# Control of Peptide Product Sizes by the Energy-Dependent Protease ClpAP†

Kee-Hyun Choi and Stuart Licht*

*Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139*

*Received March 20, 2005; Revised Manuscript Received August 19, 2005*

ABSTRACT: Processive proteases can unfold proteins and cleave them into fragments of a characteristic size. The detailed mechanism by which product sizes are controlled is still in question. One possible mechanism for the control of product sizes would be translocation of unfolded polypeptides to the protease active sites in units of defined length. We have investigated the mechanism by which ClpAP, an energy-dependent protease from *Escherichia coli*, controls the sizes of its peptide products. We show that ClpAP generates peptide products with a distribution of sizes that has a pronounced peak at a peptide length of 6−8 amino acid residues. This product size distribution, which is similar to that observed previously for the proteasome, is robust to perturbations that interfere with translocation or proteolysis. To explain these results, we propose a mechanism in which translocation alternates with proteolysis, allowing peptides of more or less uniform length to be cleaved processively from a translocating substrate. To estimate the rate and energy efficiency of ClpAP-catalyzed measurements of product sizes, we apply information theory to quantify how precisely the product sizes are controlled. This analysis may also prove to be useful in characterizing the mechanisms of other proteases and nucleases, such as the proteasome and Dicer, which control the sizes of their products.

To generate uniform product sizes, enzymes that hydrolyze polymeric substrates must be able to measure the length of a segment of substrate. Several enzymatic systems appear to carry out reasonably precise measurements. Mammalian and thermophilic proteasomes cleave protein substrates to small peptides of a characteristic size (*1*). In the case of the mammalian proteasome, controlling the size distribution of peptide products is likely to be important physiologically, since peptide products that are 8−10 amino residues long or longer can be processed further and presented as epitopes to MHC class I molecules (*2*). In fact, characterization of peptide product sizes (*3*−*5*) shows that proteasomes do control product sizes: the product peptides are log-normally distributed in size (*5*, *6*), with peaks centered at 2−3, 9−10, and 20−30 residues (*3*). In their initial study, Goldberg and co-workers also commented on the implications of the peptide size distribution for the mechanism of peptide production (*5*), noting that the breadth of the observed distribution rules out a simple "molecular ruler" mechanism in which the spacing between protease active sites uniquely determines the peptide product size. They proposed that openings in the proteasome serve as a filter, trapping large products and allowing them to be cleaved to smaller ones. Subsequent studies showed that a mutant proteasome in which the central pore of the complex is constitutively open produces products that have a median length 40% greater than the wild-type due to both increased production of larger peptides (9−10 and 20−30 residues) and decreased production of smaller peptides (2−3 residues) (*3*). Interestingly, however, while the mutation changes the relative preference for generating peptides of different defined mean sizes, it does not change the mean sizes themselves. The mechanism by which the proteasome generates peptides of a defined mean size remains an open question.

The energy-dependent processive proteases of *Escherichia coli* may also exert active control over product sizes. Structural (*7*−*9*) and biochemical studies indicate that in energy-dependent proteases such as ClpXP and ClpAP, the ATPase[1] component (ClpX or ClpA) can unfold protein substrates and translocate them through an interior channel not much larger than a single polypeptide chain (*10*, *11*). The translocation of protein substrates brings them to ClpP, which has a central proteolytic chamber 50 Å in diameter (*9*). Openings in the chamber have been proposed to allow small peptide products to exit (*9*). However, when the protease active sites of ClpP are completely inactivated (through mutagenesis or chemical modification), the undigested protein substrate becomes trapped in ClpP (*12*, *13*). While it is known that ClpAP makes ∼20−30 cuts on average in casein (a natively unfolded substrate), suggesting that the average size of peptide products is 7−8 amino acid residues (*14*), the full distribution of peptide products has not previously been reported. Thus, the degree to which ClpAP can control its product sizes is unknown: the distribution of product sizes centered at 7−8 residues might be narrow or broad.

Unlike proteasomal products, the peptide products of ClpAP are believed not to have a biological function other than as a source of amino acids. It may nonetheless be

† Funding was provided by the MIT Department of Chemistry; K.-H.C. was a Lester Wolfe Fellow.

* To whom correspondence should be addressed: Department of Chemistry, MIT, Room 16-573, Cambridge, MA 02139. E-mail: lichts@mit.edu. Phone: 617-452-3525. Fax: 617-258-7847.

[1] Abbreviations: ATP, adenosine triphosphate; ADP, adenosine diphosphate; DFP, diisopropyl fluorophosphate; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HEPES, 4-(2hydroxyethyl)-1-piperazineethanesulfonic acid; Tris, tris(hydroxymethyl)aminomethane; GFP, green fluorescent protein; LB, Luria−Bertani; IPTG, isopropyl β-D-thiogalactoside; $OD_{600}$, optical density at 600 nm.

10.1021/bi0505060 CCC: $30.25 
Published on Web 09/29/2005

important for ClpAP to control the sizes of its products. PepN, one of the major peptidases in *E. coli* (*15*), processes its substrates much faster than ClpAP (*16*), so that the overall breakdown of protein substrates to free amino acids might be fastest when ClpAP generates medium-sized products that can be rapidly hydrolyzed by PepN. Control of peptide product sizes might also be useful in smooth processing of large protein substrates, as accumulation of large peptide products in the central cavity of ClpP might hinder further proteolytic processing of the unfolded substrate. Coordination of the ClpX and ClpP activities has been proposed to prevent the translocating peptide from clogging the exit pores of the ClpP tetradecamer (*17*).

ClpAP might control product sizes by several mechanisms. As previously proposed for the proteasome (*5*), the exit of large peptides from ClpP might be hindered (*9*), thereby increasing the probability that these peptides will be cleaved again into smaller fragments. Binding requirements of the protease might also contribute to product size control. If binding of an extended region of polypeptide sequence around the cleavage site is required for efficient proteolysis, proteolysis will be disfavored when shorter segments of polypeptide substrate are bound, leading to preferential generation of longer products.

The interplay between translocation and proteolysis is another source of possible mechanisms for product size control. The relative rates of translocation and proteolysis might control product sizes. Proteolysis that was rapid compared to translocation would allow many cleavage events while the polypeptide chain was translocating through ClpP, generating small products. Alternatively, allosteric coupling between the translocation and proteolytic activities might serve as the basis of a mechanism for control of product sizes. For both ClpA (*18*) and ClpX (*17*), there is allosteric coupling between the protease and the ATPase: binding of ClpP decreases the ATPase activity. If ClpP signals ClpA to stop translocation once a sufficiently large segment of polypeptide has entered the proteolytic chamber, ClpAP will be able to exert control over the range of product sizes.

In this study, we first investigate what product size distributions are predicted by four mechanistic alternatives for control of proteolysis. Each alternative mechanism makes a different prediction about the product size distribution and its sensitivity to biochemical perturbations of the protease. In the first mechanism (referred to hereafter as mechanism 1), polypeptide translocation is independent of proteolysis (Figure 1A), and both occur with constant probability per unit time. In mechanisms 2 and 3, the rate of exit from the protease is dependent on the size of the peptide product, and peptides that do not exit the protease may be re-cleaved to smaller products (Figure 1B). These mechanisms thus postulate that the protease acts as a filter, retaining large products but allowing smaller ones to escape. Mechanism 3 includes the additional feature that the rate of proteolytic cleavage is dependent on the size of the product. In mechanism 4, translocation and proteolysis regulate each other reciprocally: activation of ClpA turns off ClpP, and vice versa (Figure 1C). This mechanism also includes the feature that entry of the translocating polypeptide into ClpP triggers the conformational switch that controls reciprocal regulation.

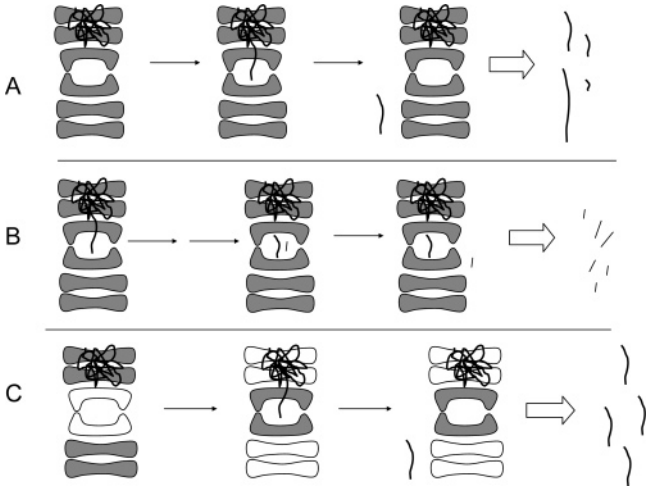


FIGURE 1: Possible mechanisms for control of product sizes by ClpAP. (A) Size control through modulation of relative rates of translocation and proteolysis (mechanism 1). ClpA and ClpP are active (gray) simultaneously, so translocation and proteolysis occur simultaneously. (B) Size control through use of a structural "filter". Small products escape from ClpP more readily than larger ones, which may be re-cleaved. Proteolysis may be random (mechanism 2) or biased toward longer products (mechanism 3). (C) Size control through allosteric coupling between translocation and proteolysis (mechanism 4). ClpA and ClpP reciprocally regulate each other's activity, so that when ClpA is active, ClpP is inactive (white). Translocation and proteolysis alternate.

While the form of the size distribution provides information about the mechanism of peptide product formation, the breadth of the size distribution of the products shows how precisely ClpAP can measure the sizes of its peptide products. To quantify the precision of product generation, we use Shannon information theory (*19*). This method of quantifying the precision of product size specification allows an estimation of the speed of measurement by the proteolytic machine and addresses the question of how much energy the proteolytic machine requires to carry out its measurement.

The results of our study provide evidence that ClpAP does exert active control over the sizes of its peptide products. The size distribution is consistent with a mechanism in which a controlled cycle of allosteric activation and inactivation of ClpP by ClpA controls the sizes of peptide products and inconsistent with the other mechanisms we examine. We also find that ClpAP exerts this control efficiently, with a rate and energy cost comparable to a rationally designed enzyme-based information processing system (*20*).

## MATERIALS AND METHODS

*ClpP-His$_6$ Purification.* A plasmid for expression of *E. coli* ClpP with a C-terminal His$_6$ tag was kindly provided by Profs. Robert Sauer and Tania Baker (MIT). ClpP-His$_6$ was expressed in *E. coli* strain DH5α/QE704/KI175 and purified as described (*13*), except that cells were grown at 37 °C and Tris buffers were replaced with HEPES buffers to avoid interference with the fluorescamine assay (see below).

*ClpA Purification.* The *clpA* gene in the overexpression vector pET9a was a generous gift from Profs. Robert Sauer and Tania Baker (MIT). The M169T mutation, which provides enhanced solubility and levels of full-length protein expression (*21*), was introduced into wild-type ClpA using

the QuikChange Site-Directed Mutagenesis Kit protocol (Stratagene), and the sequence was confirmed.

ClpA M169T protein was expressed in *E. coli* strain BL21/DE3/pLysS. Cells were grown at 37 °C in LB with kanamycin to an $OD_{600}$ of 0.6. IPTG was then added to a final concentration of 1 mM, and cells were transferred to 25 °C. After incubation for an additional 3 h at 25 °C, cells were harvested by centrifugation for 15 min at 6000*g* and purified as described (*22*) with the following modifications. All Tris buffers were replaced with HEPES buffers. Cells were lysed by sonication. After ammonium sulfate precipitation, the pellets containing ClpA were resuspended and loaded onto a Macroprep High S support (Bio-Rad) cation exchange column. An 80 mL linear gradient from 0.1 to 1 M KCl was applied. ClpA was eluted at approximately 0.6 M KCl.

*GFP-ssrA Purification.* A plasmid expressing GFP containing the S65G and S72A mutations that enhance the intensity of green fluorescence (*23*) and a C-terminal ssrA tag was kindly provided by Prof. Søren Molin (BioCentrum DTU, Denmark). Cells (*E. coli* JB401) expressing this protein were grown at 37 °C in 2YT with ampicilin to an $OD_{600}$ of 0.5. IPTG was added to a final concentration of 1 mg/L, and cells were transferred to 25 °C. After incubation for an additional 3 h at 25 °C, cells were harvested by centrifugation for 15 min at 6000*g* and purified using a published procedure (*24*), except that all Tris buffers were replaced with HEPES buffers.

*Protease Assay.* The rate of GFP-ssrA proteolysis by ClpAP was measured by monitoring the loss of GFP fluorescence (excitation at 467 nm and emission at 511 nm) using a microplate spectrofluorimeter (Molecular Devices Spectramax Gemini XS). Reactions were carried out at 37 °C using 0.2 $\mu$M hexametric ClpA ($ClpA_6$) and 0.1 $\mu$M tetradecameric ClpP ($ClpP_{14}$). The reaction buffer contained 50 mM HEPES, pH 7.5, 300 mM NaCl, 30 mM $MgCl_2$, 0.5 mM DTT, 10% glycerol, 0.32 mg/mL creatine phosphokinase (from rabbit muscle, Sigma), 30 mM phosphocreatine, and 10 mM ATP. ClpA and ClpP were incubated with ATP and all other components for 1 min on ice to enable ClpA and ClpAP complexes to assemble before addition of GFP-ssrA.

*ATPase Assay.* The rate of ATP hydrolysis by ClpA during GFP-ssrA degradation was measured using the rate of NADH oxidation (monitored by absorbance at 340 nm) coupled via pyruvate kinase and lactate dehydrogenase (*25*). Reactions were performed at 37 °C with 0.2 $\mu$M of $ClpA_6$, 0.1 $\mu$M of $ClpP_{14}$, and 10 $\mu$M GFP-ssrA in a buffer containing 50 mM HEPES, pH 7.5, 300 mM NaCl, 30 mM $MgCl_2$, 0.5 mM DTT, 10% glycerol, 0.23 mM NADH, 7.5 mM phosphoenolpyruvate, 19 U/mL pyruvate kinase (from rabbit muscle, Sigma), and 21 U/mL lactate dehydrogenase (from rabbit muscle, Sigma).

*Partial Inactivation of ClpP.* Partial inactivation of ClpP was performed by modification of the active site serine with DFP (*14*). ClpP (1.6 mg/mL) was incubated in 50 mM HEPES, pH 7.5, 200 mM KCl, 25 mM $MgCl_2$, 1 mM DTT, 0.1 mM EDTA, and 10% glycerol containing 4 mM DFP at room temperature for 90 min. Residual DFP was removed using size-exclusion chromatography (PD-10 column, Amersham Biosciences) equilibrated with the reaction buffer. The extent of labeling was determined by measuring peptidase activity with *N*-succinyl-Leu-Tyr-7-amido-4-methylcoumarin, a fluorogenic substrate. ClpP that was 72 ± 2% inactivated was used for further experiments.

*Reductive Methylation of GFP-ssrA.* Reductive methylation of GFP-ssrA was carried out as previously described (*26*) to prevent fluorescamine reaction with lysine residues (see Fluorescamine Assay for Peptide Concentration). GFP-ssrA that is completely alkylated (as determined by assay with 2,4,6-trinitrobenzenesulfonic acid (*27*)) has the same fluorescence properties and maximal degradation rate by ClpAP as unmodified GFP-ssrA (data not shown).

*Digestion of Methylated GFP-ssrA by ClpAP.* Digestion of methylated GFP-ssrA by ClpAP was carried out under the same conditions as used for the protease assays, except that the reaction mixture contained 0.1 $\mu$M $ClpA_6$, 0.05 $\mu$M $ClpP_{14}$, and 15 $\mu$M methylated GFP-ssrA. Similar experiments were carried out using α-casein (Sigma) (15 $\mu$M) as the substrate. Digestions of GFP-ssrA were also performed using subsaturating (0.1 mM) ATP or using partially inactivated ClpP. All digestions were allowed to proceed for 2−3 h. GFP-ssrA or casein degradation and ClpA autodegradation were monitored by SDS−PAGE.

*Size-Exclusion Chromatography of Peptide Products.* Peptide products were desalted using a reverse-phase cartridge (Sep-Pak C18, Waters) and concentrated by centrifugal evaporation. Desalted products were submitted for MALDI mass spectrometric analysis (MIT Biopolymers Facility). Size-exclusion HPLC was performed using a polyhydroxyethyl aspartamide column (200 mm × 4.6 mm, PolyLC) (*28*). The mobile phase was 200 mM sodium sulfate, pH 3.0 (adjusted with phosphoric acid), 5 mM potassium phosphate, and 25% acetonitrile. Peptide products were redissolved in the mobile phase and loaded onto the column. Fractions (0.5 min) were collected at a flow rate of 0.125 mL/min. To determine the apparent molecular weight of the peptides eluted, the column was calibrated with standard peptides in the 600−3500 Da range (*5*).

*Fluorescamine Assay for Peptide Concentration.* The relative amount of peptide in each fraction was measured using fluorescamine, which forms a fluorescent product on reaction with primary amines (*29*). Thirty microliters of each fraction from the polyhydroxyethyl aspartamide column was incubated with 15 $\mu$L of 100 mM sodium borate (pH 8.0), 15 $\mu$L of water, and 90 $\mu$L of 0.1% (w/v, in acetonitrile) fluorescamine for 5 min at room temperature. The fluorescamine solution was freshly prepared before use. The fluorescence emission was monitored at 510 nm at room temperature with an excitation wavelength of 380 nm.

## RESULTS

*Simulations of Product Size Distributions for Four Mechanisms of Proteolysis.* To simulate the generation of product peptides by mechanism 1, both translocation and proteolysis were modeled as stochastic processes with constant probability of occurrence per unit time (the structurally identical (*9*) protease active sites are assumed to be catalytically identical as well). At each time step in the simulation, the system might translocate one monomer unit and/or cleave the translocating chain to generate a product. The histogram of product sizes derived from this simulation (Figure 2A) shows an exponential distribution (more precisely, a geometric distribution with a large number of steps). Other

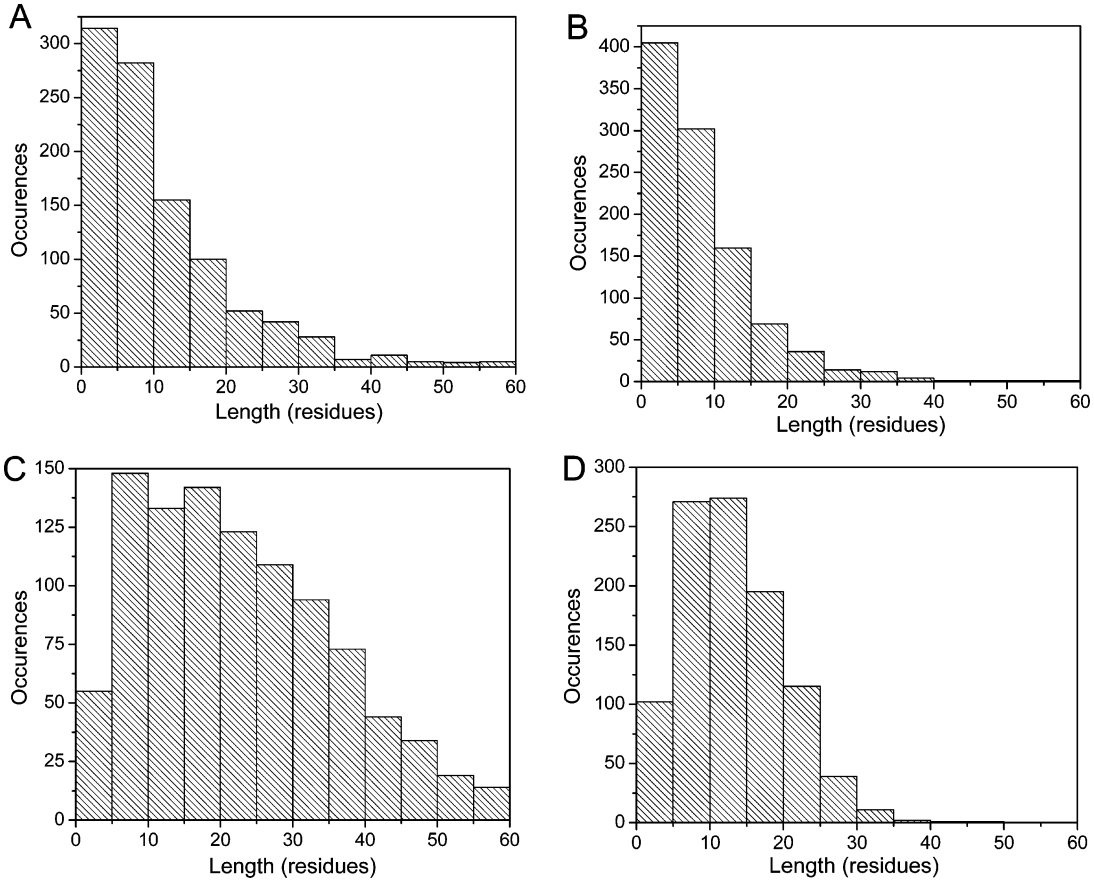


FIGURE 2: Simulations of the four mechanisms in Figure 1. In each case, ~1000 product-forming events were simulated. (A) Simulated product size distribution using mechanism 1 (translocation 10-fold faster than proteolysis); (B) simulated product size distribution using mechanism 2 (interaction parameter $\alpha = 0.08$); (C) simulated product size distribution using mechanism 3 (interaction parameter $\alpha = 0.08$, interaction parameter $\beta = 0.008$); and (D) simulated product size distribution using mechanism 4 (coupling parameter $\gamma = 80$).

variants of the independent mechanism also produce exponential product size distributions. Exponential product size distributions also result when translocation is deterministic (i.e., when a constant number of residues is translocated per unit time) but proteolysis is stochastic and when translocation and proteolysis are each partially rate-limiting (data not shown).

Mechanism 2 assumes that the larger peptides exit the protease more slowly than smaller ones and that peptide products that are retained in the protease can be re-cleaved to smaller products. The probability of retention in the protease is modeled as proportional to $\exp(\alpha L)$, where $L$ is the length of the peptide product and $\alpha$ is a parameter representing the strength of the interaction between the peptide and the protease. Numerical simulations of the second model (Figure 2B) generate product size histograms that, like those generated by the first model, have their maximum at the lowest molecular weights.

Mechanism 3 incorporates both a size-dependent rate of escape for products and a size-dependent rate of proteolytic cleavage. The product escape and proteolytic rates are modeled as proportional to $\exp(\alpha L)$ and $\exp(\beta L)$, respectively, where $L$ is the length of the peptide segment that has entered the protease and $\alpha$ and $\beta$ are parameters representing how product escape and proteolysis depend on the interaction of the peptide products/substrates with ClpAP. Numerical simulations of this model generate product size histograms that, depending on the parameters $\alpha$ and $\beta$, can have a nonzero peak and are skewed toward longer products (Figure 2C).

To simulate the generation of product peptides by mechanism 4, a switch between two states (ClpA active/ClpP inactive and ClpA inactive/ClpP active) is modeled as a stochastic process. The probability that the system will switch from the ClpA active/ClpP inactive state to the ClpA inactive/ClpP active state is modeled as proportional to $\exp(\gamma L)$, where $L$ is the number of monomer units already translocated into ClpP and $\gamma$ is a parameter representing the strength of the coupling. Numerical simulations of this model generate a product size distribution that has a nonzero peak and is skewed toward longer products (Figure 2D).

*Size Distribution of Peptides Derived from ClpAP-Catalyzed Proteolysis.* Peptide products derived from GFP-ssrA were separated using size-exclusion HPLC, and the amount of peptide in fractions corresponding to each size range was measured (*5*, *6*). Carrying out the digestion with a large excess of GFP-ssrA over ClpAP ensured that the peptide products were derived predominantly (>80%, data not shown) from GFP-ssrA rather than autodegradation (*30*, *31*) of ClpA. The observed distribution (Figure 3) is nonexponential. It has a peak at 760−900 Da (6−8 amino acid residues, assuming an average molecular weight of 119 Da/residue) and is skewed toward higher molecular weights.

Because the purification of the product peptides requires several chromatographic steps, it is possible that small peptides will be underrepresented in the observed distribu-

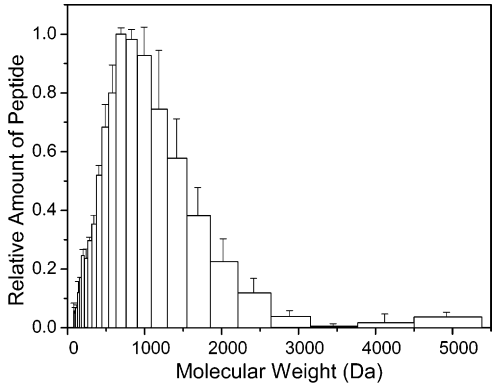


FIGURE 3: Size distribution of peptides derived from ClpAP-catalyzed proteolysis of GFP-ssrA.

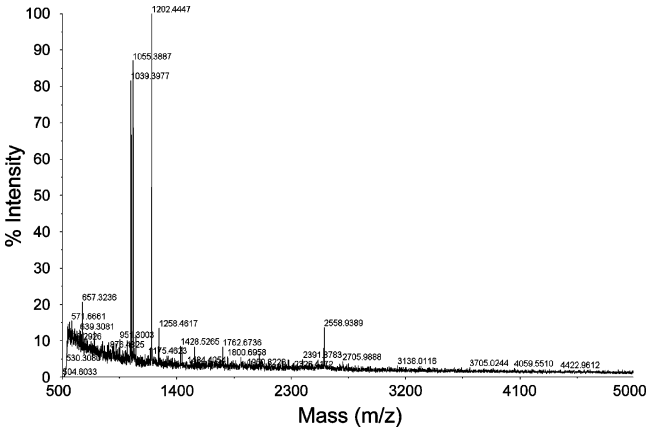


FIGURE 4: MALDI mass spectrum of peptide products.

tion. However, two lines of evidence indicate that this would be unlikely to cause an exponential distribution to appear to be nonexponential. First, studies on a variety of small peptides (three to four residues) have shown that solid-phase extraction using a C18 sorbent affords average recoveries of 85−90% (*32*). In the GFP-ssrA product mixture, peptides of molecular weight 300−400 Da are present in amounts 30−45% of the amount of peptides at the peak molecular weight, suggesting that the observed peak is not due to differential recovery of smaller peptides. Second, MALDI mass spectrometric measurements (Figure 4) qualitatively suggest that peptides ∼1000 Da dominate the product mixture, consistent with the quantitative measurement.

Sequence selectivity of proteolysis might contribute to the distribution of sizes. If sequence motifs that ClpP prefers are distributed nonuniformly throughout the sequence of the substrate, a nonuniform size distribution of product would result. To test the effect of sequence on size distribution, we measured the size distribution of peptide products derived from α-casein, a substrate unrelated to GFP-ssrA in primary sequence. The size distributions of both substrates were very similar (Figure 5), indicating that effects of primary sequence do not account for the size distributions.

*Effects of Kinetic Perturbations on Product Size Distribution.* The four mechanisms also make different predictions about the effects of artificially perturbing steps such as proteolytic cleavage or translocation. For mechanisms 1−3, the size distribution of products depends on relative rates of steps in the mechanism. In the case of mechanism 1, decreasing the rate of proteolysis relative to translocation would be expected to lead to larger products (Figure 6A),

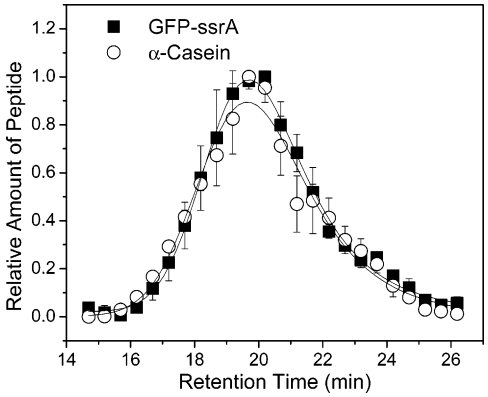


FIGURE 5: Effects of primary sequence of substrates on product size distribution. Comparison of product size distributions obtained using GFP-ssrA as substrate (filled squares) and using α-casein as substrate (open circles).

while decreasing the rate of translocation relative to proteolysis would lead to smaller products. Similarly, for mechanisms 2 and 3, the size distribution depends on the partitioning between proteolytic cleavage and escape of peptide products, so decreasing the rate of proteolysis relative to escape would be expected to generate larger products (Figure 6B,C). In contrast, mechanism 4 predicts that the size distribution of products is solely determined by the strength of the interaction that activates ClpP and inactivates ClpA, so this mechanism predicts that decreasing the rate of proteolysis will not affect the size distribution of products.

Experimentally, it is possible to decrease the rate of proteolysis by partially inactivating ClpP with the active site-directed reagent DFP (*14*). Decreasing the rate of translocation is not as straightforward, but by decreasing the concentration of ATP below its $K_m$, it is possible to decrease the translocation rate by decreasing the ATPase rate. Because natively unfolded substrates such as casein require ATP-dependent translocation (*12*) but not unfolding, the observation that the rate of proteolysis of casein is dependent on the ATP concentration (*18*, *22*) (apparent $K_m$ of 0.18 mM (*22*)) indicates that translocation will occur at a submaximal rate at subsaturating concentrations of ATP.

Neither of these experimental perturbations affects the product size distribution. Proteolysis in the presence of 0.1 mM ATP (the apparent $K_m$ for ATP in GFP degradation is 0.58 mM, as described in the next section) generates the same product size distribution as proteolysis in 10 mM ATP (Figure 7A). The peak at 23.5 min (Figure 7A) represents a small amount of contaminating adenine nucleotides, which control reactions have shown to react with fluorescamine. The material in this peak was identified as nucleotide by its UV spectrum, which exhibits a peak at 260 nm. In addition, a chromatogram of authentic nucleotides (ATP and ADP) alone shows a peak at 23.5 min of retention time (data not shown). Under conditions of subsaturating ATP, the absolute amount of peptide products (∼150 relative fluorescence units (RFU)) is about 1/8 of that under conditions of saturating ATP (∼1200 RFU), making the contribution of contaminating adenine nucleotides apparent. In the complementary experiment using partially inactivated ClpP, ClpP in which 70% of the active sites have been modified with DFP generates the same product size distribution as that observed using fully active ClpP (Figure 7B).

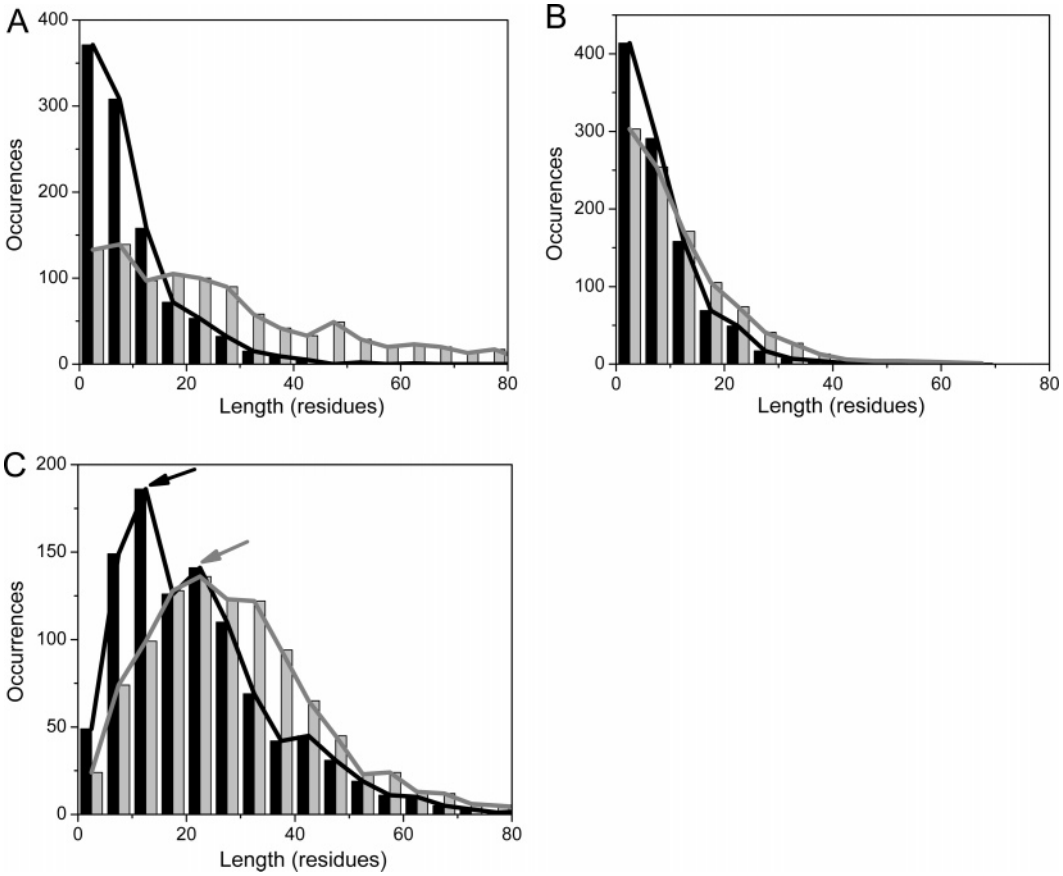


FIGURE 6: Simulated effects of decreasing proteolytic rate on the product size distribution. The rate of proteolysis is set at a reference value (0.05 per time step, histograms in black) or one-third of the reference value (histograms in gray). (A) Simulation using mechanism 1; (B) simulation using mechanism 2; and (C) simulation using mechanism 3. Arrows point to the maxima of the distributions.

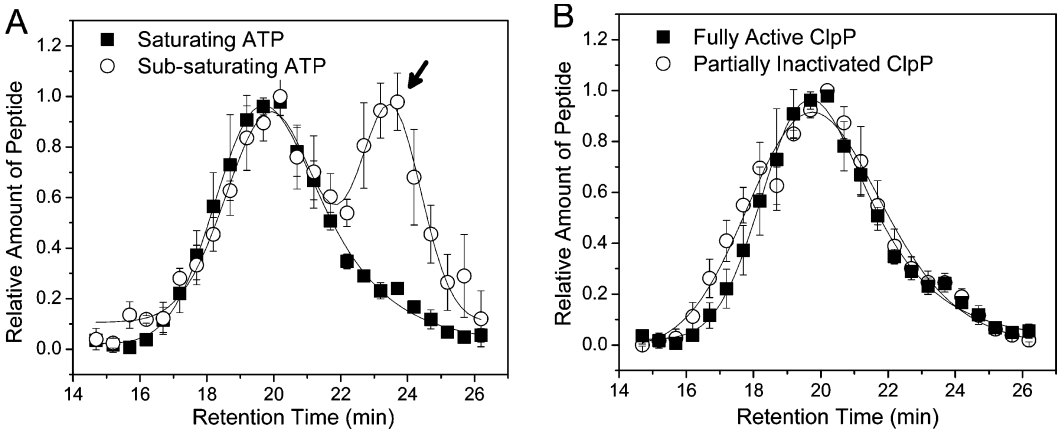


FIGURE 7: Effects of kinetic perturbations on product size distribution. (A) Comparison of product size distributions with saturating (10 mM) ATP (filled squares) and subsaturating (0.1 mM) ATP (open circles). The peak at 23.5 min retention time (arrow) is mostly contributed by nucleotides (ATP and ADP) not by the peptide products. (B) Comparison of product size distributions with fully active ClpP (filled squares) and partially inactivated ClpP (open circles).

*Kinetic Parameters for GFP Proteolysis and ATP Hydrolysis.* To help understand the rate and energy requirements for the measurement of product sizes, we measured steady-state kinetic parameters for two of the reactions that underlie this measurement: overall substrate proteolysis and ATP hydrolysis. Fitting the dependence of GFP proteolysis rate on [GFP] (see Supporting Information) to the Michaelis−Menten equation provided a $K_m$ of 4.9 ± 0.6 $\mu$M and a $k_{cat}$ (using [$ClpA_{12}P_{14}$] as the enzyme concentration, corresponding to a complex of 2 $ClpA_6$ and 1 $ClpP_{14}$ (*7*)) of 15 ± 1 $min^{-1}$. These results are similar to recently published values ($K_m$ of 5 ± 1 $\mu$M, $k_{cat}$ of 10 ± 1 $min^{-1}$) (*33*). Fitting the dependence of the ATPase rate during proteolysis of GFP on [ATP] (see Supporting Information) to the Hill equation (*34*) gave an apparent $K_m$ of 0.58 ± 0.02 mM, a $k_{cat}$ (using [$ClpA_{12}P_{14}$] as the enzyme concentration) of 44 ± 1 $s^{-1}$, and a Hill coefficient of 2.5 ± 0.2.

## DISCUSSION

We considered four mechanistic possibilities for processive proteolysis by ClpAP. In mechanism 1, the translocation of unfolded polypeptide by ClpA is independent of proteolytic

cleavage at the ClpP active sites, and size control comes from the relative rates of these processes. In mechanism 2, size control comes from size-dependent escape of products. Physically, the parameter $\alpha$ corresponding to the probability of retention of the product might correspond to some nonspecific hydrophobic interaction between the peptide product and the surface of the proteolytic chamber, where each residue of the peptide contributes a certain amount of free energy to the interaction. The relationship between $\alpha L$ and the retention probability is simulated as exponential because of the exponential relationship between activation energy and rate in activated rate theories such as Eyring or Kramers theories (*35*). In mechanism 3, size control comes from both size-dependent escape of products and a size-dependent cleavage reaction. Size-dependent cleavage would be possible if, for example, catalysis of amide bond hydrolysis depended on the presence of binding of substrate residues extending well outside the site of cleavage. Finally, in mechanism 4, allosteric communication between ClpA and ClpP allows translocation to alternate with proteolysis. This model thus assumes that each monomer unit reduces the barrier for the conformational switch by a certain amount of free energy. Physically, this might correspond to energy of binding of the unfolded polypeptide or strain energy that builds up as the unfolded polypeptide presses against the interior surface of the ClpP ring.

These mechanisms make distinct predictions about the nature of the size distribution of peptide products. The first mechanism predicts that the products will be exponentially distributed in size. If there is a single rate-limiting step in ClpAP-catalyzed proteolysis, the product size distribution would be expected to be a single exponential, because the waiting times between events are exponentially distributed for a process described by a single chemical step (*36*). Numerical simulations of this mechanism produce exponential product size distributions (Figure 2A), as expected. This mechanism is thus inconsistent with the experimental size distribution.

Mechanism 2 predicts a distribution that, like the exponential distribution, has its maximum at the lowest-sized products (Figure 2B). In this mechanism, size-dependent escape of products acts as a filter, preventing very large products from being formed. However, this filter does not prevent small products from escaping, so that nothing prevents the initial formation of small products due to random cleavages. This mechanism is thus, like the first mechanism, inconsistent with the observed product size distribution. The nature of any putative peptide filter is also somewhat problematic. Diffusion of even large polymers through narrow pores can occur on the time scale of milliseconds or less (*37*, *38*), orders of magnitude faster than the overall rate of turnover for ClpAP ($\sim$10 min$^{-1}$). While a structural filter may contribute to size control in ClpAP, it is not likely to consist of a static pore. Nonspecific binding of peptides to the inner surface of the protease might serve as an efficient filter, assuming that the residence time was dependent on hydrophobic interactions that would be larger for larger peptides. Access to the bulk solution that was gated by conformational changes (e.g., opening and closing of small apertures in the ClpP structure) might also serve as an effective filter.

To obtain a distribution of product sizes with a peak at a defined number of amino acid residues, ClpAP must have some way of cutting the polypeptide chain preferentially after translocation of a certain number of residues. One way to accomplish this would be to make the rate of proteolytic cleavage dependent on the size of the peptide segment bound to the protease active site, as postulated in mechanism 3. This mechanism includes the feature that the substrate binding site can accommodate a large number of amino acid residues extending from the site of cleavage. Mechanism 3 further postulates that when more of this extended binding site is occupied, the efficiency of cleavage is higher. Thus, this mechanism predicts preferential generation of larger products. Simulations of mechanism 3 show that preferential generation of larger products, in combination with a structural "filter," can lead to a product size distribution that, like the observed distribution, has a nonzero peak and is skewed toward higher molecular weight products (Figure 2C).

Mechanism 4 represents another way for ClpAP to control product sizes. This mechanism postulates that the active form of ClpA negatively regulates ClpP and vice versa, so that translocation alternates with proteolysis. Mechanism 4 also postulates that it is the translocation of unfolded polypeptide by the active form of ClpA into the inactive form of ClpP that triggers activation of ClpP, causing ClpA to stop translocating the substrate. When the peptide product leaves the ClpP active site, ClpP becomes inactive and ClpA becomes active, starting the cycle again. In effect, this mechanism involves ClpAP first measuring the size of the product, then cutting it.

Numerical simulations of this mechanism showed that it generates a product size distribution that qualitatively reproduces the features of the observed distribution. It has a nonzero peak and is skewed toward higher molecular weight products (Figure 2D). In this mechanism, ClpP's ability to trigger the allosteric activation of ClpP/deactivation of ClpA determines the peak size of products. Increasing the parameter $\gamma$ in the simulation, which represents ClpP's ability to trigger this allosteric activation/deactivation, leads to the formation of smaller products (data not shown).

The form of the product size distribution thus allows mechanisms 1 and 2 to be distinguished from mechanisms 3 and 4. To distinguish mechanism 4 from the other three mechanisms, we used subsaturating concentrations of ATP to hinder translocation and partial inactivation of ClpP to hinder proteolytic cleavage. In the first three mechanisms, size control comes about through control of the relative rates of translocation and proteolysis (mechanism 1) or the relative rates of proteolysis and product escape (mechanisms 2 and 3). Hindering either translocation (mechanism 1) or proteolysis (mechanisms 1, 2, and 3) is thus expected to alter the size distributions for these mechanisms (Figure 6). In contrast, hindering translocation or proteolysis individually is not expected to alter the product size distribution in mechanism 4, since the product sizes are not determined by the relative rates of translocation and proteolysis but by the ability of bound peptide to trigger the conformational switch that activates proteolysis. Mechanism 4 predicts that once substrate binding to ClpP initiates the conformational switch, translocation stops and no additional substrate enters ClpP. In this case, the peptide product formed will be of the

same size whether the proteolytic cleavage itself is slow or rapid.

Mechanism 4, unlike the simpler mechanisms considered, is consistent with all the available data. It is also consistent with previously reported results, such as the ability of ClpP to modulate the ATPase activity of ClpA (*18*) and ClpX (*17*) and the ability of active ClpA to translocate substrate into ClpP that has been inactivated by chemical modification or site-directed mutagenesis (*7*, *12*). In the latter case, trapping of large unfolded substrates in inactivated ClpP would be interpreted as resulting from the inability of the modified ClpP to assume the active conformation. Of course, the ability of mechanism 4 to account for the current experimental results does not preclude the ability of other mechanisms not considered here to do so. While mechanism 4 provides a useful starting point for mechanistic studies, further experimental work will be required to test its predictions. For example, experiments to determine whether ClpA and ClpP actually alternate activation states in the course of a single enzymatic turnover and whether translocation exhibits controlled step sizes will be required. Transient-state or single-molecule kinetic approaches may prove to be necessary to address this question, as they are sensitive to microscopic steps that are averaged out using steady-state techniques. In particular, single-molecule experiments are likely to be useful, as mechanism 4 predicts that the waiting time between proteolytic events will be nonexponentially distributed.

*Quantitative Characterization of Product Size Control by ClpAP*. ClpAP's ability to exercise control over the sizes of peptide products is analogous to carrying out a measurement. Any measurement carried out by a biomolecule will be subject to interference from thermal noise that reduces its precision (*39*). The effect of thermal noise might be reduced at the expense of slowing down the measurement by averaging it over a longer time. Expenditure of free energy is another way of reducing the effect of thermal noise (*39*). A molecular process with a sufficient number of irreversible steps can proceed in a regular, "clockwork"-like fashion (*40*). Mechanical precision of this kind could serve as the basis for the control of product sizes.

It would be thus informative to determine how ClpAP manages the compromise among precision, rate, and energy cost in controlling its product sizes. To address this question, we will need a way of quantifying the statistical uncertainty associated with a given size distribution. The amount of information required to specify the product size distribution is a measure of the precision with which the enzyme controls the size of its products.

Information theory (*19*), which was introduced by Claude Shannon and others and has since been applied to problems in biology and bioinformatics (*41*, *42*), provides a way to quantify information. The information content of a distribution expresses how much knowing the distribution reduces the uncertainty in knowing the measured quantity for an individual member of the ensemble. For example, if measurements of an experimental quantity are normally distributed with a mean of 50 and a standard deviation of 40, the distribution does not specify the value of an individual measurement precisely, while if the standard deviation is 2, the distribution contains sufficient information to specify the value of an individual measurement to a low level of uncertainty.

A decrease in uncertainty can be defined quantitatively as a decrease in statistical entropy (*19*). The uncertainty of a distribution can be defined as $\sum_i - p_i \ln(p_i)$, where $p_i$ is the probability that an individual member of the ensemble belongs to bin $i$ in a histogram (*19*). The statistical entropy of a distribution can be thus calculated as the difference between its uncertainty and that of a reference distribution. Here, we are interested in the statistical entropy of the observed size distribution compared to that of the size distribution that a completely random protease would be able to produce. We will therefore use as the reference distribution the exponential distribution, which random proteolysis would generate. The relative amount of peptide product in each bin of product sizes represents the probability of generating a product of that size, and the total uncertainty for a proteolytic cleavage is calculated by summing over all the bins in the experimental histogram. The specific reference distribution is an exponential distribution with the same mean as the experimental distribution, binned the same way as the experimental data. To quantify the uncertainty of the experimental size distribution, this distribution was fitted to a log-normal function, and the function was numerically integrated (using the quadrature function in MATLAB) over a range of 1−5000 Da. The uncertainty of an exponential distribution with the same mean size as the experimental distribution was calculated in the same way.

Calculating the statistical entropy of the experimentally observed product size distribution compared to an exponential distribution with the same mean provides a value of ~0.1 bits for the decrease in statistical entropy associated with specifying the observed distribution. The estimate of the statistical entropy of the product size distribution allows calculation of an order-of-magnitude estimate for the rate of information processing in product formation. Since the distribution of products is centered at 6−8 amino acid residues and GFP-ssrA is 251 amino acid residues long, a typical turnover will comprise about 30 proteolytic cleavages (in accord with previous estimates (*14*)). Assuming that the distribution of product sizes does not vary within a single turnover, the value of ~0.1 bits represents the information required to specify each proteolytic cleavage. Therefore, about 3 bits are required to specify the outcome of all the proteolytic cleavages in one turnover. The turnover number for GFP-ssrA is 15 $min^{-1}$, meaning that ClpAP can perform its size specification at a rate of ~1 bit/s. This estimate represents a lower limit: while a greater decrease in statistical entropy may be required to define the process that leads to product formation, a decrease of at least that magnitude is required in order to account for the observed product size distribution.

We can also ask how much free energy is dissipated in the course of specifying the product size distribution. The total free energy dissipated is an upper limit on the energetic cost of controlling product sizes. The maximal rate of ATP hydrolysis under turnover conditions is about $2 \times 10^3$ $min^{-1}$ (Figure 7B). Since hydrolysis of ATP under physiological conditions is exergonic by about 20 kT (*43*), ClpAP dissipates about $5 \times 10^4$ kT/min (hydrolysis of protein amides, which is much slower and is less exergonic,

contributes a relatively small amount to the total energy dissipation). The value of ~1 bit/s estimated above provides an estimate of $10^3$ kT/bit as a lower limit for the energy efficiency of product size specification. It is likely that the value of approximately $10^3$ kT/bit does not represent the minimum energy required for specifying the product size distribution, as mechanical steps such as unfolding will certainly consume an appreciable fraction of this energy without contributing directly to information processing.

*Biological Implications of Control of Sizes of Proteolytic Fragments.* While eukaryotic proteasomes need to control peptide product sizes because peptides below a certain length cannot be presented as antigens (*2*), there is not an obvious disadvantage to the production of very short peptides in *E. coli*. The question thus arises of whether there actually exists a selective pressure for control of peptide product sizes in *E. coli*. One possible explanation for the function of the observed size distribution of ClpAP products is that it optimizes the overall rate of degradation of misfolded proteins to free amino acids. The aminopeptidase PepN hydrolyzes short peptides to free amino acids and is responsible for most of the aminopeptidase activity in *E. coli* (*15*, *16*). Its steady-state rate of proteolysis is much greater than ClpP's for roughly comparable substrates ($k_{cat}$ of 370 $s^{-1}$ for hydrolysis of a *para*-nitroanilide substrate (*16*), compared to 2.5 $min^{-1}$ for ClpP's hydrolysis of a somewhat less activated aminomethylcoumarin substrate (*44*)). However, PepN degrades unfolded proteins very slowly (a half-time on the order of hours (*16*)). Thus, use of the slower peptidase ClpAP to digest misfolded proteins into peptides small enough for the faster peptidase PepN to accept would optimize the overall rate of processing.

It is also possible that control of product sizes is required to keep the translocating substrate from clogging the central pore of ClpAP. ClpXP (*45*) and the proteasome (*46*) can process substrates that are larger than its central pore. This ability suggests that smooth processing of large substrates will require energy-dependent proteases to have a way of preventing translocation of more unfolded polypeptide than can be productively accommodated into the central chamber of the protease. Translocation of only small segments of substrate before proteolysis and clearance of the ClpP active sites might help ClpAP avoid nonproductive binding of large segments of substrate to the interior surface of ClpP.

*General Mechanistic Features of Enzymes That Measure Product Sizes.* Other enzymes in a variety of organisms can also be thought of as "measuring" their substrates. In addition to the proteasome, which produces products with a size distribution similar to that observed for ClpAP, the nuclease Dicer also produces products with a very narrow size distribution (20−22 nucleotide residues) (*47*), and tripeptidylpeptidases cleave the first three amino acid residues from peptides (*48*). The mechanism for generation of a narrow size distribution of products may at first seem to follow trivially from the existence of multiple active sites in the enzyme structure. Dicer, for example, has two distinct nuclease active sites (*49*), suggesting that the physical spacing between them might determine the product size. Such a mechanism would be an oversimplification. Even when there is a defined spacing between active sites, the microscopic steps in the enzymatic mechanism must also be ordered in a way that allows control of product sizes. For example, a mechanism where an enzyme binds the substrate at two active sites, cleaves it at each of the sites sequentially, and releases the product will allow generation of uniformly sized products from a polymeric substrate. However, if the enzyme can deviate from the correct order of steps (e.g., if partially cleaved product can diffuse from the enzyme and re-bind at random), random cleavages will occur, decreasing the uniformity of product sizes.

One way to ensure a defined order of conformational changes is to drive them using dissipation of free energy, that is, to make each conformational change effectively irreversible. In the limit where every step is very far from equilibrium (i.e., irreversible), the enzyme always traverses the cycle in order (*40*). This would allow a complex algorithmic task to be carried out: first the enzyme does A, then B, then C, and so forth. The price paid for carrying out a complex task with high fidelity is thus the irreversible consumption of chemical energy. For proteases and nucleases, the irreversibility of the hydrolytic step may suffice to provide the free energy needed to determine the size distribution.

The phenotype of peptide product formation in ClpAP is remarkably similar to that observed for mammalian and thermophilic proteasomes (*4*, *5*). The product sizes for those enzymatic systems have a similar skew distribution and are similarly unaffected by partial inactivation of protease active sites (*4*). The proteasome may therefore have a similar mechanism of peptide product formation, in which translocation of polypeptide substrate alternates with proteolysis. There is evidence for allosteric coupling between chymotryptic sites and caspase sites in the 20S proteasome (*6*). The "measure and cut" and "bite-chew" (*6*) mechanisms are not mutually exclusive, and both may contribute to production of epitopes of the proper size and sequence.

It may be that, for proteasomes, the ability to measure product sizes serves as an integral part of the generation of epitope-like peptides from protein substrates. A protease that initiated proteolysis at one end of the substrate polypeptide and processively cleaved off peptides of a constant size from that end would generate a unique set of products even without being able to cleave specific sequences preferentially. For such an enzyme, the mechanical precision of product size measurement, combined with the ability to cleave processively without slipping, would allow preferential generation of a small subset of possible peptide products.

## NOTE ADDED IN PROOF

While this paper was under review, a phenomenological mathematical model describing product size control in the proteasome was published (Luciani, F., Kesmir, C., Mishto, M., Or-Guil, M., and de Boer, R. J. (2005) A Mathematical Model of Protein Degradation by the Proteasome, *Biophys. J. 88*, 2422−2432).

## ACKNOWLEDGMENT

We thank Bob Sauer, Tania Baker, and Søren Molin for donating plasmids; Randy Burton and Julia Flynn for advice on experimental procedures; Virginia Cornish, Catherine Drennan, and Tom Magliery for comments on the manuscript; and Muriel Medard, Desmond Lun, Ralf Koetter, and members of the Licht group for helpful discussions.

## SUPPORTING INFORMATION AVAILABLE

The steady-state kinetic data for GFP-ssrA hydrolysis and ATP hydrolysis and the MATLAB scripts used for numerical simulations are available as Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES


1. Pickart, C. M., and Cohen, R. E. (2004) Proteasomes and their kin: proteases in the machine age, *Nat. Rev. Mol. Cell Biol. 5*, 177−187.
2. Rock, K. L., York, I. A., Saric, T., and Goldberg, A. L. (2002) Protein degradation and the generation of MHC class I-presented peptides, *Adv. Immunol. 80*, 1−70.
3. Kohler, A., Cascio, P., Leggett, D. S., Woo, K. M., Goldberg, A. L., and Finley, D. (2001) The axial channel of the proteasome core particle is gated by the Rpt2 ATPase and controls both substrate entry and product release, *Mol. Cell l7*, 1143−1152.
4. Kisselev, A. F., Akopian, T. N., Woo, K. M., and Goldberg, A. L. (1999) The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes−implications for understanding the degradative mechanism and antigen presentation, *J. Biol. Chem. 274*, 3363−3371.
5. Kisselev, A. F., Akopian, T. N., and Goldberg, A. L. (1998) Range of sizes of peptide products generated during degradation of different proteins by archaeal proteasomes, *J. Biol. Chem. 273*, 1982−1989.
6. Kisselev, A. F., Akopian, T. N., Castillo, V., and Goldberg, A. L. (1999) Proteasome active sites allosterically regulate each other, suggesting a cyclical bite-chew mechanism for protein breakdown, *Mol. Cell 4*, 395−402.
7. Ishikawa, T., Beuron, F., Kessel, M., Wickner, S., Maurizi, M. R., and Steven, A. C. (2001) Translocation pathway of protein substrates in ClpAP protease, *Proc. Natl. Acad. Sci. U.S.A. 98*, 4328−4333.
8. Guo, F. S., Maurizi, M. R., Esser, L., and Xia, D. (2002) Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease, *J. Biol. Chem. 277*, 46743−46752.
9. Wang, J. M., Hartling, J. A., and Flanagan, J. M. (1997) The structure of ClpP at 2.3 angstrom resolution suggests a model for ATP-dependent proteolysis, *Cell 91*, 447−456.
10. Hoskins, J. R., Yanagihara, K., Mizuuchi, K., and Wickner, S. (2002) ClpAP and ClpXP degrade proteins with tags located in the interior of the primary sequence, *Proc. Natl. Acad. Sci. U.S.A. 99*, 11037−11042.
11. Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A., Hersch, G. L., Joshi, S. A., Kenniston, J. A., Levchenko, I., Neher, S. B., Oakes, E. S. C., Siddiqui, S. M., Wah, D. A., and Baker, T. A. (2004) Sculpting the proteome with AAA+ proteases and disassembly machines, *Cell 119*, 9−18.
12. Singh, S. K., Grimaud, R., Hoskins, J. R., Wickner, S., and Maurizi, M. R. (2000) Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP, *Proc. Natl. Acad. Sci. U.S.A. 97*, 8898−8903.
13. Kim, Y. I., Burton, R. E., Burton, B. M., Sauer, R. T., and Baker, T. A. (2000) Dynamics of substrate denaturation and translocation by the ClpXP degradation machine, *Mol. Cell 5*, 639−648.
14. Thompson, M. W., Singh, S. K., and Maurizi, M. R. (1994) Processive degradation of proteins by the ATP-dependent Clp protease from *Escherichia coli*: requirement for the multiple array of active sites in ClpP but not ATP hydrolysis, *J. Biol. Chem. 269*, 18209−18215.
15. Chandu, D., and Nandi, D. (2003) PepN is the major aminopeptidase in *Escherichia coli*: insights on substrate specificity and role during sodium-salicylate-induced stress, *Microbiology 149*, 3437−3447.
16. Chandu, D., Kumar, A., and Nandi, D. (2003) PepN, the major Suc-LLVY-AMC-hydrolyzing enzyme in *Escherichia coli*, displays functional similarity with downstream processing enzymes in archaea and eukarya−implications in cytosolic protein degradation, *J. Biol. Chem. 278*, 5548−5556.
17. Joshi, S. A., Hersch, G. L., Baker, T. A., and Sauer, R. T. (2004) Communication between ClpX and ClpP during substrate processing and degradation, *Nat. Struct. Mol. Biol. 11*, 404−411.
18. Hwang, B. J., Woo, K. M., Goldberg, A. L., and Chung, C. H. (1988) Protease Ti, a New ATP-dependent protease in *Escherichia coli*, contains protein-activated ATPase and proteolytic functions in distinct subunits, *J. Biol. Chem. 263*, 8727−8734.
19. Cover, T. M., and Thomas, J. A. (1991) *Elements of Information Theory*, Wiley, New York.
20. Benenson, Y., Paz-Elizur, T., Adar, R., Keinan, E., Livneh, Z., and Shapiro, E. (2001) Programmable and autonomous computing machine made of biomolecules, *Nature 414*, 430−434.
21. Seol, J. H., Baek, S. H., Kang, M. S., Ha, D. B., and Chung, C. H. (1995) Distinctive roles of the two ATP-binding sites in ClpA, the ATPase component of protease Ti in *Escherichia coli*, *J. Biol. Chem. 270*, 8087−8092.
22. Maurizi, M. R., Thompson, M. W., Singh, S. K., and Kim, S. H. (1994) Endopeptidase Clp: ATP-dependent Clp protease from *Escherichia coli*, *Methods Enzymol. 244*, 314−331.
23. Delagrave, S., Hawtin, R. E., Silva, C. M., Yang, M. M., and Youvan, D. C. (1995) Red-shifted excitation mutants of the green fluorescent protein, *Bio/Technology 13*, 151−154.
24. Yakhnin, A. V., Vinokurov, L. M., Surin, A. K., and Alakhov, Y. B. (1998) Green fluorescent protein purification by organic extraction, *Protein Expression Purif. 14*, 382−386.
25. Nørby, J. G. (1988) Coupled assay of $Na^+$, $K^+$-ATPase activity, *Methods Enzymol. 156*, 116−119.
26. Rypniewski, W. R., Holden, H. M., and Rayment, I. (1993) Structural consequences of reductive methylation of lysine residues in hen egg-white lysozyme: an X-ray-analysis at 1.8-Å resolution, *Biochemistry 32*, 9851−9858.
27. Habeeb, A. F. (1966) Determination of free amino groups in proteins by trinitrobenzenesulfonic acid, *Anal. Biochem. 14*, 328−336.
28. Silvestre, M. P. C., Hamon, M., and Yvon, M. (1994) Analysis of protein hydrolysates. 1. Use of poly(2-hydroxyethylaspartamide)-silica column in size-exclusion chromatography for the fractionation of casein hydrolysates, *J. Agric. Food Chem. 42*, 2778−2782.
29. Udenfriend, S., Stein, S., Bohlen, P., Dairman, W., Leimgruber, W., and Weigele, M. (1972) Fluorescamine: a reagent for assay of amino acids, peptides, proteins, and primary amines in the picomole range, *Science 178*, 871−872.
30. Gottesman, S., Clark, W. P., and Maurizi, M. R. (1990) The ATP-dependent Clp protease of *Escherichia coli*−sequence of ClpA and identification of a Clp-specific substrate, *J. Biol. Chem. 265*, 7886−7893.
31. Dougan, D. A., Reid, B. G., Horwich, A. L., and Bukau, B. (2002) ClpS, a substrate modulator of the ClpAP machine, *Mol. Cell 9*, 673−683.
32. Herraiz, T., and Casal, V. (1995) Evaluation of solid-phase extraction procedures in peptide analysis, *J. Chromatogr., A 708*, 209−221.
33. Xia, D., Esser, L., Singh, S. K., Guo, F. S., and Maurizi, M. R. (2004) Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone, *J. Struct. Biol. 146*, 166−179.
34. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, W. H. Freeman, San Francisco, CA.
35. Hänggi, P., Talkner, P., and Borkovec, M. (1990) Reaction-rate theory−50 years after Kramers, *Rev. Mod. Phys. 62*, 251−341.
36. Colquhoun, D., and Hawkes, A. G. (1995) The principles of the stochastic interpretation of ion-channel mechanisms, in *Single-Channel Recording* (Sakmann, B., and Neher, E., Eds.) pp 397−482, Plenum Press, New York.
37. Kasianowicz, J. J., Brandin, E., Branton, D., and Deamer, D. W. (1996) Characterization of individual polynucleotide molecules using a membrane channel, *Proc. Natl. Acad. Sci. U.S.A. 93*, 13770−13773.
38. Movileanu, L., and Bayley, H. (2001) Partitioning of a polymer into a nanoscopic protein pore obeys a simple scaling law, *Proc. Natl. Acad. Sci. U.S.A. 98*, 10137−10141.

39. Bennett, C. H. (1982) The thermodynamics of computation−a review, *Int. J. Theor. Phys. 21*, 905−940.
40. Svoboda, K., Mitra, P. P., and Block, S. M. (1994) Fluctuation analysis of motor protein movement and single-enzyme kinetics, *Proc. Natl. Acad. Sci. U.S.A. 91*, 11782−11786.
41. Carothers, J. M., Oestreich, S. C., Davis, J. H., and Szostak, J. W. (2004) Informational complexity and functional activity of RNA structures, *J. Am. Chem. Soc. 126*, 5130−5137.
42. Schneider, T. D. (1994) Sequence logos, machine channel capacity, Maxwell demon, and molecular computers−a review of the theory of molecular machines, *Nanotechnology 5*, 1−18.
43. Voet, D., and Voet, J. (2004) *Biochemistry*, 3rd ed., Wiley, Hoboken, NJ.
44. Thompson, M. W., and Maurizi, M. R. (1994) Activity and specificity of *Escherichia coli* ClpAP protease in cleaving model peptide substrates, *J. Biol. Chem. 269*, 18201−18208.
45. Burton, R. E., Siddiqui, S. M., Kim, Y. I., Baker, T. A., and Sauer, R. T. (2001) Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine, *EMBO J. 20*, 3092−3100.
46. Lee, C., Prakash, S., and Matouschek, A. (2002) Concurrent translocation of multiple polypeptide chains through the proteasomal degradation channel, *J. Biol. Chem. 277*, 34760−34765.
47. Zamore, P. D. (2001) RNA interference: listening to the sound of silence, *Nat. Struct. Biol. 8*, 746−750.
48. Tomkinson, B. (1999) Tripeptidyl peptidases: enzymes that count, *Trends Biochem. Sci. 24*, 355−359.
49. Zhang, H. D., Kolb, F. A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004) Single processing center models for human dicer and bacterial RNase III, *Cell 118*, 57−68.

BI0505060