doi:10.1016/j.jmb.2008.01.086 J. Mol. Biol. (2008) 377, 1607–1617

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Modeling the *in Vitro* 20S Proteasome Activity: The Effect of PA28–αβ and of the Sequence and Length of Polypeptides on the Degradation Kinetics

Michele Mishto[1,2]*, Fabio Luciani[3], Hermann-Georg Holzhütter[4], Elena Bellavista[1], Aurelia Santoro[1], Kathrin Textoris-Taube[4], Claudio Franceschi[1,2], Peter M. Kloetzel[4] and Alexey Zaikin[5]

[1]*Interdepartmental Center for Studies on Biophysics, Bioinformatics and Biocomplexity 'L. Galvani' (CIG), University of Bologna, via S. Giacomo 12, 40126 Bologna, Italy*

[2]*Department of Experimental Pathology, University of Bologna, via S. Giacomo 12, 40126 Bologna, Italy*

[3]*School of Biotechnology and Biomolecular Sciences, University of New South Wales 2052 Kensington NSW, Australia*

[4]*Institute of Biochemistry, Charité, Humboldt University, Monbijoustr. 2, 10117 Berlin, Germany*

[5]*Departments of Mathematical & Biological Sciences, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK*

Received 11 October 2007; received in revised form 25 January 2008; accepted 29 January 2008
Available online 8 February 2008

Edited by R. Huber

Proteasomes are fundamental for the degradation of intracellular proteins, having a key role in several important metabolic and signaling pathways, in the cell cycle and in antigen presentation. *In vitro* proteasomal digestion assays are widely used in molecular biology and immunology. We developed a model, ProteaMAlg (proteasome modeling algorithm) that describes the kinetics of specific protein fragments generated by 20S proteasomes in different conditions, once the substrate cleavage strengths are provided. ProteaMAlg was tested on a variety of data available in the literature as well as on new degradation experiments performed with polypeptides of different sequences and lengths. The comparison between *in vitro* and *in silico* experiments was used to quantify the effect on degradation of the sequence and the length of target polypeptides, of the presence of regulatory molecules such as PA28–αβ, and of the type of 20S proteasome (constitutive- or immunoproteasome). The model showed that the effect of the PA28 regulatory subunit results in a modification of the gating functions of the proteasome core particle. Immunoproteasome digestion experiments suggested that this form of proteasome, which is involved in generating MHC-class I epitopes, presents modified cleavage and gating activities. Our analysis improves the current understanding of the kinetics of proteasome functioning, and provides a tool to quantify and predict the effect of key parameters during *in vitro* digestion. ProteaMAlg is publicly available on the web (http://www.proteamalg.com).

Keywords: proteasome degradation; mathematical modeling; MHC class I antigen; prediction algorithm; digestion kinetics

*Corresponding author. Department of Experimental Pathology, University of Bologna, via S. Giacomo 12, 40126 Bologna, Italy. E-mail address: michele.mishto@unibo.it.

Abbreviations used: aa, amino acid; MAG, myelin associated glycoprotein; ProteaMAlg, proteasome modeling algorithm; SCS, substrate cleavage strength; DCP, double-cut product; SCP, single-cut product.

Introduction

Proteasomes are multicatalytic enzyme complexes that are responsible for the degradation of the majority of intracellular proteins, having a pivotal role in major histocompatibility complex (MHC) class I antigen processing. The 26S proteasomes consist of a catalytic 20S core (a four-ring structure with seven different subunits in each ring, arrayed as $\alpha_7\beta_7\beta_7\alpha_7$) and two 19 S regulatory complexes. Alternatively, proteasome activators PA28–αβ or –γ can bind to the end of the 20S core.[1,2] In cells exposed to IFN-γ and TNF-α, the constitutive β subunits with catalytic activity (β1, β2, and β5) are replaced by other catalytic subunits (β1i, β2i, and β5i, respectively) that are incorporated cooperatively into a newly synthesized alternative proteasome form. This isoform, named immunoproteasome, presents an altered activity of the three catalytic sites,[2] which may enhance the capacity to generate antigenic epitopes and the digestion of specific substrates.[3] In this work, we specifically studied the *in vitro* digestions by 20S proteasome in complex with PA28–αβ and investigated the effects on the kinetics of digestion. It is generally accepted that PA28–αβ, through the binding to the 20S α-rings, produces an opening of the gate and consequently increases the degradation of specific substrates.[1,3–5] Furthermore, it has been proposed that PA28–αβ could induce further conformational changes on the 20S proteasomes; [1] there have been several reports that this complex increases the processing of specific MHC class I epitopes, enhancing the cleavage strength of minor cleavage sites.[3] In this work, we address these issues by means of a mathematical modeling analysis of *in vitro* degradation experiments. In the last years, algorithms for the prediction of proteasomal cleavages and mathematical modeling of kinetics of degradation have been developed based on *in vitro* data for polypeptide degradation.[6–10] Here, we propose a new model, named ProteaMAlg (proteasome modeling algorithm), with the aim to improve the analysis and predictions of the products and their kinetics of *in vitro* proteasome degradation. We show the utility and the prediction power of our tool by fitting several selected data sets from the literature and from new *in vitro* experiments. The quantitative support of ProteaMAlg on the outcomes of *in vitro* experiments provides information to improve our understanding of the mechanisms underlying the functional diversity between constitutive- and immunoproteasomes, the effects of substrate length on the cleavage as well as the PA28-mediated modulation.

Results

The model

Two main factors characterize ProteaMAlg: the first is a kinetic description of the proteasome degradation dynamics based on a system of ordinary differential equations. The second is the use of the substrate cleavage strengths (SCSs) to assess the role of substrate specificity in the digestion experiment (for a detailed definition, see below).

The kinetic model is based on our earlier work,[10] where we proposed a model to describe the effect of the substrate length and concentration on *in vitro* degradation. Let (i,j) be a generic fragment, where i and j ($j \geq i$) denote the sequence position of the N terminus and the C terminus of the fragment within the initial substrate $(1,L)$. The concentration of fragment (i,j) inside and outside the proteasome is denoted by $n_{i,j}$ and $N_{i,j}$, respectively. The length of this fragment is $x_{i,j}=j-i+1$. The kinetic model accounts for four elementary processes to describe the kinetics of the fragment (i,j) (Fig. 1):

(1) Uptake of fragment (i,j) into the proteasome from the external environment: the corresponding rate law in the kinetic equation depends linearly upon the external concentration $N_{i,j}$ of fragment (i,j); i.e., potential competition of fragments for uptake is not considered in the model. The rate law is composed of a single rate constant (which has the same value for all fragments independent of their length and amino acid composition) and a factor that takes into account the fact that the uptake rate decreases with the increase of occupancy of the interior volume of the proteasome. This volume

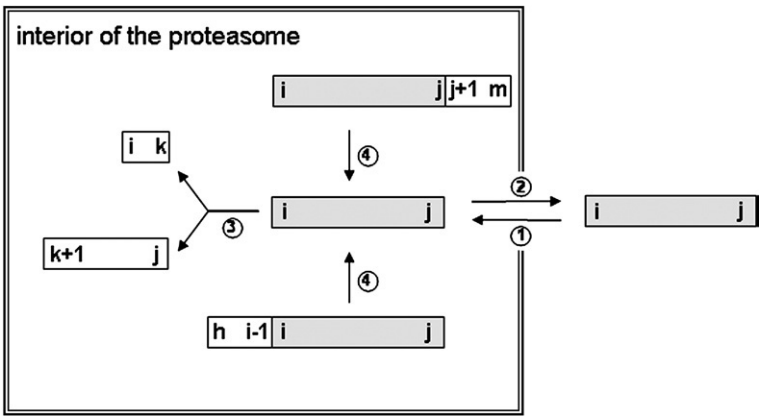

Fig. 1. Illustration of the elementary kinetic processes that an arbitrary fragment (i,j) may undergo: 1, uptake into the proteasome; 2, release from the proteasome; 3, destruction by cleavage into smaller sub-fragments; 4, generation from larger fragment having either an N- or C-terminal extension.

limiting factor is chosen such that the uptake rate is maximal for an empty proteasome and tends to zero if the number of amino acid residues (aa) A tends towards the maximal number of amino acid residues A_{max} that the proteasome may accommodate. The number of residues $A(i,j)$ contributed by fragment (i,j) residing inside the proteasome is given by $A(i,j)=x_{i,j}\,n_{i,j}$ where $x_{i,j}=j\text{-}i+1$ and $n(i,j)$ denote the length and the concentration of the fragment (i,j) within the proteasome. The total number of aa residues, A_{max}, filling the interior of the proteasome is obtained by summing over all contributions $A(i,j)$ $(i=1..L,j=i..L)$.

(2) Release of fragment (i,j) from the proteasome: the release rate depends linearly on the proteasomal concentration $n_{i,j}$ of fragment (i,j); thus, the potential competition of fragments for release is not considered in the model. The rate function $e(x_{i,j})$ for the release of fragment (i,j) is assumed to depend on the length $x_{i,j}$ of this fragment: the smaller the fragment, the larger the probability to leave the proteasome. The impact of the regulator PA28 on the release probability of fragments is taken into account by allowing the parameters determining the length-dependence of the release rate to have different values in the presence or in the absence of PA28.

(3) and (4) Proteolytic cleavage of peptides inside the proteasome: two processes account for the loss and for the gain of each fragment of length (i,j). The first process (process 3 in Fig. 1) is the loss of fragments of length (i,j) which are cut into shorter fragments, (i,k) and $(k+1,j)$; the second (process 4 in Fig. 1) is a gain process in which fragments (i,j) can be generated from larger precursor peptides ((h,j) and (i,m)). The degradation of peptides is assumed to proceed via consecutive single cleavage events; i.e., double-cleavage events, as originally proposed by the so-called molecular ruler hypothesis,[11] are not taken into account. The cleavage rate is chosen as a function of the substrate length, in order to incorporate into the model the enzyme kinetic facts known for all well-investigated proteases. The effective binding of the fragment to the active site requires a sufficiently large number of residues on both sides of the scissile bond; therefore, the longer the substrate, the more likely it is that the cleavage will occur. A detailed description and motivation of the rate laws used in the kinetic equations is given in Materials and Methods.

The second key element of our model is based on experimental evidence that the overall proteasome cleavage rate is influenced by each amino acid position of the given substrate. We therefore defined the cleavage rate of a given substrate through SCSs, namely, the relative probability that each amino acid position of the substrate is cleaved by the proteasome. This cleavage strength pattern is substrate-specific and can be obtained by the available cleavage sites prediction algorithms,[6,7,9] or by preliminary degradation experiments and a computation through the method here described (see Materials and Methods). We should recall that experimental SCSs are a function of substrate aa sequence and of the specific chemical and physical characteristic of the substrate structure, which can strongly reduce or alter the influx and the binding to the relevant proteasome subunits. We make the assumption that the polypeptides we considered in this study are easily attached to the proteasome and the translocation rate within the chamber depends on the substrate specificity.

To develop the model, we tested ProteaMAlg on several published data sets of *in vitro* degradations performed by 20S constitutive- or immunoproteasome, in the presence or in the absence of the PA28 complexes and with polypeptides differing in length, concentration and aa sequences.

Length distribution of digested fragments

First, we used ProteaMAlg to predict the length distribution of peptides generated by the 20S proteasome degradation of casein, ovalbumin and IGF-1.[12] For such prediction, the SCSs for each substrate have been obtained from ProteaSMM,[9] and a cleavage strength threshold of 0.65 was used. As reported in Fig. 2, ProteaMAlg showed an adequate fit of the experimental results, but an excess of fragments shorter than 4 aa. Such discrepancy can be explained by the conclusion of a further study from the same research group, which identified in a similar experiment a peak of 2–3mers in the length distribution of digested casein products. They revealed that the methodological

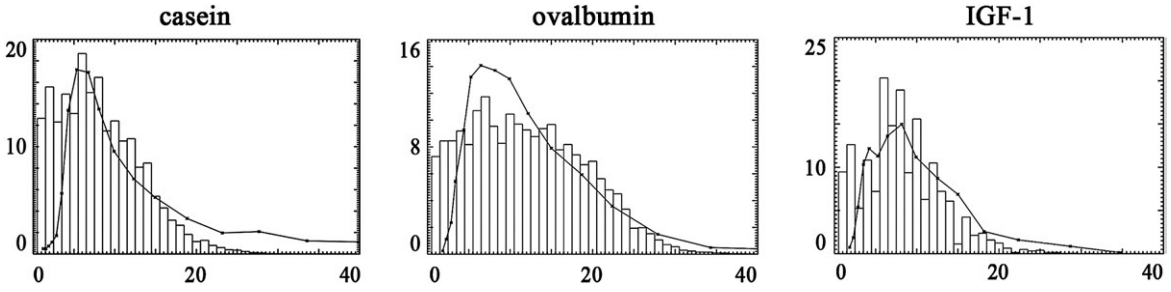

Fig. 2. Length distribution of peptides generated by the 20S degradation of casein, ovalbumin and IGF-1. The data reported by Kisselev *et al.*[12] (black line) and the *in silico* results of ProteaMAlg (histograms) are shown. The x-axis indicates peptide length (in aa), and the y-axis indicates the relative amount (percentage).

improvement of the size-exclusion chromatography technique allowed the detection of fragments shorter than 4 aa,[13] thus supporting the ProteaMAlg prediction output.

Substrate and fragment kinetics in digestions by constitutive- or immunoproteasomes

Through ProteaMAlg we followed the kinetics of each fragment produced by 20S proteasome during *in silico* degradations of the 25mer (pp89) under the conditions (e.g., SCSs, substrate and proteasome concentration, fragment re-entry) adopted by Peters et al.[8] In the experimental settings, the 20S constitutive- and immunoproteasomes were used, and the kinetics of the products were quantified by the area delimited by the highest and lowest values of each degraded fragment measured in different repetitions of the assay. Thus, the application of ProteaMAlg to these sets of experiments and the comparison with the experimental reports allowed us to test the capability of the model to predict both the fragment kinetics (either in the presence or in the absence of product re-entry) and the differences between constitutive- and immunoproteasome digestion. As shown in Fig. 3, ProteaMAlg results mimicked the experimental data, also after long digestion times when cleaved peptides likely re-entered the proteasome cavity (re-entry event). The effect of immunoproteasome on the fragment kinetics and the substrate degradation was described by tuning three parameters: increasing *e*

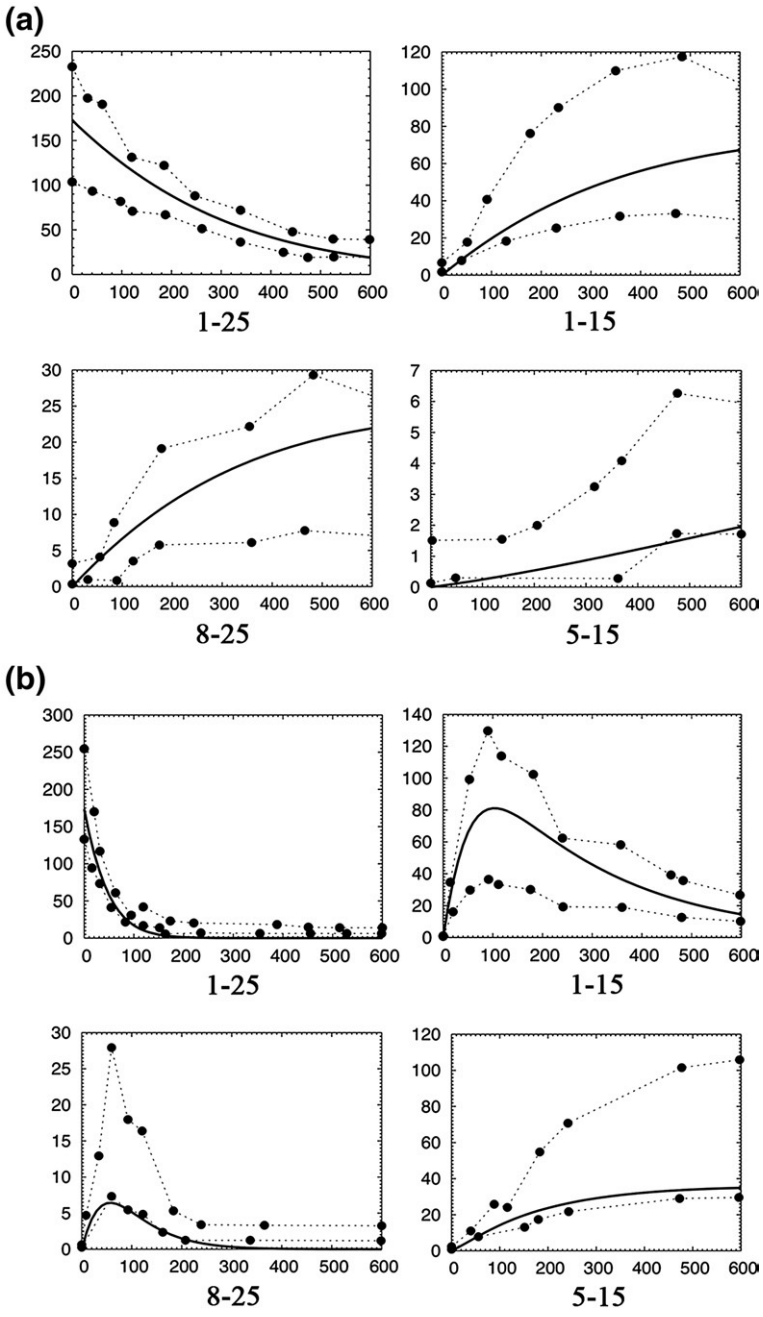

Fig. 3. pp89 25mer degradation by 20S constitutive- and immunoproteasomes. Substrate and fragment (examples) kinetics of the digestion by 20S constitutive- (a) and immunoproteasomes (b) are reported. The broken lines with circles represent the experimental value of two different experiments as reported by Peters *et al*.[8] The black line corresponds to the ProteaMAlg computation. The caption of each graph indicates the position of the peptide in the sequence of the 25mer substrate. The *x*-axis indicates the incubation time (min) and the *y*-axis indicates the amount (pmol).

(transport rate) and c (cleavage rate), and decreasing c_1 (peptide length with 50% maximum cleavage).

Effects of the PA28–αβ complex on substrate and product kinetics

One of the main challenges of this study was to predict the effects of PA28 on proteasome degradation. Therefore, we applied ProteaMAlg to the data reported by Dick *et al.*,[5] regarding the digestion of three polypeptides (pp89, JAK-1, and JAK1-D) by 20 S proteasomes in the presence or in the absence of PA28. They observed, in the presence of the regulatory complex: an increase of the decay rate for all substrates but JAK1-D; a greater amount of double-cut products (DCPs) and a concomitant decrease of single-cut products (SCPs) (Fig. 4a). *In silico* ProteaMAlg tests, performed using the same conditions and SCSs,[5] mimicked the PA28-dependent effect on the degradation rates (data not shown) and on the preferential production of DCPs over SCPs for all the substrates (Fig. 4b). Intriguingly, only two parameters were changed to imitate the effect of PA28 on the fragment kinetics and substrate degradation: e (transport rate) was increased, while θ (peptide length with 50% maximum transport) was decreased.

Effects of substrate length on the degradation rate and on fragment kinetics

There have been various reports that substrate length impinges upon the proteasome degradation rate. In this study, we tested ProteaMAlg on data published by Dolenc *et al.*,[14] which described the digestion by 20S proteasome of polypeptides with similar sequence and length between 12 aa and 30 aa. The authors observed a slight decrease in the

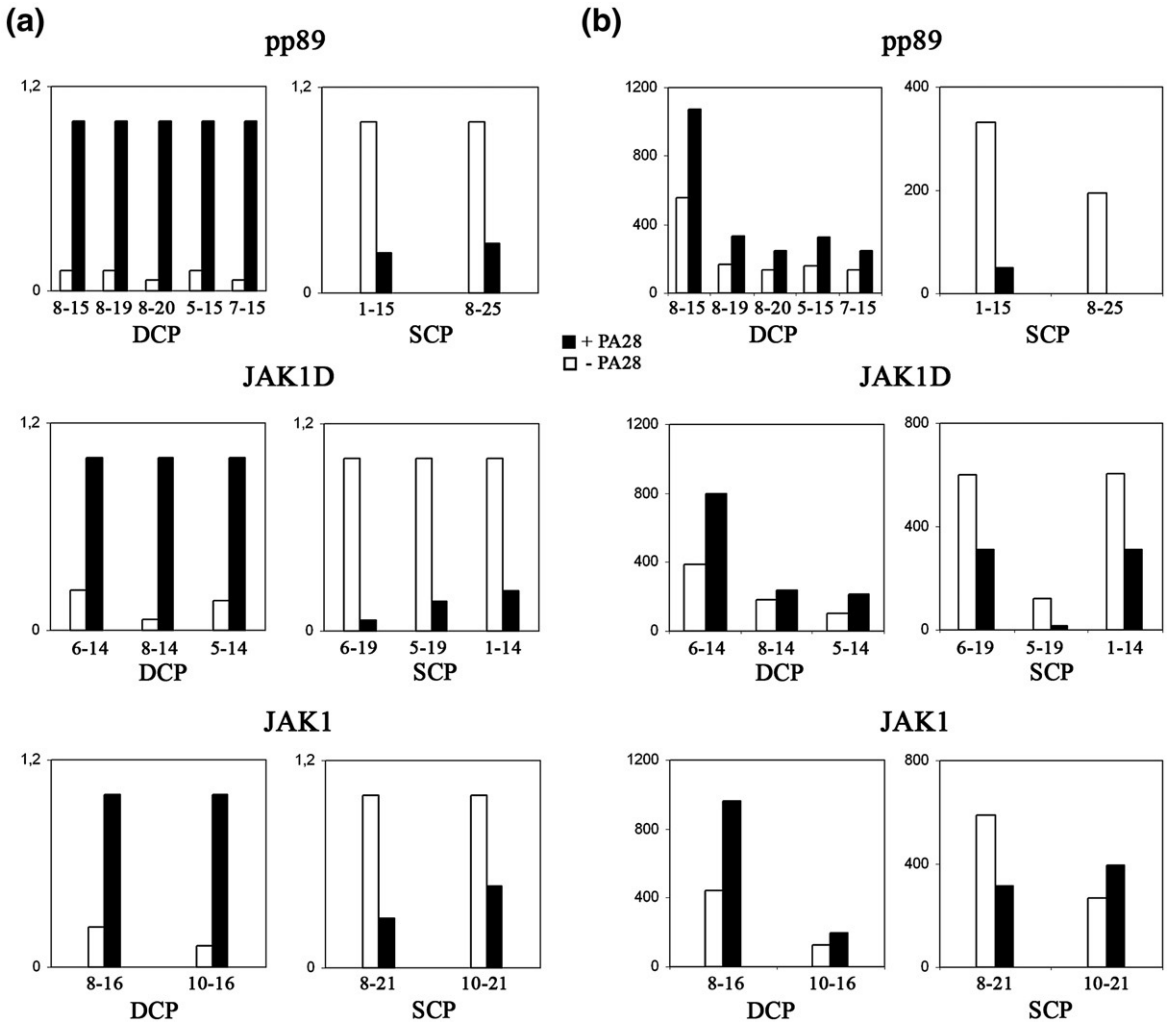

Fig. 4. PA28 increases DCPs and decreases SCPs. (a) Relative amounts of DCPs and SCPs originated from the digestion by 20S proteasomes, with or without PA28, of the three polypeptides (pp89, JAK1D and JAK1), as reported by Dick *et al.*[5] The amount of DCPs in the presence of PA28 and of SCPs in the absence of PA28 were arbitrarily set to 1 (y-axis).[5] (b) Absolute amount of DCPs and SCPs derived from the *in silico* digestion, through ProteaMAlg, of the three polypeptides by 20S proteasome with or without PA28. The product amounts were obtained after 300 min of *in silico* digestion of 9 nmol of substrate in 300 μl of solution, similar to the experimental settings.[5] The y-axis indicates the amount (pmol).

Table 1. Influence of substrate length on rate constant K_1

	12mer	14mer	16mer	23mer	30mer (insulin-β chain)
K_1	0.0028	0.0031	0.0029	0.0032	0.0035

The *in silico* digestion of substrates with different length computed through ProteaMAlg (data from Dick et al. Ref. 14). The rate constant K_1 (1/min) was calculated for each substrate. The digestion conditions were 40 nmol of substrate in a volume of 200 μl with 10 μg of 20S proteasomes.

degradation rate from 14mer to 23mer, and an increase for the longest insulin β chain (30mer). Intriguingly, the 12mer showed a very slow degradation and they inferred that proteasome has a substrate length threshold around 12–14 aa, below which the proteolysis is inefficient. In agreement with the experimental results, we predicted the highest degradation rate for the insulin β chain, the lowest for peptide p12 and an intermediate cleavage rate of the other substrates. These results were described through the computation of the rate constant K_1 for each substrate (Table 1).

In order to test the prediction power of ProteaMAlg and to achieve a better understanding of the mechanisms that regulate proteasome function, we performed new *in vitro* experiments, focusing our attention on the effect of the PA28 complex and of the substrate length on the degradation kinetics. Degradations of different substrates (see Materials and Methods) were performed with immunoproteasomes, the isoforms more involved in antigen processing,[1] to further test ProteaMAlg in experiments of immunological relevance.

Effects of substrate length on kinetics of degradation

To verify the effect of substrate length on the kinetics of degradation, four polypeptides derived from myelin associated glycoprotein (MAG) (Kloe 260, Kloe 316, Kloe 317, and Kloe 320) with overlapping sequences and different lengths were digested by 20S immunoproteasomes. The substrates showed differences in their degradation rate (Fig. 5a) and in the residue cleavage strength (Fig. 5b), although no clear length-dependent effect emerged. *In silico* ProteaMAlg experiments validated such observation, because the digestion of the longest polypeptide, Kloe 316, was faster than that for the 28mer and 29mer but slower than that for the 34mer (Fig. 5c).

Effects of the PA28–αβ complex on substrate and product kinetics

To investigate the PA28 effects, we performed *in vitro* digestion of polypeptides Kloe 256, Kloe 258, and Kloe 260 by 20S immunoproteasomes in the presence or in the absence of PA28, and we compared the experimental results with *in silico* ProteaMAlg computations. The digestion rates of polypeptides Kloe 256 and Kloe 260 were not altered by the presence of PA28 (data not shown), whereas the proteolysis of Kloe 258 was increased significantly (Fig. 6a). In the presence of the regulatory complex this polypeptide showed different SCSs and length distribution (data not shown). Because of these characteristics, we focused our attention on polypeptide Kloe 258 and on its fragment kinetics,

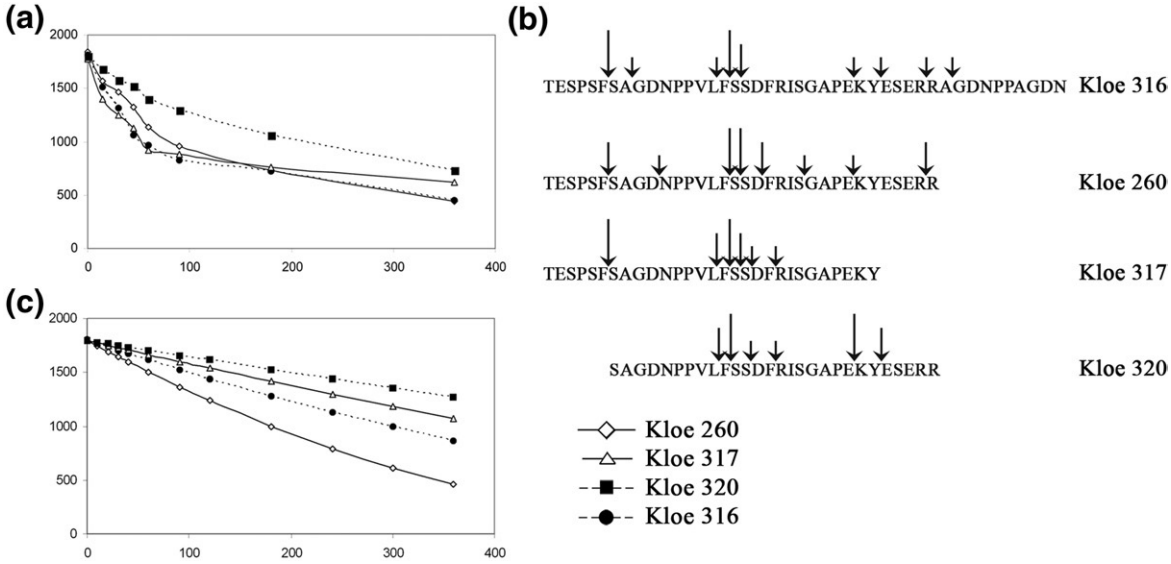

Fig. 5. Impact of substrate length on proteasomal degradation. (a) The decay rate of Kloe 260 (34mer), Kloe 316 (44mer), Kloe 317 (29mer) and Kloe 320 (28mer), digested by immunoproteasomes, are reported. (b) The amino acid sequences of the four substrates with major (long arrows) middle (middle arrows) and minor (shorter arrows) cleavage sites are shown. Both substrate length and sequence influenced the degradation rate, although no general rule has emerged. (c) The *in silico* ProteaMAlg predictions of this assay. As for the *in vitro* experiments, the degradation rate is not inversely related to the substrate length. In (a) and (c) the *y*-axis represents the substrate amount (pmol), while the *x*-axis is the incubation time (min).

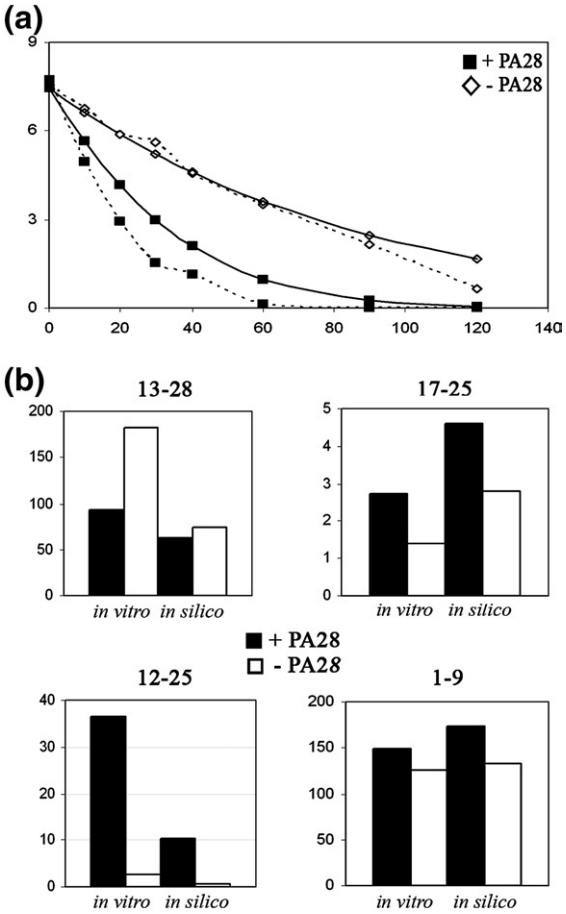

Fig. 6. Effects of PA28 on the degradation of Kloe 258 by immunoproteasomes. (a) The degradation kinetics of Kloe 258 with and without PA28. The broken and continuous lines represent the data obtained in *in vitro* and *in silico* experiments, respectively. The substrate amount (pmol) is on the *y*-axis and the incubation time (min) is on the *x*-axis. (b) Peptide amounts observed in both *in vitro* and *in silico* are reported. The *y*-axis represents the peptide amount (pmol). With PA28, fragments longer that 13 aa (such as peptide 13-28 shown here) originated from single cuts (SCPs) are generated in smaller amounts, while the double cleavage products (such as peptides 12-25 and 17-25 shown here) increase. Such an effect is less evident with fragments shorter than 14 aa: DCPs (such as peptide 17-25 shown here) are processed in greater amounts in presence of PA28, as well as SCPs (such as peptide 1-9 reported here).

adopting the following strategy: (i) in order to minimize the impact of re-entry, the amount of fragment was compared when at least 50% of the substrate was still in solution; (ii) the comparison of the amounts of fragments with same length was performed among samples where the same amount of substrate was digested; i.e., after 10 min of *in vitro* digestion in the presence of PA28 and 40 min in the absence of PA28; this last condition was chosen in order to avoid the bias due to the accelerated degradation with the PA28 complex and its impact on the average length of fragments. Through such a strategy, digestion experiments in the presence of PA28 showed an increased amount of DCPs and a reduction of SCPs for fragments longer than 13 aa, whereas, with shorter peptides a general increase of products emerged in the presence of the PA28 complex (Fig. 6b). In agreement with what was observed with the ProteaMAlg computations with data from Dick *et al.*,[5] the fit of the Kloe 258 showed an increase of e and a decrease of θ (Fig. 6). The peptide amounts, predicted by ProteaMAlg and reported in Fig. 6b, were obtained comparing the 10 min (in the presence of PA28) and 23 min (in the absence of PA28) digestions, in order to have the same amount of substrate digested.

Discussion

One of the great challenges of the current biochemistry is the comprehension of the kinetic mechanisms underlying the *in vitro* proteasome-mediated degradation of proteins and how different factors, such as the length of the substrate and the presence of regulatory molecules, influence the proteolysis. Mathematical modeling is a well-established approach to elucidating the kinetic mechanism of complex enzymes. The aims of this approach are to postulate a reaction scheme, to translate it into kinetic equations, and to compare the predicted initial rates and the time-courses of product formation with experimental measurements. In the case of the proteasome kinetic modelling, this procedure can be extremely difficult to achieve due to the complex biological events that occur during the proteolytic degradation of the protein substrate. These events include spatial processes such as the unfolding of the protein chain before entrance, the substrate translocation through the openings in the α-rings, and its subsequent movement through the interior chamber. It cannot be excluded that a partial re-folding of the substrate takes place inside the proteasome.[15] Moreover, it is very likely that there are multiple sites of the peptide chain interacting with the inner wall of the proteasome, thus influencing the binding to the active sites. These spatial processes could partially account for the still enigmatic cleavage patterns observed with different substrates. As practically nothing is known about these spatial processes, all modelling attempts — including that presented here — are based on compartment models that consider the interior of the proteasome as a homogeneous reaction space where several elementary processes occur with probabilities that depend upon the length and aa sequence of the peptide fragments.

Previously, we proposed a model of proteasome digestion dynamics, where the distribution of fragments observed experimentally was explained by assuming a cleavage mechanism that depends only upon the substrate length and concentration, and a gating function of the proteasome.[10] Here, we offer a new model (ProteaMAlg) that extends our previous work and enables the description of the dynamics of each single fragment and accounts for a

sequence-specific cleavage mechanism. Moreover, ProteaMAlg allows us to address the effects of several variables on each possible fragment arising during the *in vitro* digestion, offering to the researcher the opportunity to estimate in advance the best experimental conditions to apply. The key input of this new model is the knowledge of the SCSs. Several cleavage prediction algorithms have been proposed,[6–9] and their prediction power increase with the availability of new experimental data. Our analysis shows that the simulations with the SCSs obtained from prediction algorithms provide a satisfactory description only of the experimental distributions of fragment lengths (data from Ref.12), while a more detailed description of the observations on the fragment kinetics is reproduced correctly only when the SCSs are obtained experimentally. Hence, we suggest calculation of the SCSs patterns from preliminary digestion experiments in order to have an exhaustive description of the degradation product kinetic, at least until prediction algorithms will be improved and consequently more detailed models will be available.

Besides its utility in the design of *in vitro* experiments, ProteaMAlg offers new inputs into the understanding of the mechanisms that underlie the proteasome function. On the basis of analysis of the *in silico* and *in vitro* experiments with polypeptides of different lengths and overlapping sequences, we conclude that both the substrate length and the aa composition affect the SCSs and the overall substrate degradation rate. The aa composition has a major role in driving the kinetics of digested fragments. However, for substrates shorter than 12 aa, a well-defined decrease of the proteasome cleavage rate with the decrease of substrate length has been observed,[8,14,16] and validated here with ProteaMAlg.

PA28 generally affects proteasome cleavage activity, notwithstanding a few exceptions among the substrates tested, as reported by others.[3,5] It is not known whether the effect of PA28 on degradation is driven by substrate hydrophobicity, by specific transport sequences that influence the PA28 transport activity,[17,18] or by some other mechanism(s). Our study confirmed that DCPs, in contrast to SCPs, are favoured in the presence of PA28, although this effect is evident only for fragments longer than 13 aa. Intriguingly, with ProteaMAlg we quantify the effect of PA28 by tuning only two parameters, i.e., the transport rate coefficient e, and θ, the peptide length with 50% maximum transport; this for all the polypeptides tested. The increase in the rate e may be explained by PA28-mediated opening of the gate, as generally accepted.[1] The decrease of θ, on the contrary, is more difficult to interpret. One possible interpretation is that unknown conformational modifications of 20S proteasome due to PA28[1] promote a different transport mechanism of long fragments within the proteasome chamber.

In addition, we quantitatively predict the differences between the kinetics of constitutive- and immunoproteasome activities.[8] Modifying the value of only three parameters (e, c and c_1) we obtained the best agreement between *in silico* and *in vitro* data taken from the literature. The increase of the efflux rate e with immunoproteasomes may be interpreted as a proteasome with a larger opening of the gate, while the increase of the cleavage rate c and the decrease of c_1 may be related to the conformational modification gathered with the substitution of the catalytic subunits. Results in the literature suggest that such an increase in the degradation rate is substrate-dependent;[4] hence, the modulation of the gating and cleavage parameters in other *in silico* experiments might differ significantly. Structural and biochemical analyses linked to the ProteaMAlg outputs may allow deciphering these interesting and still unknown topics.

As we have shown, ProteaMAlg quantitatively describes the *in vitro* digestion patterns of substrates with relatively simple dynamics. It is likely that the transport within the proteasome core particle is not a limiting factor for the degradation of these substrates, and therefore our simplified assumptions are sufficient to describe the digestion dynamics. On the contrary, more complex substrates could be modeled via a more realistic description of the transport mechanisms, including the substrate–gate interaction and the dynamical opening of the gate.[19] Because proteasome translocation into 20S does not require ATP hydrolysis, a potential mechanism of translocation that could be implemented is a passive unidirectional transport. Few models have been proposed, including a model in which the transport is dictated by a spatially asymmetric proteasome–protein interaction.[20,21] We have recently discussed that transport of substrate can be described by a stochastic model where stochastic fluctuations and interaction forces between substrate and proteasome gate subunits provide the passive mechanism to translocate polypeptides within the proteasome core particle i.e., without a requirement for external energy.[20] Experimental studies suggested this type of mechanism in many biological systems involving protein transport.[22] A stochastic approach allows us to model the degradation dynamics assuming a specific transport function, without the requirement for detailed information on the translocation mechanisms.[21] On the other hand, a detailed description of the proteasome translocation functioning could probably be very important, especially in the understanding of proteasome impairment and inhibition. An attempt to reconstruct the substrate transport rate from the experimental data without resolving the unknown mechanism of the peptide translocation might be a further step of model improvement: as soon as more data are available, the universality of the mechanistic approach in the ProteAMalg will enable us to merge the translocation function into the influx coefficients and investigate the potential influence of the proteasome transport properties.

It has been proposed recently that proteasome starts degrading large substrates (polypeptides) from their most flexible domains.[23] Depending on

the substrate, this flexible initiation site can be situated at the C or N terminus of the polypeptide, or internally. The latter implies that the substrate may enter the proteasome core particle also via loop structures, eventually with the terminal sites remaining outside the proteasome. For these more complex phenomena of substrate translocation, a more comprehensive study is required to assess the effect of the specific flexible sites on the estimate of SCSs. If these types of substrate were to be considered in an *in silico* simulation, the internal flexibility could be taken into account by modifying the SCSs and the transport rates according to the observed experimental data.

In summary, we tested ProteaMAlg in several conditions of digestion, predicting the differences between constitutive- and immunoproteasome and the effect of PA28–αβ; only four parameters emerged as pivotal (e, θ, c and c_1). The parameters e_1 and c_2 vary only between two values, while the influx rate, a, and the volume of the internal chamber, v, are kept constant. We propose as default conditions, variation of parameters to predict the effect of PA28 on the digestion to increase e up to sixfold and to decrease θ by ~30%. The software implementation of ProteaMAlg is publicly available on the web†.

Materials and Methods

The model

ProteaMAlg describes the *in vitro* degradation of a substrate of length L characterized by a defined probability F_i ($i=1..L$) that each aa position is cleaved by purified proteasomes. Suppose that F_i is the cleavage strength at position i, and $N_{i,j}(t)$ is the concentration of the fragment (i,j) outside the proteasome that starts at position i and ends at position j at any time t. The length of this fragment is denoted by $x_{i,j}=j-i+1$. Taking the initial substrate concentration $N_{1,L}(0)$ and its cleavage pattern F_i as input, we predict the kinetics of the digested substrate $N_{1,L}(t)$ and of all possible fragments generated during degradation $N_{i,j}(t)$. The predictions are obtained by the solution of $L^*(L+1)$ first order kinetic equations:

$$\frac{dN_{i,j}}{dt} = -a\left[1 - v\sum_{k=1}^{L}\sum_{l=k}^{L} x_{k,l}n_{k,l}\right]N_{i,j} + e(x_{i,j})n_{i,j}$$

$$\frac{dn_{ij}}{dt} = a\left[1 - v\sum_{k=1}^{L}\sum_{l=k}^{L} x_{k,l}n_{k,l}\right]N_{i,j} - e(x_{i,j})n_{i,j} - c(x_{i,j})n_{i,j}\sum_{k=i}^{j-1}F_k + \quad (1)$$

$$F_j\sum_{k=j+1}^{L} c(x_{i,k})n_{i,k} + F_{i-1}\sum_{k=1}^{i-1} c(x_{k,j})n_{k,j} \quad (2)$$

where $n_{i,j}$ is the concentration of the fragment (i,j) in the proteasome chamber and $i=1..L$, $j=i..L$ are the N terminus and the C terminus of the fragment, respectively. The first term in both equations describes the uptake of the fragment into the proteasome chamber. The parameter a represents the transport rate of the substrate within the proteasome chamber and in this work is assumed to be a constant. The second factor in the term uptake function describes the filling of the proteolytic chamber: the uptake of fragments decreases with the filling of the proteolytic chamber and becomes zero when the internal amount of fragments is equal to the volume $1/v$. The second term in both Eqs. (1) and (2) stands for the efflux of fragments. We assume that long peptides have more residues than the number that can bind to the 20S proteasome,[14] thus impairing their passage through the narrow pore: a natural assumption for the efflux rate can be a steep Hill function, e.g.,:

$$e(x_{i,j}) = \frac{e}{1 + (x_{i,j}/\theta)^{e_1}} \quad (3)$$

where the rate is maximum (e) for the shortest fragments and decreases to zero with the increase of the fragment length $x_{i,j}=j-i+1$. The parameter θ is the critical length at which the efflux rate is half of the maximum e (Fig. 7a), while the index e_1 shapes the slope of the curve $e(x_{i,j})$: the higher the value of e_1 the steepest is the decline of the curve around the length θ. See Ref. 10 for a more detailed description of the influx and efflux functions.

The last three terms in Eq. (2) describe the cleavage mechanism. The first term is the loss of the fragment (i,j) that is cleaved at position k with cleavage strength F_k, while the last two terms represent the gain in the concentration of the fragment (i,j) due to the cleavage of longer fragments that include the amino acid sequence (i,j). The cleavage rate $c(x_{i,j})$ is assumed to have the following dependence on the fragment length $x_{i,j}$:

$$c(x_{i,j}) = \frac{(c)(x_{i,j}/c_1)^{c_2}}{(1 + (x_{i,j}/c_1)^{c_2})} \quad (4)$$

This dependence is motivated by the fact that the fragment should be of a certain minimal length to dock to the binding groove of the catalytic sites.[8,13–16] Thus, in Eq. (4) we propose:(i) the parameter c_1 representing the

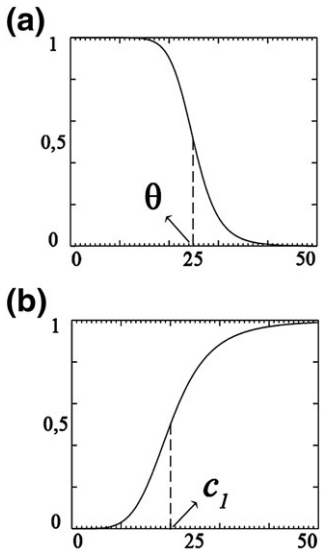

Fig. 7. The transport and cleavage rate of the ProteaMalg algorithm. The transport rate (a) and the cleavage rate (b) coefficients are on the y-axis as a function of the fragment length in aa (x-axis). θ (a) and c_1 (b) are shown graphically.

† http://www.proteamalg.com

substrate length at which the binding affinity between the catalytic site of the proteasome and the fragment is half of the maximum cleavage rate (c); (ii) the parameter c_2 that shapes the slope of the curve representing the relationship between the cleavage rate and the fragment length (Fig. 7b). For an overview of the simulation algorithm parameters see Table 2. The terms F_i are SCSs that may be obtained from experimental evaluation or from cleavage sites prediction algorithms such as ProteasMM,[9] Paproc,[6] or Netchop.[7] Furthermore, we assume that any fragment generated will have the SCSs as the initial substrate from which it has been generated. Finally, we assume that cleavages are only single-cut events; with this assumption, double-cleavage products (DCPs) can be explained as fragments generated by two single cuts on the same fragment. The re-entry, the degradation of previously generated fragments, can be forbidden by setting the transport rate a positive only for the substrate (and zero for all other fragments).

Materials and Chemicals

The polypeptides (synthesized by Charité, Berlin, Germany) represent the following sequences:

Kloe 256, sequence 44–70 of Matrix-1 of the influenza virus H1N1
TRPILSPLTKGILGFVFTLTVPSERGLQR;
Kloe 258, sequence 101–128 of human myelin basic protein
PSQGKGRGLSLSRFSWGAEGQRPGFGYG;
Kloe 260, sequence 543–576 of human MAG
TESPSFSAGDNPPVLFSSDFRISGAPEKYESERR;
Kloe 316 (Kloe 260 modified), sequence 543–571 of MAG
TESPSFSAGDNPPVLFSSDFRISGAPEKYESER-RAGDNPPAGDN;
Kloe 317
TESPSFSAGDNPPVLFSSDFRISGAPEKY;
Kloe 320, sequence 549–576 of MAG
SAGDNPPVLFSSDFRISGAPEKYESERR.
Cell cultures

Lymphoblastoid cell lines (LcLs) are B lymphocytes immortalized with Epstein Barr virus (EBV), as described.[24]

Table 2. List of the parameters present in ProteaMAlg, their description and units

Parameter	Description	Dimension
a	Uptake rate	time^{-1}
e	Maximum transport rate	time^{-1}
θ	Peptide length with 50% max. transport[a]	aa
e_1	Slope of efflux rate	–
c	Maximum cleavage rate	time^{-1}
c_1	Peptide length with 50% max. cleavage[b]	aa
c_2	Slope of cleavage rate	–
v	Proteasome volume factor	aa^{-1}

[a] Critical length at which substrates are transported at 50% of the maximum transport rate e.
[b] Critical length at which 50% of the maximum cleavage rate c is reached.

The 20S proteasome and PA28–αβ purifications

Proteasomes and PA28 complex were purified from LcLs as described.[24]

In vitro digestion of long substrates and peptide quantification

Synthetic polypeptides were digested with 0.25–1 μg of 20S with or without 3–12 μl of PA28 and samples were analyzed as described.[24] The proteolytic inactivity of PA28 preparation and the pureness of 20S preparation were tested. For each experiment, digestions were repeated at least twice. The amount of substrate in the digestion solution was computed from a calibration curve constructed with different amounts of substrate. The amount of peptide product was calculated using the mass balance method.[8]

Substrate cleavage strength computation

In order to calculate the cleavage strengths of the substrates tested in the new biochemical experiments, we used the following method: for each residue of the substrate (e.g., residue 7), we summed the amount (in pmol) of each peptide that had this residue at the C terminus (e.g., peptide 1-7) or the following residue at the N terminus (e.g., peptide 8-21). Since the production of DCPs depends on the cleavage of both peptide ends. For DCPs, we added to the sum only half of the value (in pmol), while for the SCPs we summed the entire amount. Hence, the residue sum was divided by the total amount of peptides, giving the cleavage strength value (from 0 to 1). We summed the amount of peptides at a specific time-point, when less than 50% of substrate was digested, to avoid the re-entry effect.

Acknowledgements

This work was financed, in part, by a grant from PROTEOMAGE project, sponsored by the European Commission – 6th Framework Program. We thank the Volkswagen-Stiftung (Germany), BSC (HPC-Europa Transnational Access program), the European Union through the Network of Excellence BioSim, contract no. LSHB-CT-2004-005137, the Ministerio de Educacion y Ciencia (Spain) - FEDER (projects BFM2003-07850 and FIS2006-11452) and the Australian Research Council (ARC) through the Discovery scheme (DP0556732).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.01.086

References

1. Kloetzel, P.-M. (2001). Antigen processing by the proteasome. *Nature Rev.* **2**, 179–187.
2. Rock, K.-L. & Goldberg, A.-L. (1999). Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu. Rev. Immunol.* **17**, 739–779.
3. Strehl, B., Seifert, U., Kruger, E., Heink, S., Kuckelkorn, U. *et al.* (2005). Interferon-gamma, the functional plasticity of the ubiquitin-proteasome system, and MHC class I antigen processing. *Immunol. Rev.* **207**, 19–30.
4. Kloetzel, P.-M. (2004). Generation of major histocompatibility complex class I antigens: functional interplay between proteasomes and TPPII. *Nature Immunol.* **5**, 661–669.
5. Dick, T.-P., Ruppert, T., Groettrup, M., Kloetzel, P.-M., Kuehn, L. *et al.* (1996). Coordinated dual cleavages induced by the proteasome regulator PA28 lead to dominant MHC ligands. *Cell*, **8**, 253–262.
6. Nussbaum, A.-K., Kuttler, C., Hadeler, K.-P., Rammensee, H.-G. & Schild, H. (2001). PAProC: a prediction algorithm for proteasomal cleavages available on the WWW. *Immunogenetics*, **53**, 87–94.
7. Kesmir, C., Nussbaum, A.-K., Schild, H., Detours, V. & Brunak, S. (2002). Prediction of proteasome cleavage motifs by neural networks. *Protein Eng.* **15**, 287–296.
8. Peters, B., Janek, K., Kuckelkorn, U. & Holzhütter, H.-G. (2002). Assessment of proteasomal cleavage probabilities from kinetic analysis of time-dependent product formation. *J. Mol. Biol.* **318**, 847–862.
9. Tenzer, S., Peters, B., Bulik, S., Schoor, O., Lemmel, C. *et al.* (2005). Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. *Cell Mol. Life Sci.* **62**, 1025–1037.
10. Luciani, F., Kesmir, C., Mishto, M., Or-Guil, M. & de Boer, R.-J. (2005). A mathematical model of protein degradation by the proteasome. *Biophys. J.* **88**, 2422–2432.
11. Wenzel, T., Eckerskorn, C., Lottspeich, F. & Baumeister, W. (1994). Existence of a molecular ruler in proteasomes suggested by analysis of degradation products. *FEBS Lett.* **349**, 205–209.
12. Kisselev, A.-F., Akopian, T.-N., Woo, K.-M. & Goldberg, A.-L. (1999). The sizes of peptides generated from protein by mammalian 26 and 20S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J. Biol. Chem.* **274**, 3363–3371.
13. Kohler, A., Cascio, P., Leggett, D.-S., Woo, K.-M., Goldberg, A.-L. *et al.* (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*, **7**, 1143–1152.
14. Dolenc, I., Seemuller, E. & Baumeister, W. (1998). Decelerated degradation of short peptides by the 20S proteasome. *FEBS Lett.* **434**, 357–361.
15. Sharon, M., Witt, S., Felderer, K., Rockel, B., Baumeister, W. & Robinson, C. V. (2006). 20S proteasomes have the potential to keep substrates in store for continual degradation. *J. Biol. Chem.* **281**, 9569–9575.
16. Holzhütter, H.-G. & Kloetzel, P.-M. (2000). A kinetic model of vertebrate 20S proteasome accounting for the generation of major proteolytic fragments from oligomeric peptide substrates. *Biophys. J.* **79**, 1196–1205.
17. Groll, M. & Huber, R. (2003). Substrate access and processing by the 20S proteasome core particle. *Int. J. Biochem. Cell Biol.* **35**, 606–616.
18. Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. (1998). Proteasome: paradigm of a self-compartmentalizing protease. *Cell*, **9**, 367–380.
19. Smith, D. M., Chang, S. C., Park, S., Finley, D., Cheng, Y. *et al.* (2007). Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry. *Mol. Cell*, **27**, 731–744.
20. Zaikin, A. & Pöschel, T. (2005). Peptide size dependent active transport in the proteasome. *Europhys. Lett.* **69**, 725–731.
21. Zaikin, A., Mitra, A. K., Goldobin, D. S. & Kurths, J. (2006). Influence of transport rates on the protein degradation by proteasomes. *Biophys. Rev. Lett.* **1**, 375–386.
22. Smith, D. M., Benaroudj, N. & Goldberg, A. (2006). Proteasomes and their associated ATPases: a destructive combination. *J. Struct. Biol.* **156**, 72–83.
23. Piwko, W. & Jentsch, S. (2006). Proteasome-mediated protein processing by bidirectional degradation initiated from an internal site. *Nature Struct. Mol. Biol.* **13**, 691–697.
24. Mishto, M., Santoro, A., Bellavista, E., Sessions, R., Textoris-Taube, K. *et al.* (2006). A structural model of 20 S immunoproteasomes: effect of LMP2 codon 60 polymorphism on expression, activity, intracellular localization and insight into the regulatory mechanisms. *Biol. Chem.* **387**, 417–429.