Scanning Transmission Electron Microscopy and Small-Angle Scattering Provide Evidence That Native *Escherichia coli* ClpP Is a Tetradecamer with an Axial Pore†

John M. Flanagan,* Joseph S. Wall, Malcolm S. Capel, Dieter K. Schneider, and John Shanklin*

Department of Biology, Brookhaven National Laboratory, Upton, New York 11973

Received March 1, 1995; Revised Manuscript Received June 1, 1995®

ABSTRACT: The *Escherichia coli* ATP-dependent caseinolytic protease (Clp) is composed of two distinct subunits; protease, ClpP, and ATPase, ClpA. Active ClpP has been overexpressed to approximately 50% of soluble protein in *E. coli*, and purified to homogeneity. Direct mass determination of individual particles using scanning transmission electron microscopy (STEM) yields a mean native molecular mass of 305 ± 9 kDa for the ClpP oligomer, suggesting that it has a tetradecameric structure. Small-angle X-ray scattering (SAXS) curves were determined for ClpP in solution at concentrations of 1−10 mg/mL. A combination of STEM and SAXS data was used to derive a model for ClpP, comprising a cylindrical oligomer about 100 Å in diameter and about 75 Å in height with an axial pore about 32−36 Å in diameter. The volume of the pore is estimated to be ~70 000 Å^3, similar in size to those found in chaperone proteins, and is large enough to accommodate unfolded polypeptide chains, although most globular folded proteins would be excluded.

Regulation of the accumulation of specific proteins in the cells depends on the rate of both synthesis and degradation of the proteins. In *Escherichia coli*, the bulk of protein turnover is coupled to ATP[1] hydrolysis (Maurizi, 1992). This bacterium has two well-characterized proteases that mediate energy-dependent degradation of methylcasein *in vitro*, Lon and Clp (also referred to as Ti). Homologs of both archetypes have been identified in eukaryotic organisms, but their functions are not understood (Maurizi et al., 1990b). Lon was the first ATP-dependent protease to be purified and characterized from *E. coli*. It is composed of a bifunctional peptide, with both protease and ATPase functions (Charette et al., 1984; Chung & Goldberg, 1981). A second ATP-dependent protease, Clp, originally was purified as residual caseinolytic activity present in Lon$^-$ cells (Hwang et al., 1987; Katayama-Fujimura et al., 1987). In contrast to the Lon protease, Clp is composed of two nonhomologous subunits, a protease ClpP (subunit M_r 21.5 kDa) and an ATPase ClpA (subunit M_r 87 kDa), both of which are required for ATP-dependent degradation of methylcasein (Katayama-Fujimura et al., 1987; Gottesman et al., 1990). In this system, the substrate cannot be directly hydrolyzed by the protease, but must first associate with a presenting subunit. Different substrates require different presenting subunits; for example, λ O protein is degraded by ClpX−ClpP, but not by ClpA−ClpP (Wojtkowiak et al., 1993). Recently, ClpA was shown to function also as a chaperone protein, suggesting a possible link between folding and degradation pathways (Wickner et al., 1994). A second link between these processes is suggested by the GroEL−GroES-dependent degradation of the synthetic substrate CRAG by ClpP in the absence of ClpA or ClpX (Kandror et al., 1994).

The oligomeric structure of purified ClpP was reported to be between a decamer and a dodecamer, based on size-exclusion chromatography (Hwang et al., 1988; Maurizi et al., 1990a). Electron microscopy of negatively stained images of purified ClpP were interpreted to show two stacked rings composed of irregular hexagons, with a central depression or cavity ~110−120 Å in diameter and ~100 Å in height (Maurizi et al., 1990a).

We extended these studies by determining the mass of large numbers of individual ClpP molecules using STEM and show that the native oligomeric structure is a tetradecamer. We have also used SAXS and SANS to probe the low-resolution structure of ClpP in solution. A hollow cylinder model ~75 Å high with a diameter of ~100 Å and a large 32−36-Å axial pore best fits these data and is consistent with the STEM images of negatively stained specimens.

EXPERIMENTAL PROCEDURES

Materials. Enzymes for nucleic acid modifications were purchased from either Boehringer Mannheim or New England Biolabs. Oligonucleotides with the desired sequences were synthesized with an Applied Biosystems DNA synthesizer. The fluorescent peptide *N*-Suc-LY-AMC was purchased from Sigma. All other reagents were of analytical grade.

Cloning and Expression of E. coli ClpP. The ClpP gene was cloned by PCR amplification of *E. coli* genomic DNA

† This work was supported by grants from the Office of Basic Energy Sciences, U.S. Department of Energy (to J.S.), Brookhaven National Laboratory Directed Research and Development Program no. 93−32. The STEM facility is supported by NIH Biotechnology Resource Grant No. RR01777 (to J.W.), and the National Synchrotron Light Source beam lines U9 and X12B by the Office of Health and Environmental Research, Department of Energy Grant Nos. KP04-01 and BO-34.

* Authors to whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts,* August 1, 1995.

[1] Abbreviations: ATP, adenosine 5′-triphosphate; CD, circular dichroism spectroscopy; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; EM, electron microscopy; IPTG, isopropyl β-thiogalactoside; N-Suc-LY-AMC, *N*-succinyl-Leu-Tyr-7-amido-4-methylcoumarin; OAc, acetate; PCR, polymerase chain reaction; PMSF, phenylmethanesulfonyl fluoride; MES, 4-morpholineethanesulfonic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid, SANS, small-angle neutron scattering; SAXS, small-angle X-ray scattering; SEC, size-exclusion chromatography; SDS, sodium dodecyl sulfate; STEM, scanning transmission electron microscopy; Tris,tris(hydroxymethyl)-aminomethane.

0006-2960/95/0434-10910$09.00/0

using the following oligonucleotide primers: 5′ oligonucleotide, 5′-CAGGAGACGGTCATGAGTTACAGCGGCGAA-3′; 3′ oligonucleotide, 5′-ACAGTTGCGCGGATCCCAT-CAATTACGATG-3′, derived from the published sequence of ClpP (Maurizi et al., 1990a). The amplified DNA was purified and digested with *Bsp*HI and *Bam*HI for cloning into the *Nco*I/*Bam*HI sites of the pET3d cloning vector. The authenticity of the insert DNA was determined by double-stranded sequencing. To express ClpP, the plasmid (pCLPP3d) was transformed into BL21(DE3) cells.

Purification of ClpP from BL21(DE3)-pCLPP3d. Nine liters of BNTA (10 g of NZ-amine and 5 g of NaCl per liter) + 100 mg/mL of ampicillin were inoculated with cells from a frozen glycerol stock of BL21(DE3)-pCLPP3d. The culture was grown until it reached stationary phase and then harvested by centrifugation. The cell pellets were pooled and stored frozen at −20 °C until they were used. To purify ClpP, the cell pellet was resuspended in 1/40th of the original culture volume in 100 mM Tris·HCl, pH 8.0, 1.0 mM EDTA, 1.0 mM DTT, 0.1% Triton X-100, 0.2 mg/mL hen egg white lysozyme, and 0.02 mg/mL DNase I. The suspension was incubated on ice for 30 min, transferred to a 37 °C water bath for 10 min, and then incubated on ice for an additional 10 min. The viscosity of the solution was reduced by adding $MgCl_2$ to a final concentration of 10 mM before incubating for a further 30 min on ice. Cellular debris was removed by centrifugation at 185000*g* for 40 min, and the supernatant was then applied to a 200-mL Q fast-flow (Pharmacia-LKB) column equilibrated with 50 mM Tris·HCl pH 7.5, 1 mM DTT, and 50 mM NaCl. Protein was eluted using a linear salt gradient from 50 to 600 mM NaCl. Fractions enriched in ClpP were pooled and then further purified by chromatography on a hydroxyapatite (HAP) column using a linear phosphate gradient from 1 to 250 mM $NaKPO_4$, pH 7.5, in 200 mM NaCl and 1.0 mM DTT; ClpP eluted as the major 280-nm absorbing peak. The most highly enriched fractions were pooled and concentrated to approximately 15 mg/mL before size-exclusion chromatography. The final purification step was a TSK-250 size-exclusion column (Bio-Rad). The column was equilibrated in 10 mM MOPS, pH 6.8, 200 mM KCl, 0.1 mM EDTA, and 1.0 mM DTT. The peak fractions from this column contained apparently homogeneous ClpP and were pooled, concentrated to 20 mg/mL, dialyzed against 10 mM Tris·HCl, pH 7.5, 200 mM KCl, 0.1 mM EDTA, and 1.0 mM DTT, and stored at −80 °C.

Kinetic Studies of Peptide and Casein Hydrolysis by ClpP. The peptidase activity of ClpP was assayed at 25 °C by continuously recording the increase in fluorescence due to the released fluorophore as described previously (Woo et al., 1989) in a reaction buffer containing 50 mM Tris·HCl, pH 7.5, 1.0 mM DTT, 1 mM N-Suc-LY-AMC, and various concentrations of purified ClpP. The wavelengths of excitation and emission were 380 and 460 nm, respectively. The values reproted for V_{max} and K_m were obtained from a double-reciprocal plot by linear regression.

The ability of the recombinant ClpP to hydrolyze casein was assayed in the presence of ClpA and ATP by measuring the release of TCA-soluble counts from [^{3}H]methylcasein as described previously (Katayama-Fujimura et al., 1987). The 0.1-mL reaction mixture contained 50 mM Tris·HCl, pH 8.0, 1.0 mM DTT, 25 mM $MgCl_2$, 4 mM ATP, 50 mM KCl, 24 mg (~60 000 cpm) of [*methyl*-^{3}H]methylcasein, saturating amounts of ClpA, and various concentrations of ClpP.

Amino-Terminal Sequencing of Purified ClpP. Purified ClpP was run on a 12% SDS−polyacrylamide gel. The protein band was electroblotted to a ProBlot (ABI) membrane and visualized by Ponceau-S staining. The ClpP band underwent automatic Edman degradation to determine the sequence of amino-terminal residues.

Scanning Transmission Electron Microscopy. STEM was performed using the Brookhaven Biotechnology Resource Instrument (Wall, 1979; Wall & Hainfeld, 1986). The instrument was operated at a voltage of 40 keV, and all specimens were maintained at −150 °C on the liquid N_2-cooled stage. Freeze-dried specimens were prepared as described (Makhov et al., 1993; Mosesson et al., 1981). In summary, a 20−30 Å thick carbon film was deposited from a carbon arc to single-crystal rock salt in an ion-pumped bell jar. This carbon was floated off the rock salt onto distilled water and picked up from the air side by touching it with a Ti grid coated with a holey carbon film. TMV (3 μL of a 100 μg/mL stock) was injected into the drop of water retained on the surface, washed with 20 mM NH_4OAc, and a 3-μL sample, containing 1−5 μg/mL ClpP, was added to a drop of 20 mM NH_4OAc on this film. In some cases the carbon films were pretreated with detergent, either SDS or β-octyl glucoside, and then washed several times with distilled water to remove excess detergent (Wall & Hainfeld, 1986). No glow-discharge treatment was used. Treating the carbon films with detergents alters the affinity of some proteins for the carbon's surface (M. Simon and J. S. Wall, unpublished data). After allowing 1 min for attachment, the grids were washed 10−15 times by partial wicking with filter paper and adding 5 μL of buffer. Finally, the excess fluid was removed by pinching the grid between pieces of filter paper and then rapidly immersing it in liquid nitrogen slush to freeze the sample. The grids were then dried under vacuum facing a liquid nitrogen cold trap for 6−8 h while warming from −140 to −90 °C at a rate of ~8°/h. Negatively stained specimens were prepared similarly, except that the samples were air-dried after staining with 2% (w/v) vanadate methylamine (nanoVan, Nanoprobes Inc.) or 2% (w/v) uranyl acetate. Mass analysis of the image data was carried out as described by Wall and Hainfeld (1986) or with an automated mass analysis program (J. S. Wall, unpublished data). The only departure from the previous mass-measurement protocol was our use of an absolute STEM mass calibration rather than the TMV control particles in each field.

Small-Angle Scattering Measurement and Analysis. SAXS data were collected on beam line X12B at the National Synchrotron Light Source (NSLS) using a 2D multiwire detector system (Capel et al., 1995). The correction for the dependence of the scattering curves on protein concentration to obtain the scattering curve at infinite dilution was described previously (Kataoka et al., 1989). In all cases, the final dialysate was used as a blank.

Purified ClpP used in the SAXS experiments was dialyzed into a buffer containing 50 mM Tris·HCl, pH 7.5, 200 mM KCl, 1 mM DTT, 0.1 mM EDTA, and 10% glycerol. The sample holders of 1 mm path length were made of KelF, and the windows were made of Kaptan. The samples were maintained at 25 °C throughout the experiment with a water-jacketed sample holder and a circulating water bath in a

helium-purged shroud to help control background due to gas scatter. The integration time for all samples was 15 min.

The scattering data are presented as the relative scattering intensity $I(Q)$, where the momentum transfer, Q, is defined as

$$Q = 4\pi \sin \theta/\lambda$$

where λ is the wavelength (1.38 Å), and 2θ is the full scattering angle. Following point-by-point extrapolation to zero protein concentration at each Q, Guinier analysis (Guinier & Fournet, 1955) was performed on the buffer-subtracted curve to estimate the radius of gyration, R_g. Both the linearity of the resultant Guinier plots and the scattered intensity at zero angle, $I(0)$, were used to verify that samples were free of aggregates. We have also analyzed the scattering data using the radial distance distribution function, $P(r)$. The $P(r)$ functions were calculated from the real-space algorithms of both Glatter (1980) and Svergun (1992), or according to

$$P(r) = \left(\frac{1}{2\pi^2}\right) \int I(Q)(Qr) \sin(Qr)\, dQ$$

using the indirect Fourier inversion algorithm described by Moore (1980). The $P(r)$ curve calculated by this method requires as input an estimate of the longest intramolecular vector, d_{max}. For ClpP, d_{max} was chosen as described previously (Flanagan et al., 1993). The R_g of the particle was also estimated from the second moment of the $P(r)$ function by the equation

$$R_g^2 = \frac{\int r^2 P(r)\, dr}{2\int P(r)\, dr}$$

Small-angle neutron scattering experiments were performed on the H9B beam line at the Brookhaven's HFBR (Schneider & Schoenborn, 1984). A wavelength of 4.8 Å was used, and the beam was collimated to a diameter of 6 mm at the sample. Protein solutions, *E. coli* ClpP, horse spleen apoferritin, and β-amylase, were prepared as described for SAXS experiments, and the samples were contained at 7 °C in spectroscopic quartz cuvettes of 1 mm path length; their absorbance at 280 nm was measured. Before SANS measurements, each protein was repurified by chromatography on a TSK-250 SEC column to eliminate any high molecular weight aggregates. The scattering of the protein solutions and the buffer was recorded with an area detector in incremental cyclic exposures for 3 h per sample. After buffer subtraction, the net forward scattering intensity, $I(0)$, was about 400 counts per unit concentration of 1 mg/mL. The Guinier analysis was performed with data in the range of Q between 0.02 and 0.04 Å^{-1}.

RESULTS

Purification and Characterization of ClpP Protein Produced in BL21(DE3)-pCLPP3a. In the absence of an exogenous inducer, ClpP expressed in Bl21(DE3)-pCLPP3a accumulated to 40−50% of the total soluble protein and was purified to apparent homogeneity, as described under Experimental Procedures, with an overall yield of greater than 40%. *E. Coli* ClpP is synthesized with a 14 amino acid precursor peptide at the amino terminus that is autoproteolytically removed (Maurizi et al., 1990a,b). Amino-terminal sequence analysis of the purified protein verified that the 14-residue amino-terminal leader sequence was correctly processed (the sequence of the first 10 amino acids was ALVPMVIEQT) and is an indication that ClpP is active. The best peptide substrate for ClpP is N-Suc-LY-AMC (Woo et al., 1989; Arribas & Castano, 1993). In agreement with the previously reported kinetic parameters for this substrate (Woo et al., 1989; Arribas & Castano, 1993), our purified ClpP had a K_m value of 655 ± 50 μM with a V_{max} of 420 ± 50 pmol min^{-1} (μg of protein)$^{-1}$. Moverover, this ClpP was also catalytically active, in a ClpA- and ATP-dependent manner, against [^{3}H]methylcasein. The apparent native molecular mass of ClpP was determined by size-exclusion chromatography to be ~250 kDa, in agreement with several other studies (Hwang et al., 1988; Maurizi et al., 1990a; Arribas & Castano 1993). The CD spectrum of ClpP shows several prominent features in the range between 260 and 178 nm, such as the presence of a large maximum in ellipticity at 190 nm and two minima at 208 and 222 nm suggesting a high degree of α-helical secondary structure. The secondary structure content of ClpP, determined from this spectrum by the methods of Hennessey and Johnson (1981) and Johnson (1988), indicate an α-helix content of 50%, a β-sheet content of 14%, and a turn content of 13%, with the remainder, 20%, being assigned to nonstandard structure. These biochemical data, together with the sequencing of the cloned DNA used for expression, assure that the protein we are overexpressing and purifying is the authentic ClpP protein.

Determination of the Molecular Mass and Oligomeric State of Native ClpP. Previous studies suggested that the native ClpP exists as a dodecamer, on the basis of its elution volume from SEC (Hwang et al., 1988; Maurizi et al., 1990a). Using SEC, a protein's elution volume depends upon its Stokes radius, which often is directly related to its molecular weight. However, for proteins whose structures are elongated or hollow or those which interact with the column matrix, the molecular weight estimated by SEC can differ significantly from the true molecular weight. To further characterize the native oligomeric state of ClpP, we measured the mass of the particle in a freeze-dried state by STEM. This technique allows the mass analysis of single molecules, as opposed to methods of mass measurement which assess the average mass of an ensemble of particles. In STEM, the mass of unstained particles is determined from the intensity of scattered electrons, which is directly proportional to the total number of atoms in the electron beam, weighted by their atomic number (Wall & Hainfeld, 1986).

Initial STEM experiments for *E. coli* ClpP indicated that the protein interacts strongly with the carbon substrate, causing overcrowding of the grid and the presence of an amorphous denatured protein film, even when deposited at a concentration of 10 μg/mL. To improve the quality of the STEM images, we examined the effects of various solvents, pretreatments of the carbon grids, and fixation procedures. The most significant improvement was obtained when ClpP was deposited at a concentration of 1-5 μg/mL from a buffer containing 50 mM triethanolamine·HCl, pH 7.5, and 200 mM KCl. In the first successful specimen, 100 ClpP particles that appeared to be intact (i.e., having clean sharp edges), were well isolated, and were in regions of the grid free of denatured protein were subjected to mass analysis, giving a mean mass of 303 ± 27 kDa. On the

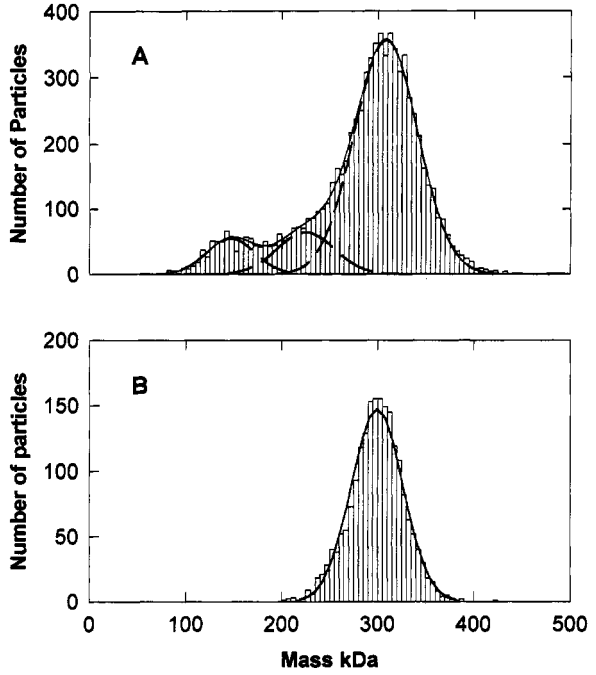

FIGURE 1: Mass distribution of ClpP particles determined by STEM. (A) The mass distribution of 8020 independent particles was determined using an automated mass analysis program. The entire mass distribution could be described by three Gaussian components at 308 ± 28 kDa (80%), 224 ± 28 kDa (11%), and 148 ± 25 kDa (9%) (dashed lines). The solid line shows the curve described by the sum of these Gaussian components. (B) Mass distribution of 1100 ClpP molecules selected on the basis of clear edge definition. The ClpP was deposited from a buffer containing 50 mM triethanolamine·HCl, pH 7.5, 200 mM KCl. The grids were prepared as described in the text.

basis of a molecular weight of 21.5 kDa for the ClpP monomers, this result suggests that native *E. coli* ClpP is a tetradecamer (14-mer) rather than a dodecamer (12-mer) or a decamer (10-mer). To eliminate the possibility that a bias toward higher molecular weight species was introduced by the selection criteria, we determined the mass of 8020 particles using an automated procedure. The only selection criterion used to determine suitability was that the particles were well isolated; Figure 1A shows the resulting mass distribution. The most prominent feature is a large peak centered at 308 kDa. The entire mass distribution could be analyzed in terms of three Gaussian components, one at 308 ± 28 kDa (80% of the total particles), a second at 224 ± 28 kDa (11%), and a third at 148 ± 25 kDa (9%). The individual Gaussians are shown as dashed lines, and the sum of the three components is shown as a solid line. The majority of the particles corresponded to a tetradecamer (~300 kDa), although two minor species corresponding to a decamer (~220 kDa) and a heptamer (~150 kDa) were also observed. Examination of images corresponding to the ~150-kDa species revealed that they are similar in shape (roughly circular) and dimension to the majority of particles found in each field, except that their intensity was approximately half of the value of the other particles. To determine whether the ~220-kDa species is a degradation product of the ~300-kDa species produced in preparing the STEM samples, we selected 1100 well-isolated particles solely on the basis of the criterion of clear edge definition. In all cases, between one-half and two-thirds of all the particles in a given field passed the selection criterion, comparable to the number that passed using the automated

Table 1: STEM Mass Analysis of 11 Replicate *E. coli* ClpP Samples

TMV mass per unit length (kDa/Å)	ClpP mass (kDa)
13.24 ± 0.12 (10)[a]	301.1 ± 21.3 (100)
12.94 ± 0.15 (15)	287.9 ± 20.5 (100)
13.18 ± 0.23 (9)	315.8 ± 24.0 (99)
13.14 ± 0.26 (13)	301.7 ± 24.5 (100)
13.15 ± 0.17 (12)	306.6 ± 25.6 (100)
12.88 ± 0.18 (9)	301.0 ± 23.8 (100)
12.85 ± 0.26 (21)	310.1 ± 23.8 (100)
12.92 ± 0.38 (7)	304.4 ± 21.2 (100)
13.14 ± 0.08 (12)	314.3 ± 25.5 (100)
13.01 ± 0.21 (8)	306.3 ± 19.4 (110)
12.82 ± 0.31 (10)	305.0 ± 24.5 (100)
average of means = 13.02	average of means = 305.0
SD of means = 0.15	SD of means = 8.7

[a] The quantity in parentheses refers to the number of observations.

mass determination (Figure 1B). In contrast to the mass distribution generated without any selection criteria (Figure 1A), a single peak at 298 ± 27 kDa corresponding to a tetradecamer emerged from this analysis, suggesting that the subsidiary peak at ~220 kDa corresponds to poorly defined particles, which probably are degraded. Moreover, three independent lines of evidence suggest that purified ClpP is homogeneous in solution and that the heterogeneity observed in Figure 1B is introduced during deposition on the carbon film. First, purified ClpP elutes as a single sharp peak in HPLC size-exclusion chromatography, and STEM analysis of column fractions eluting from the leading edge, center, or trailing edge of this peak gave identical mass distributions (J. M. Flanagan, unpublished data). Second, the SAXS and SANS curves of ClpP show a well-defined subsidiary maximum and minimum which would not be present in a heterogeneous population (see below). Third, dynamic light scattering measurements of ClpP are consistent with a monodispersed solution (data not shown).

The errors associated with the mean mass of ClpP for single-particle measurements and not the standard error of the mean are quoted above. The standard error of the mean for the distribution in Figure 1B is ~1 kDa ($\sigma/N^{1/2}$, where N = 1100). Using the standard error is only valid if the contribution of systematic errors to the total error is small. To independently estimate the error associated with the mean, we analyzed the mass distribution obtained from 11 replicate samples (each a separate grid). The replicates each contained many high-quality TMV images used as an internal calibration check and showed very little denatured protein in the background. Individual particles were picked only on the basis of sharp edge definition. Table 1 shows the results of STEM mass analysis for ClpP and TMV. The average of the individual means for ClpP was 305 kDa, and the standard deviation of the means was 9 kDa. For data with minimal systematic errors, the error in the mean should be ~2.3 kDa ($1/N^{1/2}$, where N = 100 measurements). The observed standard deviation of the means is ~9 kDa, compared with ~23 kDa for the standard deviation of the individual measurements, and is consistent with a substantial random component and a smaller nonrandom component of error. The average mass per unit length of the TMV was 13.02 kDa/Å and compared favorably with the known value of 13.1 kDa/Å. The standard deviation of the means for the mass per unit length of TMV was 0.15, or about 1.2% of the mass per unit length. From this analysis, we believe that the

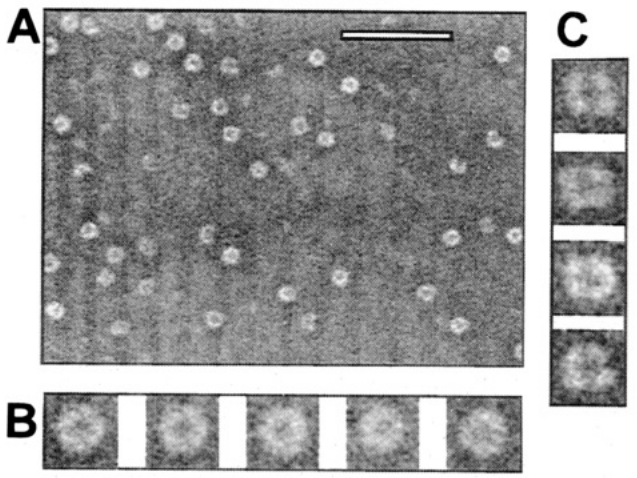

FIGURE 2: STEM micrograph of native *E. coli* ClpP in vanadate negative stain. (A) Field of view is 2650 × 1900 Å; reference bar, 500 Å. (B) End-on views, 146 × 146 Å field. (C) Side-on views, 146 × 146 Å field.

accuracy of the STEM mass measurements is closer to the ~9 kDa observed in the replicate specimens than to the likely underestimate given by the standard error of the mean or the overestimate from the standard deviations of 23 kDa from individual particle analysis. A value of 305 ± 9 kDa rules out the possibility that native ClpP is a 13-mer or a 15-mer (both are 2σ from the mean) and strongly supports an oligomeric state of 14.

Structural Characterization of Vanadate-Stained Images of ClpP. Earlier electron micrograms of ClpP negatively stained with uranyl formate suggest that the molecule is roughly cylindrical with a height of 100 Å and a diameter of 110−120 Å (Maurizi et al., 1990a). End-on views, looking down the axis of the cylinder, suggested that the molecule may have a central cavity or depression and possessed a 6-fold axis of symmetry. To further investigate the structure of the molecule, we obtained images of purified ClpP negatively stained with vanadate, pH 7.0 (Figure 2). These images were chosen for analysis since they were generally of better quality than those obtained from uranyl acetate stained ClpP (pH 3). Most showed a stain-excluding annulus 100 ± 10 Å in diameter. A gallery of typical images is shown in Figure 2B. In the center of these presumed end-on (top) views is a central, roughly circular stain-penetrating pore. Analysis of end-on views of ClpP, in terms of a power spectrum, revealed a preference for 7-fold axial symmetry, but the noise was so severe that this was not judged significant (data not shown). An analysis of symmetry from single particles this small requires prominent features, which ClpP did not have. To characterize the central, stained region, we calculated the radial intensity distribution profile for several images. In all cases, the radial signal showed a peak with a relatively sharp outer edge and a central trough that varied in depth. In several cases, the trough returned to the baseline, indicating that the central portion of ClpP is hollow. From the radial profiles we derived an average pore diameter of 32 ± 2 Å from the width of the trough at half-height. A few images (<10%) appeared rectangular, measuring about 100 × 75 Å. These views also showed a central stain-penetrable region (Figure 2C). We presume that the more rectangular-shaped particles are side views of the same molecules, where the longer axis (100 Å) corresponds to the diameter of the cylinder, and the shorter axis corresponds to its height. In the shorter dimension, the central band probably corresponds to the axial pore, seen in the end-on views, and in the longer dimension, the presence of a band suggests either that the structure is composed of two layers of subunits or that each subunit is bilobed. These interpretations are consistent with those from a previous study which suggested that the ClpP oligomer was cylindrical (Maurizi et al., 1990a), although we infer that the height of the cylinder is 75 Å, about 25 Å shorter than that previously reported.

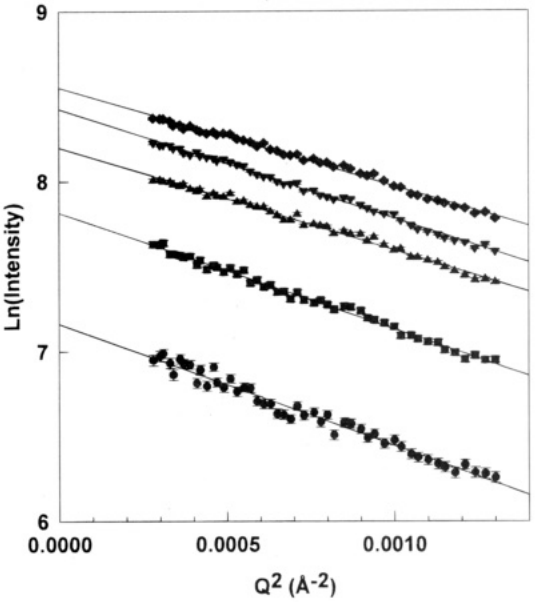

FIGURE 3: Guinier plots of ClpP as a function of protein concentration. The SAXS curves for *E. coli* ClpP were determined at 1 (●), 2 (■), 4 (▲), 8 (▼), and 10 mg/mL (◆).

Small-Angle Scattering of ClpP. Without crystals suitable for X-ray structure analysis, solution scattering of *E. coli* ClpP provides a means of obtaining direct structural information about its molecular arrangement in solution. In contrast to X-ray crystallography, a solution-scattering profile provides information at low resolution from orientationally averaged particles (Glatter, 1982), providing information complementary to other techniques, such as electron microscopy.

In a SAXS experiment, two model-independent parameters can be determined: the forward scatter, $I(0)$, which is proportional to the square of the molecular weight and directly proportional to the concentration of the complex, and the radius of gyration, R_g, which measures the size of the molecule. Figure 3 shows Guinier plots for *E. coli* ClpP as a function of protein concentration between 1 and 10 mg/mL. The linearity of these plots, particularly the absence of upward curvature in the small-angle region, indicates that ClpP solutions are monodisperse. The slope of the Guinier curves decreases with increasing protein concentration, a phenomenon induced by multiple scattering effects and interparticle interference. To minimize these effects upon the measured R_g value, we performed Guinier analysis on a scattering curve obtained by point-by-point extrapolation of the scattering data to zero protein concentration (Kataoka et al., 1989); the R_g for ClpP then is 44.2 ± 0.2 Å.

The molecular weight of the ClpP oligomer was also determined by its small-angle neutron scattering from its $I(0)$ value. Compared with X-ray radiation, neutrons have several advantages in determining the molecular weight of proteins;

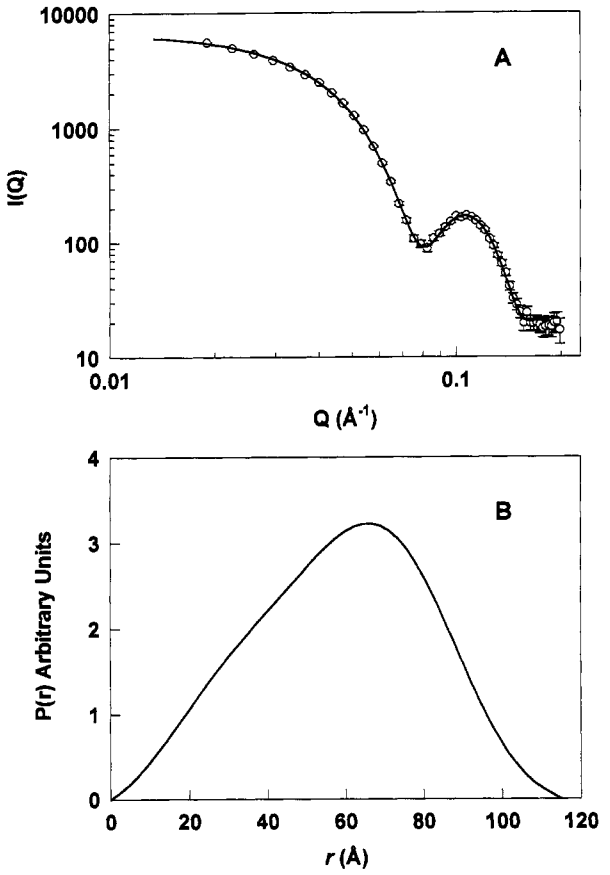

FIGURE 4: Solution scattering curve of ClpP. (A) Comparison of the SAXS curve for a 5 mg/mL ClpP sample (O) with the curve calculated for a hollow cylinder having a height of 75 Å, a diameter of 98 Å, and an axial pore of 35 Å (—). (B) Radial distance distribution function $P(r)$ from SAXS data of ClpP. $P(r)$ curves were calculated with the method of Moore (1980) from data extrapolated to infinite dilution. The radius of gyration obtained as the second moment of the $P(r)$ function was 44.8 Å.

in particular, neutron experiments in ^1H_2O are not sensitive to assumptions about the partial specific volume (Jacrot & Zaccai, 1981), and neutron sources are inherently more stable than synchrotron X-ray sources. The molecular weight of ClpP was obtained using a modification of the method of Jacrot and Zaccai (1981). Here, we eliminated the contributions of instrumental factors using standard proteins of known molecular mass bracketing the expected mass of ClpP (apoferritin, 475 kDa, and β-amylase, 220 kDa). The molecular mass of ClpP was found to be 285−310 kDa; this range is consistent with the ClpP complex containing 13 or 14 subunits. Although the precision of this measurement is not sufficient to distinguish between oligomers containing 13 or 14 subunits, the mass determined in this experiment is consistent with the results of our STEM analysis and is significantly larger than that suggested by SEC (Hwang et al., 1988; Maurizi et al., 1990a).

The modulations of the scattering curve observed at higher angles provide information on the overall structure of the molecule in solution. Figure 4A shows the scattering curve of ClpP collected in the range of Q between 0.02 and 0.13 Å^{-1}; this range corresponds to real spacings between 300 and 50 Å. Its most striking features are the presence of a strong subsidiary maximum centered at a Q of ~0.11 Å^{-1} and a well-defined minimum at $Q \sim 0.08$ Å^{-1}. The presence of the subsidiary maximum and its height relative to the central one are consistent with the existence of an axial pore within ClpP extending through the entire molecule (see below). Moreover, a well-defined minimum and a subsidiary maximum in the scattering curve show that the ClpP preparation is homogeneous for size and structure because these features are highly sensitive to heterogeneity in the sample.

Table 2: Geometrical Parameters and Reduced χ^2 for Best Fits of Analytical Models to the Scattering Data for ClpP at 5 mg/mL.

Hollow Cylinder				
height	radius	pore radius	R_g	reduced χ^2
75.8 ± 0.1 Å	49.1 ± 0.1 Å	18.1 ± 0.1 Å	43.6 ± 0.1 Å	1.9
Hollow Sphere				
radius	pore radius	R_g	reduced χ^2	
49.9 ± 0.1 Å	23.0 ± 0.1 Å	40.3 ± 0.1 Å	18.1	
Solid Cylinder				
height	radius	R_g	reduced χ^2	
82.6 ± 0.1 Å	46.6 ± 0.1 Å	41.1 ± 0.1 Å	57.2	
Solid Ellipsoid of Revolution				
radius of rotation	semiaxis	R_g	reduced χ^2	
50.5 ± 0.1 Å	56.7 ± 0.1 Å	40.7 ± 0.1 Å	64.9	

The information contained in the scattering curve can also be represented as the radial Patterson curve [$P(r)$], obtained by Sine−Fourier inversion of the scattered intensity. The information in this representation is the distribution of lengths of atom-weighted interatomic vectors within the particle. The $P(r)$ profile for ClpP is shown in Figure 4B and was calculated using the algorithm of Moore (1980) with a d_{max} of 118 Å. Identical $P(r)$ curves were calculated using the algorithms of both Glatter (1980) and Svergun (1992), suggesting that this curve accurately reflects the scattering data for ClpP (data not shown). The $P(r)$ curve for ClpP indicates that the particle is roughly isometric. The $P(r)$ curve for a particle elongated in one dimension is quite distinct. For example, the curve for a long rod will have a peak at small r corresponding to the cross-sectional diameter of the rod, while at long distances $P(r)$ will fall off approximately linearly. For ClpP, the $P(r)$ curve falls off smoothly, and there is no tail at large r. In addition, the bias of the $P(r)$ curve toward longer radii is consistent with the presence of a central pore within the molecule. For a solid object with an axial ratio of 1:0.75, the distribution of vectors through the $P(r)$ curve should be roughly symmetric about the most probable vector. The distribution of vectors in the curve for ClpP is quite asymmetric, with the proportion of short vectors being substantially less than that of the longer vectors. The radius of gyration determined from the $P(r)$ function was 44.8 Å, in good agreement with the R_g of 44.2 Å obtained from Guinier plots.

The Scattering Profile of ClpP Is Best Represented by a Hollow Cylinder Model. Further information about the shape and dimensions of ClpP in solution was obtained by comparing the experimental scattered intensity with that calculated from simple geometrical models. The calculated intensities of a sphere, cylinder, ellipsoid of revolution, hollow cylinder, and hollow sphere (Glatter, 1982) were fitted to the solution scattering data (Q of 0.02−0.13 Å^{-1}) for ClpP. The dimensions of each model (2 or 3 parameters) and the intensity scaling factors were optimized by nonlinear least squares analysis (Table 2). The reduced χ^2 values varied from 1.9 for the hollow cylinder to ~65 for the solid ellipsoid

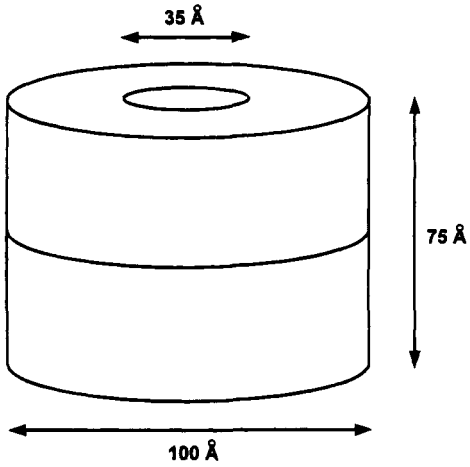

FIGURE 5: Principal dimensions on a schematic diagram of the *E. coli* ClpP cylinder. The dimensions are the average of those obtained from our EM and SAXS models. We have depicted ClpP as being composed of two rings of seven monomers each.

of revolution. In terms of the χ^2 criterion, the scattering curve of ClpP is best described by a hollow cylinder. There is excellent agreement of the experimental scattering intensity with the curve calculated for a hollow cylinder with a height of 98 Å, a diameter of 75 Å, and an axial pore 35 Å in diameter as shown in Figure 4A. Furthermore, the radius of gyration of the model (43.6 Å) is in excellent agreement with that from Guinier plots (44.2 Å). All of the other models gave R_g values that were too small and that, more importantly, did not accurately reproduce the intensity and position of the subsidiary maximum in the experimental scattering curve. The errors in the fitted parameters shown in Table 2 represent the shape of the χ^2 surface near its minima; they have little relationship to the accuracy of the overall fit. To obtain better estimates for the error associated with the fit, the scattering curves of 10 independent data sets varying in protein concentration from 2 to 10 mg/mL were fit with the hollow cylinder model. From the variation in the fitted parameters from these data sets, the hollow cylinder was found to have a mean height of 75 ± 3 Å, a diameter of 98 ± 4 Å, and an axial pore of 35 ± 2 Å. These values are in agreement with the dimensions estimated by electron microscopy of vanadate-stained ClpP as described above.

DISCUSSION

The combination of EM and solution small-angle scattering has allowed us to propose a low-resolution structural model for the *E. coli* ClpP subunit of the Clp ATP-dependent protease (see Figure 5). Our model differs from previous models in two key aspects. First, using SANS and STEM, we demonstrated that ClpP is a tetradecamer rather than a decamer or dodecamer as previously reported (Hwang et al., 1988; Maurizi et al., 1990a). The previous molecular weight of ClpP was based upon the elution behavior of ClpP in SEC. Our SANS results show that in solution ClpP is larger than a decamer or dodecamer, although we cannot distinguish between oligomers containing 13 and 14 subunits. The molecular weight obtained from SANS does not depend upon assumptions about the shape of the molecule, as in the case of SEC. The STEM data provides direct evidence that ClpP is a tetradecamer. Further, after submitting the manuscript for review, we became aware of a study revealing the presence of 7-fold symmetry in *E. coli* ClpP (Kocsis et al., 1995). These authors developed several new statistical tests for rotational symmetry that pool the information content of entire populations of particles which were required to identify symmetry in ClpP. Second, using a combination of solution SAXS and EM of vanadate-stained images, we demonstrated the presence of a large axial pore in ClpP. Our estimates for its diameter using two independent methods are within experimental error of each other with 35 ± 2 Å for SAXS and 32 ± 2 Å for EM. Both values represent an average diameter, although the nature of averaging is quite different. In electron microscopy, the pore diameter is related to the projection of stain density along the axis of the cylinder, and thus represents the average of projected stain density along this axis. In contrast, in SAXS, the pore diameter is a mass-averaged diameter that depends upon how well the model actually represents the "true" structure of the particle. In our low-resolution model of ClpP, we assume a pore of uniform diameter. However, side-on views of ClpP suggest that the pore widens in the center of the molecule. The solution-scattering data presented in this study cannot be used to verify this possibility, since it was not collected to high enough resolution. Additional studies are underway to better model the interior of the ClpP molecule.

The organization of the 14 ClpP monomers in the native oligomer cannot be determined directly from our data. However, in light of the recently described 7-fold symmetry of *E. coli* ClpP (Kocsis et al., 1995), the most likely arrangement is two stacked rings, each of seven monomers. Consistent with this interpretation, side-on images of vanadate-stained ClpP show striations along both axes; in the short dimension (~75 Å) the striation is probably due to the central pore, while in the longer dimension (~100 Å) it reflects the two subunit rings. Our occasional observation of "half-molecules" (~10%) in freeze-dried ClpP specimens examined by STEM supports this model. The half-molecules (~150 kDa) have a similar shape to the more prevalent images of intact ClpP (~300 kDa), as would be expected if they were formed by two stacked rings separating during deposition on the carbon grid. At the current resolution of our model, it is not clear whether monomers from adjacent rings are stacked directly on top of one another or the monomers in one ring are staggered with respect to those in the other ring.

The proposed structure for the native ClpP oligomer has several important implications. The volume of the central pore in ClpP is quite large (~70 000 Å^3); in fact, it is nearly as large as the cavity found in one ring of GroEL monomers (Braig et al., 1994; Chen et al., 1994). Unfolded substrate proteins are thought to bind in, or near, the entrance of the pore formed by one ring of GroEL subunits (Langer et al., 1992; Chen et al., 1994). If the active sites for the ClpP protease are located within the central pore, ClpP could accommodate a relatively large unfolded protein but would exclude even relatively small folded proteins. Interestingly, in a recent study, ClpP was shown to degrade a substrate (CRAG) that is stably bound to *E. coli* GroEL in a ClpA- and ClpX-independent system (Kandror et al., 1994), suggesting that both ClpP and GroEL recognize substrates similarly. Given the function of ClpP in degrading proteins and its requirement for a presenting molecule such as ClpA or ClpX to hydrolyze unfolded protein substrates, it seems likely that the active sites for proteolysis lie within the central pore. This arrangement could provide a high degree of

selectivity for unfolded versus folded substrates.

It is tempting to speculate that the *E. coli* ClpP/ClpA system is a structural and functional homolog of the 26S proteaseome found in the cytosol of eukaryotes. As in the ClpP/ClpA system, ATP-dependent proteolysis by the 26S proteasome involves two components: the 20S protease and the 19S ATPase, which is also thought to confer substrate specificity. There is little sequence homology between either component of the 26S proteasome and ClpP or ClpA. Moreover, the 20S protease and the 19S ATPase are composed of multiple, nonidentical subunits. In the simplest case of the *Thermoplasma acidophilum* 20S proteasome, there are two different subunits, α and β, in four stacked seven-membered rings (Lowe et al., 1995). The complex formed between the 20S and 19S components contains a 2-fold axis of symmetry, and the 19S component is distinctly asymmetric (Peters et al., 1993). The overall structure of the 26S proteasome probably reflects the fact that both components are made up of multiple, nonidentical subunits. In contrast, ClpA and ClpP are both homooligomers, and thus, their complex is likely to have a distinctly different organization. There is one potentially important similarity between the 20S protease component and ClpP. At low resolution, both the 20S proteasome and ClpP can be described as hollow cylinders, and in the case of the *T. acidophilum* 20S protease, there is a clear 7-fold symmetry to the particle (Pulher et al., 1992; Lowe et al., 1995). The subunits of the 20S protease are arranged into four stacked rings; in *T. acidophilum* this arrangement corresponds to a molecular organization of $\alpha_7-\beta_7-\beta_7-\alpha_7$. Like *E. coli* ClpP, the *T. acidophilum* proteasome has an axial pore. The pore, formed by the central two rings of β_7 subunits in the 20S proteasome, houses the 14 catalytic sites (Lowe et al., 1995; Seemuller et al., 1995). Interestingly, the size of the pore in *E. coli* ClpP determined by STEM and solution scattering is similar to that found in the interior of the $\beta_7-\beta_7$ rings in the *T. acidophilum* 20S proteasome.

Another biologically important issue raised by our findings concerns the ClpP/ClpA complex. Previous studies suggest that *E. coli* ClpA exists as a hexamer (Maurizi, 1991), while our results indicate that ClpP is a tetradecamer. It is difficult to imagine how a protein possessing 6-fold symmetry interacts with one having 7-fold symmetry. As in the case of *E. coli* ClpP, the oligomeric state of ClpA is based upon its elution volume in size-exclusion chromatography, which can be misleading as discussed in the context of ClpP. Further, ClpA disassembles into monomers and dimers in the absence of ATP, complicating the characterization of its molecular weight. Additional studies of ClpA alone and on the ClpP/ClpA complex will be required to resolve these issues.

ACKNOWLEDGMENT

We thank Dr. J. Sutherland for use of the circular dichroism spectrometer on beam line U9. We would like to thank Beth Lin for preparing the ClpP STEM samples and Dr. Martha Simon for running the scanning transmission electron microscope.

REFERENCES

Arribas, J., & Castano, J. G. (1993) *J. Biol. Chem. 268*, 21165−21171.

Braig, K., Otwinowski, Z., Hedge, R., Boisvert, D. C., Joachimiak, A. J., Horwich, A. L. & Sigler, P. B. (1994) *Nature 371*, 578−586.

Capel, M. S., Smith, G. C., & Yu, B. (1995) *Rev. Sci. Instrum.* (in press).

Charette, M. F., Henderson, G. W., Doane, L. L., & Markovitz, A. (1984) *J. Bacteriol. 158*, 195−201.

Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., Clarke, A. R., & Saibil, H. R. (1994) *Nature 371*, 261−264.

Chung, C. H., & Goldbert, A. L. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4931−4935.

Flanagan, J. M., Kataoka, M., Fujisawa, T., & Engelman, D. (1993) *Biochemistry 32*, 10359−10370.

Glatter, O. (1980) *J. Appl. Crystallogr. 13*, 7−11.

Glatter, O. (1982) in *Small Angle X-ray Scattering* (Glatter, O., & Kratky, O., Eds.) pp 167−196, Academic Press, London.

Gottesman, S., Clark, W. P., & Maurizi, M. R. (1990) *J. Biol. Chem. 265*, 7886−7893.

Guinier, A., & Fournet, B. (1955) *Small Angle X-ray Scattering*, pp 5−82, John Wiley, New York.

Hennessey, J. P., Jr., & Johnson, W. C., Jr. (1981) *Biochemistry 20*, 1085−1094.

Hwang, B. J., Park, W. J., Chung, C. H., & Goldberg, A. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5550−5554.

Hwang, B. J., Woo, K. M., Goldberg, A. L., & Chung, C. H. (1988) *J. Biol. Chem. 263*, 8727−8734.

Jacrot, B., & Zaccai, G. (1981) *Biopolymers 20*, 2413−2426.

Johnson, C. W. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 145−167.

Kandror, O., Busconi, L., Sherman, M., & Goldberg, A. L. (1994) *J. Biol. Chem. 269*, 23577−23582.

Kataoka, M., Head, J. F., Seaton, B. A., & Engelman, D. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6944−6948.

Katayama-Fujimura, Y., Gottesman, S., & Maurizi, M. R. (1987) *J. Biol. Chem. 262*, 4477−4485.

Kocsis, E., Cerritelli, M. E., Trus, B. L., Cheng, N., & Steven, A. C. (1995) *Ultramicroscopy 58* (in press).

Langer, T., Pfeifer, G., Martin, J., Baumeister, W., & Hartl, F.-U. (1992) *EMBO J. 11*, 4757−4765.

Lowe, J., Stock, D., Jap, D., Zwickl, P., Baumeister, W., & Huber, R. (1995) *Science 268*, 533−539.

Makhov, A. M., Trus, B. L., Conway, J. F., Simon, M. N., Zurabishvili, T. G., Meshyanzhinov, V. V., & Stevens, A. C. (1993) *Virology 194*, 117−127.

Maurizi, M. R. (1991) *Biochem. Soc. Trans. 19*, 719−723.

Maurizi, M. R. (1992) *Experientia 48*, 178−201.

Maurizi, M. R., Clark, W. P., Katayama-Fujimura, Y., Rudikoff, S., Pumphrey, J., Bowers, B., & Gottesman, S. (1990a) *J. Biol. Chem. 265*, 12536−12545.

Maurizi, M. R., Clark, W. P., Kim. S.-H. & Gotteman, S. (1990b) *J. Biol. Chem. 265*, 12546−12552.

Moore, P. B. (1980) *J. Appl. Crystallogr. 13*, 168−175.

Mosesson, M. W., Hainfeld, J. F., Wall, J. S., & Haschemeyer, R. H. (1981) *J. Mol. Biol. 215*, 589−596.

Peters, J.-M., Cejka, S., Harris, J. R., Kleinschmidt, J. A., & Baumeister, W. (1993) *J. Mol. Biol. 234*, 932−937.

Pulher, G., Weinkauf, S., Bachmann, L., Muller, S., Engel, A., Hegerl, A., & Baumeister, W. (1992) *EMBO J. 11*, 1607−1616.

Schneider, D. K., & Schoenborn, B. P. (1984) in *Neutrons in Biology* (Schoenborn, B. P., Ed.) Vol. 27, pp 119−142, Plenum Press, New York.

Seemuller, E., Lupas, A., Stock, D., Lowe, J., Huber, R., & Baumeister, W. (1995) *Science 268*, 579−582.

Svergun, D. I. (1992) *J. Appl. Crystallogr. 25*, 495−503.

Wall, J. S. (1979) in *Introduction to Analytical Electron Microscopy* (Hren, J. J., Goldstein, J. I., & Joy, D. C., Eds.), pp 333−342, Plenum Publishing, New York.

Wall, J. S., & Hainfeld, J. F. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 335−376.

Wickner, S., Gottesman, S., Skowyra, D., Hoskins, J., McKenney, K., & Maurizi, M. R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 12218−12222.

Wotjkowiak, D., Georgopoulos, C., & Zylicz, M. (1993) *J. Biol. Chem. 268*, 22609−22617.

Woo, K. M., Chung, W. J., Ha, D. B., Goldberg, A. L., & Chung, C. H. (1989) *J. Biol. Chem. 264*, 2088−2091.

BI950473W