Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 156 (2006) 165–174
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

# Crystal structure at 1.9 Å of *E. coli* ClpP with a peptide covalently bound at the active site

Agnieszka Szyk, Michael R. Maurizi *

*Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA*

Received 29 December 2005; received in revised form 2 March 2006; accepted 15 March 2006
Available online 21 April 2006

**Abstract**

ClpP, the proteolytic component of the ATP-dependent ClpAP and ClpXP chaperone/protease complexes, has 14 identical subunits organized in two stacked heptameric rings. The active sites are in an interior aqueous chamber accessible through axial channels. We have determined a 1.9 Å crystal structure of *Escherichia coli* ClpP with benzyloxycarbonyl-leucyltyrosine chloromethyl ketone (Z-LY-CMK) bound at each active site. The complex mimics a tetrahedral intermediate during peptide cleavage, with the inhibitor covalently linked to the active site residues, Ser97 and His122. Binding is further stabilized by six hydrogen bonds between backbone atoms of the peptide and ClpP as well as by hydrophobic binding of the phenolic ring of tyrosine in the S1 pocket. The peptide portion of Z-LY-CMK displaces three water molecules in the native enzyme resulting in little change in the conformation of the peptide binding groove. The heptameric rings of ClpP-CMK are slightly more compact than in native ClpP, but overall structural changes were minimal (rmsd ~ 0.5 Å). The side chain of Ser97 is rotated ~90° in forming the covalent adduct with Z-LY-CMK, indicating that rearrangement of the active site residues to a active configuration occurs upon substrate binding. The N-terminal peptide of ClpP-CMK is stabilized in a β-hairpin conformation with the proximal N-terminal residues lining the axial channel and the loop extending beyond the apical surface of the heptameric ring. The lack of major substrate-induced conformational changes suggests that changes in ClpP structure needed to facilitate substrate entry or product release must be limited to rigid body motions affecting subunit packing or contacts between ClpP rings.


*Keywords:* ClpP; *Escherichia coli* ClpP; Peptide binding; Chloromethyl ketone; Chloromethyl ketone inhibitor; X-ray crystallography; β-hairpin; Axial channel

## 1. Introduction

Intracellular protein degradation plays a vital role in protein quality control and in the regulation of virtually all cellular functions and is, itself, a highly regulated process requiring the expenditure of metabolic energy in the form of ATP (Gottesman et al., 1997; Sauer et al., 2004). Most intracellular protein degradation is carried out by non-specific proteases that are structurally designed to control their activities so that only appropriate substrates are degraded (Lupas et al., 1997). Typical of such proteases are the proteasomes, which are responsible for cytosolic protein degradation in eukaryotic cells and in archaebacteria (Pickart and Cohen, 2004; Schmidt et al., 1999), and ClpP, which functions in eubacterial cells and in the organelles of higher eukaryotes (Adam and Clarke, 2002; Corydon et al., 1998; Kang et al., 2002). *Escherichia coli* ClpP serves as a simple model for the more complex systems because it recapitulates two basic biochemical features of those systems: ClpP subunits are arranged in two heptameric rings that assemble into a barrel-shaped tetradecamer with the active sites located within an internal chamber (Kessel et al., 1995; Wang et al., 1997), and the ClpP tetradecamer forms a stable complex with hexameric ATP-dependent chaperones (ClpA or ClpX) that are required for rapid degradation of stable folded proteins (Grimaud et al., 1998; Kessel et al., 1995). Substrate selection is the function of ClpA and ClpX, which either directly bind specific protein substrates or bind adaptor proteins that interact with the specific substrates (Flynn et al., 2003; Gottesman et al., 1993; Mogk et al., 2004). ClpA and ClpX catalyze ATP-dependent protein unfolding

* Corresponding author. Fax: +1 301 480 2284.
*E-mail address:* mmaurizi@helix.nih.gov (M.R. Maurizi).

1047-8477/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2006.03.013

and translocation allowing them to deliver the bound substrates to the degradative chamber (Hoskins et al., 2000; Kenniston et al., 2003; Singh et al., 2000; Weber-Ban et al., 1999).

The work of unfolding and translocating substrate proteins during degradation is accomplished through ATP-dependent conformational changes and structural motions in ClpA or ClpX, but it is likely that these changes also affect the complex between the chaperones and ClpP and cause some dynamic changes in ClpP itself. The stability of ClpX–ClpP or ClpA–ClpP interactions is dependent on the nucleotide state of the chaperone (Singh et al., 1999) indicating that contacts between the different subunits change during cycles of ATP hydrolysis. The sevenfold symmetry of ClpP and the sixfold symmetry of ClpA and ClpX dictate that, at any given time, none of the subunit interactions within a complex will be identical and, further, that allosteric effects exerted between the respective components will be different during the reaction cycle (Beuron et al., 1998). The highly mobile N-terminal region of ClpP (Kang et al., 2004; Wang et al., 1997), which is not directly involved in the primary interface with ClpA or ClpX, appears to directly or indirectly affect binding of the chaperones and is in position to make contact with substrates as they are translocated from the chaperone (Gribun et al., 2004; Kang et al., 2004). The interface between the two ClpP rings also undergoes conformational change in response to binding of the chaperone in the distal surface (Kang et al., 2005), and such allosteric communication through the ring would be expected to affect holoenzyme activity. A dramatic illustration of conformational variability in ClpP comes from a recent report that a new class of antibiotic peptides can bind to ClpP and increase its activity causing harmful intracellular proteolysis that can block cell growth (Brotz-Oesterhelt et al., 2005). Structural information about conformational states of ClpP is vital to our understanding these diverse properties and functions of ClpP in ATP-dependent protease complexes and bound to other biological effectors as well.

In native ClpP, face-to-face joining of the two rings enclose an aqueous chamber ~50 Å in diameter and with a volume sufficient to hold about 50 kDa of protein (Kessel et al., 1995; Wang et al., 1997). The axial pores are relatively narrow (~10 Å) and cannot admit folded proteins but will allow passage of unstructured peptides. ClpP belongs to the serine protease family; in the *E. coli* enzyme, the catalytic triad consists of Ser97, His122, and Asp171 (Maurizi et al., 1990b; Wang et al., 1997). On its own, ClpP has limited peptidase activity against short unstructured peptides (Thompson and Maurizi, 1994), but ClpAP and ClpXP complexes can process longer peptides and folded native proteins. Peptide bond cleavage does not show strict sequence specificity, although substrates are cleaved preferentially after non-polar residues (Thompson et al., 1994). The size distribution of peptide products is independent of the rate of translocation (Choi and Licht, 2005), and is independent of which chaperone component (ClpA or ClpX) delivers the substrate (S.G. Kang, R. Sivendran, M.R. Maurizi, unpublished). These results suggest that rate at which products exit the chamber is relatively fast compared to translocation or cleavage.

To examine the conformation of substrates bound at the active site of ClpP and to determine what parts of the ClpP structure are affected by substrate binding, we crystallized *E. coli* ClpP bound to the peptidyl inhibitor, benzyloxycarbonyl leucyltyrosine chloromethylketone (Z-LY-CMK). Chloromethyl ketones are irreversible transition-state inhibitors and have been used for mapping the interactions between the side chains of substrate molecule and subsites in many serine and cysteine proteases (Brandstetter et al., 2001; Powers et al., 1977). The structure of the ClpP-CMK complex determined at a resolution of 1.9 Å shows formation of two covalent bonds between the inhibitor molecule and catalytic serine and histidine residues and provides details on the interactions of the peptide with the extended active site of ClpP. Our results confirm that the *E. coli* ClpP N-terminal peptide is located within the axial pore and that this putative substrate engagement loop maintains a dynamic conformation at the apical surface where substrate and chaperone interactions occur.

## 2. Experimental procedures

### 2.1. Expression, assay, and inactivation of ClpP

ClpP was over expressed and purified as described previously (Maurizi et al., 1994). Samples of pure enzyme were chemically inactivated by an addition of Z-LY-CMK (Bachem). A solution of ClpP (2.5 mg/ml; 0.12 mM) in 50 mM Tris, pH 8.0, containing 0.2 M KCl was incubated at room temperature for 30 min with a 2.5 molar excess of Z-LY-CMK. Completeness of inhibition was monitored by assaying protease activity in the presence of ClpA. Covalently modified ClpP (ClpP-CMK) was separated from excess inhibitor by acetone precipitation and subsequent gel filtration on a HiTrap™ G-25 desalting column (Amersham Biosciences). The ClpP-CMK was further purified using a Superdex 200 column (Amersham Biosciences) equilibrated in 50 mM Tris, pH 7.5, 0.2 M KCl, and 10% glycerol.

ClpP peptidase and protease activities were measured before and after treatment with Z-LY-CMK. Peptidase activity of ClpP alone was assayed using *N*-succinyl-Leu-Tyr-(7-amido-4-methyl coumarin) and monitoring the release of fluorescent 7-amino-4-methyl coumarin (Maurizi et al., 1994). Activated peptidase activity was measured in the presence of ClpA and adenosine-5′-(3-thiotriphoshate) with the decapeptide, FAPHMALVPV. Peptide products were separated and quantitated by reverse phase chromatography as described (Thompson and Maurizi, 1994). Protease activity was measured by ATP-dependent acid-solubilization of [$^3$H]α-casein in the presence of ClpA as described (Maurizi et al., 1994).

### 2.2. Crystallization and data collection

For crystallization, ClpP-CMK was dialyzed against 10 mM Hepes/NaOH pH 7.5 containing 0.2 M NaCl, and concentrated to 15 mg/ml by ultrafiltration using a Microcon 50 cartridge (Millipore). Crystals were obtained by hanging drop vapor diffusion at room temperature after mixing equal volumes of the protein solution and the reservoir solution (100 mM tri-sodium citrate, pH 5.6, 150 mM ammonium acetate, and 30% PEG 4000). Rectangular plate crystals, suitable for X-ray diffraction, reached their maximum size within 15–20 days. For the diffraction experiments, crystals were flash-frozen in the stream of nitrogen without cryoprotectant. The measurements were performed using a synchrotron radiation source at beamline X9B of the National Synchrotron Light Source at Brookhaven National Laboratory (wavelength 1.25 Å). The intensities were recorded on Quantum 4 CCD detector. Data processing and scaling, done with the program, HKL2000 (Otwinowski and Minor, 1997), yielded a complete dataset extending to a resolution of 1.9 Å. Crystals of the ClpP-CMK complex belong to the space group C2 ($a = 190.7$ Å, $b = 101.0$ Å, $c = 155.4$ Å, $\beta = 99.0°$) and 14 monomers of protein forming the biological unit are present in the asymmetric unit. The data collection statistics are summarized in Table 1.

### 2.3. Structure solution and refinement

The initial model of the ClpP-CMK complex was obtained by the molecular replacement method using the program, AMoRe (Navaza and Saludjian, 1997) with the search probe represented by coordinates of the Cα-atoms taken from the native ClpP tetradecamer structure (PDB code: 1TYF) (Wang et al., 1997). The complete set of coordinates, corresponding to the protein component of the native ClpP model was then aligned on the easily identifiable unique solution (correlation coefficient 0.75).

Initially, all structural refinement was conducted with program, CNS (Brunger et al., 1998). After the rigid body refinement and NCS-restrained simulated annealing, in which all experimental data within the resolution ranges 12.0–2.7 Å were used, values of the crystallographic $R$ factor and free $R$ were lowered to 0.29 and 0.32, respectively. Values of all B-factors were set to arbitrary 30 Å$^2$. After inspection of the $F_o - F_c$ and $2F_o - F_c$ electron density maps, the 18 N-terminal and three C-terminal residues were removed from the model. In subsequent calculations, interspersed by the manual corrections to the model using the program, O (Jones et al., 1991), resolution was gradually extended to 2.3 Å, at which point NCS-restrained refinement of the individual B-factors was included in the protocol. Subsequently, after gradual extension of the resolution to the limit of the experimental data, a careful inspection of the $F_o - F_c$ electron density maps in the active site region clearly showed the presence of the CMK molecules covalently attached to all 14 monomers of ClpP. The model of the CMK moiety was generated using the program, InsightII (http://www.accelrys.com/products/insight) and necessary parameters for the CNS library were generated by a similarity to such for existing fragments. After modeling the inhibitor into appropriate electron density peaks and following refinement at the resolution 12–1.9 Å, values of $R$ factor and free $R$ dropped to 0.24 and 0.27, respectively. The resulting electron density maps confirmed the correctness of the CMK-modeling, including presence of the covalent bonds between the Z-LY fragment of the inhibitor and the Ser97 and His122 from ClpP. At this stage, the model was missing 21 terminal residues of ClpP as well as whole solvent structure. Further model improvement was conducted with the program, Refmac 5 (Murshudov et al., 1997), and the program ARP, (Perrakis et al., 2001), was used for the automatic generation of an initial solvent structure. All solvent sites were visually inspected, and solvent atoms located near the termini of ClpP, in the active site region, as well as inside larger electron density peaks were removed prior subsequent refinement and manual correction steps. At this stage, the structural refinement was continued without NCS restrains, as indicated by the conformational heterogeneity of the side chains atoms in numerous residues as well as the terminal fragments of ClpP. Although lack of interpretable electron density peaks indicated substantial dynamic disorder of the N-terminal regions, it was possible to model several additional residues to each of 14 protein chains. In the case of one monomer (here described as B) the $2F_o - F_c$ electron density peaks were interpretable sufficiently to determine the structure of all but the first and last residues. During the last steps of model building, additional molecules of water, glycerol, and polyethylene glycol (PEG) segments were modeled into

Table 1
Data collection and refinement statistics

| | |
|---|---|
| *Data collection* | |
| Space group | C2 |
| Cell dimensions (Å) | $a = 190.7$, $b = 101.0$, $c = 155.4$, $\beta = 99.0°$ |
| Resolution range (Å)[a] | 50.0–1.9 (1.97–1.90) |
| Completeness (%) | 94.0 (95.1) |
| Observed reflections | 753 821 |
| Unique reflections | 215 884 |
| $R_{merge}$ | 0.051 (0.373) |
| Mosaicity (°) | 0.8 |
| $I/\sigma(I)$ | 21.5 (2.8) |
| *Data refinement* | |
| Resolution range (Å) | 15.0–1.9 |
| No. of reflections used for refinement | 203 390 |
| No. of protein atoms | 21 054 |
| No. of solvent molecules | 2881 |
| Average B-factor (Å$^2$) | 36.0 |
| Rmsd from ideal values | |
| Bond lengths (Å) | 0.019 |
| Bond angels (°) | 1.79 |
| Crystallographic $R$ factor | 0.173 |
| Free $R$(5% data) | 0.233 |

[a] Values in parenthesis are for the highest resolution shell.

electron density peaks. Values of the crystallographic *R* factor and free *R* obtained during the refinement of the final model are 0.173 and 0.233, respectively. In addition to the atoms forming the ClpP-CMK complex, the model describes 2855 water sites, 11 glycerol molecules, and 15 partial PEG molecules. The refinement statistics for the final structure are shown in Table 1.

The coordinate and structure factors have been deposited in the protein Data Bank with entry code 2FZS. All structure figures were prepared using the programs: PyMol (DeLano, 2002), Ribbons (Carson, 1997), and Molscript (Kraulis, 1991).

## 3. Results

### 3.1. Chemical modification of ClpP with Z-LY-CMK

Reaction of Z-LY-CMK with ClpP leads to complete loss of ClpA-dependent protease activity (Fig. 1) and of peptidase activity (Fig. 1, inset). Analysis of the modified ClpP by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate followed by electrospray ionization mass spectrometry confirmed the presence of a single protein species with a molecular weight of 21997.2 corresponding to the mass of a ClpP monomer with covalently bound inhibitor (theoretical $M_W$, 22004.2). ClpP-CMK was digested by incubation with lysyl endopeptidase C and the resulting peptides were separated by reverse phase chromatography, collected, and sequenced by Edman degradation. Native (unreacted) ClpP was analyzed in parallel. A single peptide from the native ClpP eluting at 24 min in the gradient was missing in the digest of ClpP-CMK, and a new peptide appeared at ~29 min. No other changes were seen in the peptide profiles. The peptide from native ClpP had the N-terminal sequence, Arg-Phe-Xxx-Leu-Pro-Asx-Ser-Arg-Val-Met-Ile-His-Gln, corresponding to residues 111–123 of ClpP. The peptide from ClpP-CMK had the same sequence, but no amino acid was detected in the cycle where histidine occurs in native ClpP. These data indicate that the chloromethyl ketone formed a stable adduct exclusively and quantitatively with His122 of ClpP.

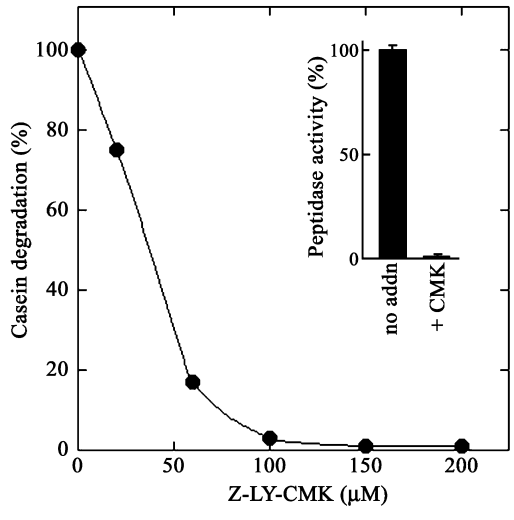


Fig. 1. Inactivation of ClpP with the peptide chloromethyl ketone, Z-LY-CMK. ClpP was treated with the indicated amounts of Z-LY-CMK for 30 min at 22 °C and assayed for casein degrading activity in the presence of ClpA and ATP. Inset, The sample of ClpP treated with 200 μM Z-LY-CMK was assayed for peptidase activity using *N*-succinyl-Leu-Tyr-(7-amido-4-methyl coumarin). Activities were normalized to the activity of ClpP incubated without the inhibitor.

### 3.2. Overall structure of ClpP-CMK

ClpP modified with Z-LY-CMK formed crystals belonging to the same C2 space group as the native enzyme (Wang et al., 1997) and had nearly identical unit cell parameters. A single tetradecamer containing two interdigitated heptameric rings was present in the asymmetric unit, and the overall structure was very similar to that of the native enzyme. When ClpP-CMK and native ClpP were subjected to a global structural alignment, the root-mean-square deviation (rmsd) was 0.44 Å after superposition of 2420 equivalent Cα atoms (figure not shown). As seen with native ClpP, ClpP-CMK monomers folded into a globular head domain created by repeated α/β units and a long handle consisting of strand β9 and α-helix E (Fig. 2). Depending on the choice of monomers, the rmsd for 172 equivalent Cα atoms (corresponding to residues 18–189) varied between 0.25 and 0.45 Å, similar to variations seen between the monomers in ClpP-CMK itself. The catalytic site, which is located in a cleft at the junction of the head and the handle, is essentially unaltered except for a change in the position of the catalytic Ser97 and in the number of structural water molecules (see below). The most significant conformational differences between the models of the ClpP-CMK and native ClpP subunits were seen in the regions of the N- and C-termini (Fig. 2).

The electron density for the N-terminal region of ClpP-CMK was somewhat broken and variable among the 14 subunits. In one monomer it was possible to build a complete model of the N-terminal carbon backbone (Fig. 3A), while for the remaining monomers only partial models were possible (Fig. 3B). In all subunits, the N-terminal polypeptide adopts a β-hairpin conformation, with residues 1–7 lying within the axial pore and residues 8–17 forming a loop that extends out from the pore above the heptameric ring surface. This model is consistent with recently described models of the N-terminal parts of human ClpP and *Streptococcus pneumoniae* ClpP (Gribun et al., 2004; Kang et al., 2004) and differs from the original model for *E. coli* ClpP in which the terminal density was assigned to residues 11–18 and residues 1–10 were not included in the model (Wang et al., 1997). Recently, a higher resolution structure of native ClpP (Bewley et al., 2006) revealed that the N-terminal peptides could adopt both configurations within the same tetradecamer. The N-terminus will be addressed in more detail below.

The C-terminal residue, Asn193, was visible in each of the subunits of ClpP-CMK, although it was relatively

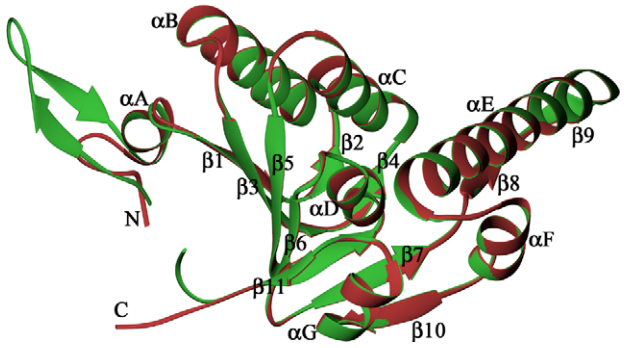


Fig. 2. Superposition of Cα atoms of single monomers from *E. coli* ClpP and ClpP-CMK. Native ClpP (PDB 1TYF, monomer N) is drawn as a red ribbon and ClpP-CMK (PDB 2FZS, monomer B) is drawn as a green ribbon. The rmsd for superposition of residues 18–189 was 0.27 Å.

mobile in native ClpP. Asn193 makes salt bridge with Lys84 from the adjacent subunit, and thus contributes to structural stability of the heptamer (Fig. S1). In human ClpP, the residues in positions 192 and 193 also made contact with the adjacent subunit and affected the folding and stability of the heptamer (Gribun et al., 2004; Kang et al., 2004). In the ClpP-CMK crystal, the C-terminal region appears to be stabilized by a molecule of PEG, incorporated during crystallization, for which electron density corresponding to the three repeating units was clearly defined in each of fourteen monomers (figure not shown). The PEG interacts through a van der Waals contact with the phenolic ring of Tyr60 and forms hydrogen bond to the guanidinium group of Arg192 and the hydroxyl group of Tyr62.

The individual subunits ClpP-CMK and native ClpP aligned very closely, but there was a small deviation in the alignment of the two tetradecamers. When the ClpP-CMK and native ClpP tetradecamers were superimposed by first optimally aligning a single subunit from each, several of the remaining subunits were slightly displaced from each other (Fig. S2). In the affected ClpP-CMK subunits, the β-sheet, which contributes to the interface between the two rings, and α-helix B, which is exposed to the surface of the degradation chamber, were displaced toward the center. As a result of the slightly denser packing, the diameter of the ClpP-CMK chamber is reduced by ~1.5 Å. Little if any displacement was observed between the two apposing heptamers, indicating that the ring interfaces are not greatly altered as a result of the peptide in the active site.

### 3.3. Active site structure and inhibitor binding

Electron density was present in nearly equal intensity in the active sites of all 14 monomers, consistent with essentially complete occupancy by the inhibitor. The inhibitor is covalently linked to Ser97 and His122 (Figs. 4A and B). All atoms of the CMK molecule had well-defined positions, and only the phenyl ring of the protecting benzyloxycarbonyl group showed evidence of disorder in weaker density peaks and elevated B-factors. Upon addition of several water molecules, the model of the active site in the ClpP-CMK complex proved to be relatively complete as indicated by the lack of the positive and negative peaks in the $F_o - F_c$ electron density maps contoured at the 2.7 and −2.7 σ levels, respectively.

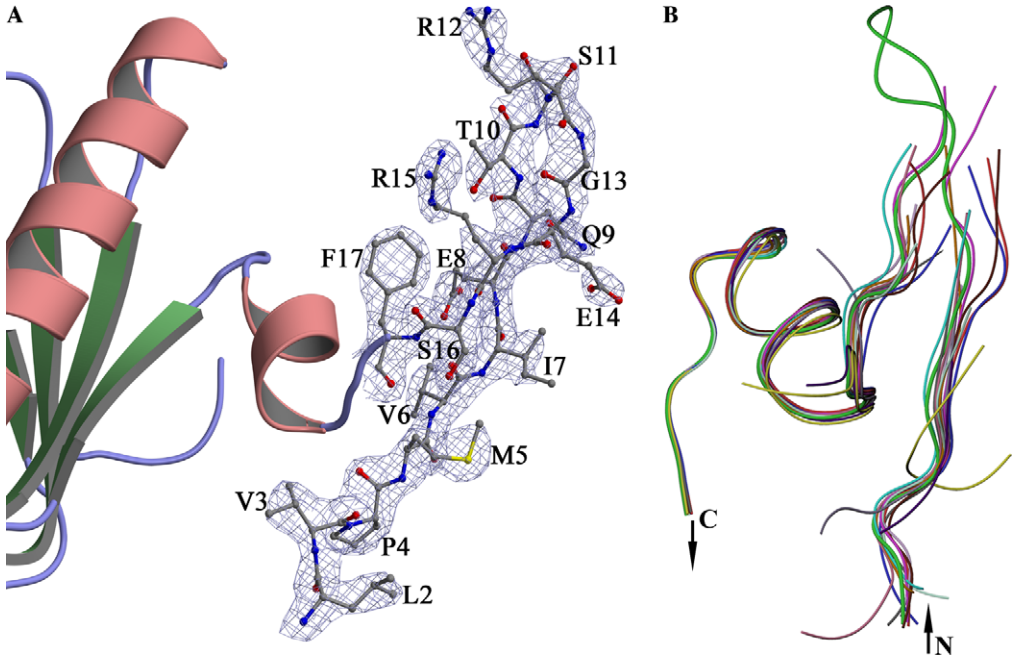


Fig. 3. The N-terminal residues 2–17 ClpP-CMK form a β-hairpin. (A) A $2F_o - F_c$ electron density map was contoured at the level of 1.0 σ for the N-terminal residues 2–17 of monomer B of ClpP-CMK. (B) An overlay was made of the carbon backbone of the N-terminal loops for all 14 monomers from the two rings of ClpP-CMK. The complete backbone was traced only in monomer B (green). The partial electron density for the other monomers indicates the high mobility of this region.

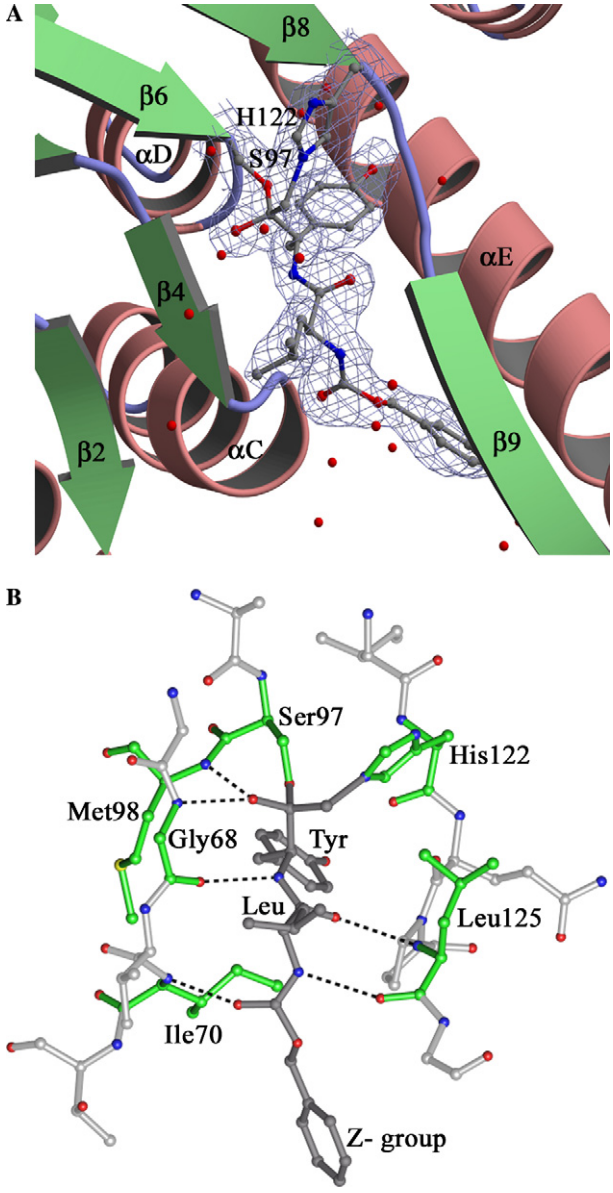


Fig. 4. Structure of the inhibitor, Z-LY-CMK, bound to ClpP. (A) The Z-LY-CMK adduct is shown in ball and stick representation and ClpP is represented in a ribbon diagram. The inhibitor is lodged in the cleft between the globular head domain and the α/β handle and forms two covalent bonds with Ser97 and His122 for which the side chains are drawn as balls and sticks. A $2F_o - F_c$ electron density map contoured at level 1.0 σ and an $F_o - F_c$ electron density map contoured at level 2.7 and −2.7 σ are shown for the inhibitor molecule and the side chains of Ser97 and His122, respectively. (B) The multiple bonding interactions between Z-LY-CMK and ClpP are shown. Hydrogen bonds between Z-Leu-Tyr-CMK and β-strand 4 and the N-terminal portion of β-strand 9 of ClpP are shown as dashed lines. Gly68, Ile70, and Leu125 create four hydrogen bonds with the peptide backbone of the inhibitor, while the backbone amide nitrogens of Gly68 and Met98 form an "oxyanion hole" and make hydrogen bonds to the hemiketal oxyanion of the inhibitor. The inhibitor is attached to ClpP by two covalent bonds, one between the carbonyl carbon of the inhibitor and Oγ of Ser97 and the other between the methylene group of the inhibitor and Nε2 of His122.

In the active site, the hydroxyl group of Ser97 forms a tetrahedral adduct with the carbonyl carbon of the inhibitor. To react with the ketone carbonyl carbon, the side chain of Ser97 is rotated ~90° towards relative to its position in the native state of ClpP. The Nε2 imidazole nitrogen from His122 is covalently joined to the methylene moiety of the CMK. Completing the tetrahedral complex, the hemiketal oxygen atom of the inhibitor is positioned in the oxyanion hole formed by the backbone amide nitrogens of Gly68 and Met98. Both main chain nitrogen atoms form hydrogen bonds with the oxyanion in the tetrahedral complex. The third catalytic residue, Asp171, makes a hydrogen bond with Nδ1 of His122, and an additional hydrogen bond is formed with Nε2 of His138 from the adjacent monomer.

The peptide backbone of the inhibitor is located between the short strand β4 and the N-terminal portion of strand β9 (Fig. 4A). In all, five hydrogen bonds are formed between the inhibitor molecule and the backbone atoms of Gly68, Ile70, and Leu125 (Fig. 4B). These interactions preserve a hydrogen bond network present in the native state of ClpP, in which the substrate binding sites are occupied by water molecules that are displaced upon binding of the inhibitor. The leucine side chain of the inhibitor is rotated away from the surface of ClpP and makes a relatively weak hydrophobic interaction with the side chain of Val69. The phenyl group of the benzyloxycarbonyl moiety is somewhat mobile but makes van der Waals contacts with Ile70, Ile142, Val145, and the side chain carbons of Glu141. These residues combine to produce a hydrophobic platform that provides stabilizing interactions with

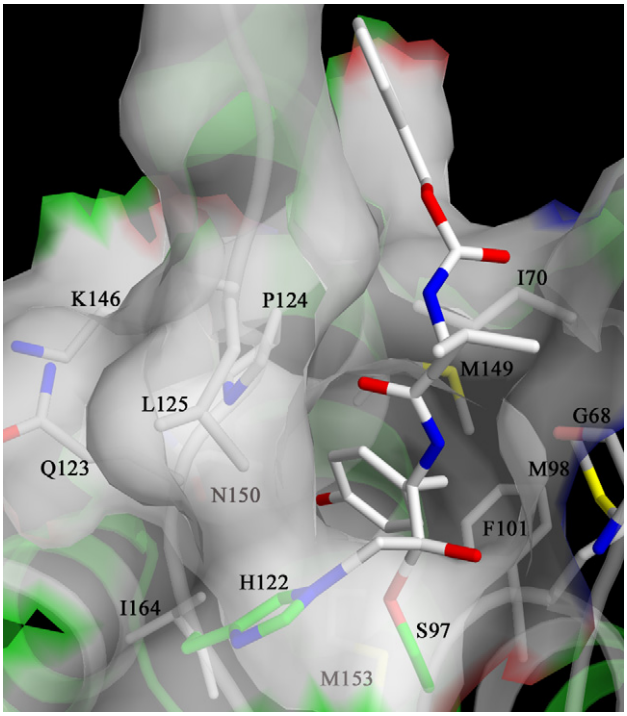


Fig. 5. The S1 substrate-binding pocket of ClpP. The S1 binding pocket of ClpP is depicted as a semi-transparent gray surface with the residues contributing to the pocket shown in the background as sticks. The inhibitor, Z-Leu-Tyr-CMK, is represented as a white stick, and active site residues Ser97 and His122 are represented as green sticks.

substrates through the side chains of the P3 and P4 residues of substrates.

### 3.4. The S1 substrate pocket

The position and conformation of the inhibitor is further stabilized by hydrophobic interactions between the P1 tyrosine residue and the S1 substrate pocket of ClpP. The phenolic ring of tyrosine fits into a deep hydrophobic pocket (Fig. 5). The entrance to the pocket is an oval constriction lined with side chain atoms from Ile70, His122, Pro124, and Leu125 and backbone atoms from Gly68, Ser97, and Met98. The pocket opens up into a large hydrophobic cavity formed primarily from the hydrophobic portions of the side chains of Phe101, Gln123, Pro124, Lys146, Met149, Asn150, Met153, and Ile164. The tyrosine ring is stabilized in the pocket by a hydrogen bond to a water molecule that makes a bridge to the amide nitrogen of Asn150 lying at the bottom of the pocket. The pocket is spacious enough to accommodate the side chain of any amino acid, and there is a marked absence of charged groups within the pocket.

### 3.5. The N-terminal hairpin

The electron density in the area of the N-termini of ClpP varied greatly between individual subunits. The positions and conformations of residues 3–5 were reasonably well defined for all the 14 monomers, but the polypeptide segments between residues 7 and 17 were more fragmented. A complete backbone was traced in only one monomer (chain B) in which the electron density peaks could be interpreted for all but the first residue (Fig. 3A). The N-terminal loop of monomer B is stabilized by the hydrogen bonding to monomer K from symmetry related ClpP-CMK molecule (Fig. S3). In this subunit, the N-terminal residues Leu2 through Ile7 interact with the inner wall of axial channel, while residues Glu8 to Arg15 extend beyond the apical surface of tetradecamer. Residues 5 through 17 adopt a β-hairpin conformation, in which a series of anti-parallel hydrogen bonds is formed between the two segments, Met5-Gln9 and Glu14-Phe17. The turn is made by residues Ser11, Arg12, and Gly13, resulting in the positively charged guanidino group being positioned at the apex of a loop protruding above the ring surface and toward the axial pore (Fig. 3A). Asp18 and Ile19 connect the hairpin to α-helix A. The partially determined β-hairpins in the other subunits align quite well with this region of chain B (Fig. 3B), suggesting that the N-terminal β-hairpin is adopted in all subunits. The lack of completely defined density in most of the subunits indicates that the loops are mobile.

## 4. Discussion

### 4.1. Global structural changes in ClpP-CMK

The ClpP chamber is large enough to accommodate ~450 residues of polypeptide (Kessel et al., 1995; Wang et al., 1997), but, with active ClpP, it is unlikely that intact proteins would accumulate within the chamber, because peptide bond cleavage rates can exceed $10^4\,min^{-1}$ per tetradecamer ClpP (Thompson and Maurizi, 1994), whereas unfolding and translocation rates are much lower (Kenniston et al., 2003, 2005; Singh et al., 2000, 1999; Weber-Ban et al., 1999). In the current model of ClpP action, polypeptides enter the chamber through the axial channels and are cleaved rapidly, followed by rapid diffusion of peptide products out of the chamber. There is still little known about binding of polypeptides in the chamber or at the active sites, conformational changes caused by polypeptide binding, rates of substrate entry and product release, or the channels by which peptide substrates enter and products exit the chamber.

This study was initiated to learn whether substrate binding at the active sites in ClpP affects the structure or properties of the ClpP degradation chamber or alters the size and accessibility of the entry or exit channels. We obtained a high-resolution structure of ClpP with a peptide substrate covalently bound at the active site. Remarkably, ClpP subunit conformation and interactions between ClpP subunits in the tetradecamer are largely unaffected when the equivalent of a tripeptide substrate is bound at all 14 active sites. One difference observed in the ClpP-CMK compared to earlier structures of native ClpP (Wang et al., 1997) is in the N-terminal region, where a β-hairpin that contributes to the substrate entry channel is well defined. Changes at the active site itself affect side chain positions rather than backbone configurations, and there are no displacements in the regions linking the substrate binding sites to the axial channel and the N-terminal loops.

One subtle difference is a denser subunit packing in ClpP-CMK compared to native ClpP. Subunits of ClpP-CMK are displaced ~1.5 Å toward the center of the chamber, most notable for α-helix B, which lines the surface of the chamber. Another change is seen at the C-terminus, where Asn193 makes hydrogen bonds or salt bridges with Lys84 from the adjacent subunit at five or six of the seven subunit interfaces in each ring. The presence of substrate thus appears to favor a more collapsed state of the rings, which could affect the hydration of the molecule and contribute to the overall stabilization of the solvent exposed secondary elements, such as the N-terminal β-hairpin. None of the changes suggests an obvious mechanism for peptide product release.

### 4.2. Residues 7–17 form a loop surrounding the axial channel

An important result of this study is the confirmation of the structure of the N-terminus of ClpP. The β-hairpin structure observed in ClpP-CMK is similar to that described for the human (Kang et al., 2004) and *S. pneumonia* (Gribun et al., 2004) ClpPs. The Cα atoms of the corresponding fragments in *S. pneumonia* (monomer D, residues Met1-Ser10, Glu13-Met29) and hClpP (monomer E, residues Pro1-Val7, Tyr17-Cys30) superimpose with the rmsd

values of 0.46 Å and 0.72 Å, respectively. The finding of similar conformations in the N-termini of ClpP from three different sources indicates that this structure is highly conserved. This structure combined with the high degree of amino acid sequence homology in the N-terminal loop of ClpP (Kang et al., 2004) suggests that the β-hairpin plays an essential role in the biochemical function of these enzymes. The flexible nature of the N-terminal β-hairpin is consistent with a role in altering the structural properties of the axial pore in response to changes in interactions with the associated ClpA and ClpX chaperones.

Studies with the *S. pneumoniae* ClpP (Gribun et al., 2004) provide evidence that the β-hairpin of ClpP can have a gating function controlling the rate of peptide entry. The question arises whether this property is similar to the gating function played in the proteasome by the N-terminal peptides of the alpha subunits (Groll et al., 2000). In response to binding of the PA26 activator, the N-terminal part of the alpha subunit chains radically change position, moving outward to provide a larger channel for peptides (Whitby et al., 2000). In the proteasome, non-identical N-terminal polypeptides interact with one another and undergo a concerted conformational change in response to activator binding. In ClpP-CMK, the N-termini do not interact and bind stably to the channel walls, so it is possible that they do not undergo the kind of concerted motion seen in the proteasome. Interestingly, a recent crystallographic model of ClpP (Bewley et al., 2006) shows the N-terminal loops from the two heptamers in two different states; in one ring, most of the loops are out as in ClpP-CMK, and, in the other ring, the loops are tucked into the channel as in the original model of ClpP (Wang et al., 1997). Those data indicate that the N-terminal loops are indeed dynamic and suggest that movement of the N-terminal peptides may occur during a protein degradation cycle. Deleting portions of the N-terminus ClpP reduced binding to ClpX (Gribun et al., 2004; Kang et al., 2004), raising the possibility that the N-terminus might interact with ClpX when it is in the out configuration, but additional studies are needed to rule out indirect perturbation of portions of the ClpP surface that interact with ClpX.

While peptide products might exit the ClpP chamber through the axial channels, as proposed for the proteasome (Groll et al., 2000; Whitby et al., 2000), another possible exit mode is through the side walls of the ClpP chamber. It was shown with human ClpP (Kang et al., 2005) that ClpX could allosterically regulate interaction between the two rings. More recently, NMR studies of *E. coli* ClpP have also suggested conformational flexibility in the wall of the chamber (Sprangers et al., 2005). The structure of ClpP-CMK shows no effects of substrate binding on the ring interface or on the conformations of residues that make up the chamber wall, and, thus, the current data do not answer the question whether substrate-induced conformational changes play a role in promoting product release from ClpP. An important consideration is that asymmetric occupancy of the active sites could produce sufficient structural strain to open channels or loosen contacts between the rings.

### 4.3. The substrate binding pocket and active site of ClpP

The observed interactions between the peptide portion of the inhibitor and ClpP are consistent with the hydrogen bonding model of substrate binding predicted by Wang et al. (1997). The peptide lies roughly parallel to the equator and is oriented N to C in a clockwise direction viewed from the axial channel. The peptide forms an anti-parallel β-sheet with β-strands 4 and 9 of ClpP, and the side chain of leucine and the phenyl ring of the benzyloxycarbonyl group make van der Waals contacts with relatively featureless hydrophobic surfaces. The orientation of the peptide is consistent with substrates entering the chamber in a C to N direction and directly interacting with one of the active sites in the ring through which they have passed. The side chain of the tyrosine moiety of the inhibitor is stabilized inside the pocket exclusively by hydrophobic interactions. As pointed out by Wang et al. (1997) who docked the side chain of Trp the S1 pocket, the properties, shape, and size of the pocket accommodate a wide variety of side chains in the P1-residue of the substrate. Nevertheless, ClpP does not cleave randomly and shows a significant degree of discrimination in cleaving after various P1 residues that are charged or have branches at the β-carbon position (R. Sivendran and M.R. Maurizi, unpublished). Also, in the peptide substrate, FAPHMALVPV, cleavage by ClpP is very rapid when the P1 position is occupied by methionine (Maurizi et al., 1990a; Thompson et al., 1994), and, while variants with leucine or tryptophan in the P1 position were also cleaved at a lower rate, no cleavage occurred with other amino acids, such as glutamate, serine, lysine, isoleucine, and quite unexpectedly, tyrosine in the P1 position (Thompson et al., 1994). This binding discrimination is not completely explained by the observed structure around the P1 site in the crystal, suggesting that this region might display conformational variability that influences the manner in which the peptide is positioned in the binding site. One possibility is that the mode of the binding in the tetrahedral intermediate does not reflect the initial substrate–enzyme complex. Because residues in the P′ positions, C-terminal to the scissile bond, are known to affect the cleavage of peptide substrates (Thompson et al., 1994), we are in the process of testing peptide ligands with residues on both sides of the scissile bond that will bind at the active site of ClpP.

ClpP is a unique family of serine proteases (Maurizi et al., 1990b; Pandit and Srinivasan, 2004; Wang et al., 1997), and this structure shows that tetrahedral intermediates at the active site in this family are similar to those observed in other serine proteases. The structure clearly defines the catalytic center and the orientation of the catalytic residues during peptide cleavage. In particular, the important role of the amide nitrogens of Gly68 and Met98 in forming the pocket for stabilizing the oxyanion is confirmed. The overlap in

structures of the catalytic centers of native ClpP and ClpP-CMK suggests that the conformation of the catalytic center itself is maintained in the assembled tetradecamer and not subject to allosteric effects. Changes in ClpP peptidase and protease activity are most likely due to changes in substrate accessibility to the chamber or to binding interactions within the chamber that prevent polypeptides from being positioned in the extended active site.

## Acknowledgments

We thank S.K. Singh for generously providing reagents and ClpA protein for biochemical assays and for help and advice in preparing and characterizing ClpP; Hwa-Yong Yong for preparing the inactivated ClpP used for analysis of the site of modification and confirming the kinetics of inactivation of ClpP with Z-LY-CMK; and Bijan Ahvazi for providing advice on various aspects of the structure refinement. We are grateful to Susan Gottesman and members of her laboratory, as well as to other members of the Maurizi laboratory, for constant discussions and suggestions. This study was funded by the intramural program of the National Cancer Institute, Center for Cancer Research, Bethesda, MD.

## Appendix A. Supplementary data

Supplementary data (Figs. S1–S3) associated with this article can be found, in the online version, at doi:10.1016/j.jsb.2006.03.013.

## References

Adam, Z., Clarke, A.K., 2002. Cutting edge of chloroplast proteolysis. Trends Plant Sci. 7, 451–456.

Beuron, F., Maurizi, M.R., Belnap, D.M., Kocsis, E., Booy, F.P., Kessel, M., Steven, A.C., 1998. At sixes and sevens: characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease. J. Struct. Biol. 123, 248–259.

Bewley, M.C., Graziano, V., Griffin, K., Flanagan, J.M., 2006. The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. J. Struct. Biol. 153, 113–128.

Brandstetter, H., Kim, J.S., Groll, M., Huber, R., 2001. Crystal structure of the tricorn protease reveals a protein disassembly line. Nature 414, 466–470.

Brotz-Oesterhelt, H., Beyer, D., Kroll, H.P., Endermann, R., Ladel, C., Schroeder, W., Hinzen, B., Raddatz, S., Paulsen, H., Henninger, K., et al., 2005. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. Nat. Med. 11, 1082–1087.

Brunger, 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., 1998. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D54, 905–921.

Carson, M., 1997. Ribbons. Methods Enzymol. 277, 493–505.

Choi, K.H., Licht, S., 2005. Control of peptide product sizes by the energy-dependent protease ClpAP. Biochemistry 44, 13921–13931.

Corydon, T.J., Bross, P., Holst, H.U., 1998. A human homologue of *Escherichia coli* ClpP caseinolytic protease: recombinant expression, intracellular processing and subcellular localization. Biochem. J. 331, 309–316.

DeLano, W.L., 2002. DeLano Scientific San Carlos, CA, USA, <http://www.pymol.org/>.

Flynn, J.M., Neher, S.B., Kim, Y.I., Sauer, R.T., Baker, T.A., 2003. Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. Mol. Cell 11, 671–683.

Gottesman, S., Clark, W.P., de Crecy-Lagard, V., Maurizi, M.R., 1993. ClpX, an alternative subunit for the ATP-dependent Clp protease of *Escherichia coli*. Sequence and in vivo activities. J. Biol. Chem. 268, 22618–22626.

Gottesman, S., Wickner, S., Maurizi, M.R., 1997. Protein quality control: triage by chaperones and proteases. Genes Dev. 11, 815–823.

Gribun, A., Kimber, M.S., Ching, R., Sprangers, R., Fiebig, K.M., Houry, W.A., 2004. The ClpP double ring tetradecameric protease exhibits plastic ring–ring interactions, and the N Termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation. J. Biol. Chem. 280, 16185–16196.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., Maurizi, M.R., 1998. Enzymatic and structural similarities between the *Escherichia coli* ATP-dependent proteases, ClpXP and ClpAP. J. Biol. Chem. 273, 12476–12481.

Groll, M., Bajorek, M., Kohler, A., Moroder, L., Rubin, D.M., Huber, R., Glickman, M.H., Finley, D., 2000. A gated channel into the proteasome core particle. Nat. Struct. Biol. 7, 1062–1067.

Hoskins, J.R., Singh, S.K., Maurizi, M.R., Wickner, S., 2000. Protein binding and unfolding by the chaperone ClpA and degradation by the protease ClpAP. Proc. Natl. Acad. Sci. USA 97, 8892–8897.

Jones, T.A., Zou, J.Y., Cowan, S.W., Kjeldgaard, M., 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A47, 110–119.

Kang, S.G., Dimitrova, M.N., Ortega, J., Ginsburg, A., Maurizi, M.R., 2005. Human mitochondrial ClpP is a stable heptamer that assembles into a tetradecamer in the presence of ClpX. J. Biol. Chem. 280, 35424–35432.

Kang, S.G., Maurizi, M.R., Thompson, M., Mueser, T., Ahvazi, B., 2004. Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. J. Struct. Biol. 148, 338–352.

Kang, S.G., Ortega, J., Singh, S.K., Wang, N., Huang, N.N., Steven, A.C., Maurizi, M.R., 2002. Functional proteolytic complexes of the human mitochondrial ATP-dependent protease, hClpXP. J. Biol. Chem. 277, 21095–21102.

Kenniston, J.A., Baker, T.A., Fernandez, J.M., Sauer, R.T., 2003. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an $AAA^+$ degradation machine. Cell 114, 511–520.

Kenniston, J.A., Baker, T.A., Sauer, R.T., 2005. Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing. Proc. Natl. Acad. Sci. USA 102, 1390–1395.

Kessel, M., Maurizi, M.R., Kim, B., Kocsis, E., Trus, B.L., Singh, S.K., Steven, A.C., 1995. Homology in structural organization between *E. coli* ClpAP protease and the eukaryotic 26 S proteasome. J. Mol. Biol. 250, 587–594.

Kraulis, P.J., 1991. MOLSCRIPT—A program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. 24, 946–950.

Lupas, A., Flanagan, J.M., Tamura, T., Baumeister, W., 1997. Self-compartmentalizing proteases. Trends Biochem. Sci. 22, 399–404.

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

Maurizi, M.R., Clark, W.P., Kim, S.H., Gottesman, S., 1990b. Clp P represents a unique family of serine proteases. J. Biol. Chem. 265, 12546–12552.

Maurizi, M.R., Thompson, M.W., Singh, S.K., Kim, S.H., 1994. Endopeptidase Clp: ATP-dependent Clp protease from *Escherichia coli*. Methods Enzymol. 244, 314–331.

Mogk, A., Dougan, D., Weibezahn, J., Schlieker, C., Turgay, K., Bukau, B., 2004. Broad yet high substrate specificity: the challenge of $AAA^+$ proteins. J. Struct. Biol. 146, 90–98.

Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D 53, 240–255.

Navaza, J., Saludjian, P., 1997. aMoRe: an automated molecular replacement program package. Methods Enzymol. 276, 581–594.

Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collecting in oscillation mode. Methods Enzymol. 276, 307–326.

Pandit, S.B., Srinivasan, N., 2004. Identification and analysis of a new family of bacterial serine proteinases. In Silico Biol. 4, 563–572.

Perrakis, A., Harkiolaki, M., Wilson, K.S., Lamzin, V.S., 2001. ARP/wARP and molecular replacement. Acta Crystallogr. D: Biol. Crystallogr. 57, 1445–1450.

Pickart, C.M., Cohen, R.E., 2004. Proteasomes and their kin: proteases in the machine age. Nat. Rev. Mol. Cell Biol. 5, 177–187.

Powers, J.C., Lively 3rd, M.O., Tippett, J.T., 1977. Inhibition of subtilisin BPN′ with peptide chloromethyl ketones. Biochim. Biophys. Acta 480, 246–261.

Sauer, R.T., Bolon, D.N., Burton, B.M., Burton, R.E., Flynn, J.M., Grant, R.A., Hersch, G.L., Joshi, S.A., Kenniston, J.A., Levchenko, I., 2004. Sculpting the proteome with AAA(+) proteases and disassembly machines. Cell 119, 9–18.

Schmidt, M., Lupas, A.N., Finley, D., 1999. Structure and mechanism of ATP-dependent proteases. Curr. Opin. Chem. Biol. 3, 584–591.

Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S., Maurizi, M.R., 2000. Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. Proc. Natl. Acad. Sci. USA 97, 8898–8903.

Singh, S.K., Guo, F., Maurizi, M.R., 1999. ClpA and ClpP remain associated during multiple rounds of ATP-dependent protein degradation by ClpAP protease. Biochemistry 38, 14906–14915.

Sprangers, R., Gribun, A., Hwang, P.M., Houry, W.A., Kay, L.E., 2005. Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. Proc. Natl. Acad. Sci. USA 102, 16678–16683.

Thompson, M.W., Maurizi, M.R., 1994. Activity and specificity of *Escherichia coli* ClpAP protease in cleaving model peptide substrates. J. Biol. Chem. 269, 18201–18208.

Thompson, M.W., Singh, S.K., Maurizi, M.R., 1994. Processive degradation of proteins by the ATP-dependent Clp protease from *Escherichia coli*. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis. J. Biol. Chem. 269, 18209–18215.

Wang, J., Hartling, J.A., Flanagan, J.M., 1997. The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. Cell 91, 447–456.

Weber-Ban, E.U., Reid, B.G., Miranker, A.D., Horwich, A.L., 1999. Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. Nature 401, 90–93.

Whitby, F.G., Masters, E.I., Kramer, L., Knowlton, J.R., Yao, Y., Wang, C.C., Hill, C.P., 2000. Structural basis for the activation of 20S proteasomes by 11S regulators. Nature 408, 115–120.