová<sup>‡§§</sup>,  
and Carolyn K. Suzuki<sup>¶¶</sup>

From the <sup>‡</sup>Institute of Molecular Biology, Slovak Academy of Sciences, 845 51 Bratislava, Slovak Republic, <sup>¶</sup>University of Medicine and Dentistry of New Jersey (UMDNJ), New Jersey Medical School, Department of Biochemistry and Molecular Biology, Newark, New Jersey 07103, <sup>||</sup>Institute of Microbiology, Academy of Sciences of the Czech Republic, 142 20 Prague 4, Czech Republic, and the <sup>‡‡</sup>Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

**Mechanistic studies of ATP-dependent proteolysis demonstrate that substrate unfolding is a prerequisite for processive peptide bond hydrolysis. We show that mitochondrial Lon also degrades folded proteins and initiates substrate cleavage non-processively. Two mitochondrial substrates with known or homology-derived three-dimensional structures were used: the mitochondrial processing peptidase  $\alpha$ -subunit (MPP $\alpha$ ) and the steroidogenic acute regulatory protein (StAR). Peptides generated during a time course of Lon-mediated proteolysis were identified and mapped within the primary, secondary, and tertiary structure of the substrate. Initiating cleavages occurred preferentially between hydrophobic amino acids located within highly charged environments at the surface of the folded protein. Subsequent cleavages proceeded sequentially along the primary polypeptide sequence. We propose that Lon recognizes specific surface determinants or folds, initiates proteolysis at solvent-accessible sites, and generates unfolded polypeptides that are then processively degraded.**

A mechanistic understanding of energy-dependent proteolysis has developed primarily from studies analyzing the 26 S proteasome or the ClpP holoenzyme and substrates carrying degradation sequences or tags (1–3). In general, the 26 S proteasome degrades substrates that carry multiple ubiquitin moieties (4). Similarly, for bacterial ATP-dependent proteases, many substrates are targeted for degradation by discrete pep-

tide sequences or tags (5–7). The N-end rule established that the identity of the first amino acid of a polypeptide within the context of an unstructured amino terminus determines the half-life of a protein (8, 9). In addition, an 11-amino-acid SsrA tag may be added co-translationally to the carboxyl terminus of a nascent polypeptide when its translation is stalled on the ribosome resulting in proteolysis (7, 10). Other amino- or carboxyl-terminal sequences that are both necessary and sufficient for targeting substrates to bacterial ATP-dependent proteases have also been identified (11). Once a degradation sequence is recognized, the substrate is engaged at an unstructured or loosely folded region, after which it is unfolded and translocated to the proteolytic active site, where it is degraded processively, generally from the amino to the carboxyl terminus or vice versa (1, 11–21).

Mitochondrial ATP-dependent proteases are functionally similar to those in bacteria (22–24). In human mitochondria, homologs of bacterial Lon and ClpP reside within the matrix (25–27), and at least three FtsH homologs are localized to the inner membrane with their active sites located in either the matrix or the intermembrane space (28–30). Discrete degradation signals such as ubiquitin or SsrA have not been identified for substrates of mitochondrial ATP-dependent proteases, which are generally unassembled, misfolded, or oxidatively damaged proteins (31–35).

To investigate cleavage site selection by a mitochondrial ATP-dependent protease, we used purified mitochondrial Lon and two substrates, the  $\alpha$ -subunit of the mitochondrial processing peptidase (MPP $\alpha$ )<sup>1</sup> and the steroidogenic acute regulatory protein (StAR). MPP $\alpha$  assembles with MPP $\beta$  into a heterodimer that removes the amino-terminal targeting presequences of proteins imported into the mitochondrial matrix (36, 37). Previous work has shown that unassembled subunits of MPP are degraded by mitochondrial Lon *in vivo* (32). StAR is a monomeric protein responsible for translocating cholesterol to the mitochondrial inner membrane, which is the rate-limiting step in steroid hormone biosynthesis. StAR carries an amino-terminal mitochondrial targeting presequence and is composed of a single cholesterol-binding domain referred to as the StAR-related lipid transfer (START) domain (38, 39).

\* This study was supported by grants from the National Institutes of Health, the Basil O'Connor Scholars Award – March of Dimes, and the Foundation of UMDNJ (to C. K. S.); from the Slovak Grant Agency and the Science and Technology Assistance Agency (to E. K.); from the Twinning Program–National Science Foundation/National Academy of Sciences (to C. K. S. and E. K.); and from the Israel Science Foundation (to J. O.). 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.

<sup>§</sup> The on-line version of this article (available at <http://www.jbc.org>) contains eight supplemental figures.

<sup>¶</sup> Both authors contributed equally to this work

<sup>\*\*</sup> Present address: Depts. of Pediatric and Adolescent Medicine and Biochemistry and Molecular Biology, Mayo Clinic College of Medicine, Rochester, MN 55905.

<sup>§§</sup> To whom correspondence may be addressed. Tel.: 421-2-59307442; Fax: 421-2-59307416; E-mail: Eva.Kutejova@savba.sk.

<sup>¶¶</sup> To whom correspondence may be addressed. Tel.: 973-972-1555; Fax: 973-972-5594; E-mail: suzuki@umdnj.edu.

<sup>1</sup> The abbreviations used are: MPP, mitochondrial processing peptidase; MDH, malate dehydrogenase; MS, mass spectrometry; MS/MS, tandem MS; MALDI-TOF, matrix-assisted laser desorption/ionization-time-of-flight; TOF/TOF, tandem TOF; StAR, steroidogenic acute regulatory protein; START domain, StAR-related lipid transfer domain; NTA, nitrilotriacetic acid; AMP-PNP, 5'-adenylyl- $\beta$ , $\gamma$ -imidodiphosphate.

Cell-based assays show that StAR is degraded in a two-step process within the mitochondrial matrix (40, 41). *In vitro* results presented here suggest that mitochondrial Lon is one of the proteases responsible for StAR degradation.

To determine the initiation and progression of Lon-mediated proteolysis, peptides produced in a time course of degradation were identified by mass spectrometry. MPP $\alpha$  and StAR were used as substrates in this analysis; these purified proteins are folded, and their respective three-dimensional structures either are known or can be derived by homology-based modeling. Initial Lon cleavages generally occurred between hydrophobic amino acids positioned at internal sites within the primary sequence of the protein, without apparent preference for secondary structure. Initiating cleavages occurred within  $\alpha$  helices,  $\beta$ -sheets, or unstructured regions. Although the initiating Lon cleavage sites were hydrophobic, they were exposed at the surface of the protein and surrounded by strongly polar environments. Subsequent cleavages occurred processively in a linear manner along the resulting polypeptide chains. Taken together, these results provide insight into an alternate mechanism of ATP-dependent proteolysis.

#### EXPERIMENTAL PROCEDURES

**Purification of MPP $\alpha$  and MPP $\beta$** —Yeast MPP $\alpha$  and MPP $\beta$  were produced and purified recombinantly using previously described protocols (42). Briefly, *Escherichia coli* producing MPP $\alpha$  were resuspended in HN buffer (20 mM HEPES, 20 mM NaCl, pH 7.4), sonicated on ice, and centrifuged at  $10,000 \times g$  for 15 min at 4 °C. The supernatant was applied onto a DE-52 cellulose column (Whatman), washed with HN buffer, and eluted with a linear NaCl gradient (0.02–0.3 M NaCl, 20 mM HEPES, pH 7.4). Fractions were assayed for MPP $\alpha$  by immunoblotting and for contaminating ATPase activity using the previously described method of Lanzetta *et al.* (43). Those fractions containing MPP $\alpha$  and having low ATPase activity were combined and diluted four times with HN buffer and applied to a Mono Q HR5/5 column, washed with HN buffer, and eluted with a linear NaCl gradient (0.02–0.3 M NaCl, 20 mM HEPES, pH 7.4). MPP $\beta$  produced in *E. coli* was purified from inclusion bodies as described previously (42). Bacteria were resuspended in HN buffer, sonicated on ice, and centrifuged at  $10,000 \times g$  for 25 min at 4 °C. The pellet was washed with HN buffer and resuspended in 20 mM HEPES, pH 7.4, 0.6 M sorbitol, 7.5 M urea, 1% Tween 20. The sample was incubated at 30 °C for 40 min and then centrifuged at  $10,000 \times g$  for 10 min at 4 °C. To renature yeast MPP $\beta$ , the protein present in the supernatant was dialyzed extensively against HN buffer and centrifuged at  $10,000 \times g$  for 30 min. The activity of purified MPP $\alpha$  and MPP $\beta$  was determined by testing their ability to form an active heterodimeric complex that cleaved off the mitochondrial targeting presequence of the full-length malate dehydrogenase (MDH) protein. MDH (0.8  $\mu$ M) was incubated with MPP $\alpha$ , MPP $\beta$ , or MPP $\alpha$  combined with MPP $\beta$  (2  $\mu$ M) in a reaction mixture (20 mM HEPES, pH 7.5, 20 mM NaCl, 1 mM MnCl<sub>2</sub>) for 30 min at 30 °C and analyzed by SDS-PAGE.

**Purification of Mitochondrial Lon and StAR**—Yeast Lon carrying a carboxyl-terminal hexahistidine tag was overproduced and purified from isolated yeast mitochondria as described previously (35). Briefly, mitochondria (10 mg/ml) were suspended in solubilization buffer (20 mM HEPES, pH 8.0, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, 20% glycerol) containing 1 mM ATP, 0.1 mM *N*- $\alpha$ -tosyl-L-lysyl-chloromethyl ketone, 0.1 mM *N*-*p*-tosyl-L-phenylalanine chloromethyl ketone, and 1.6 mg of Lubrol/mg of mitochondrial protein. The lysate was centrifuged at  $200,000 \times g$  at 4 °C for 10 min, and the supernatant was applied to a Ni<sup>2+</sup>-NTA agarose column. The column was washed with 5 volumes of solubilization buffer containing 40 mM imidazole. Bound protein was eluted with solubilization buffer containing 0.2 M imidazole and concentrated on a Centricon-100 filter (Millipore). Human Lon was overproduced and purified from the Rosetta *E. coli* strain (Novagen) as described previously (44). Briefly, bacteria producing an amino-terminal hexahistidine-tagged form of Lon lacking its predicted mitochondrial targeting sequence were resuspended in Solution A (0.2 M NaCl, 25 mM Tris buffer, pH 7.5, 20% glycerol, 2 mM  $\beta$ -mercaptoethanol) and sonicated on ice. The lysate was centrifuged, and the supernatant was loaded onto a Ni<sup>2+</sup>-NTA column (Qiagen). The column was washed three times with NTA buffer (0.5 M NaCl, 25 mM Tris, pH 7.5, 20% glycerol) and five times with Solution A containing 40  $\mu$ M imidazole. Lon bound to the column was eluted in three steps with Solution A

containing 0.1, 0.2, or 0.3 M imidazole. Recombinant mouse StAR was produced and purified as a carboxyl-terminal hexahistidine-tagged form of the protein lacking its predicted mitochondrial targeting sequence (amino acids 1–47) in the Rosetta *E. coli* strain (Novagen), using the same protocol as described for human Lon.

**ATP-dependent Degradation of MPP or StAR by Lon**—Yeast MPP $\alpha$ , MPP $\beta$ , or MPP $\alpha\beta$  (0.8  $\mu$ M) was combined with human mitochondrial Lon (0.8  $\mu$ M monomer) containing 20 mM Tris, pH 7.9, 10 mM MgCl<sub>2</sub> in the presence or absence of 2 mM ATP on ice. An aliquot was taken representing time 0, and the remaining sample was incubated at 37 °C for the indicated time periods. The samples were subject to SDS-PAGE and Coomassie Brilliant Blue staining or to mass spectrometry.

Mouse StAR (3  $\mu$ M) and human Lon (0.8  $\mu$ M monomer) were incubated in a reaction containing 25 mM Tris, pH 7.9, and 10 mM MgCl<sub>2</sub> in the presence or absence of 2 mM ATP for the time periods as indicated at 37 °C. The reactions were analyzed by SDS-PAGE and Coomassie Brilliant Blue staining or by mass spectrometry.

**Mass Spectrometry (MS)**—Yeast MPP $\alpha$  (3  $\mu$ M) and human Lon (0.8  $\mu$ M) were incubated in a reaction containing 25 mM Tris, pH 7.9, 2.5 mM ATP, 10 mM MgCl<sub>2</sub> at 37 °C; aliquots were withdrawn at the time points indicated. The reaction was terminated by adding trifluoroacetic acid in acetonitrile (0.1% final concentration). StAR (3  $\mu$ M) and human Lon (0.3  $\mu$ M) were incubated in a reaction containing 25 mM Tris, pH 7.9, 10 mM MgCl<sub>2</sub>, and 2.5 mM ATP 37 °C; aliquots were withdrawn at the time points indicated. The reaction was terminated by adding trifluoroacetic acid in acetonitrile (0.1% final concentration). The resulting peptide products were purified and concentrated using a C<sub>18</sub> ZipTip (Millipore) following the manufacturer's directions for peptides. A mass error tolerance of 20 ppm was used to tentatively assign peptide identity in the MS mode. A more definitive assignment was made by MS/MS fragmentation; all major fragment ions were accounted for.

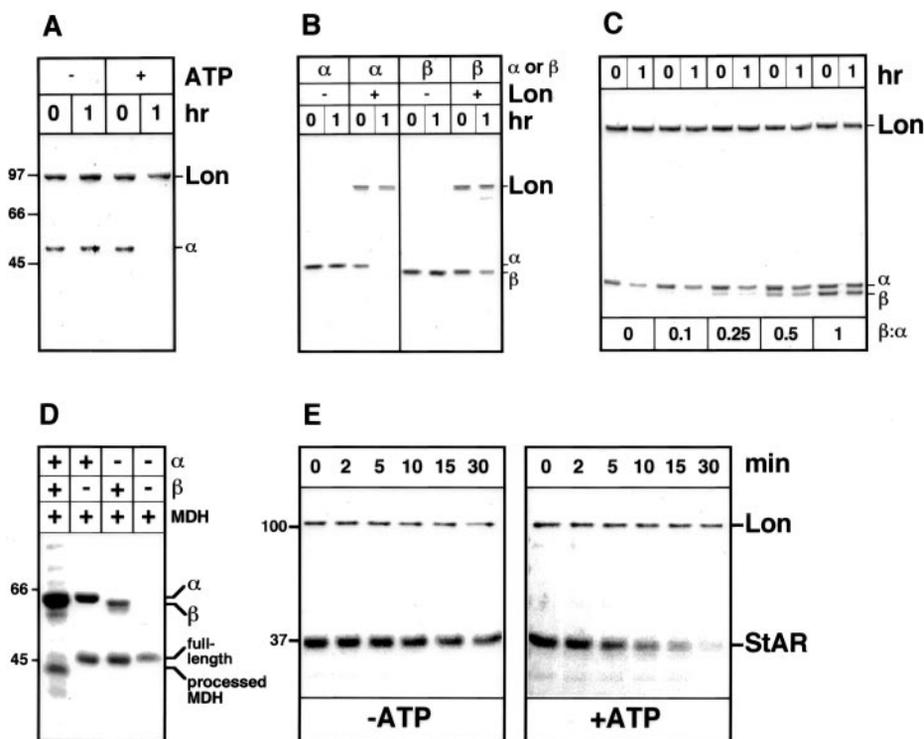
All MALDI-TOF mass spectra were acquired in positive ion reflector mode on a 4700 proteomics analyzer with TOF/TOF optics (Applied Biosystems, Framingham, MA). The recrystallized  $\alpha$ -cyano-4-hydroxycinnamic acid was used as matrix at a concentration of 5 mg/ml in 50% water/acetonitrile and 0.1% trifluoroacetic acid. To reduce the matrix adducts, 10 mM ammonium monobasic phosphate was included in the matrix. Matrix and sample were mixed in a 1:1 (v/v) ratio and spotted onto a stainless steel sample plate for MS and MS/MS sequencing. The MALDI mass spectra obtained were internally calibrated using angiotensin II and ACTH (fragments 18–39) as standards (peptide MALDI-MS calibration kit, Sigma). For fragment ion analysis in MS/MS mode, the collision energy was 1 keV, and the collision gas was air. Analysis of MS data were performed with ExpASY (www.expasy.org), and Data Explorer™ (Applied Biosystems, Framingham, MA) was used for fragment ion analysis.

**Bioinformatics of Substrate Cleavage Sites**—Substrate peptides were matched to the respective protein sequences using regular expression functions in PERL. A PERL script was also used to count amino acids surrounding the cleavage site. Hidden Markov model analysis of digested peptides and neighboring protein sequences was performed using the HMMER software. Heat maps were generated by the "image" function in the program R.

**Molecular Modeling**—The coordinates of backbone and side chain atoms of MPP were imported from the Protein Data Bank (1HR6, Research Collaboratory in Structural Bioinformatics (RCSB), Rutgers, NJ). The molecular modeling programs SYBYL (Tripos Associates, St. Louis, MO); MolMol (60); and MOE (Chemical Computing Group, Montreal, Canada) were used for display and surface property determinations. The secondary structure shown was defined in the RCSB coordinate file. The molecular surface of MPP $\alpha$  shows the Gaussian-Connelly surface with grid spacing of 0.75 and probe radius of 1.4 Å; the AMBER-like well depth parameters with AMBER force field and charges were used to solve the non-linear Poisson-Boltzman equation. The dielectric constant of the interior of protein was set to 2, whereas that of solvent was 80, and a salt concentration of 0.145 M was set to determine Debye radius. Calculations of the exposed surface area show that Met-398, Ser-138, Phe-312, and Leu-362 have 75, 64, 58, and 47% of their side chains exposed to the solvent, respectively.

The molecular model of StAR was constructed using the homology modeling protocol SEGMOD utility of GeneMine software (Department of Chemistry, UCLA). The 2.2 Å resolution crystal structure of the START domain of human MLN64 (Protein Data Bank accession number 1EM2) was used as a template structure to produce the homology model of StAR. The sequence alignment derived from superimposing the structures of the MLN64 START domain and the corresponding region of StAR is shown (Supplemental Fig. 1); no manipulations were made to generate the sequence alignment of StAR and the START

**FIG. 1. Mitochondrial Lon selectively degrades unassembled MPP $\alpha$  and StAR.** In A–E, samples were analyzed by SDS-PAGE and Coomassie Brilliant Blue staining. A and B, MPP $\alpha$  (in A) and MPP $\alpha$  or MPP $\beta$  (in B) (0.8  $\mu$ M) were incubated with Lon (0.8  $\mu$ M) in the presence or absence of 2 mM ATP at 37 °C. C, MPP $\alpha$  (0.8  $\mu$ M) and MPP $\beta$  (0, 0.08, 0.2, 0.4, and 0.8  $\mu$ M) were incubated with Lon (0.8  $\mu$ M) in the presence of 2 mM ATP at 37 °C. D, MPP $\alpha$  assembled with MPP $\beta$  forms an active peptidase. MPP $\alpha$  (2  $\mu$ M, lanes 1 and 2) and MPP $\beta$  (2  $\mu$ M, lanes 1 and 3) were incubated with MDH (0.8  $\mu$ M, lanes 1–4) at 30 °C for 30 min. Full-length and processed forms of MDH are indicated. In A–D, the yeast proteins MPP, Lon, and MDH were used. E, mouse StAR (3  $\mu$ M) was incubated with human Lon (0.8  $\mu$ M) at 37 °C, and aliquots were withdrawn at the time periods indicated.



domain of MLN64 as there was significant sequence identity (37%) and sequence homology (44%) between the two proteins (Supplemental Fig. 2). The SEGMOD utility proceeded as follows. First, the sequence was divided into short fragments, and corresponding structural fragments in the Protein Data Bank were then matched to the sequence. Finally, the fragments were fitted onto the framework of the template structure. The process was reiterated to generate 10 independent model structures; the average of these 10 structures was used to produce the final model. The stereochemical quality of the final model was verified using ProCheck. The root mean deviation between the MLN64 START domain structure and the final model structure for 170 paired C $\alpha$  atoms of StAR is 1.07 Å.

The START domain of StAR is an  $\alpha/\beta$  type structure containing a central twisted anti-parallel  $\beta$ -sheet consisting of nine  $\beta$ -strands (see Fig. 5D). The  $\beta$ -sheet is twisted in such a way that it forms two symmetrical U-shaped structures that are sandwiched between two long  $\alpha$ -helices; one section of both helices is at the bottom of the  $\beta$ -sheet, and the other section is in the cavity of the U-shaped  $\beta$ -sheets. These long  $\alpha$ -helices are located at the amino and carboxyl termini of the protein. Calculations of the exposed surface area show that Val-155, Ala-172, and Leu-226 have 57, 45, and 35% of their side chains exposed to the solvent, respectively.

## RESULTS

**Folded Substrates Are Degraded by Mitochondrial Lon—**MPP $\alpha$  and StAR were degraded by Lon in an ATP-dependent manner as demonstrated by SDS-PAGE (Fig. 1A, B, and E) and mass spectrometry (Fig. 2). MPP $\alpha$  was more sensitive to Lon-mediated proteolysis than MPP $\beta$  (Fig. 1B); the formation of  $\beta\beta$  homodimers likely explains the relative stability of MPP $\beta$  (Supplemental Fig. 3). However, when MPP $\alpha$  was assembled into a heterodimeric complex with MPP $\beta$ , it was resistant to Lon-mediated proteolysis; increasing molar ratios of  $\beta$ : $\alpha$  from 0.1 to 1 led to increased MPP $\alpha$  stability (Fig. 1C). The assembled MPP $\alpha\beta$  heterodimer formed an active peptidase that removed the mitochondrial targeting presequence of MDH (Fig. 1D); neither subunit alone cleaved off the MDH targeting presequence. In addition, experiments showed that unassembled MPP $\alpha$  was resistant to partial digestion with trypsin (Supplemental Fig. 4A). Thus, MPP $\alpha$  that is degraded by Lon is in a folded state, as it is both trypsin-resistant and competent for assembly into an active enzyme. Human mitochondrial Lon also degraded mouse StAR in the presence of ATP but not in its

absence (Fig. 1E). StAR was in a folded state as it was resistant to trypsin digestion (Supplemental Fig. 4A). In addition, previous work has demonstrated that purified StAR is active as a cholesterol-binding protein that preferentially enhances sterol transfer in mitochondria isolated from steroidogenic cells (45).

**Initiation and Progression of Lon-mediated Proteolysis—**Peptides generated by Lon during a time course of MPP $\alpha$  and StAR degradation were identified by tandem mass spectrometry (MS/MS) (Figs. 2 and 4). Representative MS/MS spectra of the singly charged ions of peptides derived from MPP $\alpha$  (Fig. 2A) and from StAR (Fig. 2B) at the longest time points are shown (120 and 30 min, respectively). Examples of MS/MS sequencing spectra of singly charged ions derived from MPP $\alpha$  or StAR peptides are also shown. In all mass spectrometry experiments, the background level of peptide production was determined by incubating MPP $\alpha$ , StAR or Lon alone in the presence or absence of ATP (Supplemental Fig. 5). At the longest time point, few if any peptides were detected in reactions containing only MPP $\alpha$  or StAR either in the presence or in the absence of ATP; only a low level of peptide production was observed in samples containing only Lon and nucleotide, presumably as a result of autoproteolysis.

The amino acid frequency at Lon cleavage sites within MPP $\alpha$  and StAR is summarized (Fig. 3). The 5 amino acids preceding and following all cleavage sites were counted by a PERL script (Fig. 3, P(-5) to P(+5)). The amino acids at P(-1) and P(+1) in both MPP $\alpha$  and StAR were most commonly leucine  $\rightarrow$  alanine  $\rightarrow$  serine. Peptide bond hydrolysis did not occur, however, after every leucine, alanine, and serine residue within the substrate sequences. Other amino acids such as methionine, phenylalanine, threonine, and valine were present at P(-1) and/or P(+1) in either StAR or MPP $\alpha$ . The frequency of amino acids adjacent to one another at P(-1) and P(+1) was also determined (Supplemental Fig. 6). No common motifs or features within the peptides derived from either MPP $\alpha$  or StAR were identified using hidden Markov model analysis of peptide sequences identified by MS/MS.

The location and sequence of MPP $\alpha$  and StAR peptides identified at each time point of Lon-mediated degradation are

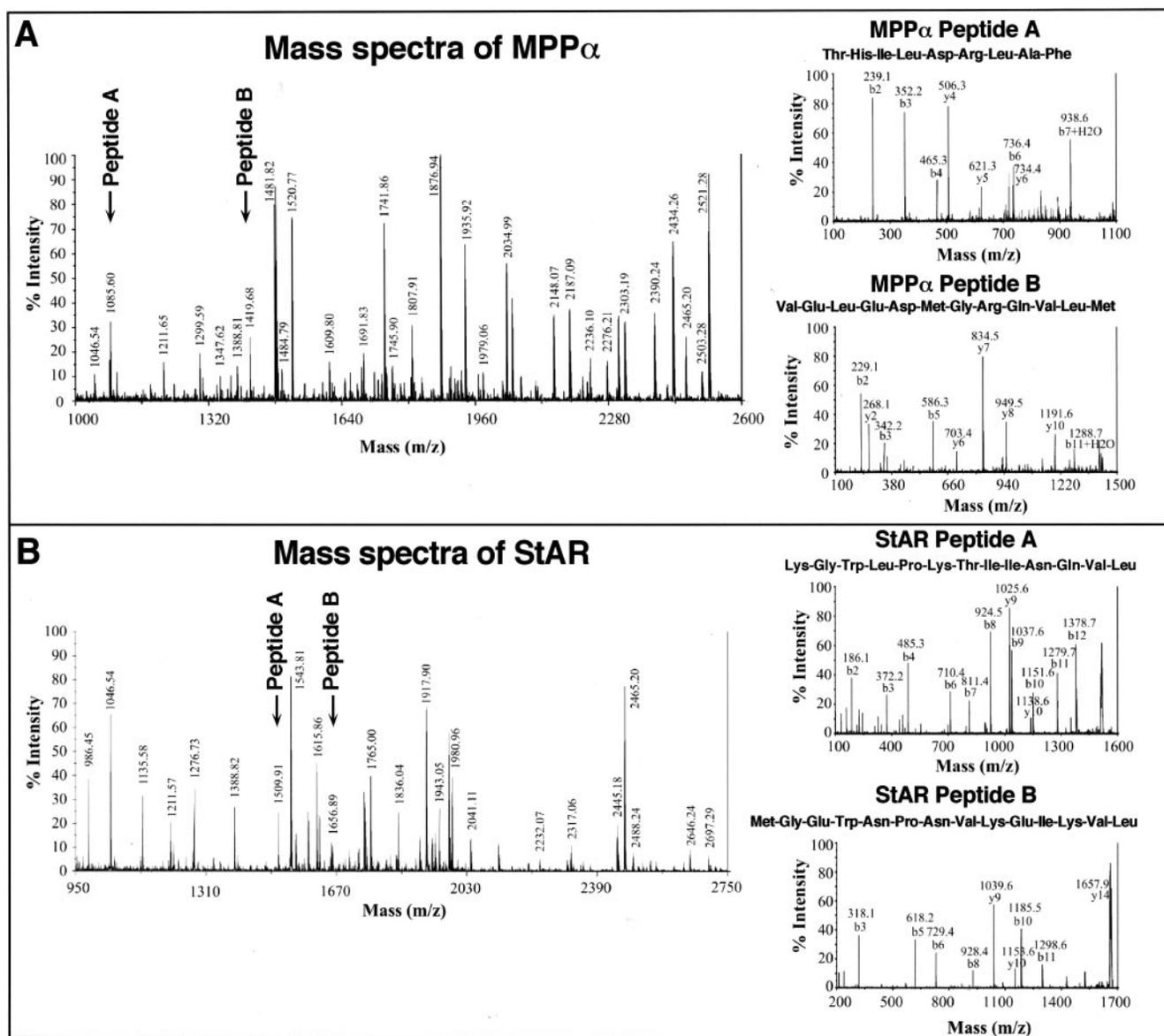


FIG. 2. Mass spectra of yeast MPP $\alpha$  and mouse StAR peptides produced by human Lon after 120 and 30 min (A and B, respectively). Yeast MPP $\alpha$  (3  $\mu$ M) was incubated with human Lon (0.8  $\mu$ M); mouse StAR (3  $\mu$ M) was incubated with human Lon (0.3  $\mu$ M). Examples of MS/MS sequencing spectra of singly charged ions derived from peptides A and B of MPP $\alpha$  or StAR are also shown.

shown in Fig. 4. The time point at which a peptide first appeared is indicated by a color-coded line beneath the amino acid sequence; once a peptide was detected, it was consistently observed at all subsequent time points. The first time point at which MPP $\alpha$  peptides generated by Lon were detected was  $T = 2$  min; MS/MS sequencing identified six peptides that all mapped to internal sites within the primary sequence of the substrate (Fig. 4A,  $T = 2$  min, red). No apparent preference for secondary structure was observed as cleavages occurred within  $\alpha$ -helices and unstructured regions and between  $\alpha$ -helices and/or  $\beta$ -sheets. At the following time point,  $T = 5$  min (yellow), the peptides identified were either adjacent to those detected at  $T = 2$  min (Fig. 4A, lines B, C, E, and F) or were located at the opposite end of the proteolytic fragment generated at an earlier time point (Fig. 4A, lines F and G; Supplemental Fig. 7). Subsequent MPP $\alpha$  cleavages at  $T = 10$ , 20, and 30 min (pink, blue, and green, respectively) occurred in a processive sequential manner (Fig. 4A, line B to line A and line G to line F). Peptides identified at the longest time points,  $T = 60$  and 120 min (brown and black, respectively), were located at internal as well as terminal regions of the protein. The identity of MPP $\alpha$  pep-

ptides generated by Lon at each time point was highly reproducible. The absence of identified peptides in certain regions of the protein was consistently observed and may be attributed to the limitations of MS such as ion suppression and the inability to detect small peptides or to peptide instability. No stable intermediates of MPP $\alpha$  cleavage were detected at any time point either by immunoblotting with various polyclonal antisera or by mass spectrometry (data not shown).

Similar results were obtained in a time course of Lon-mediated degradation of StAR (Fig. 4B). The earliest time point at which peptide products were detected was  $T = 1$  min (Fig. 4B, red); three peptides were identified that all mapped to internal regions within the primary sequence of StAR. Initial cleavages occurred within  $\beta$ -sheets and within unstructured regions of the protein. At the next time point,  $T = 2$  min (yellow), the peptides identified were either adjacent to those from the previous time point (Fig. 4B, lines C and D) or located at the opposite end of a larger proteolytic fragment generated earlier (Fig. 4B, line B). Subsequent peptide bond hydrolysis at  $T = 5$  (blue) and  $T = 30$  min (green) occurred processively and sequentially, as most clearly shown in Fig. 4B, line B.

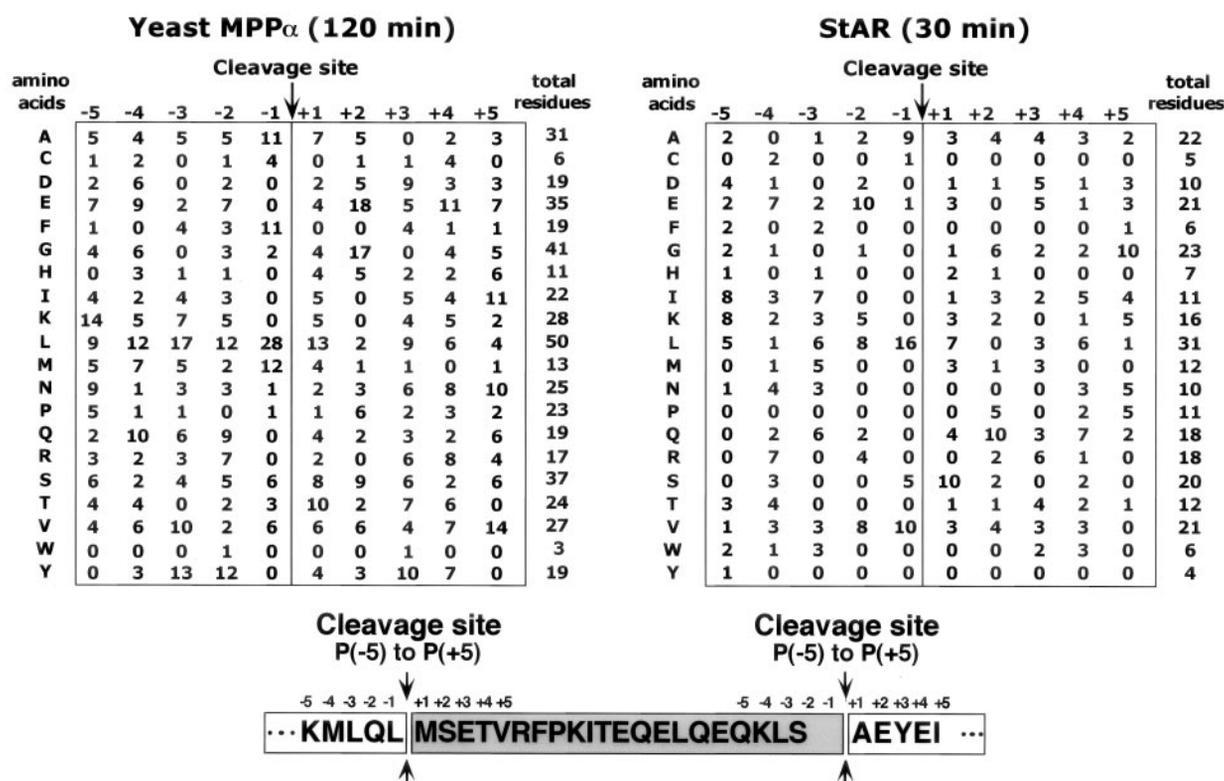


FIG. 3. Amino acid frequency at Lon cleavage sites within MPP $\alpha$  and StAR. A PERL script was used to count the amino acids at the P(-5) to P(+5) position of Lon cleavage sites; the total number of amino acid residues within MPP $\alpha$  or StAR is also shown in tabular form. Shown also is the position of P(-1) and P(+1) residues at the cleavage sites flanking a peptide.

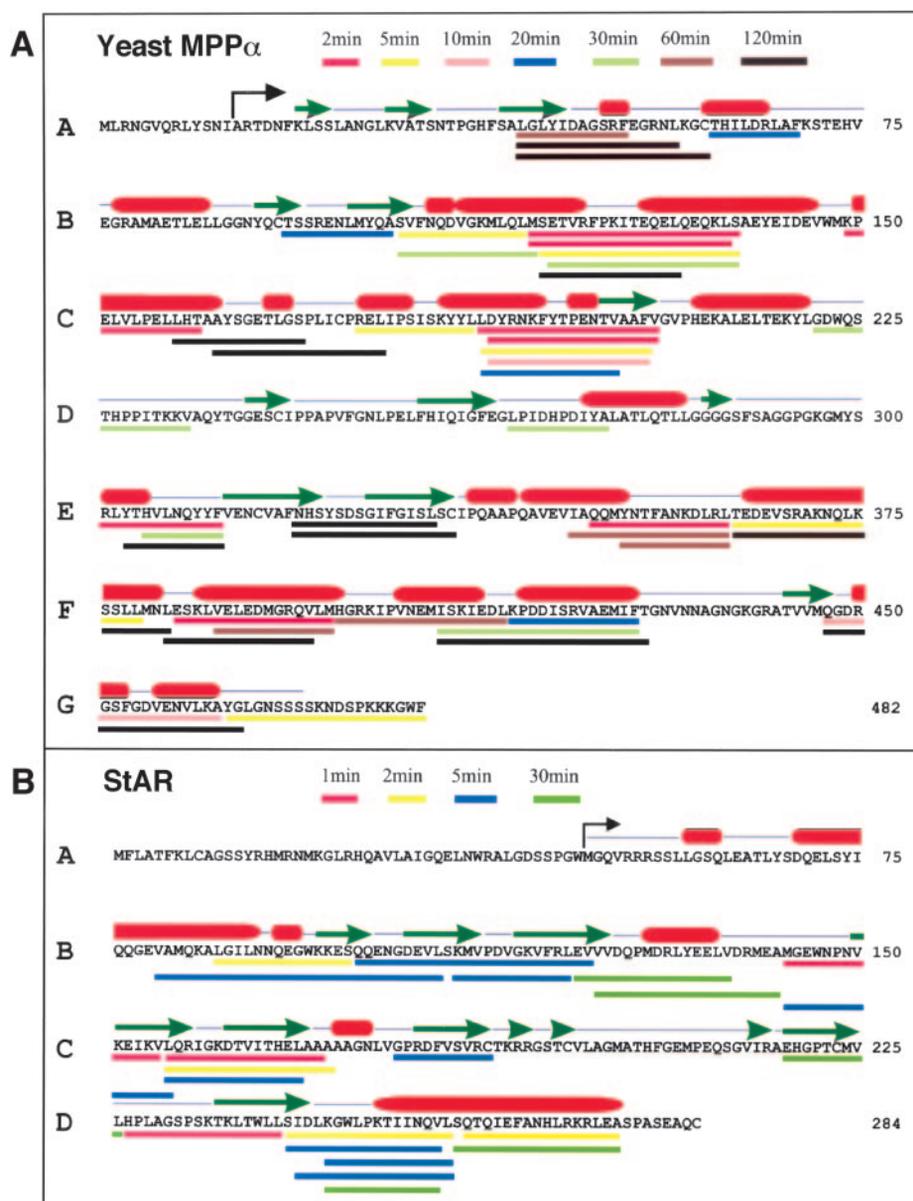
*Mitochondrial Lon Initiates Cleavage at Hydrophobic Sites at the Surface of MPP $\alpha$  and StAR*—To gain further insight into the initiation of substrate degradation, the amino acids at the P(-1) position of initial Lon cleavage sites were mapped onto the crystal structure of MPP solved at 2.5 Å (46) and onto a homology-based model of StAR derived from the crystal structure of the START domain of MLN64 solved at 2.2 Å (47) (Fig. 5). MPP $\alpha$  consists of two domains that each contain six  $\beta$ -strands and eight  $\alpha$ -helices; these domains exhibit a 2-fold symmetry. Fig. 5A shows the P(-1) residue at the carboxyl terminus of MPP $\alpha$  peptides generated by Lon at  $T = 2$  min. Thr-160 is located on an  $\alpha$ -helix that is packed against a twisted  $\beta$ -plate within the amino-terminal domain; Met-398 exhibits a 2-fold symmetry with respect to Thr-160, which is located in a corresponding site within the carboxyl-terminal domain. Ser-138 and Val-205 in the amino-terminal domain and Phe-312 and Leu-362 in the carboxyl-terminal domain also show a similar 2-fold symmetry. The symmetric location of these cleavage sites suggests that certain structural folds or surface properties are recognized by mitochondrial Lon. The environment surrounding the side chains at initial cleavage sites was determined by calculating the molecular surface of MPP $\alpha$  (Fig. 5B). Hydrophilic regions are shown in red, neutral regions are shown in gray, and hydrophobic regions are shown in blue. The side chain atoms of the hydrophobic amino acids at P(-1) of initiating Lon cleavage sites are on the surface of MPP $\alpha$  surrounded by a highly polar medium; these side chains are rendered as green space-filled balls (Fig. 5B). None of the P(-1) residues have side chains oriented within the hydrophobic region of the molecule or at the interface between MPP $\alpha$  and  $\beta$  subunits. As an example, the amino acids forming a hydrophilic environment surrounding Met-398 are shown in Fig. 5C; Lys-29, Asp-47, Glu-98, Asn-199, Lys-232, Asp-269, Asp-325, Arg-394, Gln-395, and His-399 (shown in red) surround Met-398 (shown in green).

The homology model of StAR was produced using the crystal structure of the MLN64 START domain as a template (Fig. 5D). The START domain of StAR consists of a central twisted anti-parallel  $\beta$ -sheet that shows an apparent 2-fold symmetry (Fig. 5D); it contains a hydrophobic central tunnel with an opening on either side of the protein. The P(-1) amino acids at the carboxyl terminus of StAR peptides generated by Lon at  $T = 1$  min (Val-155, Ala-172, and Leu-226) were mapped on the ribbon model and on the molecular surface model of StAR (Fig. 5, D and E, respectively). Although hydrophobic, Val-155 and Leu-226, which are located at the outer surface of the hydrophobic tunnel, and Ala-172, which is positioned between an  $\alpha$ -helix and  $\beta$ -sheet, are all surrounded by charged or polar residues. As an example, the hydrophilic environment surrounding Val-155 is shown in Fig. 5F; Glu-141, Lys-154, Val-155, Gln-157, and Arg-158 (shown in red) encircle Val-155 (shown in green). Further analysis of the homology model of StAR shows the presence of 33 basic and 29 acidic amino acids. Only 14 of the basic residues and 15 of the acidic residues participate in ion pair formation; nearly half of the charged residues are not neutralized, thereby resulting in an extremely polar surface. These unengaged polar residues on the surface of StAR may function as recognition patches for Lon. Taken together, our results raise the possibility that folded substrates are degraded by ATP-dependent proteases *in vivo*.

#### DISCUSSION

In the prevailing model of ATP-dependent proteolysis, protein degradation proceeds by four steps involving substrate 1) recognition; 2) engagement; 3) unfolding and translocation; and 4) peptide bond hydrolysis. Unlike the 26 S proteasome and bacterial ATP-dependent proteases that generally recognize their substrates by degradation sequences or tags, mitochondrial ATP-dependent proteases appear to selectively degrade untagged substrates. The vast majority of mitochondrial

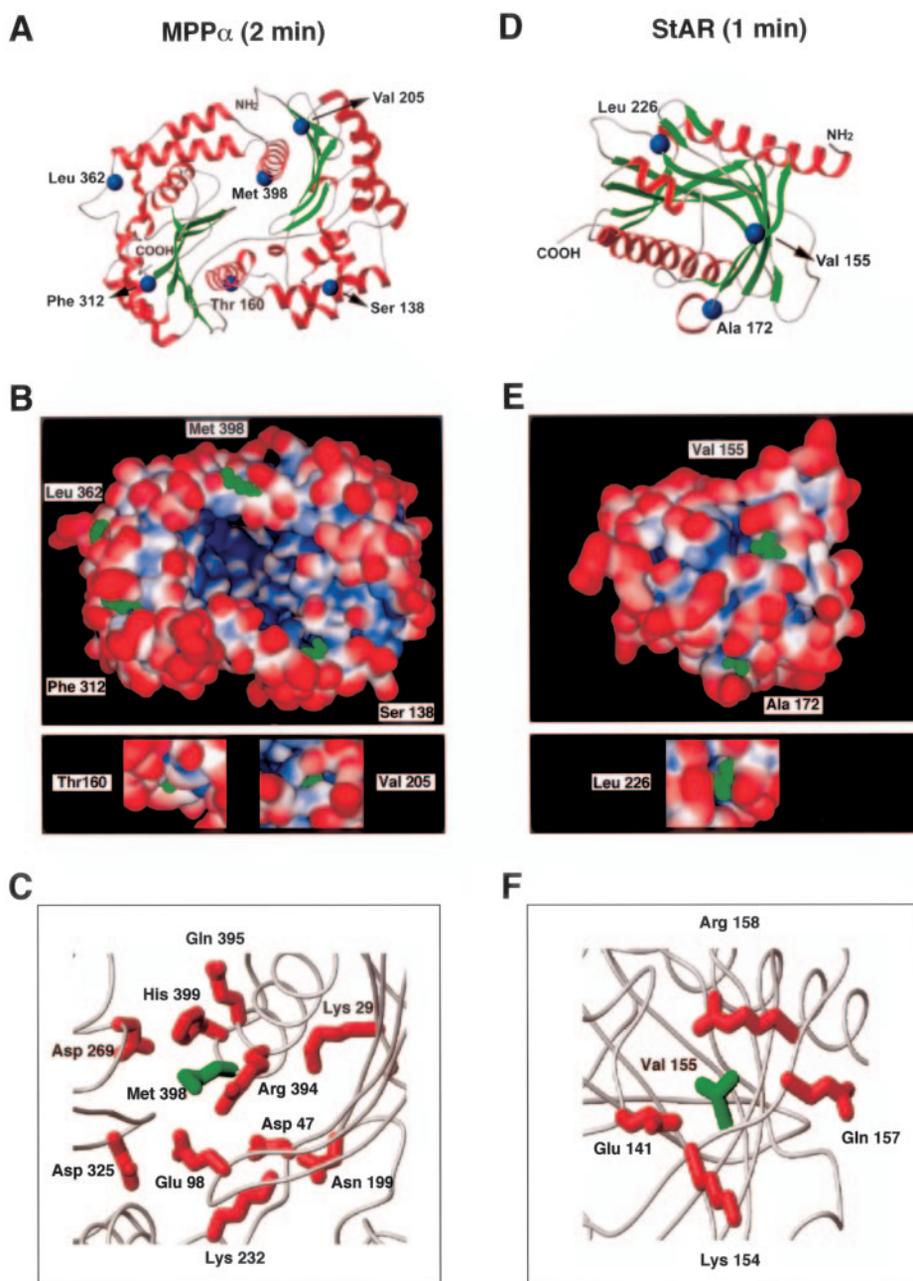
**FIG. 4. MPP $\alpha$  and StAR peptides generated during a time course of Lon-mediated degradation were mapped on the primary sequence of the proteins (A and B, respectively). Symbols above the amino acid sequences are as follows. Black right angled arrows indicate the start of the recombinant proteins produced without their respective mitochondrial targeting presequence; green arrows denote  $\beta$ -sheets; thick red lines denote  $\alpha$ -helices. Symbols beneath the amino acid sequences are as follows. Color-coded lines indicate the time point at which a peptide first appeared. For MPP $\alpha$  peptides,  $T = 2$  (red), 5 (yellow), 10 (pink), 20 (blue), 30 (green), 60 (brown), and 120 (black) min; for StAR peptides,  $T = 1$  (red), 2 (yellow), 5 (blue), and 30 (green) min.**



matrix proteins function as multisubunit complexes, thus the proper assembly state of these proteins is likely monitored by the quality control system governed by ATP-dependent proteases. The extent to which unassembled proteins are misfolded or unfolded *in vivo* is not known; it is possible that for unassembled proteins, exposed hydrophobic or buried regions function as recognition elements for ATP-dependent proteases. However, the results presented here show that purified unassembled MPP $\alpha$  is in a folded state as it is trypsin-resistant and capable of associating with MPP $\beta$  into an active peptidase (Fig. 1D and Supplemental Fig. 4). Lon selectively degrades unassembled but not assembled MPP $\alpha$ , although the sites of initial cleavage on the  $\alpha$ -subunit are exposed in the heterodimeric complex. In fact, proper folding of MPP $\alpha$  appears to be a requirement for Lon-mediated degradation. A trypsin-sensitive and assembly-incompetent form of MPP $\alpha$  that lacks 17 carboxyl-terminal amino acids (48) is not degraded by Lon; however, when MPP $\alpha$ (-17) is in an assembly-competent state, it is degraded (Supplemental Fig. 4). Attempts to unfold wild type MPP $\alpha$  by urea or heat denaturation leads to aggregation and resistance to degradation by Lon (Supplemental Fig. 8). Therefore we were not able to determine whether the initiation and progression of proteolysis are different when MPP $\alpha$  is unfolded.

The recognition element(s) that target StAR and unassembled MPP $\alpha$  for Lon-mediated proteolysis are unknown; however, such determinants may be identical to, or associated with, the sites for initial cleavage within each respective substrate (discussed below).

Previous work shows that once a protein substrate carrying a degradation sequence or tag is recognized by an ATP-dependent protease, it is engaged at an unstructured or loosely folded region and actively unfolded; protein unfolding is a prerequisite for translocation through a relatively narrow channel leading to the proteolytic core of the enzyme (49, 50). The degradation of folded proteins by mitochondrial Lon raises the question of how such substrates gain access to the active site of the protease. As the diameter of the central channel of yeast mitochondrial Lon is 2.5 nm (51) and that of the isolated bacterial Lon protease domain complex is 1.8 nm (52), a ring-opening mechanism within Lon is required for the degradation of folded proteins such as MPP $\alpha$  (6.4  $\times$  5.0 nm) or StAR (5.0  $\times$  3.6 nm). Cryo-electron microscopy demonstrates that mitochondrial Lon from *Saccharomyces cerevisiae* is a homo-heptameric ring-shaped complex composed of flexible subunits; data suggest that ATP-binding influences conformational changes within the holoenzyme (51). In the absence of added ATP or non-



**FIG. 5. Location of initiating Lon cleavage sites within MPP $\alpha$  and StAR.** A and D, ribbon models of the MPP $\alpha$  crystal structure and the StAR START domain derived by homology modeling. The locations of P(-1) amino acids at the carboxyl terminus of MPP $\alpha$  peptides ( $T = 2$  min) are as follows: Ser-138, Thr-160, Val-205, Phe-312, Leu-362, and Met-398. For StAR peptides ( $T = 1$  min), the locations are as follows: Val-155, Ala-172, and Leu-226. B and E, surface map of MPP $\alpha$  and StAR showing solvent accessibility of these P(-1) residues (green). Red, hydrophilic regions; gray, neutral regions; blue, hydrophobic regions. Lower panels show Thr-160 and Val-205 within MPP $\alpha$ , which are located on the opposite face, and Leu-226 within StAR, which is located on the opposite face. C and F, amino acids coordinated around Met-398 of MPP $\alpha$  and Val-155 of StAR. Coordinates of backbone and side chain atoms of MPP $\alpha$  and StAR were obtained from Protein Data Bank accession numbers 1HR6 and 1EM2, respectively.

hydrolyzable ATP analogs, asymmetric ring-like structures are observed in which two adjacent subunits are in an extended conformation, suggesting the possibility of ring opening. The addition of ATP or AMP-PNP increases the percentage of closed symmetric ring-shaped particles. Ring opening and changes in channel diameter have been demonstrated for hexameric helicases such as the bacterial Rho transcription terminator and bacteriophage T7 helicase (53, 54). Crystallographic and pre-steady state kinetic analyses suggest that the open ring state of Rho and T7 helicase is poised for loading onto mRNA and single-stranded DNA, respectively.

These helicases, like Lon and other ATPase components of energy-dependent proteases, belong to the AAA<sup>+</sup> family of proteins (ATPases associated with diverse cellular activities) (55, 56); thus, they may operate by similar mechanisms. A mechanistic similarity between hexameric helicases and the ATP-dependent proteolysis of unfolded polypeptides has already been suggested (57) and may be applicable as well to the degradation of folded protein substrates reported in this study. As mitochondrial Lon also binds to single-stranded DNA and RNA (44,

58, 59), it is tempting to speculate that the mechanism(s) underlying its proteolytic and nucleic acid binding activities are functionally and structurally integrated. ATP-dependent ring opening may be involved in accommodating nucleic acid or folded protein substrate into the central core of Lon, and cycles of ATP-hydrolysis may function in translocating the enzyme either along nucleic acid or along unfolded regions of the substrate once the folded protein has been initially cleaved. The initial Lon cleavage sites within MPP $\alpha$  and StAR appear to be hydrophobic residues that are surface-exposed and surrounded by a highly charged environment; such sites may function as recognition elements or patches for the protease. It is striking that the initial cleavage sites within MPP $\alpha$  have structural symmetry (Fig. 5, A and B), suggesting that surface properties of the folded substrate are specifically recognized and cleaved by mitochondrial Lon. In addition, a size limit is likely imposed upon folded proteins that enter into the proteolytic core of the Lon holoenzyme. This is consistent with the observation that MPP $\alpha$  complexed with MPP $\beta$  is not degraded, although initiating Lon cleavage sites are exposed and not masked by sub-

unit interactions. Assembled MPP $\alpha$  may be resistant to degradation because Lon cannot accommodate the MPP $\alpha\beta$  complex into its central channel. Thus, the selective degradation of folded proteins is determined perhaps not only by structural features at the protein surface but also by protein size. Once the folded protein is initially cleaved, resulting in proteolytic fragments that are most likely unstructured, processive degradation occurs from the amino- to the carboxyl-terminal ends or *vice versa*. Further studies are required to arrive at a mechanistic understanding of how folded substrates are recognized, initially cleaved, and subsequently degraded by mitochondrial Lon and to determine whether ring opening and closing underlie the proteolytic and nucleic acid binding function of the holoenzyme.

**Acknowledgments**—We thank Drs. S. S. Patel and H. Stahlberg for critical review of this manuscript and N. Parkhomenko and H. Kolesarova for technical assistance.

#### REFERENCES

- Prakash, S., and Matouschek, A. (2004) *Trends Biochem. Sci.* **29**, 593–600
- Pickart, C. M., and Cohen, R. E. (2004) *Nat. Rev. Mol. Cell. Biol.* **5**, 177–187
- Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J., Grant, R. A., Hersch, G. L., Joshi, S. A., Kenniston, J. A., Levchenko, I., Neher, S. B., Oakes, E. S., Siddiqui, S. M., Wah, D. A., and Baker, T. A. (2004) *Cell* **119**, 9–18
- Pickart, C. M., and Eddins, M. J. (2004) *Biochim. Biophys. Acta* **1695**, 55–72
- Gonzalez, M., Frank, E. G., Levine, A. S., and Woodgate, R. (1998) *Genes Dev.* **12**, 3889–3899
- Hoskins, J. R., Singh, S. K., Maurizi, M. R., and Wickner, S. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8892–8897
- Karzai, A. W., Roche, E. D., and Sauer, R. T. (2000) *Nat. Struct. Biol.* **7**, 449–455
- Varshavsky, A. (2003) *Nat. Cell Biol.* **5**, 373–376
- Varshavsky, A. (2000) *Harvey Lect.* **96**, 93–116
- Withey, J. H., and Friedman, D. I. (2003) *Annu. Rev. Microbiol.* **57**, 101–123
- Hoskins, J. R., Yanagihara, K., Mizuuchi, K., and Wickner, S. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11037–11042
- Benaroudj, N., Zwickl, P., Seemuller, E., Baumeister, W., and Goldberg, A. L. (2003) *Mol. Cell* **11**, 69–78
- Burton, R. E., Siddiqui, S. M., Kim, Y. I., Baker, T. A., and Sauer, R. T. (2001) *EMBO J.* **20**, 3092–3100
- Kenniston, J. A., Baker, T. A., Fernandez, J. M., and Sauer, R. T. (2003) *Cell* **114**, 511–520
- Kim, Y. I., Burton, R. E., Burton, B. M., Sauer, R. T., and Baker, T. A. (2000) *Mol. Cell* **5**, 639–648
- Lam, Y. A., Lawson, T. G., Velayutham, M., Zweier, J. L., and Pickart, C. M. (2002) *Nature* **416**, 763–767
- Lee, C., Schwartz, M. P., Prakash, S., Iwakura, M., and Matouschek, A. (2001) *Mol. Cell* **7**, 627–637
- Navon, A., and Goldberg, A. L. (2001) *Mol. Cell* **8**, 1339–1349
- Ortega, J., Lee, H. S., Maurizi, M. R., and Steven, A. C. (2002) *EMBO J.* **21**, 4938–4949
- Reid, B. G., Fenton, W. A., Horwich, A. L., and Weber-Ban, E. U. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3768–3772
- Singh, S. K., Grimaud, R., Hoskins, J. R., Wickner, S., and Maurizi, M. R. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8898–8903
- Langer, T. (2000) *Trends Biochem. Sci.* **25**, 247–251
- Arnold, I., and Langer, T. (2002) *Biochim. Biophys. Acta* **1592**, 89–96
- Suzuki, C. K., Rep, M., van Dijl, J. M., Suda, K., Grivell, L. A., and Schatz, G. (1997) *Trends Biochem. Sci.* **22**, 118–123
- Bross, P., Andresen, B. S., Knudsen, I., Kruse, T. A., and Gregsen, N. (1995) *FEBS Lett.* **377**, 249–252
- Kang, S. G., Ortega, J., Singh, S. K., Wang, N., Huang, N. N., Steven, A. C., and Maurizi, M. R. (2002) *J. Biol. Chem.* **277**, 21095–21102
- 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
- Casari, G., Fusco, M. D., Ciarmatori, S., Zeviani, M., Mora, M., Fernandez, P., Michele, G. D., Filla, A., Coccozza, S., Marconi, R., Durr, A., Fontaine, B., and Ballabio, A. (1998) *Cell* **93**, 973–983
- Coppola, M., Pizzigoni, A., Banfi, S., Bassi, M. T., Casari, G., and Incerti, B. (2000) *Genomics* **66**, 48–54
- Banfi, S., Bassi, M. T., Andolfi, G., Marchitello, A., Zanotta, S., Ballabio, A., Casari, G., and Franco, B. (1999) *Genomics* **59**, 51–58
- Wagner, I., Arlt, H., van Dyck, L., Langer, T., and Neupert, W. (1994) *EMBO J.* **13**, 5135–5145
- Suzuki, C. K., Suda, K., Wang, N., and Schatz, G. (1994) *Science* **264**, 273–276; Correction, *Science* **264**, 891
- Savel'ev, A. S., Novikova, L. A., Kovaleva, I. E., Luzikov, V. N., Neupert, W., and Langer, T. (1998) *J. Biol. Chem.* **273**, 20596–20602
- Bota, D. A., and Davies, K. J. A. (2002) *Nat. Cell Biol.* **4**, 674–680
- van Dijl, J. M., Kutejová, E., Suda, K., Perecko, D., Schatz, G., and Suzuki, C. K. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10584–10589
- Gakh, O., Obsil, T., Adamec, J., Spizek, J., Amler, E., Janata, J., and Kalousek, F. (2001) *Arch. Biochem. Biophys.* **385**, 392–396
- Gakh, O., Cavadini, P., and Isaya, G. (2002) *Biochim. Biophys. Acta* **1592**, 63–77
- Strauss, J. F., III, Kishida, T., Christenson, L. K., Fujimoto, T., and Hiroi, H. (2003) *Mol. Cell. Endocrinol.* **202**, 59–65
- Soccio, R. E., and Breslow, J. L. (2003) *J. Biol. Chem.* **278**, 22183–22186
- Stocco, D. M. (2001) *Annu. Rev. Physiol.* **63**, 193–213
- Granot, Z., Geiss-Friedlander, R., Melamed-Book, N., Eimerl, S., Timberg, R., Weiss, A. M., Hales, K. H., Hales, D. B., Stocco, D. M., and Orly, J. (2003) *Mol. Endocrinol.* **17**, 2461–2476
- Geli, V., Yang, M. J., Suda, K., Lustig, A., and Schatz, G. (1990) *J. Biol. Chem.* **265**, 19216–19222
- Lanzetta, P. A., Alvarez, L. J., Reinach, P. S., and Candia, O. A. (1979) *Anal. Biochem.* **100**, 95–97
- Liu, T., Lu, B., Lee, I., Ondrovicova, G., Kutejova, E., and Suzuki, C. K. (2004) *J. Biol. Chem.* **279**, 13902–13910
- Petrescu, A. D., Gallegos, A. M., Okamura, Y., Strauss, J. F., III, and Schroeder, F. (2001) *J. Biol. Chem.* **276**, 36970–36982
- Taylor, A. B., Smith, B. S., Kitada, S., Kojima, K., Miyaura, H., Otwinowski, Z., Ito, A., and Deisenhofer, J. (2001) *Structure (Camb.)* **9**, 615–625
- Tsujishita, Y., and Hurley, J. H. (2000) *Nat. Struct. Biol.* **7**, 408–414
- Janata, J., Hola, K., Kubala, M., Gakh, O., Parkhomenko, N., Matuskova, A., Kutejova, E., and Amler, E. (2004) *Biochem. Biophys. Res. Commun.* **316**, 211–217
- Ishikawa, T., Beuron, F., Kessel, M., Wickner, S., Maurizi, M. R., and Steven, A. C. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 4328–4333
- Voges, D., Zwickl, P., and Baumeister, W. (1999) *Ann. Rev. Biochem.* **68**, 1015–1068
- 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
- Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulova, F., Dauter, Z., Maurizi, M. R., Rotanova, T. V., Wlodawer, A., and Gustchina, A. (2004) *J. Biol. Chem.* **279**, 8140–8148
- Ahnert, P., Picha, K. M., and Patel, S. S. (2000) *EMBO J.* **19**, 3418–3427
- Gai, D., Zhao, R., Li, D., Finkielstein, C. V., and Chen, X. S. (2004) *Cell* **119**, 47–60
- Lupas, A. N., and Martin, J. (2002) *Curr. Opin. Struct. Biol.* **12**, 746–753
- Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. (1999) *Genome Res.* **9**, 27–43
- Prakash, S., Tian, L., Ratliff, K. S., Lehotzky, R. E., and Matouschek, A. (2004) *Nat. Struct. Mol. Biol.* **11**, 830–837
- Fu, G. K., and Markovitz, D. M. (1998) *Biochemistry* **37**, 1905–1909
- Lu, B., Liu, T., Crosby, J. A., Thomas-Wohlever, J., Lee, I., and Suzuki, C. K. (2003) *Gene (Amst.)* **306**, 45–55
- Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graph.* **14**, 51–55