Decelerated degradation of short peptides by the 20S proteasome

Iztok Dolenc[a,b], Erika Seemüller[a], Wolfgang Baumeister[a,*]

[a] *Max-Planck-Institut fur Biochemie, Am Klopferspitz 18a, Martinsried D-82152, Germany*
[b] *Department of Biochemistry and Molecular Biology, J. Stefan Institute, Ljubljana, Slovenia*

Received 31 July 1998

Abstract Based on a twelve residue master peptide comprising all five specific cleavage sites defined for the proteasome, a set of variant peptides was generated in order to probe specificity and to elucidate the mechanism which determines product size. It is shown that the rate of degradation by the 20S proteasome from *Thermoplasma acidophilum* depends critically on the length of the peptide substrate. Peptides of 14 residues and longer are degraded much faster than shorter peptides although the sites of cleavage remain unchanged. The decelerated degradation of peptides shorter than 14 residues explains the accumulation of products with an average length of seven to nine residues.

Key words: Proteasome; Molecular ruler; Protein degradation; *Thermoplasma acidophilum*

1. Introduction

Self-compartmentalization has begun to emerge as a regulatory principle common to several proteolytic enzymes (for recent reviews see [1,2]). The proteasome became the paradigm of a class of proteases unrelated in sequence and fold which have converged towards a common architecture: the rather large homo- or heterooligomeric assemblies enclose inner cavities, several nanometers in size, which allow them to confine the proteolytic action. Access to these nanocompartments is restricted to unfolded polypeptides which can pass through the narrow gates guarding the entrance. Therefore, these proteases must interact with a machinery capable of recognizing target proteins and rendering them degradable, i.e. presenting them in an unfolded form to the protease. In eukaryotic proteasomes these functions are provided by the 19S cap complexes which associate with the 20S proteolytic core complex to form the 26S proteasome. Related but much simpler precursor complexes were recently described for bacteria [3] and archaea (Zwickl, P., personal communication). The 20S complex is a barrel-shaped structure of four stacked rings. Each outer ring contains either seven identical (bacteria, archaea) or seven distinct but related (eukaryotes) α-type subunits; the two adjacent inner rings contain seven identical or related β-type subunits. Jointly, the two β-rings form the central cavity from where the active sites are accessible. Proteasomes degrade substrates by a novel catalytic mechanism, in which the hydroxyl group of the N-terminal threonine of a β-type subunit serves as the active site nucleophile [4,5]. While in archaeal and bacterial proteasomes all 14 β-type subunits are proteolytically active, in eukaryotic proteasomes only 6 of the 14 β-type subunits have an N-terminal threonine and thus contribute to the proteolytic activity. In order to access the proteolytic central cavity of the 20S complex substrate proteins must pass the narrow constrictions, an outer one formed by the α-subunit ring and an inner one formed by the β-subunit ring. Beyond allowing to control proteolysis, the confinement of the proteolytic action to the central cavity also provides the basis for the proteasome's processive mode of action: it does not release peptides after a single cleavage but proceeds to make multiple cleavages before discharging the degradation products [6,7]. When the specificity of proteasomes is probed with short fluorogenic peptides three major peptidase activities are found in eukaryotic 20S proteasomes: a chymotrypsin-like (P1 residue in substrate is hydrophobic or aromatic amino acid), a trypsin-like (P1 is basic amino acid) and a peptidylglutamyl-peptide hydrolyzing (P1 is glutamate) activity. Additionally, a branched-chain amino acid preferring and a small neutral amino acid preferring activity have been defined [8]. With the same type of assay, archaeal and bacterial proteasomes exhibit only chymotrypsin-like activity [9,10]. However, when longer peptides or proteins are used as substrates the very complex cleavage patterns do not correlate with the aforementioned specificities [11,12]. Obviously an assignment of specificities using the residue directly adjacent to the cleavage site (the P1 position) falls short of reality. Recent studies, using irreversible inhibitors of proteasome activity indicate that the P4 position has a critical role in defining specificity [13].

It is intriguing to observe that, in spite of cleaving protein substrates in an apparently non-specific manner, the peptide products fall into a relatively narrow size range, averaging around 7–9 residues. It is this feature, which predisposed the proteasome for a role it assumed in the course of evolution, namely the generation of immunocompetent peptides [14–16]. The observation that peptide products have a restricted range of sizes led to the proposal that proteasomes may possess an intrinsic molecular ruler. One of the options considered at the time was that the distance between active sites, acting in concert, could provide the mechanistic basis for such a ruler [11]. When the crystal structure of the *Thermoplasma* proteasome revealed a distance of 2.8 nm between neighboring active sites, which corresponds to a hepta- or octapeptide in an extended conformation, this seemed to provide strong evidence in support of such a molecular ruler hypothesis [4]. On the other hand, recent more quantitative analyses of product length, while in agreement with an average length of 8 ± 1 residues, showed larger size variations, which are difficult to reconcile with a purely geometry-based ruler which should yield products more focused in length [17].

In this communication we report on degradation experi-

*Corresponding author. Fax: (49) (89) 8578 2641.

Abbreviations: AMC, 7-amino-4-methylcoumarin; HPLC, high performance liquid chromatography; IPTG, isopropyl-1-thio-β-D-galactopyranoside; Ni-NTA, nickel-nitrilotriacetic acid; RP, reverse phase; Suc, succinyl

PII: S0014-5793(98)01010-2

ments performed with *Thermoplasma* proteasomes and a set of synthetic peptide substrates which were designed to identify features which (co)determine cleavage sites and product size. To this end we have chosen a master peptide of 12 residues (Fig. 1) which contains all five of the aforementioned specific cleavage sites defined for proteasomes at least once. This peptide is derived from the human hemoglobin A chain and includes residues 23–31. We have mutated and permutated this sequence and extended it in length up to 23 residues by (partial) sequence duplication; thus while changing the length of the peptide, the cleavage sites remain the same (Fig. 2). Surprisingly, it turns out that the rate of degradation depends critically on the length of the substrate peptides.

2. Materials and methods

2.1. Materials

BL21(DE3) cells were obtained from Stratagene (Germany) and Ni-NTA resin from Qiagen (Germany), and oxidized insulin β-chain was from Sigma (Germany). Peptides were synthesized on a multiple peptide synthesizer (MultiSynTech) by conventional fluorenylmethyloxycarbonyl chemistry and the purity was determined by reverse phase HPLC. All other chemicals were of analytical grade.

2.2. Expression and purification of recombinant Thermoplasma proteasomes

BL21(DE3) cells were transformed with plasmid pRSET5a, containing the genes for the wild-type α-subunit and β-subunit of the *Thermoplasma acidophilum* proteasome with a His_6-tag at the C-terminus of the β-subunit. Cells were grown in SOB medium to mid-log phase at 37°C. Five hours after induction with 1 mM IPTG cells were harvested, suspended in sonication buffer (50 mM Na-phosphate, 300 mM NaCl, pH 8.0), and sonicated for 10 min (Sonifier 250, Branson). The lysate was centrifuged (30 000×*g*, 1 h), and the supernatant was applied to a Ni-NTA affinity column (0.65×4 cm), equilibrated in sonication buffer. Unbound proteins were removed with wash buffer (50 mM Na-phosphate, 300 mM NaCl, pH 6.0) and His_6-tagged proteasomes were eluted with a gradient of 0–500 mM imidazole in wash buffer. Fractions containing pure protein were pooled and dialyzed against 50 mM Tris-HCl buffer with 1 mM EDTA, pH 7.5. The enzyme was stored at 4°C.

2.3. Hydrolysis of substrates and analysis of products

Degradation experiments with single peptide substrate or a mixture of different peptide substrates were performed in 100 mM Tris-HCl buffer with 1 mM EDTA, pH 7.5. The reaction volume was kept at 200 μl, with a final substrate concentration of 200–400 μM, except in determination of k_{cat}/K_M where the concentration was 50 μM. The mixture was preheated to 60°C before initiation of the reaction by addition of 10 μg recombinant proteasomes. At various time points 10-μl aliquots were removed and the reaction was stopped with 110 μl 0.03% trifluoroacetic acid and the sample stored at −20°C. The consumption of substrates was observed by application of samples to an RP-HPLC column.

Products of peptide hydrolysis were analyzed by HPLC (Gold System; Beckman) containing the programmable solvent module 125 and a diode array detector module 168. For peptide separation, a 2-mm column (C18, LiChroCart 125-2) from Merck (Germany) was used. The column was equilibrated in 0.1% (v/v) trifluoroacetic acid in water. Peptides were eluted from the RP column with a linear gradient of 0–60% acetonitrile in 0.1% trifluoroacetic acid at a flow rate of 0.3 ml/min over a 60-min period. Absorption of the eluted peptides was monitored at 206 nm and at 280 nm. Fractions coinciding with absorption maxima were manually collected, concentrated by evaporation and analyzed by Edman degradation on a gas-phase sequencer 477A equiped with PTH-analyzer 120A (both Applied Biosystems) and/or by mass spectrometry. Masses were determined with a tandem quadrupole instrument API III equipped with an atmospheric pressure ionization source (Sciex, Thornhill, Ont., Canada). The instrument *m/z* scale was calibrated with ammonium adduct ions of polypropylene glycol.

2.4. Determination of k_{cat}/K_M

Kinetic data for substrate hydrolysis were determined from substrate consumption derived from HPLC chromatograms. k_{cat}/K_M was calculated from the velocity relationship $v=k_{cat}/K_M[E_o][S]$ (working at low substrate concentration), where $[E_o]$ represents the total enzyme concentration and [S] the substrate concentration. For k_{cat}/K_M determination only measurements were used, where maximally 20% of substrate was consumed.

3. Results

As shown in Fig. 1 the 12-residue peptide HEYGAEALERAG (P12a) contains all five cleavage sites defined for eukaryotic proteasomes including a 'chymotryptic' site (Y3-G4), which according to assays with fluorogenic peptides is the preferred cleavage site for archaeal and bacterial proteasomes. With this peptide we have performed degradation experiments as described in Section 2; at several time points the formation of degradation products was monitored by reversed phase HPLC (Fig. 3). When approximately 50% of the peptides were consumed ($t_{1/2}$) the products were analyzed by mass spectrometry and, when necessary for an unambiguous identification, also by N-terminal sequencing. Degradation of peptide P12a yielded two predominant products, HEYGA and EALERAG, implying that the bond between A5 and E6 is the preferred cleavage site. In addition, two minor cleavage sites have been identified between A7 and L8 and between R10 and A11, respectively. From the estimated ratio 100:10:1 for cleavages of (A5-E6)/(A7-L8)/(R10-A11), the dominance of the A5-E6 cleavage site is obvious. Thus the principal cleavage site corresponds neither to the 'chymotryptic' nor to any of the other four cleavage sites described for proteasomes. To find out whether the preference for the A5-E6 site is due to the alanine at position P1, this residue was replaced by aspartic acid in peptide P12b (Fig. 2). This substitution in fact rendered the peptide resistant to degradation, while exchange of the residue in the P1′ position (E to L) had no effect at all, i.e. efficient cleavage occurred between A5 and L6 (peptide P12c) (Fig. 3B,C). This is important because the sequon AL when occurring in a different context (for example in the same peptide (A7-L8)) or in the insulin β-chain (see below) is not a major cleavage site. When put into a completely different context, such as peptides IQVYSRHPAENGK or IEKELRHGKSAEAT the sequon AE is not used at all as a cleavage site (data not shown). Obviously the P1 and P1′ positions are insufficient as determinants of cleavage sites. In order to examine whether the location of sequon AE within the peptide had an effect, peptides P12d

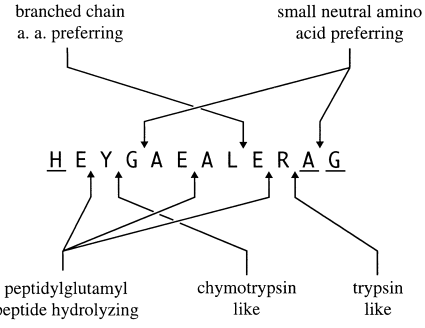

Fig. 1. Master peptide of 12 residues. The sequence is derived from the human hemoglobin A chain and contains all five of the specific cleavage sites defined for proteasomes.

and P12e (Fig. 2) were designed in which this sequon is shifted towards the C-terminus. The HPLC chromatograms show that the changed location of AE within the peptide had no effect on its preferred cleavage (Fig. 3D,E).

For a geometry-based molecular ruler one would predict that distance is a more important criterion for selection of a cleavage site than sequence. We have therefore extended the length of the master peptide P12a N-terminally. A phenylalanine was inserted into the −1 position to facilitate the identification of products; otherwise we have repeated sequence motifs from the master peptide in order to keep potential cleavage sites and their context unchanged. In the longest peptide P23 the core motif EYGAEALER occurs therefore twice in a tandem arrangement (Fig. 2). Degradation of P23 yielded three major products (Fig. 3F) and their analysis revealed that the preferred cleavage sites correspond to those identified previously in peptide P12a, i.e. irrespective of the location in the peptide the two AE sequons are the principal cleavage sites. Also in peptides P14 and P16 with shorter N-terminal extensions AE remains the principal cleavage site.

The insulin β-chain has been used as a model substrate previously with *Thermoplasma* [11] as well as with eukaryotic proteasomes [12]. In most of these studies the degradation was allowed to proceed over rather long periods of time in order to characterize the end products. In the context of this study we were interested in determining the preferred cleavage sites. Degradation products of insulin β-chain at $t_{1/2}$ are shown in Fig. 2. The dominant cleavage sites are between G8 and S9, between E13 and A14 and between V18 and C19. Minor sites are between A14 and L15, between L17 and V18, between F24 and F25 and between F25 and Y26.

When monitoring degradation of the aforementioned substrates in a time-dependent manner by reverse phase HPLC we observed that the longer substrates were degraded more rapidly than the shorter ones. To obtain more quantitative insight into this phenomenon we determined apparent second order rate constants k_{cat}/K_M for each of the substrates (Table 1). A comparison of the k_{cat}/K_M values shows a large difference between the shortest peptide P12a and the other peptides. The value of approx. 350 M^{-1} s^{-1} for peptide P12a is almost 20 times lower than for peptide P14 which is two residues longer, although the cleavage sites are unchanged. Interestingly, it is in the same range as the values typically found with short fluorogenic peptides such as Suc-LLVY-AMC which are commonly used for assaying proteasome activity (see e.g. [9]). When increasing the length of the substrates further (14–23 residues) a marginal decrease of velocity is observed. The rate constant for the degradation of the oxidized insulin β-chain is approx. 25 000 M^{-1} s^{-1}, i.e. this substrate is degraded about 70 times faster than P12a and about 4 times faster than the P14–P23 group of substrates. From the ratios of k_{cat}/K_M one would predict a competition between

Table 1
k_{cat}/K_M values for degradation of peptide substrates of different length by the proteasome from *Thermoplasma acidophilum*

Substrate	k_{cat}/K_M (M^{-1} s^{-1})
Insulin β-chain	25 000
P23	5500
P16	6000
P14	6500
P12a	350

Fig. 2. The set of peptides used in degradation experiments. The master peptide P12a was modified as indicated by boxes. Cleavages obtained by proteasome degradation are marked by arrows. Minor cleavage sites in oxidized insulin β-chain are marked by small arrows.

substrates differing in length, in which the longer ones are degraded preferentially.

To test this prediction experimentally, we performed competition experiments with two or three different substrates (Fig. 4). As shown in the time-dependent elution patterns it is indeed always the longer peptide which is degraded preferentially when equimolar mixtures of peptides are exposed to the proteasome. When P12a and P23 are mixed, P23 is degraded much faster than P12a (Fig. 4A). Addition of another longer peptide (insulin β-chain at the same molar concentration) changes the order of degradation: now the insulin β-chain is the fastest to be degraded, followed by degradation of P23; the concentration of the shortest substrate (P12a) remains almost unchanged during the experiment (Fig. 4B). Repeating the experiment in the absence of the shortest peptide (P12a) did not result in changes in the rate of insulin β-chain degradation (Fig. 4C).

4. Discussion

It is obvious from the experiments reported in this communication that the *Thermoplasma* proteasome cleaves peptide substrates in a specific manner. In the master peptide P12a, which contains all five cleavage sites defined for proteasome, it is surprisingly the bond between A5 and E6 which is cleaved preferentially, and not the 'chymotryptic' site Y3-G4, as one might expect from assays with short fluorogenic peptides [9]. The sequon AE remains the preferred cleavage site when it is moved towards the C-terminus or when it occurs twice in a 23-residue peptide in which the core of the master peptide is tandemly arranged. However, when placed in a different context, such as the peptides IQVYSRHPAENGK or IE-

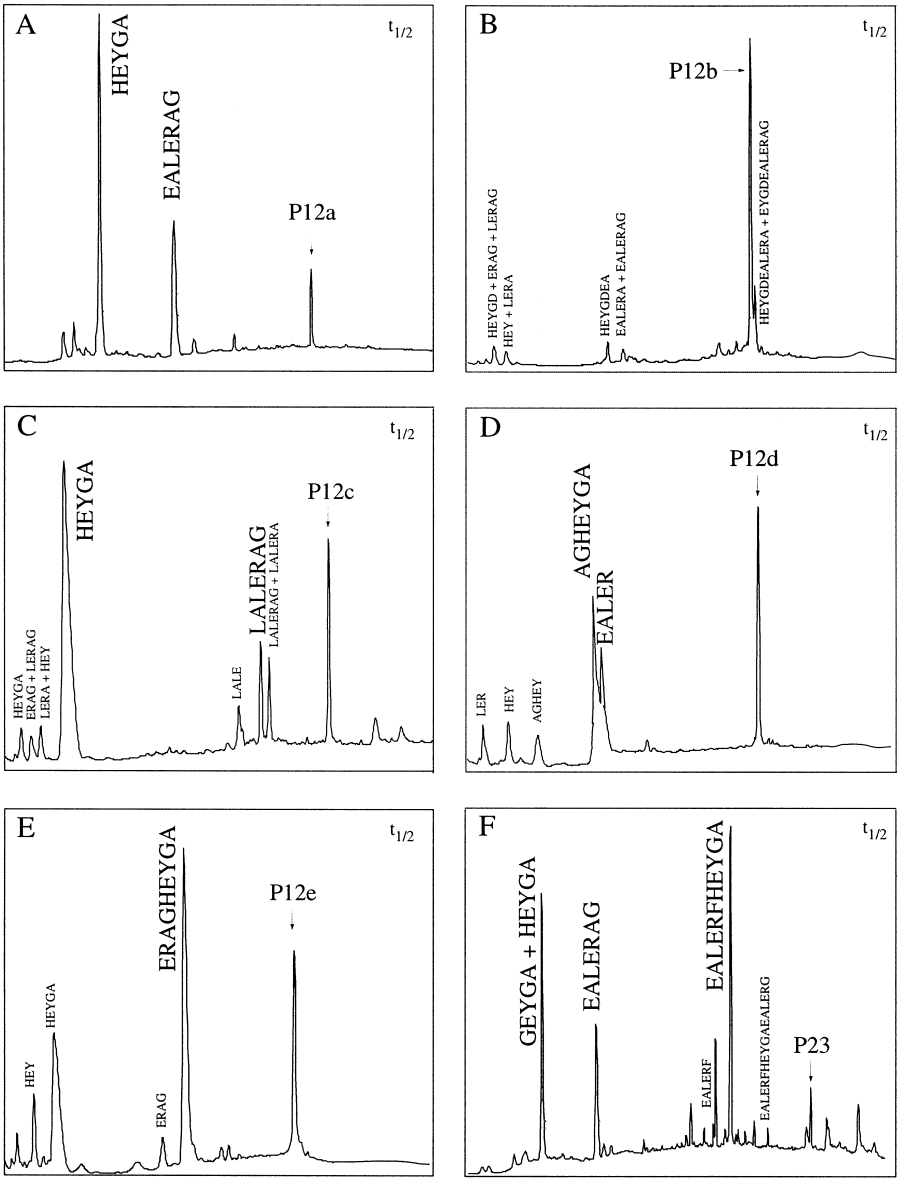

Fig. 3. RP-HPLC elution chromatograms of cleavage products produced by the proteasome at half time of substrate consumption. Substrates are described in Fig. 2. Conditions for degradations are described in Section 2.

KELRHGKSAEAT the sequon AE is no longer a cleavage site. Some changes of residues next to the scissile bond are tolerated; for example when AE is changed to AL (P12c), without otherwise changing the context A5L6 remains the dominant cleavage site. Again, when the sequon AL is placed in a different context (A7L8 in the same peptide or in insulin β-chain) it is no longer a preferred cleavage site. Other changes, such as the replacement of the alanine (A5) by an aspartic acid residue renders the peptide undegradable. We conclude that both the residues next to the scissile bond and their wider context determine whether or not a site is disposed to cleavage. Given the limited set of peptides used in this study it is clearly not feasible to unravel the non-trivial rules which determine the proteasome specificity.

Surprisingly, the rate of degradation depends critically on the length of the peptide substrate. When peptides displaying the same array of cleavage sites are extended in the length, the k_{cat}/K_M values change significantly. The *Thermoplasma* proteasome is a rather slow enzyme for peptides of 12 residues and shorter with k_{cat}/K_M values of approx. 350 M^{-1} s^{-1}. k_{cat}/K_M values in the same range are measured with short fluorogenic peptides. Peptides of 14 residues and longer are degraded 20 times faster although the sites of cleavage remain unchanged. The 30-residue insulin β-chain is degraded approx. 70 times faster than the 12-mer. Consequently, one would expect that in a competition experiment in which peptides of different lengths are exposed to the proteasome, the longer peptides are degraded before the shortest ones. This is exactly what we observed experimentally (see Fig. 4).

It is well established now that the 20S proteasome is a processive enzyme, i.e. a polypeptide chain which enters the central proteolytic cavity undergoes multiple cleavages before the degradation products are released [6,7]. It is interesting to note that the vast majority (~95%) of the peptides which exit the proteasome is 14 residues and shorter (see figure 7 in [17]). The average length of peptides was found to be 8 ± 1 residues.

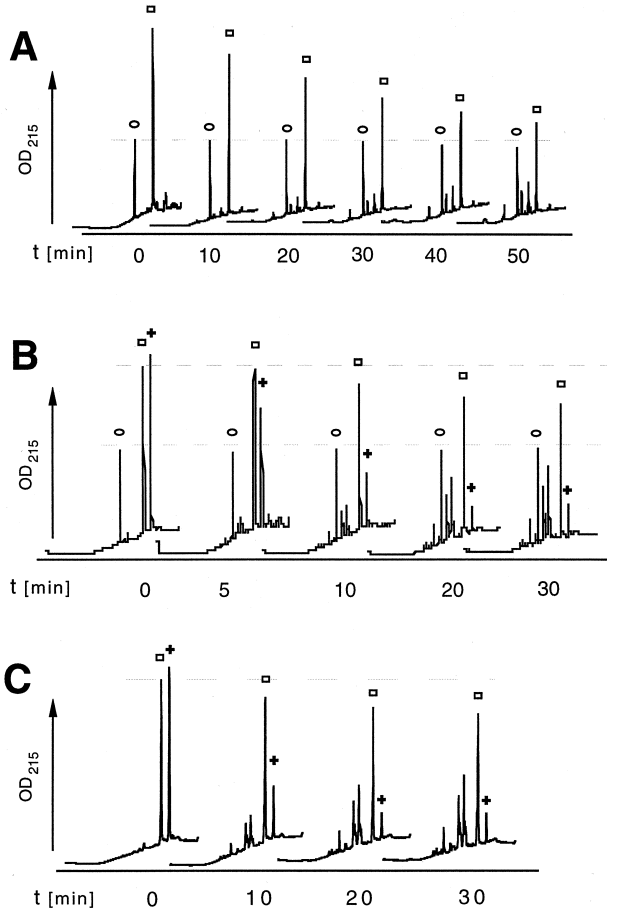

Fig. 4. Time dependent RP-HPLC elution profiles for competition experiments. The chromatogram represents the action of proteasome in the mixture of substrates: (A) P12 and P23, (B) P12, P23 and oxidized insulin β-chain, (C) P23 and oxidized insulin β-chain. Substrates are as follows: circle represents P12; square represents P23; cross represents oxidized insulin β-chain.

The upper limit in product length of the residues coincides with the sharp drop in the velocity of degradation which we observed for peptides shorter than 14 residues. We conclude that these peptides are not efficiently withheld by the proteasome and have a high probability to exit; this is equivalent to a low affinity for peptides <14 residues. Upon prolonged incubation they have a chance to re-enter, but this is an inefficient and therefore a slow process.

These considerations have consequences for the concept of the 'molecular ruler'. Although the mean length of the peptide products (8 ± 1 residues) corresponds to the distance between active sites in the *Thermoplasma* proteasome (2.8 nm) – assuming that peptides are in extended conformation – this seems to be merely coincidental. In fact, a purely geometry-based ruler should yield products more focused in length than is observed experimentally [11], and distance should be the primary determinant of a cleavage site rather than specific sequons. However, this is not what we find. For example the sequon AE embedded in a specific context remains the cleavage site irrespective of its location in the peptide. If it occurs twice in a peptide (P23) with an intervening sequence of 11 residues it will remain the preferred cleavage site. Another indication that the distance between two subunits is not the basis for a molecular ruler is noticed from digestion experiments using proteasomes with decreased numbers of active subunits. Cleavage sites of substrates are found to be the same as with wild-type proteasome (Dolenc, in preparation). Thus the proteasome cleaves polypeptides specifically, but the rules remain to be elucidated. Below a certain length (<14 residues) degradation products have a high probability to exit the proteasome. Although they might re-enter and be degraded further, this will be a slow and inefficient process and therefore degradation products <14 residues will accumulate. It appears that the length distribution of the peptides is kinetically or thermodynamically controlled. It should be mentioned here that the lower 'affinity' of the proteasome for shorter peptides was considered as an alternative explanation when the molecular ruler hypothesis was originally proposed [11].

Acknowledgements: We thank Peter Zwickl (Max-Planck Institute, Martinsried, Germany) and Ralf Golbik (Martin Luther University, Halle, Germany) for critical discussion of the manuscript. We are grateful to Frank Siedler for peptide synthesis and to Friedrich Lottspeich and Christoph Eckerskorn (Max-Planck Institute, Martinsried, Germany) for peptide sequencing and mass spectrometry measurements.

References

[1] Baumeister, W., Walz, J., Zuhl, F. and Seemuller, E. (1998) Cell 92, 367–380.
[2] Lupas, A., Flanagan, J.M., Tamura, T. and Baumeister, W. (1997) Trends Biochem. Sci. 22, 399–404.
[3] Wolf, S., Nagy, I., Lupas, A., Pfeifer, G., Cejka, Z., Muller, S.A., Engel, A., De Mot, R. and Baumeister, W. (1998) J. Mol. Biol. 277, 13–25.
[4] Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. and Huber, R. (1995) Science 268, 533–539.
[5] Seemuller, E., Lupas, A., Stock, D., Lowe, J., Huber, R. and Baumeister, W. (1995) Science 268, 579–582.
[6] Dick, L.R., Moomaw, C.R., DeMartino, G.N. and Slaughter, C.A. (1991) Biochemistry 30, 2725–2734.
[7] Akopian, T.N., Kisselev, A.F. and Goldberg, A.L. (1997) J. Biol. Chem. 272, 1791–1798.
[8] Orlowski, M., Cardozo, C. and Michaud, C. (1993) Biochemistry 32, 1563–1572.
[9] Dahlmann, B., Kuehn, L., Grziwa, A., Zwickl, P. and Baumeister, W. (1992) Eur. J. Biochem. 208, 789–797.
[10] Tamura, T., Nagy, I., Lupas, A., Lottspeich, F., Cejka, Z., Schoofs, G., Tanaka, K., De Mot, R. and Baumeister, W. (1995) Curr. Biol. 5, 766–774.
[11] Wenzel, T., Eckerskorn, C., Lottspeich, F. and Baumeister, W. (1994) FEBS Lett. 349, 205–209.
[12] Ehring, B., Meyer, T.H., Eckerskorn, C., Lottspeich, F. and Tampe, R. (1996) Eur. J. Biochem. 235, 404–415.
[13] Bogyo, M., Shin, S., McMaster, J.S. and Ploegh, H.L. (1998) Chem. Biol. 5, 307–320.
[14] Goldberg, A.L., Gaczynska, M., Grant, E., Michalek, M. and Rock, K.L. (1995) Cold Spring Harbor Symp. Quant. Biol. 60, 479–490.
[15] Rock, K.L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D. and Goldberg, A.L. (1994) Cell 78, 761–771.
[16] Heemels, M.T. and Ploegh, H. (1995) Annu. Rev. Biochem. 64, 463–491.
[17] Kisselev, A.F., Akopian, T.N. and Goldberg, A.L. (1998) J. Biol. Chem. 273, 1982–1989.