The 20 S proteasome processively degrades cell proteins to peptides. Information on the sizes and nature of these products is essential for understanding the proteasome’s degradative mechanism, the subsequent steps in protein turnover, and major histocompatibility complex class I antigen presentation. Using proteasomes from *Thermoplasma acidophilum* and four unfolded polypeptides as substrates (insulin-like growth factor, lactalbumin, casein, and alkaline phosphatase, whose lengths range from 71 to 471 residues), we demonstrate that the number of cuts made in a polypeptide and the time needed to degrade it increase with length. The average size of peptides generated from these four polypeptides was $8 \pm 1$ residues, but ranged from 6 to 10 residues, depending on the protein, as determined by two new independent methods.

However, the individual peptide products ranged in length from approximately 3 to 30 residues, as demonstrated by mass spectrometry and size-exclusion chromatography. The sizes of individual peptides fit a log-normal distribution. No length was predominant, and more than half were shorter than 10 residues. Peptide abundance decreased with increasing length, and less than 10% exceeded 20 residues. These findings indicate that: 1) the proteasome does not generate peptides according to the “molecular ruler” hypothesis, and 2) other peptidases must function after the proteasome to complete the turnover of cell proteins to amino acids.

The 20 S proteasome is a 700-kDa proteolytic complex that is present in all eukaryotic cells, archaea, and certain bacteria (1, 2). In eukaryotes, the proteasome is an essential component of the ATP-ubiquitin-dependent pathway for protein degradation (3, 4). The 20 S particle functions as the proteolytic core of the 26 S proteasome complex, which catalyzes the rapid degradation of many short lived regulatory proteins and of proteins with abnormal conformation (1, 4–6). In mammals, the proteasome is also responsible for the breakdown of most long lived cell proteins and for the generation of most peptides presented to the immune system on MHC $^\text{I}$ class I molecules (5, 7).

The 20 S particle is a barrel-shaped structure composed of four stacked rings. Each outer ring contains seven related $\alpha$-subunits, and each of the inner rings seven related $\beta$-subunits, which catalyze peptide bond cleavage. The active sites are located within the central chamber of the 20 S particle, into which protein substrates must enter through a narrow openings in the $\alpha$- and $\beta$-rings. Proteasomes cleave peptide bonds by a novel catalytic mechanism, in which the hydroxyl group of the threonine on the N terminus of the $\beta$-subunit serves as the active site nucleophile (8–12).

Recently, we have shown that the 20 S proteasome degrades proteins by a highly processive mechanism to oligopeptides (10). A typical protease makes a single cleavage in a polypeptide substrate and then releases the fragments, which may be cleaved to smaller products in subsequent enzymatic rounds. In contrast, the proteasome appears to make many cleavages in the polypeptide and to digest it to small products without dissociation of the partially degraded substrates. This novel mode of degradation appears highly important for an intracellular proteolytic system, since the release of large protein fragments could interfere with cell function and regulation. However, definitive proof of this mode of degradation requires knowledge of the number of cleavages made in a protein substrate and the sizes of peptides generated. If the proteasome makes repeated cuts processively along the length of the polypeptide, one would predict that the enzyme should make a greater number of cleavages and take more time to digest longer polypeptides than shorter substrates. One goal of the present study was to test these predictions.

Knowledge about the size distribution of peptides produced by the proteasome is important for understanding the subsequent steps in the protein degradative pathway and the process of MHC class I antigen presentation. *In vivo*, most of the peptides generated by the proteasome must be rapidly hydrolyzed to amino acids, which are utilized in synthesis of new proteins or in intermediary metabolism. In mammalian cells, some of these peptides are utilized in antigen presentation, possibly after further proteolytic processing to the final 8–9-residue peptides presented on the cell surface (13). These latter steps are poorly understood, in part because of a lack of precise information on the sizes of the peptides released by the proteasome during protein breakdown.

It is widely believed that the proteasome degrades polypeptides according to a “molecular ruler” to yield products of rather uniform size (8, 12, 14, 15), as first proposed by Wenzel et al. (16). This group reported that a large fraction of the peptides generated during the breakdown of hemoglobin or insulin $\beta$-chains by archaeal proteasomes were between 7 and 9 residues long (16). Since the distance between adjacent active sites corresponded to an octapeptide in an extended conformation (8, 16, 17), it was proposed that peptides of 7–9 residues were routinely generated as a result of coordinated cleavages by neighboring active sites. However, evidence for such a molecular ruler is quite limited. In the study by Wenzel et al. (16) or
in other studies of peptides generated by the proteasome (18–23), the relative amounts of peptides of different sizes were not quantified. In addition, Wenzel et al. (16) analyzed peptide products after prolonged incubations, during which the products were likely to undergo repetitive cleavage by the proteasome (20, 22, 23). The present studies were undertaken to determine the mean size of peptides generated from full-length proteins, to measure the relative amounts of peptides of different sizes, and thus to critically test the molecular ruler model.

Most prior studies have focused on peptides ranging from 4 to 44 residues in length, which may be degraded differently from proteins. In contrast, in this study, we have investigated the digestion of proteins of different sizes ranging in length from 70 to 471 residues. We have introduced several new approaches to evaluate the mean number of cuts made in each protein substrate, the mean sizes of the peptides generated by proteasome, and the relative amounts of products of different lengths. Since only unfolded molecules can enter the 20 S particle and be degraded (24), the present studies utilized denatured polypeptides. Conditions were chosen where the protein substrate was present in large excess, such that the released peptides were not digested further. We have employed 20 S proteasomes from the archaea Thermoplasma acidophilum, because this particle is structurally simpler than the eukaryotic 20 S proteasome. It contains seven identical α- and β-subunits, and thus seven identical active sites, located at equal distances from each other (8), unlike the eukaryotic particle which contains seven different α-subunits and seven different β-subunits, and has three active sites with different specificities (12). Therefore, the factors determining product size and rate of proteolysis should be easier to elucidate and the data easier to interpret with archaeal particle.

EXPERIMENTAL PROCEDURES

Materials—The Thermoplasma proteasome was expressed in Escherichia coli and purified as described previously (10). Recombinant human IGF was a kind gift of Dr. W. Prouty (El Lilly, Indianapolis, IN), bovine α-lactalbumin and β-casein, alkaline phosphatase from E. coli, Leu-enkephalin amide, substance P, and oxidized A- and B-chains of bovine insulin were purchased from Sigma. Synthetic peptides SHNFKEK, YPHPARIGL, TYQTRRALV, and YSSEDMDTM were kindly provided by Dr. K. Rock (University of Massachusetts Medical Center, Worcester, MA). All other peptides were from Bachem AG (Switzerland).

For use as substrates, IGF and lactalbumin were unfolded by reduction of disulfide bonds and carbamoylation as described previously (10). The IGF, lactalbumin, and casein were exhaustively, reductively methylated (10) to prevent the extent of the reaction of the undegraded molecule with fluorescamine, and the modified polypeptide purified by HPLC prior to use. To denature alkaline phosphatase, it was oxidized with performic acid (25). The concentrations of modified IGF, lactalbumin, and casein were determined by UV absorption at 280 nm. Calculations of the amount of peptide (i.e. the amount of peptides) were measured by the reaction with fluorescamine at pH 6.8 (10). These peptides were then hydrolyzed completely to amino acids with 6 M HCl in sealed ampoules for 24 h at 108 °C, and the amount of amino acids after hydrolysis was measured by the fluorescamine assay in 0.2 M phosphate buffer (pH 6.8) (10). These peptides were then hydrolyzed completely to amino acids with 6 M HCl in sealed ampoules for 24 h at 108 °C, and the amount of amino acids after hydrolysis was measured by the fluorescamine assay in 0.2 M phosphate buffer (pH 6.8). A standard mixture of amino acids, treated in the same fashion as the samples, was used for calibration of the amino acid assay. Then, the mean size was determined by dividing the molar amount of amino acids found after acid treatment by the molar amount of peptides before the treatment.

Size-exclusion Chromatography of Degradation Products—Size-exclusion chromatography was performed on polyhydroxyethyl aspartamide column (0.46 × 20 cm, Poly LC, Columbia, MD), using a HP1090 chromatographer (Hewlett-Packard). The mobile phase was 0.2 M sodium sulfate, 25% acetonitrile, pH 3.0 (adjusted with phosphoric acid), and the flow rate was 0.125 ml/min. To determine the apparent molecular mass of the peptides eluted, the column was calibrated each time before use with 8–10 standard peptides in the 550–3500 Da range. The pool of proteasome’s products (the same as in acid hydrolysis method) containing 5–10 nmol of peptides was dissolved in 50 μl of the mobile phase and loaded onto the column. Fractions (0.5 min) were collected, and the molar amount of peptides in each fraction was measured with the fluorescamine assay as described above. The corresponding fraction of the control mixture, in which the substrate was incubated without proteasome, was also run on the size-exclusion column. No fluorescamine-reactive material was found in the fractions of this run.

RESULTS

Number of Cuts Made in Substrate Molecules—To ensure rapid hydrolysis and to eliminate possible complications due to substrate folding, three substrates (alkaline phosphatase, IGF, and lactalbumin) were denatured prior to study, as described above. The other substrate studied, casein, had little or no tertiary structure and did not require denaturation to be degraded rapidly by the proteasome. To determine the number of cuts made by the proteasome in these different polypeptides, we measured the rate of disappearance of substrate molecules and the rate of appearance of the new amino groups during the same incubation. Denatured IGF, lactalbumin, casein, and alkaline phosphatase were incubated at 53 °C and pH 7.5 with highly purified recombinant Thermoplasma proteasomes. Initial substrate concentrations were high enough to ensure a constant rate of degradation during the entire incubation period. At different times, aliquots were removed, and one portion was analyzed by HPLC to determine the amount of substrate consumed (by measuring the amount of undegraded substrate by integration of its peak area), another portion was used to...
determine the amount of peptides produced by assaying them as the number of new primary amino groups generated that react with fluorescamine. As found previously (10), the rates of accumulation of peptide products and of the disappearance of the substrates paralleled each other. The ratio of the amount of new products generated to the amount of protein molecules degraded did not change with time, indicating a processive mechanism. Moreover, the products once released by the proteasome did not get cleaved again at later times under these incubation conditions, where there was a large molar excess of the substrate.

If the proteasome makes $n$ cuts in a protein molecule and the products generated do not undergo further cleavages, there should always be $n$-fold more new amino groups than substrate molecules consumed. Therefore, the number of cuts made per protein can be determined by dividing the rate of peptide product accumulation by the rate of substrate disappearance (Fig. 1). This value increased with the length of the substrate, ranging from 11 cuts in IGF, which is 70-residues long, to 71 cuts in alkaline phosphatase, which contains 471 residues (Table I). The demonstration that a large number of cuts are made in a single substrate molecule, together with the previous finding that the proteasome does not release the substrate until all these cuts are made (10), is clear evidence of a highly processive mechanism for protein degradation (Fig. 2).

**Time Necessary to Degrade Different Protein Molecules**—This highly processive mechanism suggests that the proteasome should require more time to degrade a longer polypeptide than a short one, provided that it moves along the substrate at a relatively constant rate. These measurements of the rate of disappearance of the different substrates (Table I) allowed us also to determine the time needed to degrade a substrate molecule, because these experiments were performed at $V_{\text{max}}$. The number of protein molecules degraded per min by a 20 S particle was calculated by dividing the rate of substrate disappearance by the molar concentration of the proteasome (Table I). The reciprocal therefore represents the time that the proteasome takes to degrade one substrate molecule, assuming it degrades only one substrate molecule at a time.

The time required to degrade each of these polypeptides was a characteristic feature of the substrate. As shown in Fig. 3, the time for degradation by the proteasome depended on the polypeptide’s length. At 53 °C, the enzyme required 10 s to digest one IGF molecule (70 residues, 11 cuts) and 50 s to degrade one casein molecule (209 residues, about 20 cuts).

![Fig. 1. Two methods to determine the mean size of the peptides generated by the proteasome and the number of cuts in a single polypeptide.](image)

**“Two Rates” Method**

\[
\text{Number of cuts} = \frac{\text{New Amino Groups (moles)}}{\text{Substrate Molecules Degraded (moles)}}
\]

**“Acid Hydrolysis” Method**

\[
\text{Length of Substrate (residues)} = \frac{\text{Mean Length (residues) \times Number of cuts + 1}}{\text{Amino Acids (N-groups after hydrolysis, moles)}}
\]

**Kinetic parameters of protein degradation by the proteasome**

The number of protein molecules degraded and peptide bonds cut by a single proteasome particle per min were determined by division of the corresponding cleavage rate (as described under “Experimental Procedures”) by the proteasome concentration measured by the methods of Bradford or Lowry with similar results. The number of cuts in the molecule made (column 5) was determined by dividing the number peptide bonds cleaved (column 4) by the number of substrate molecules degraded (column 3). See Fig. 1 for an explanation of the calculations.

![Fig. 2. The number of cuts made by the 20 S proteasome in a protein depends on its length.](image)

When casein was modified by fluorescein isothiocyanate, the proteasome still took approximately 1 min to degrade it. Thus, with smaller polypeptides, there was almost a linear relationship between substrate length and duration of the degradative process. In contrast, almost 5 min were required to digest alkaline phosphatase (471 residues, 71 cuts). The disproportionately long time required for degradation of alkaline phos-
HPLC on a reverse-phase column. The amount of peptides proteasomes were separated from the undigested substrate by peptide products, another approach was developed, which we mean (see below).

II represent the range of the mean sizes obtained in independ-

ducibly (Table II). The standard errors on these values in Table 
of peptides generated from different proteins were found repro-

ducible (Table II). The standard errors on these values in Table II represent the range of the mean sizes obtained in independent experiments, rather than the variation in length about the mean (see below).

To obtain an independent measure of the mean sizes of these peptide products, another approach was developed, which we call the acid hydrolysis method. The peptides generated by the proteasomes were separated from the undigested substrate by HPLC on a reverse-phase column. The amount of peptides produced was measured with fluorescamine, they were then hydrolyzed completely to amino acids by acid treatment, and the amount of amino acids was measured with fluorescamine. If an individual peptide containing $n$ residues is hydrolyzed completely, the molar amount of amino acids produced should be $n$-fold greater than the amount of peptides present initially. Therefore, dividing the amount of amino acids after acid hy-
drolysis by the amount of peptides before this treatment gives the length of peptide, and with a mixture of peptides, this approach gives the weighted average of the lengths of peptides, i.e. their mean size (Fig. 1).

The values obtained by this method for the mean sizes of the peptides generated from the different substrates (Table II) ranged between 7 and 9 residues and resembled closely the values obtained by the two rates method for the same substrate (Table II). Thus, with proteins that differ in length by 7-fold and more than 20-fold in the time required for their degrada-
tion, the mean product length was close to 8 residues, although slightly smaller values were consistently found with IGF (6 to 7 residues) and longer values with casein (9–11 residues). This finding of a mean size of 8 residues is consistent with the suggestion that a major determinant of product size is the distance between active site threonines, which corresponds to 8 residues of an extended polypeptide chain. However, the demon-
stration of similar mean sizes does not imply that individual peptide products are of uniform lengths, as would be expected if the proteasome cleaves substrates according to a molecular ruler model (8, 16).

**Size Distribution of the Peptide Products**—To understand how proteins are digested within the proteasome's central chamber and to investigate the metabolic fates of peptides generated, it is essential to determine the actual size distribution of the products produced from a specific substrate. Two approaches were used to characterize the lengths of products generated during degradation of IGF. In the first one, these peptides were separated by HPLC on a C18-reverse phase column (Fig. 4A), and all peaks detected with UV were collected individually. The molar amounts of the peptides in each peak were then measured with the fluorescamine assay. These measurements demonstrated that individual peptides appeared in nonstochiometric amounts, and that the molar amounts of the majority of individual products were less than the molar amount of substrate degraded. Thus, although the proteasome always degraded IGF molecules processively to peptides of small size, it did not cut all IGF molecules at identical places. As a result, the particle generated at least 50 different products (as judged from the number of distinct HPLC peaks), while cleaving on the average eleven bonds in each IGF molecule (Table I).

| Substrate      | Two rates method | Acid hydrolysis method | Size-exclusion chromatography |
|----------------|------------------|------------------------|-------------------------------|
| Alkaline phosphatase | 6.5 ± 0.5       | 9.3 ± 1.0              | 10.2                          |
| Casein         | 10.6 ± 3.1       | 9.0 ± 1.3              | 11.2                          |
| Lactalbumin    | 8.1 ± 0.5        | 7.7 ± 1.4              | 9.1                           |
| IGF            | 5.9 ± 0.8        | 6.7 ± 0.3              | 7.0                           |
| Average        | 7.8 ± 1.8        | 8.2 ± 1.0              | 9.4                           |

**Fig. 3.** The time needed by the 20 S proteasome to degrade a polypeptide depends on its length. The time for degradation was calculated as the inverse of the number of protein molecules degraded/min at $V_{max}$. It was presumed that only one polypeptide is degraded by a proteasome particle at a time. The proteins are: 1, IGF; 2, lactal-
bumin; 3, casein; 4, alkaline phosphatase; and 5, fluorescein isothiocyanate-casein.

**Table II**

Mean size of peptides generated by the proteasome

The values in the table are means ± S.E. of at least three experiments (for the acid hydrolysis method) or two experiments (for the two rates method). In the two rates method, the mean size was determined by dividing the number of amino acids in a protein by the number of cuts plus one (see Table I and Fig. 1). In the acid hydrolysis method, the pooled products were hydrolyzed to amino acids, and mean peptide size was obtained by dividing the amount of amino acids after acid hydro-
ysis by the amount of peptides before it (Fig. 1).
The molecular weights of the peptides present in the 27 major peaks were determined by mass spectrometry. Some peaks, which appeared to be impure by mass spectrometry, were analyzed by microsequencing to establish the nature and to show a linear dependence on the logarithm of their molecular weights. The great majority of these peptides ranged in molecular mass from below 500 to 2000 Da, and relatively few were detected in the 2000–10,000 Da range (Fig. 4B). Together these peaks comprised 75% of the total molar amount of peptides generated. Assuming that the average molecular mass of an amino acid in IGF is 116 Da, the majority of the products were between 4 and 12 residues long, although one major peptide of 413 Da on sequencing appeared to be a tripeptide. The mean molecular mass of analyzed peptides for this distribution was 800 Da (7 residues), which equals the mean length obtained by the acid hydrolysis method (Table II).

II. However, peptides of this size, 7 ± 1 residues, were not a predominant species, and the size distribution was highly asymmetric. Peptides of 5 residues and shorter were most abundant and the abundance of the products tended to decrease as their size increased.

This HPLC-mass spectrometry approach was not applicable to the longer substrates. Since there were more than 50 peptides generated from IGF (Fig. 4A), it seemed likely that almost 100 would be generated from lactalbumin and several hundreds from casein and alkaline phosphatase. Even if such a large number of peptides could be resolved on HPLC, it would be impractical to analyze these products individually, as was done with IGF-derived peptides. Moreover, analysis of IGF degradation products could not include the least abundant 25% peptides. Therefore, an additional method to analyze the size distribution of peptides was developed, in which product length was assayed using size-exclusion chromatography on polyhydroxyethyl aspartamide HPLC column. This column had been reported to fractionate peptides in the range of 500–10,000 Da, which should include the great majority of proteasome products, and has been used to separate by size peptides generated from casein by trypsin (27). In initial control studies, the retention times of 13 randomly selected synthetic peptides with very different sequences were found to be highly reproducible and to show a linear dependence on the logarithm of their molecular weights (Fig. 5).

To analyze the sizes of the proteasome products, these peptides were first separated from the undegraded protein substrate by HPLC on a reverse phase column, as described above. The combined products were then fractionated on the size-exclusion column. The molar amounts of the peptides in each fraction were determined by the fluorescamine assay and graphed against the elution time (Fig. 6). The resulting curves were highly reproducible (see, for example, the two separate runs for casein) and differed with each substrate. However, the general pattern was similar for the peptides generated from all four proteins. With the longer substrates, casein and alkaline phosphatase, the abundance of peptides fit closely a normal distribution when graphed against the logarithms of their molecular weights (Fig. 6), i.e. against the retention time on the size exclusion column, which is proportional to the log(molecular weight). In other words, on such a logarithmic scale, the amounts of peptides generated by proteasome seem to be nor-

---

**Fig. 4. Distribution of peptides generated by proteasomes from IGF.** A, HPLC separation of a digest. IGF was incubated with Thermoplasma proteasomes for 4 h as described under “Experimental Procedures.” At this time ~45% of the substrate was consumed. Peptides were separated on a C_{18} (5 μm) column using the buffer system described under “Experimental Procedures.” The flow rate was 0.75 ml/min, and the gradient of acetonitrile was from 0 to 8% in 60 min, from 8 to 28% in 100 min, and from 28 to 38% in 20 min. UV absorption at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides. The large peak, eluting at 214 nm was used for detection of peptides.
mally distributed, indicating a log-normal distribution. With IGF and lactalbumin, the abundance of peptides of different sizes also appeared to fit a log-normal distribution, but with greater deviation, probably because of the smaller sampling size (Fig. 6).

In contrast, when relative amounts of the products were graphed against molecular weight on a linear scale (Fig. 4C), this distribution was clearly asymmetric. The decrease in amount of products with increasing size is similar to the finding we obtained in the mass spectroscopy analysis of the same IGF-derived peptides (Fig. 4B). In addition, both methods gave the same mean molecular masses of 800 Da (7 residues) for the products, which agreed well with the mean size value obtained by the acid hydrolysis method.

As shown in Fig. 6, the sizes of the products varied widely, and a fraction was shorter than 500 Da (4–5 residues), but the exact size of the shortest peptides could not be determined, since it was impossible to calibrate the size-exclusion column for peptides smaller than 500 Da. On the other hand, peptides reached up to 2300 Da (about 20 residues) when the shorter substrates, IGF and lactalbumin, were degraded. With the longer substrates, casein and alkaline phosphatase, some peptides ranged up to 5000 Da or ~45 residues. This broad range in the sizes of products and the asymmetric log-normal distribution are not consistent with the proteasome’s cleaving proteins according to a molecular ruler mechanism.

Another informative way to analyze the data on size distribution of products (Fig. 6) was by calculating the cumulative frequency of peptides of a given size, i.e. the fraction of all peptides having a molecular weight equal or less than the peptide of interest (Fig. 7). Fig. 7A demonstrates that such curves for the products of IGF degradation obtained by mass spectrometry and size-exclusion chromatography were practically indistinguishable. Thus, these two different methods yielded a very similar or identical size distribution of peptide products. Plotting the data from the gel filtration column for products from the other polypeptide substrates (Fig. 7B) also showed a smooth increase in cumulative frequency with increasing peptide size. The lack of an inflection point in these cumulative frequency plots is clearly not consistent with a normal distribution and consistent with a log-normal distribution. This type of analysis also clearly demonstrates that proteasomes generate a significant fraction of products shorter than pentapeptides. The fraction of these small peptides ranged from 15% of the products generated from lactalbumin up to 40% of the IGF-derived peptides. With all substrates, 40–50% of the products fell between 6 and 10 residues.

Of the peptides generated from IGF and lactalbumin, all were shorter than 20 residues (2300 Da). However, a small fraction (<10%) of the peptides generated from the longer proteins, casein and alkaline phosphatase, were between 20 and 30 residues long. Some peptides (<5%) generated from casein even appeared to be longer than 30 residues. This high value, however, may be an artifact due to the unusual primary sequence of casein, in which 15% of all residues are prolines.
Proteasomes are highly processive degradation systems that cleave proteins into small peptides. The mechanism involves multiple active sites within the proteasome, which can degrade proteins at different rates. The proteasome has three types of activity: endo-, exo-, and dipeptidyl peptidase activities. The endo-activity breaks down proteins into peptides, while the exo-activity further cleaves peptides into smaller fragments.

### Discussion

**Mode of Protein Degradation—**The present findings and related observations (10) clearly show that the proteasome digests proteins in a fundamentally different manner from the great majority of proteolytic enzymes. Generally, an endoprotease or exoprotease dissociates from its substrate after each cleavage. In contrast, the proteasome degrades proteins in a highly processive manner (10), i.e. the enzyme makes multiple cuts in a protein and converts it to oligopeptides before attacking the next substrate molecule. For example, the archaeal proteasome made about 70 cleavages in degrading alkaline phosphatase (Fig. 2, Table I) and converted it into peptides ranging from about 4 to 30 residues in length. Further evidence for a highly processive mechanism was the finding that the number of cuts in a polypeptide and the time needed to degrade it increase with the length of the substrate. Interestingly, as the length of the substrate increased, the number of peptide bonds cleaved per min decreased (Table I). With the longer substrates more time is probably necessary for a cleavable bond on the substrate to reach the active site(s). For polypeptides shorter than 200 amino acids, the time required to degrade them was directly proportional to their length, but increased steeply and nonlinearly with alkaline phosphatase, a 471-residue protein. Perhaps the alkaline phosphatase, although initially denatured, retained more secondary structure or reversibly aggregated; either mechanism would retard its diffusion into the central chamber or proteolytic attack.

This highly processive degradation to small peptides must mean that the protein substrate remains associated with the proteasome until the degradative process is complete. Such a mechanism would appear to be highly advantageous to the cell, since it ensures rapid elimination of proteins targeted for degradation and prevents the accumulation of partially degraded proteins, which could be highly toxic. Presumably, this behavior is due to the particle's complex architecture. To reach the active sites in the proteasome's central chamber, substrates must first traverse the narrow opening in the α-ring, the outer chamber, and finally the opening in the β-ring. These multiple chambers probably help prevent the premature release of substrates until the products are small enough to exit the particle.

**Rates of Protein Degradation—**Under the present conditions (53 °C, pH 7.5), which were chosen to ensure high activity of the Thermoplasma proteasome, the particle took about 50 s to degrade casein and 4.5 min to digest alkaline phosphatase. At the optimal growth temperature for this organism (60 °C), degradation of casein occurred approximately 20–25% faster, but the proteasome still took approximately 40 s to degrade casein (not shown). These values were calculated on the assumption that at Vmax only one polypeptide molecule was degraded at a time by the proteasome. However, it is unclear whether one proteasome particle degrades only 1 or 2 substrate molecules at a time. Two Nanogold-modified insulin molecules were found associated with each end of a single proteasome (24). If 2 molecules can be degraded simultaneously, the time required for degradation of a single polypeptide should be twice as long as values shown here (Fig. 3).

In either case, the times necessary to digest proteins appear surprisingly long. If these rates are similar to the rates in vivo, then it would suggest that the proteasome may take a longer time to degrade a protein than the ribosome takes to synthesize it. In E. coli, the ribosome requires about 10 s to synthesize a protein the size of casein and 24 s for one the size of alkaline phosphatase, which are much shorter than the times for proteolysis measured here. The eukaryotic proteasome (12) may well digest proteins at different rates, since it contains only three active sites with distinct specificities, and within the 26 S complex, its function is linked to ATP hydrolysis, which may enhance the rate of protein degradation.

**Size of Degradation Products—**Using either of the new methods, the average size of the products of the proteasome was found to be 8 residues (Table II). Similar sized peptides were generated from proteins that varied 7-fold in length, had markedly different sequences, and differed up to 25-fold in the times required for their degradation. Small differences, however, in the mean size were seen repeatedly with different substrates (Table II), apparently because some longer peptides were generated from the longer polypeptides (Fig. 6). An average size of 8 residues indicates that the proteasome cuts only 10–15% of the peptide bonds present in the full-length protein. In vivo, the remaining 85–90% of the peptide bonds must be cleaved by other proteolytic enzymes, which remain to be identified. Since peptides of this size cannot be found in the cytosol, they must be rapidly digested and tend to be short-lived in cell extracts. Thus, intracellular proteolysis seems to require the concerted activity of the 20 S complex and multiple endo- and exopeptidases.

As shown by size-exclusion chromatography and mass spectrometry, the sizes of peptides produced by the proteasome are broadly and asymmetrically distributed around the mean. This pattern and the results obtained by plotting cumulative frequencies of individual peptide sizes are not consistent with a Gaussian or normal distribution (e.g. in a cumulative frequency plot, a Gaussian distribution would have shown an inflection point at 50%). However, the product sizes seem to fit a log-normal distribution, such that the relative amount of the peptides decreases as their length increases. This fit was particularly strong with casein and alkaline phosphatase, presumably because of the larger number of peptides generated from these substrates. It is noteworthy that (except for peptides generated from casein) no product was longer than 30 residues, and in all cases, at least 90% of the peptide products were shorter than 20 residues. While peptides longer than 20 residues comprised fewer than 10% of the peptides produced from casein and

---

2 T. N. Akopian, unpublished results.
alkaline phosphatase, these products actually contain 20–30% of the amino acids and mass of the original proteins.

These findings therefore are not consistent with the idea that the proteasome digests substrates according to a precise molecular ruler, as had been first proposed (16) and as has been widely assumed (8, 12, 14, 15). On the contrary, octapeptides were not even a major species. In fact, less than 15% of the peptides fell within the 7–9-residue range, and products were not uniformly distributed about this size. This size distribution is also not consistent with the limited data on product size published previously (16, 18–23). Several studies (16, 20–23) had suggested that the peptides generated by proteasomes were distributed more symmetrically and tightly about the mean. Several reasons may account for these differences. 1) Short peptides are more likely to be missed in eluates from reverse-phase column used in all previous studies, because of their low UV absorption. By contrast, we used size-exclusion column from which peptides were eluted as pools, and not as individual peaks, and also fluorescamine, which detects shorter peptides reliably and quantitatively. 2) The absolute amounts of these products were not quantified previously. 3) Most prior studies used substrates containing less than 44 residues, which provide too small a sampling size for accurate conclusions about the size-distribution of the products of large proteins. In addition, Wenzel et al. (16) did not find products longer than 14 residues, probably because the products were analyzed after a 24-h incubation, when the proteins had been digested completely and when peptides released from the proteasome are cleaved further by the particle (20, 22, 23).

The mechanisms which determine the size of products and this log-normal distribution are quite unclear. It was proposed that the main factor determining product size (the molecular ruler) was the distance between adjacent active sites, which corresponds to a 7- or 8-residue peptide in an extended conformation (8, 16). Our finding of a mean product size of 8 suggests that this distance is one factor influencing product size. However, since less than 15% of the products were of this size, other factors must also influence product size. The processive cleavage of a polypeptide presumably results from the coordinated action of multiple active sites. Huber and coworkers (8, 17) proposed that the substrate remains covalently attached to one active site threonine residue, while it is attacked by an active site on an adjacent subunit. Such a mechanism would generate predominantly 7–8-residue peptides. However, if cleavages are made by nonadjacent active sites, longer products will be generated. Peptides smaller than 7 residues are probably produced if the cleaved fragments diffuse to and are cut by other active sites (18). The probability of such an additional attack must depend on the peptide’s sequence. Once generated, the shorter peptides will have a higher chance to exit the particle, especially since smaller peptides tend to bind less tightly to proteolytic sites than longer ones (16). A final determinant of product size may simply be the time that a fragment remains within the β-chamber (with longer occupancy, there should be more opportunities for additional cleavages).

Presumably, similar factors determine product size in more complex eukaryotic proteasomes, but it is impossible to extrapolate from these data on the products of the archaeal particle to those of the 20 or 26 S eukaryotic proteasome that do not contain seven identical, symmetrically distributed active sites. Analogous studies of product size with eukaryotic proteasomes are also of particular interest because of their role in generating the 8–9-residue peptides used in MHC class 1 antigen presentation (13).

Acknowledgments—We are grateful to Drs. M. Zelen (Dana Farber Cancer Institute) for advice in statistical analysis of data, E. Seemüller (Max-Planck Institute for Biochemistry, Germany) for providing proteasome clones, O. Coux for preparation of Fig. 1, R. Moerschell for assistance with HPLC experiments, O. Kandror, P. Zwickl, V. Solomon, O. Coux, and H. C. Huang for thoughtfull criticisms, and A. Scott for assistance in the preparation of this manuscript.

REFERENCES
1. Coux, O., Tanaka, K., and Goldberg, A. L. (1996) Annu. Rev. Biochem. 65, 801–847
2. Tamura, T., Nagy, I., Lupas, A., Lottspeich, F., Cejka, Z., Schofs, G., Tanaka, K., Denout, R., and Baumeister, W. (1995) Curr. Biol. 5, 766–774
3. Cecchonover, A. (1994) Cell 79, 13–21
4. Hochstrasser, M. (1996) Annu. Rev. Genet. 30, 405–439
5. Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A. L. (1994) Cell 78, 761–771
6. Lee, D. H., and Goldberg, A. L. (1996) J. Biol. Chem. 271, 27280–27284
7. Craiu, A., Gaczynska, M., Akopian, T., Gramm, C. F., Fenteany, G., Goldberg, A. L., and Rock, K. L. (1997) J. Biol. Chem. 272, 13437–13445
8. Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) Science 268, 533–539
9. Seemüller, E., Lupas, A., Stock, D., Lowe, J., Huber, R., and Baumeister, W. (1995) Science 268, 579–582
10. Akopian, T. N., Kisseliev, A. F., and Goldberg, A. L. (1997) J. Biol. Chem. 272, 1771–1778
11. Fenteany, G., Standaert, R. F., Lane, W. S., Choi, S., Carey, E. J., and Schreiber, S. L. (1995) Science 268, 726–731
12. Groll, M., Ditzel, I., Lowe, J., Stock, D., Bochtler, M., Bartunik, H., and Huber, R. (1997) Nature 386, 463–471
13. Goldberg, A. L., Gaczynska, M., Grant, E., Michalek, M., and Rock, K. L. (1995) Cold Spring Harbor Symp. Quant. Biol. 60, 479–490
14. Stuart, D. I., and Jones, E. Y. (1997) Nature 386, 437–438
15. Baumeister, W., and Lupas, A. (1997) Curr. Opin. Struct. Biol. 7, 273–278
16. Wenzel, T., Eckerskorn, C., Lottspeich, F., and Baumeister, W. (1994) FEBS Lett. 349, 205–209
17. Stock, D., Ditzel, L., Baumeister, W., Huber, R., and Lowe, J. (1995) Cold Spring Harbor Symp. Quant. Biol. 60, 525–532
18. Dick, I. R., Moosm, C. R., DeMartino, G. N., and Slaughter, C. A. (1991) Biochemistry 30, 2725–2734
19. Dick, I. R., Aldrich, C., Jameson, S. C., Moosm, C. R., Pamanek, B. C., Doyle, C. K., DeMartino, G. N., Bevan, M. J., Forman, J. M., and Slaughter, C. A. (1994) J. Immunol. 152, 3884–3894
20. Ehring, B., Meyer, T., Eckerskorn, C., Lottspeich, F., and Tampe, R. (1996) Eur. J. Biochem. 235, 404–415
21. Groettrup, M., Ruppert, T., Kuehn, L., Seeger, M., Standera, S., Koszinowski, U., and Klotetzki, P. M. (1995) J. Biol. Chem. 270, 23808–23815
22. Niedermann, G., King, G., Butz, S., Birner, U., Grimm, R., Shabanowitz, J., Hunt, D. F., and Eichmann, K. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 8572–8577
23. Dick, T. P., Ruppert, T., Groettrup, M., Klotetzki, P. M., Kuehn, L., Koszinowski, U. H., Stevanovic, S., Hild, H., and Ramnsee, H. G. (1996) Cell 86, 253–262
24. Wenzel, T., and Baumeister, W. (1995) Nat. Struct. Biol. 2, 199–204
25. Hirs, C. H. W. (1967) Methods Enzymol. 11, 197–198
26. Gill, S. C., and von Hippel, P. H. (1989) Anal. Biochem. 182, 319–326
27. Silvestre, M. P. C., Hamon, M., and Yvon, M. (1994) J. Agric. Food Chem. 42, 2778–2782
28. Gornick, L., and Smith, W. (1993) The Cartoon Guide to Statistics, Harper Collins Publishers, New York