Constructing a Virtual Proteasome

Alexey Zaikin[1], Fabio Luciani[2], and Juergen Kurths[1]

[1] University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany
[2] School of Biotechnology and Biomolecular Sciences, University of New South Wales, 2026 Sydney, Australia

Abstract. The proteasome is a barrel-shaped multi-subunit protease involved in the degradation of the ubiquitin-tagged proteins. Proteasomes play many important roles in the cellular metabolism, from regulation of the important proteins to generating antigenic peptides for the immune system. The prediction of the proteasome function can help in the treatment of cancer, immune system disorders and neurodegenerative diseases, hence the constructing of a virtual proteasome is a necessary task. Here we review several model approaches which allow the analytical or numerical prediction of the proteasome functioning. The review shed light on the-state-of-the-art of the proteasome modelling and discusses the perspectives of this research.

Proteasomes are multicatalytic enzyme complexes that are responsible for degradation of the majority of the intracellular proteins into smaller peptides. They are present in all eukaryotic cells, archaea, and certain bacteria [7,46,20]. Proteasomes are absolutely essential for the homeostasis because the removal of proteasome genes in eukaryotes is lethal [14]. Proteasomes have been found in the form of different, although similar molecular complexes that consist of the central part, the 20S proteasome, and regulating caps, the 19S [7] and PA28 particles [38,22]. The most important 26S complex, which degrades ubiquitinated proteins, contains in addition to the 20S proteasome a 19S regulatory complex composed of multiple ATPases and components necessary for binding protein substrates [7]. The 20S proteasome is a barrel-shaped structure composed of four stacked rings of 28 subunits [37,7]. The active cleavage sites are located within the central chamber of the 20S proteasome, into which protein substrates must enter through narrow openings of outer rings. The 20S proteasome degrades proteins by a highly processive mechanism [2], making many cleavages of the protein and digesting it into small products. This is important for the intracellular proteolytic system because the release of large protein fragments could interfere with the cell function and regulation [19]. Proteasomes can be found in its usual form as a constitutive proteasome, or as an immunoproteasome with modified cleavage centers [20,21].

Many roles in the cell's metabolism are played by proteasomes: they destroy abnormal and misfolded proteins tagged with ubiquitin and are an essential component of the ATP-ubiquitin-dependent pathway for protein degradation [6]. Proteasomes play an important role in the immune system by generating antigenic peptides of 8-12 residues to be presented by the MHC class I molecules and, hence, are the main supplier of peptides for its recognition by killer T-cells

[21,26,43,12]. Recently, proteasome inhibition has been suggested as a promising new target for cancer treatment [34,1,9]. The proteasome function has also been linked directly to the pathophysiology of malignancies, neurodegenerative disorders, type I diabetes, cachexia [40,11] and to ageing [53].

Here we review several recently developed mathematical models which describe the influence of a length dependent cleavage, in- and efflux rates, and transport processes on the outcome of a digestion experiment. We start with short introduction to available experimental results and proceed with a general description of aspects necessary for the modelling of the proteasome. Then we consider a microscopic model of the proteasome, review a possible translocation mechanism inside the proteasome [52] and show how one can compute the length dependent transport rates analytically or numerically. Next we review the transport model of the proteasome and discuss the influence of transport rates on the proteasome output [50,51]. This approach is based on the idea that proteasome output significantly depends on the length dependent transport. After that we show how one can describe the kinetics of the protein degradation on the mesoscopic level without details of the transport mechanism [27,28]. In all these approaches we neglect the sequence specificity of the incoming protein making the degradation dependent mainly on the protein length. Finally we discuss the future of the modelling approaches and show how one can include the sequence specificity in these models and how one can merge the models describing the protein translocation with models which described the in- and out flux.

1 Experiment and modelling

Due to its significance in the cellular metabolism, the simulation of the proteasome function is the central task in the building of a virtual immune system [29]. On the long road from the initial idea to the pharmacy and the start of a new drug the simulation and prediction of the proteasome function seem to be possible now only on its early stages, i.e., for the experiments *in vitro*. These experiments study the digestion of different substrates by the proteasome and the temporal dynamics of the fragment concentration in the course of time by mixing of purified proteasomes and different substrates. The experimental results are analysed, e.g., by mass spectroscopy methods and provide us with information about the substrate cleavage pattern and the quantity of different fragments cut from the initial substrate. Even in these first stages of drug design, however, the simulation of the proteasome could significantly decrease the experiment costs through the identification and prediction of proper parameter ranges.

A principle scheme of the experiment *in vitro* is shown in the Fig. 1. The solution contains the mixture of constitutive or immuno-proteasomes with substrate or mix of substrates. The solution is analysed in different moments of time hence giving the information how much substrate has been degraded and how much and which fragments have been produced. One of the most important experimental result is the cleavage pattern of the particular protein, which shows where protein has been cleaved. Additionally the analysis of the fragment

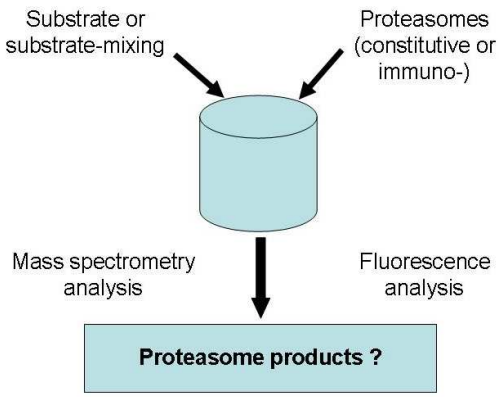

Fig. 1. A principal scheme of the experiment *in vitro*. The solution, where the substrate is mixed with proteasome, is analysed for different moments of time. Mass spectroscopy or fluorescence analysis then shows how much substrate has been degraded, and which fragments have been produced.

amount produced from this cleavage site gives the estimation of the cleavage strength at this protein position. Especially interesting is the production of a certain sequences, epitopes, which are then requested by the immune system.

Next important experimental result that describes the proteasome function is a length distribution of the fragments obtained in vitro experiments. For long substrates it was found that this dependence typically is nonmonotonous and has a single peak around the length of 7-12 amino acids (aa) for practically all types of the proteasome [32,31,19,5]. This length of peptides is the most requested length for a normal functioning of the immune system. The mechanism behind such a length distribution is not completely clear. It was widely believed that the proteasome degrades proteins according to the "molecular ruler" to yield products of rather uniform size. It was proposed [49] that peptides of 7–9 residues were generated as a result of coordinated cleavages by neighboring active sites. However, evidence for the molecular ruler is quite limited because the maximum in the length distribution is smoothed and not pronounced as a sharp peak [19]. It is also interesting to note that in some experiments three peak length distributions have been found [23]. Several theoretical models for the kinetics of proteasome degradation have been published. Some of the models describe the degradation of short peptides with qualitatively different kinetics [44,45,41] or small number of cleavage positions [16]. The theoretical model [16,35] for the degradation of long substrates is applied to specific proteins with predefined cleavage sites and is fitted to experimental data describing the fragment quantity after proteasomal degradation.

Finally the amount of the fragments produced versus time evolution provides the information about the dynamics of the fragment concentration. Interestingly,

the concentration of a certain fragment can first increase and then decrease. This is connected with the fact that the fragment produced can compete with the initial substrate and reenter the proteasome, providing the decrease of the initial substrate degradation.

To model the proteasome mechanism one should adequately describe three essential processes involved in the proteasome function: selection of cleavage sites, kinetics of generated fragments, and a peptide translocation inside the proteasome (see Fig. 2). At the moment there is no model which describe adequately all these factors. In this paper we describe these modelling approaches and discuss how one can merge them. It is important to note also that for different situations different models should be used in dependence on the fact what factor provides the largest influence on a phenomenon studied.

2 Finding the cleavage pattern

Several proteasomal cleavage prediction methods have been published. The first method, FragPredict, was developed by Holzhutter et al. [15] and is publicly available as a part of MAPPP service (www.mpiib-berlin.mpg.de/MAPPP/). It combines proteasomal cleavage prediction with MHC and TAP binding prediction. FragPredict consists of two algorithms. The first algorithm uses a statistical analysis of cleavage-enhancing and -inhibiting amino acid motifs to predict potential proteasomal cleavage sites [15]. The second algorithm, which uses the results of the first algorithm as an input, predicts which fragments are most likely to be generated. This model takes the time-dependent degradation into account based on a kinetic model of the 20S proteasome [16]. At the moment, FragPredict is the only method that can predict fragments, instead of only possible cleavage sites. PAProC (www.paproc.de) is a prediction method for cleavages by human as well as wild-type and mutant yeast proteasomes. The influences of different amino acids at different positions are determined by using a stochastic hill-climbing algorithm [24] based on experimentally in vitro verified cleavage and non-cleavage sites [33]. Recently Tenzer et al. [47] have published a method for predicting which peptides can be presented by the MHC class I pathway. In this work they characterise the cleavage specificity of the proteasome in terms of a stabilised matrix method (SMM) defining the specificity of the constitutive and immunoproteasome separately. This method is available at www.mhcpathway.net. All these methods make use of limited in vitro data for characterising the specificity of the proteasome. Moreover, both FragPredict and the matrix based methods by Tenzer et al. [47] are linear methods, and may not capture the non-linear features of the specificity of the proteasome. Another prediction method called NetChop [18,30] has two important extensions: first, the prediction system is trained on multilayered artificial neural networks. This allows the method to incorporate higher order sequence correlations in the prediction scheme, making it potentially more powerful than both PAProC and the matrix based methods, which use a linear method to predict proteasome cleavage. Second, the method

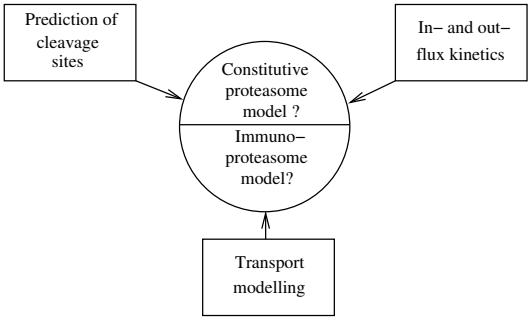

Fig. 2. Different factors which should be taken into account to model the proteasome function.

is trained to predict proteasomal cleavage on in vitro digest data, similarly to the other previous methods, and using naturally processed MHC class I ligands.

In large scale benchmark calculations the predictive performance of the different methods are compared [47,30]. According to these tests, at the moment, NetChop and SMM methods provide most reliable predictions of proteasomal cleavage.

3 Possible translocation mechanism

Considering the highly processive mechanism of the protein degradation by the proteasome, a question naturally arises: what is a mechanism behind such translocation rates? Let us discuss one of the possible translocation mechanisms. In [52] we have assumed that the proteasome has a fluctuationally driven transport mechanism and shown that such a mechanism generally results in a nonmonotonous translocation rate. Since the proteasome has a symmetric structure, three ingredients are required for fluctuationally-driven translocation: the anisotropy of the proteasome-protein interaction potential, thermal noise in the interaction centers and the energy input. Under the assumption that the protein potential is asymmetric and periodic, and that the energy input is modeled with a periodic force or coloured noise, one can even obtain nonmonotonous translocation rates analytically [52]. Here we review these results and show that for different forms of the proteasome-protein interaction potential one gets translocation rates.

We assume that after the protein has entered the proteasome, the protein-proteasome interaction is characterised by a spatially periodic asymmetric potential $U(x)$ with the period P equal to the distance between amino acids in the protein. In reality there is a basic periodicity, namely the periodicity of the protein (or peptide) backbone, that is superposed by a non-periodic (in our sense irregular) part that is attributed to the amino-acid-specific residues. Below we consider also the influence of the nonperiodic constituent. The spatial asymmetry is resulted from breaking the symmetry by entering the proteasome from one

end as well as from the $C-N$ asymmetry of the protein (or peptide) backbone. On the Fig. 4, left we have plotted several examples of such asymmetric periodic potential. The detailed form of the asymmetric periodic interaction potential is of less importance for this qualitative study.

The proteasome acts upon the protein by a certain number of equidistant interaction centers. The dynamics of the protein inside the proteasome is, hence, governed by l interactions centers, where l is the number of protein elements (amino acids or multiplicatives of it). There appear the following forces: potential force (protein-proteasome interaction) $-l\partial U(x)/\partial x$, fluctuations with collective $lF(t)$ and individual components $f_1(t) + ... + f_l(t)$, and protein friction forces $l\beta\dot{x}$ [4], where x is the coordinate of the protein with respect to the proteasome and β is the coefficient of friction. Due to small size of all protein particles, moving in the liquid cytosol, the motion occurs in the overdamped realm [3], hence we neglect inertia forces. Note that transport is possible only in the case of nonequilibrium fluctuations. In the simplified case, when fluctuations can be represented by a sum of a collective periodic force and individual for every protein residue thermal noise, the model is analytically tractable, predicting the velocity dependence on the peptide size. Normalising all forces by friction and taking $\beta = 1$, the translocation of a protein in the proteasome is governed then by

$$\frac{\partial x}{\partial t} = -\frac{\partial U(x)}{\partial x} + F(t) + \frac{1}{l}(f_1(t) + ... + f_l(t)). \tag{1}$$

Analytical results are possible if we assume collective oscillations of the peptide elements, e.g., $F(t) = A\cos(\omega t)$, where A and ω stand for the amplitude and frequency of this oscillations. Additionally, each interaction center undergoes local thermal fluctuations, represented by mutually uncorrelated white noise of intensity σ^2: $f_i(t) = \xi_i(t)$, where $\langle \xi_i(t)\xi_j(t')\rangle = \sigma^2\delta(t-t')\delta_{ij}$. In this case the stochastic term in Eq. 1 is white noise of intensity σ^2/l. The Fokker-Planck equation for the peptide coordinate probability distribution $w(x,t)$ associated with Eq. 1 is

$$\frac{\partial w}{\partial t} = -\frac{\partial}{\partial x}\left[\left(F(t) - \frac{\partial U}{\partial x}\right) w(x,t)\right] + \frac{\sigma^2}{2l}\frac{\partial^2 w(x,t)}{\partial x^2}, \tag{2}$$

which may be solved in quasi-stationary adiabatic approximation $\partial w/\partial t = 0$ [17]. We obtain

$$\frac{\sigma^2}{2l}\frac{\partial w(x,F)}{\partial x} - \left(F - \frac{\partial U}{\partial x}\right) w(x,F) = -G(F), \tag{3}$$

where $G(F)$ is the probability flux. For any periodic potential $U(x)$ the quasi-stationary solution of Eq. 3 is

$$w(x,t) = \left[C(F) - \frac{2G(F)}{\sigma^2/l}\int_0^x \exp\left(\frac{U(x') - Fx'}{\sigma^2/2l}\right) dx'\right] \exp\left(-\frac{U(x) - Fx}{\sigma^2/2l}\right), \tag{4}$$

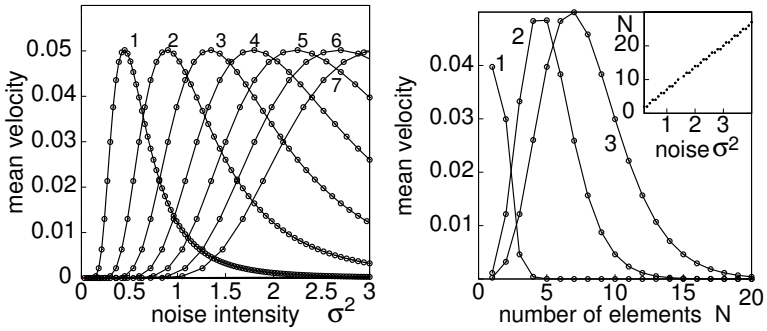

Fig. 3. Left: mean velocity of the protein as a function of noise intensity for different lengths of the protein (shown by numbers). The transport is possible only for a certain noise intensity. Right: Mean velocity as a function of the peptide length for different noise intensities $\sigma^2 = 0.6$ (curve 1), 2 (curve 2), and 3 (curve 3). The inset plot shows the depth of the protein penetration for the velocity equal to 0.0005.

where $C(F(t))$ and $G(F(t))$ are unknown functions of t. Using the periodicity condition $w(0,t) = w(P,t)$ and the normalisation of $w(x,t)$ we get $G(F)$. If the amplitude A meets the condition $LA \ll \sigma^2/l$, one can expand $G(F)$ and obtain

$$G(F) \approx G_{01}F + G_{02}F^2 \tag{5}$$

with the expansion coefficients $G_{01} = P/(I_{10}I_{20})$,

$$G_{02} = G_{01}\left(\frac{I_{11}}{I_{10}} - \frac{I_{21}}{I_{20}} - \frac{lP}{\sigma^2}\left(1 - \frac{2I_{30}}{I_{10}I_{20}}\right)\right), \quad I_{10} = \int_0^P e^{U'(x)}\,dx, \tag{6}$$

$$I_{20} = \int_0^P e^{\left(-U'(x)\right)}\,dx, \quad I_{11} = \frac{2l}{\sigma^2}\int_0^P xe^{U'(x)}\,dx,\ U'(x) = \frac{2lU(x)}{\sigma^2}. \tag{7}$$

$$I_{21} = \frac{2l}{\sigma^2}\int_0^P xe^{\left(-U'(x)\right)}\,dx\ \ I_{30} = \int_0^P \int_0^x e^{\left(U'(x')-U'(x)\right)}\,dx'\,dx. \tag{8}$$

Substituting Eq. 5 into $\overline{\langle \dot{x} \rangle} = \int_0^P \overline{G(x,t)}\,dx$, where $\overline{(\cdot)}$ denotes time averaging, we obtain the average protein transport velocity or the translocation, as a function of the noise intensity σ^2 and the peptide size l

$$R_t \approx \overline{\langle \dot{x} \rangle} \approx \frac{P^2A^2}{2I_{10}I_{20}}\left[\frac{I_{11}}{I_{10}} - \frac{I_{21}}{I_{20}} - \frac{lP}{\sigma^2}\left(1 - \frac{2I_{30}}{I_{10}I_{20}}\right)\right]. \tag{9}$$

Using these formula one compute the velocity of the protein translocation as a function of the noise intensity, or, more improtantly, as a function of the protein length (see Fig. 3). It is clearly seen that assuming the fluctuationally driven transport the lentgh can crucially change the velocity of the protein translocation. This function can be monotonous or nonmonotonous one. The results are

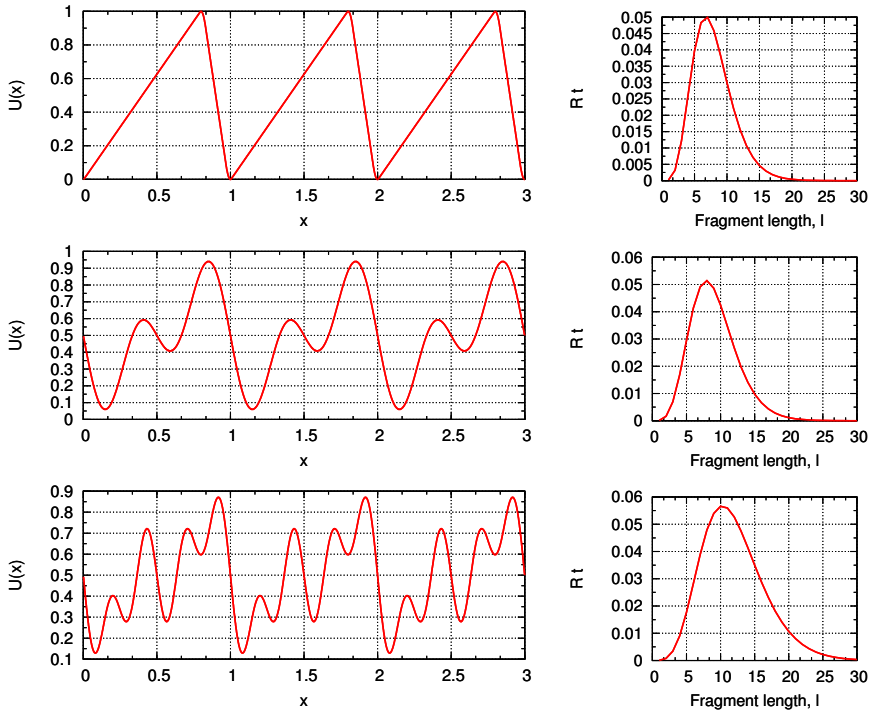

Fig. 4. Left: Different examples of asymmetric periodic potential $U(x)$ with the period $P = 1$. Only three periods are shown. From top to bottom: $U(x)$ is a saw-tooth function with smoothed angles (for details see [25]; $U(x) = -(0.25\sin(2\pi x/P) + 0.25\sin(4\pi x/P) + 0.5$; $U(x) = -(0.166\sin(2\pi x/P) + 0.166\sin(4\pi x/P) + 0.166\sin(8\pi x/P)) + 0.5$. Right: the corresponding dependencies of the velocity rate R_t on the peptide length l expressed in amino acids and $A = 1.15$.

quite general, for example, if the noise intencity is large enough, the nonmonotous transport rates can be obtained for a large variety of assymetric periodic potentials. These dependencies of the translocation rate on the peptide length, computed using Eq. 9, are shown in Fig. 4 (right) for different interaction potentials shown in the same Fig. These results show that very different forms of the interaction potential, if this potential fullfils the conditions of the fluctuationally driven transport, result in the nonmonotonuos translocation rate functions.

4 Transport model and influence of transport rates on the protein degradation

4.1 The transport model

Next let us show how one can compute the proteasome output if the transport rates are given. In our model we assume that the proteasome has a single chan-

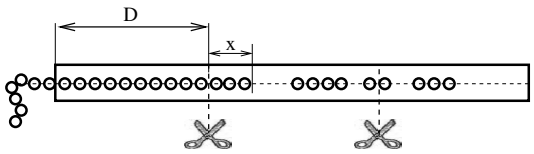

Fig. 5. Schematic diagram of the protein degradation by the proteasome. The protein strand (denoted by '0-0-0-0') enters the proteasome from the left to be cleaved by the cleavage centers (denoted by the scissors). The length of the strand after crossing the cleavage center is denoted by x.

nel for the entry of the substrate with two cleavage centers present at the same distance from the ends, yielding in a symmetric structure as confirmed by experimental studies of its structure. In reality a proteasome has six cleavage sites distributed spatially around its central channel. However, due to the geometry of its locations, we believe that a translocated protein meets only two of them. Whether the strand is indeed transported or cleaved at a particular position is a stochastic process with certain probabilities (see Fig. 5).

Parameter	Description	Dimension	Default Value
l	Protein Length	Amino acids	200
L	Proteasome Length	Amino acids	80
D	Distance between a cleavage center	Amino acids	15
	and the proteasome end		
N	Number of degraded proteins	-	1000
γ	Cleavage rate	-	0.001

Table 1. Parameter values used in the model

The protein strand can be cleaved if it lies close to the cleavage center or it could be transported forward by one amino-acid. We assume that the probability of transport depends only on the length of the strand inside the proteasome. The probability of transport is, therefore, given by a translocation rate function, $v(x+D)$ where $x+D$ is the length of the strand inside the proteasome in terms of amino acids. The probability of cleavage is assumed to be a constant, denoted by γ. We also assume that the degradation of proteins by the proteasome is a highly *processive* mechanism [2], i.e., in other words, the protein is not released by the proteasome until it is completely processed. This leads to the possibility of the proteasome making several cuts in the same protein strand and in the formation of greater number of smaller length peptides. The model also does not allow cleavage products to overlap or outrun their predecessors. The standard parameters we have used for this particular model are given in Table 1.

4.2 Analytics - Distribution of peptide lengths

To derive analytically the proteasome length distribution, we apply the master equation approach for the distribution of the coordinate of the front end of the strand and find its stationary solution as well as the distribution of peptide length. We also assume that the products of cleavage leave the chamber immediately after cleavage. This assumption can be justified by the fact that the products of cleavage have rather small size and can really left the proteasome very fast. We note that this assumption will be also motivated if the transport rate function is a monotonously decaying one, and the characteristic peptide length is small in comparison to the distance between the cleavage center and the proteasome end. In this case most of the peptides move faster than the incoming protein strand, and do not block it. Later on, using numerical simulations we compare whether this assumption can significantly change the results.

One cleavage center. We take the offset of the coordinate x (measured in amino acids) along the proteasome at the first cleavage center (see Fig. 5). During the time interval dt, the protein strand can move by one amino-acid with the probability $v(x+D)dt$ and can be cut with the probability γdt.

Let us have an ensemble of identical proteasomes with proteins inside, and $w(x)$ be the distribution of the proteins with the coordinate x. Then, at the x-th bond, the change in $dw(x)$ consists of the increase $v(x+D-1)w(x-1)dt$ due to the movement of protein ends from $x-1$ to x, the decrease $-v(x+D)w(x)dt$ due to the movement of protein ends from x to $x+1$, and the loss $-\gamma w(x)dt$ due to cleavage:

$$dw(x) = \left[v(x+D-1)\, w(x-1) - v(x+D)\, w(x)\right] dt - \gamma\, w(x)dt. \tag{10}$$

At $x=0$ (the protein ends at the cleavage center) we have to set the boundary condition: no strands are to be cut, and no strands move from $x=-1$ because no proteins end there, but there is a gain $\sum_{x=1}^{\infty} \gamma\, w(x)dt$ from cleavage, i.e.

$$dw(0) = \Big[\gamma \sum_{x=1}^{\infty} w(x) - v(D)\, w(0)\Big] dt. \tag{11}$$

Using the standard master equation technique, we consider the continuous limit, supposing that $w(x)$ and $v(x)$ vary slightly from one bond to another. Then, $\sum_{x=1}^{\infty} w(x) \approx \int_0^{\infty} w(x)dx = 1$ (the last being the normalisation condition), and the discrete derivative in Eq. (10) may be replaced with the continuous derivative:

$$\dot{w}(x,t) = -\frac{\partial}{\partial x}\big(v(x+D)\, w(x,t)\big) - \gamma\, w(x,t), \tag{12}$$

with the boundary condition Eq. (11), $\dot{w}(0) = \gamma - v(D)\, w(0)$, i.e. for stationary case

$$w(0) = \frac{\gamma}{v(D)}. \tag{13}$$

The stationary solution of Eq. (12) with the boundary condition (13) gives the asymptotic distribution of protein coordinates which coincides with the length distribution of peptides $\rho(x)$ (because probability of a strand to be cut is independent of its coordinate). The problem Eqs. (12),(13) admits the stationary solution

$$\rho(x) = w(x) = \frac{\gamma}{v(x+D)} \exp\Big[-\gamma \int_0^x \frac{dx'}{v(x'+D)}\Big], \tag{14}$$

thus providing us with the analytically found proteasome product length distribution.

Two cleavage centers. Let us now have a second cleavage center at $x_2 = L - 2D$ (see Fig. 5). If γdt is the probability of cleavage by a center during the time interval dt, then the probability of a peptide of length x to be cut by any center is $\gamma(x)dt$:

$$\gamma(x) = \begin{cases} \gamma, & 0 < x < x_2; \\ 2\gamma, & x \geq x_2. \end{cases}$$

In this case, the length distribution does not coincide with the distribution of protein coordinates. Let us start with the distribution of protein coordinates $w(x)$. Beyond the cleavage centers it obeys the equation which is similar to Eq. (12) :

$$\dot{w}(x,t) = -\frac{\partial}{\partial x}\big(v(x+D)\, w(x,t)\big) - \gamma(x)\, w(x,t). \tag{15}$$

At the first cleavage center, again, $w(0) = \gamma/v(D)$; at the second one ($x = x_2$)

$$\dot{w}(x_2) = \gamma \int_{x_2}^{\infty} w(x)dx + v(x_2 + D - 1)\, w(x_2 - 1) - v(x_2 + D)\, w(x_2) - \gamma w(x_2),$$

what provides

$$w(x_2) = w(x_2 - 1) + \frac{u}{v(x_2 + D)}, \tag{16}$$

where $u \equiv \gamma \int_{x_2}^{\infty} w(x)dx$.

So, for $x \mathrm{in} (0, x_2)$ the protein coordinate distribution is:

$$w(x) = w_1(x) = \frac{\gamma}{v(x+D)} \exp\Big[-\gamma \int_0^x \frac{dx'}{v(x'+D)}\Big], \tag{17}$$

and for $x \mathrm{in} [x_2, \infty)$, one can obtain $u = v(x_2 + D)w_1(x_2)$,

$$w_2(x) = \frac{2\gamma}{v(x+D)} \exp\Big[-\int_0^x \frac{\gamma(x')\, dx'}{v(x'+D)}\Big]. \tag{18}$$

To note, the formulae (17) and (18) may be combined:

$$w(x) = \frac{\gamma(x)}{v(x+D)} \exp\Big[-\int_0^x \frac{\gamma(x')\, dx'}{v(x'+D)}\Big]. \tag{19}$$

In its turn, the distribution of peptide lengths $\rho(x)$ is proportional to a superposition of $w(x)$ and shifted $w_2(x)$, i.e., $\rho(x) \propto w(x) + w_2(x + x_2)$, that gives us after normalisation:

$$\rho(x) = \frac{w(x) + w_2(x + x_2)}{1 + e^{-\gamma \int_0^{x_2} \frac{dx'}{v(x'+D)}}}. \tag{20}$$

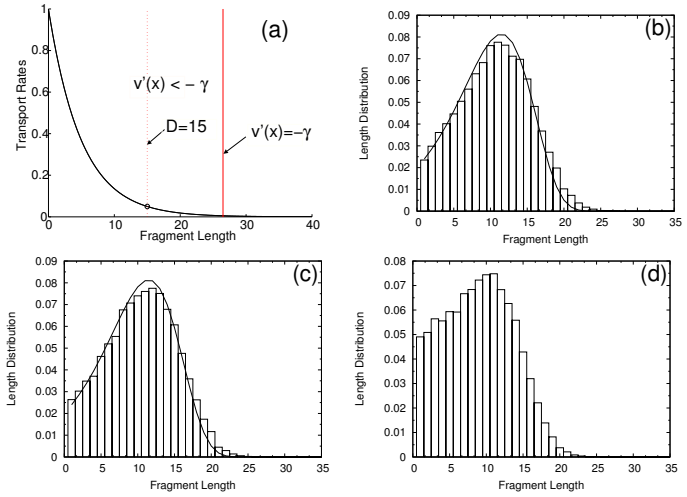

Fig. 6. (a) Monotonously decreasing transport rate function $v(x) = v_1(x)$. The vertical lines show the location of the point where the condition for the maximum in the length distribution (Eq. 22) holds and the location of the cleavage centre. The parameters are $D = 15$ and $\gamma = 0.001$. As predicted by theory (solid line) the numerically computed peptide length distribution (denoted by boxes) has a maximum for one cleavage center (b), two cleavage centers with immediate disappearance of cleavage products (c), and in the case where the cleavage products do not disappear (d).

Maximum in peptide length distribution As shown below, the presence of a second cleavage centre does not significantly change the results. Hence, to find the conditions for a maximum in the peptide length distribution, we use the analytical expressions for the case of one cleavage centre. This peptide length distribution has extrema at the points where

$$0 = \frac{dw(x)}{dx} = \frac{\gamma}{v^2(x+D)} \exp\Big[-\gamma \int_0^x \frac{dx'}{v(x'+D)}\Big]\Big(-\frac{dv(x+D)}{dx} - \gamma\Big), \tag{21}$$

what gives the condition for extremum,

$$\frac{dv(x+D)}{dx} = -\gamma. \tag{22}$$

This equation should be fulfilled at least in one point for $x + D > 0$. To note γ is here the rate of cleavage. Hence, there are no limitations for its value. Eq. (22) shows that the condition for obtaining a maximum when we have a single cleavage center is independent of the actual form of the transport rate function, but is rather dependent on its slope. This would suggest that it is possible to obtain a peak in the length distribution even when the transport rate function is monotonically decreasing, and indeed that is what we find from our numerical simulations.

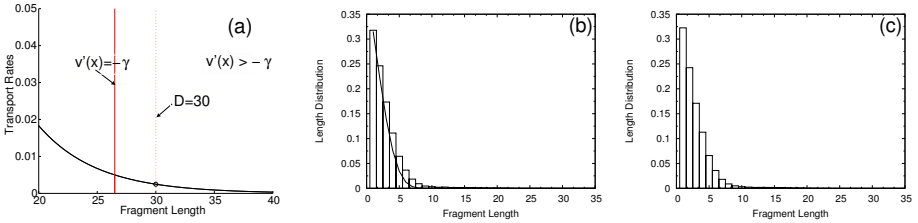

Fig. 7. (a) The case when D is large and the maximum condition holds nowhere. As predicted by the analytical theory (solid line) the length distribution does not have peaks under the assumption that the cleavage products disappear immediately (b) or without this assumption (c).

4.3 Comparison with numerical results

We adapt our model for numerical simulation with the help of Gillespie algorithm [10] which enables the system to jump to the next event via the calculation of the waiting time before any event will occur. Following the approach suggested by us in [50], we stochastically model the system where several events can happen with different probabilities. Suppose that on some moment of time we have a set of N probable events with rates R_i, where the $i-th$ event has the rate R_i and $i = 1..N$. Then by generating two uniformly distributed in (0, 1) random numbers RN_1 and RN_2, we estimate the time T after which the next event would occur as

$$T = -\frac{\log(RN_1)}{\sum_{i=1}^{N} R_i}. \tag{23}$$

The concrete event k that occurs after this time can be then found from:

$$\sum_{i=1}^{k} R_i < \sum_{i=1}^{N} R_i RN_2 \leq \sum_{i=1}^{k+1} R_i. \tag{24}$$

The peptide or its part inside the proteasome can either be shifted by one amino acid or it can be cleaved if it is located near the cleavage center. Inside the proteasome, the translocation rates of the substrate or fragments depend only

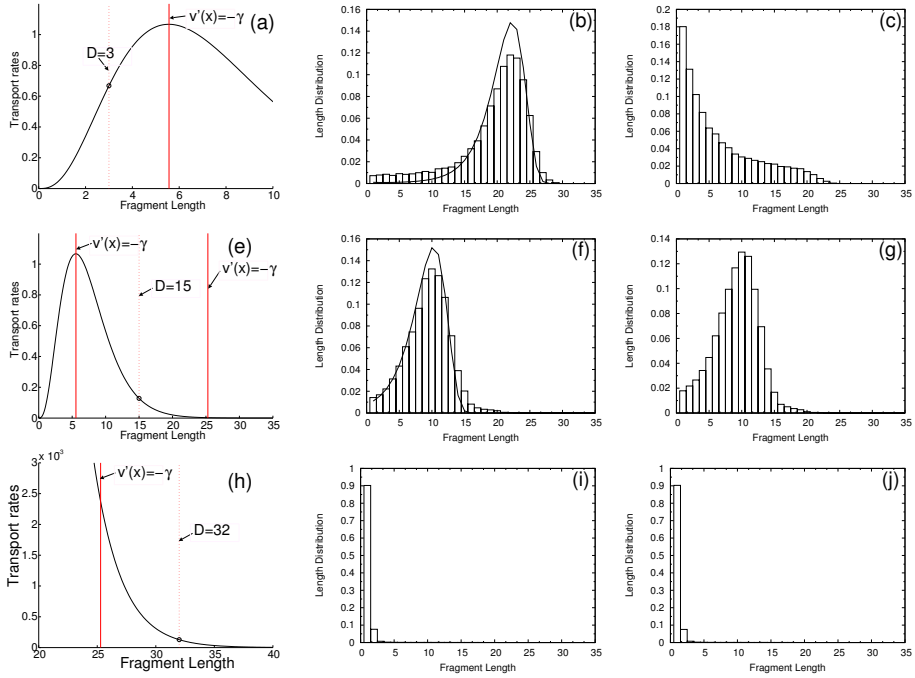

Fig. 8. The case of non-monotonous transport rate function $v_2(x)$, shown by solid line in (a,e,h). From top to bottom: different location of the cleavage centre $D = 3, 15, 32$ (shown the vertical line) with respect to two points where the condition of maximum holds (shown vertical lines); (b,f,i) the corresponding length distributions for the case of two cleavage centres. Numerics (boxes) is always in good agreement with theoretical predictions (solid line); (c,g,j) the corresponding length distribution if computed without the assumption of immediate disappearance. Only in one case (c) of rather unrealistic too small D this length distribution changes a lot from distributions (b,f,i).

on their lengths and are described by the translocation rate function $v(x+D)$ (see Fig. 5). The probability of cleavage is described by the function $R_c(p)$, where p is the position in the substrate sequence. We set the constant cleavage probability $R_c(p) = \gamma$ and discuss another situation later. When the protein is degraded, its fragments lengths are counted in the length distribution. The reliability of the results is ensured as we conduct the study over a large number of proteins N so that the trends observed in the length distribution have the basis in statistical results. To consider different situations, we study different relations between the geometry of the proteasome (the parameter D) and two qualitatively different transport rate functions, monotonous and non-monotonous one.

For the sake of generality we choose the following qualitatively different transport rate function:

$$v_1(x) = e^{-0.2x}, \quad v_2(x) = 0.125e^{-\alpha x}x^3, \tag{25}$$

where x is the peptide length and $\alpha = 0.54$ is a constant. It is important to note that such monotonous and non-monotonous transport functions can be obtained also if we assume that the protein translocation is driven by nonequilibrium fluctuations, as we have assumed it in [52].

Monotonously decreasing transport rates. First we consider a case of the monotonously decreasing transport rate function v_1. This form of function (see Fig. 6 (a)) may correspond to two different situations: when the condition (Eq. 22) can be fulfilled at some point and when it cannot. The first case occurs if the parameter D is smaller than the point where the derivative of the transport function is equal to the value $-\gamma$. As predicted by the theory in this case one should observe the maximum in the length distribution. This is indeed the case, as can be shown by a comparison of the numerical results with the theoretical curve for one cleavage centre (see Fig. 6 (b)) or for two cleavage centres (see Fig. 6 (c)). It should be noted that addition of the second cleavage centre does not change the results both as predicted from the theory or computed with the Gillespie algorithm. This holds for all situations considered in this paper, hence below we plot only the results for two cleavages centers. These plots clearly show a good matching between the analytical theory developed and the results of numerical simulations. To remind, both the theory and numerical simulations have been made under the assumption that the cleavage products disappear immediately after the cut. The numerical simulations allow also to check what happens if this assumptions does not hold and the cleavage product will still move along the channel. As expected, for a monotonously decreasing transport function, the length distribution does not significantly change (see Fig. 6 (d)). This happens because the shorter cleavage products have larger probability of transport and do not stack the incoming protein.

The situation qualitatively differs if after the cleavage centre the condition of the maximum is never fulfilled. This happens if D is relatively large (see Fig. 7 (a)). In this case the length distribution would not have a maximum which is in good correspondence to the result predicted by the theory. The Fig. 7 (b) illustrates that also in this case analytical results are in good matching with numerics. The length distributions also practically do not change if the cleavage products do not disappear after the cleavage but are transported with the same rules as the incoming protein (see Fig. 7 (c)).

Non-monotonous transport rates. To cover several possible forms of the transport function, next we analyse the transport function $v_2(x)$ with one peak (see Fig. 8 (e)). The derivative of this function can be equal to $-\gamma$ in two points dividing all possible coordinates into three regions where the cleavage centre D can be located. Let us study these three possibilities.

First, let us consider the case when the location of the cleavage centre allows the maximum condition to be fulfilled in two points (see Fig. 8 (a)). In this case the theory predicts the non-monotonous length distribution, as confirmed also by numerics both for one or two cleavage centres (see Fig. 8 (b)). However, if we let the cleavage products do not disappear, then the length distribution is monotonously decreasing, as found by numerical simulations (see Fig. 8 (c)). This is the only case when using this assumption changes the results significantly and the theory does not predict the form of length distribution without this assumption. One should note that this case of $D = 3$ seems to be rather unrealistic, because it means that the cleavage centre is too close to the entrance what is not the case in the reality [37,7].

Next, we analyse the case, when the cleavage centre $D = 15$ is between two points where the derivative is equal to $-\gamma$, and hence the maximum condition can be fulfilled in one point (see Fig. 8 (e)). As predicted by theory and confirmed by numerics, the length distribution is to have a maximum in this case (see Fig. 8 (f)). Noteworthy, the theory in this case works sufficiently also in the case if the cleavage products do not disappear immediately (see Fig. 8 (g)). If we believe in the non-monotonous transport rate function hypothesis [50], this case seems to be the most adequate for protein degradation by the proteasome. In the last case, the cleavage centre is so deep in the proteasome $D = 32$ that the maximum condition can be fulfilled nowhere. As expected, the length distribution has no maxima in all cases computed for this relation between the transport function and the geometry of the proteasome.

5 Kinetic model of the the proteasome

The model. Here we review the mathematical description of the proteasome degradation in the case if the translocation plays not the main role and can be neglected in comparison to the kinetics of the in- and outflux rates. The model describes the rates at which the concentrations of fragments of length k change over time. The concentrations change by proteasomal cleavage, making two short fragments out of a long one, and by the influx and efflux of fragments through the gates. The dynamics does not depend on the amino acid sequence and orientation of the fragment, but only depends on the length of the fragment. Let n_k and N_k be the concentration of fragments of length k inside and outside the proteolytic chambers. Then,

$$\frac{dN_k}{dt} = -a(k)\Big[1 - v\sum_{j=1}^{L} jn_j\Big]N_k + e(k)n_k \ , \tag{26}$$

$$\begin{aligned}\frac{dn_k}{dt} &= a(k)\Big[1 - v\sum_{j=1}^{L} jn_j\Big]N_k - e(k)n_k \\ &\quad -c\sum_{i=1}^{k-1} F_{k,i}n_k + c\sum_{j=k+1}^{L}(F_{j,k} + F_{j,j-k})n_j \ ,\end{aligned} \tag{27}$$

Parameter	Description	Dimension	Default value
L	Length of the Substrate	amino acids	100
$N_L(0)$	Initial substrate concentration	mol	100
$\hat{a}$	Rate of influx	time^{-1}	0.1
$\hat{e}$	Rate of efflux	time^{-1}	1
c	Cleavage rate	time^{-1}	1
θ	Critical fragment length	amino acids	25
μ	Preferred cleavage position	amino acids	9
σ	Std of cleavage position	amino acids	3
v	Scaling factor	-	1/200

Table 2. Parameters values of the kinetic model.

for $k = 1, 2, \ldots, L$. The substrate N_L is an outside fragment of length $k = L$. The first term of Eq. (26) describes the influx of fragments into the proteasome. For the influx function $a(k)$ we consider the case where there is no re-entry of fragments other than the substrate, i.e., we set $a(k) = \hat{a}$ for $k = L$, and $a(k) = 0$ otherwise. The influx of substrate into the proteolytic chambers is a rate limiting factor in protein degradation. Experimental works have suggested that the influx is limited by the maximum amount of amino acids that can be accommodated in the proteasome (see [27] and refs therein). In our model the influx rate therefore decreases when the total amount of amino acids inside, $\sum_{k=1}^{L} kn_k$, increases. The maximum filling of the proteasome is normalised to one by a scaling parameter v determining the maximum number of amino acids that can be accommodated within the CP.

We assume that the influx does not strongly depend on the amino acid composition of the substrate. Based on the intuition that each peptide binds with a probability p to the gate subunits, hence impairing the passage through the narrow pore, it has been proposed that the efflux rate is a negative exponent of the length $exp(-\gamma n)$ where $\gamma = \frac{1}{1-p}$ and n is the fragment's length [16]. On the other hand, the analysis of *in vitro* digestion of 25 and 27 aa long substrates performed with the 20S proteasome suggests a length dependent reprocessing rate which decreases with the increase of the substrate length. These data suggest an increasing Hill function with high exponent to describe the reduced cleavage rate for short substrates [36]. Hence, we describe the efflux rate with a phenomenological Hill function with high exponent and a critical length $\theta = 25$ aa $e(k) = \hat{e}/(1 + (k/\theta)^{10})$. The efflux rate switches at a fragment length of $k \simeq \theta$ from the maximal efflux rate $\hat{e} = 1$ for short fragments to an efflux close to zero for long fragments (see 9A).

The first two terms of Eq. (27) are the same influx and efflux terms as discussed above. The last terms describe the cleavage machinery located in the core of the proteasome. Fragments of length k are cut at a maximum rate c and with probability $0 < F_{k,i} < 1$ into two fragments of length i and $k - i$. Two terms account for the loss and for the gain of each fragment of length k. The negative term corresponds to a loss for fragments of length k which are cut in

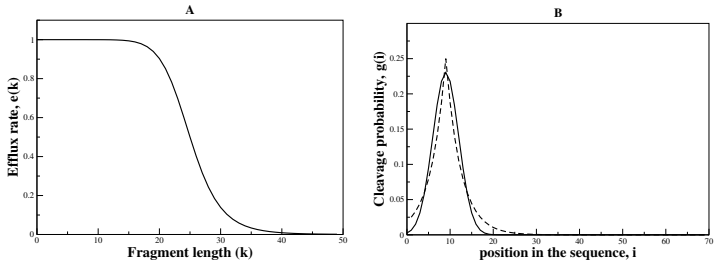

Fig. 9. A. The efflux rate. **B.** Binomial (dashed lines) and gaussian (solid lines) distributions for the cleavage probability.

shorter fragments, and the positive term is a gain because fragments of length $j > k$ can be cleaved into a fragment of length k. For parameters see Table 2.

The main assumption for the cleavage mechanism is that the proteasome cleaves proteins starting around their N-termini or C-termini. It has been suggested that there is a preferred length of 7-9 aa for an optimal docking of the substrate with the binding grooves during the cleavage process inside the CP [13]. We therefore assume that the proteasome starts at a distance $m \simeq 9$ from one end of the protein/peptide, and scans the substrate chain in both directions until a cleavage site is found. One can model [27] the cleavage probability with a phenomenological binomial or a Gaussian distribution, as here.

The model has three rate parameters: the cleavage rate c, the maximum influx rate $\hat{a}$, and the maximum efflux rate $\hat{e}$. A normal time scale of proteasome experiments is minutes. However, experimental results on proteasome degradation are typically compared for a certain level of substrate degradation, rather than at a specific point in time. Since time is not an important issue, one can always rescale the time such that $c = 1$ per time unit. Increasing the cleavage rate will therefore be the same as decreasing the flux through the gates (i.e., as decreasing $\hat{a}$ and $\hat{e}$). For details of the simulation algorithm see [27].

Kinetics. Experimental data suggest that the *in vitro* degradation rate of substrates by the proteasome obeys Michaelis-Menten kinetics (see [27] and refs therein). For long substrates the maximum degradation rate and the Michaelis-Menten constant are known to decrease with the length of the substrate. Our model also exhibits Michaelis-Menten kinetics (see 10). For various initial substrate concentrations, 10 depicts the depletion of the substrate ($L = 100$) in the solution (Panel A), and the corresponding filling of the proteasome (Panel B). There is a rapid initial phase during which the proteasome fills up by influx of the substrate. At the very early stage of degradation, due to the filling of the proteasome, the substrate loss is not linear. When the initial substrate concentration is low this initial phase accounts for a significant depletion of the substrate concentration N_L (see 10A). Otherwise, the substrate concentration remains high and the filling of the proteasome approaches a quasi steady state corresponding to a maximum degradation rate.

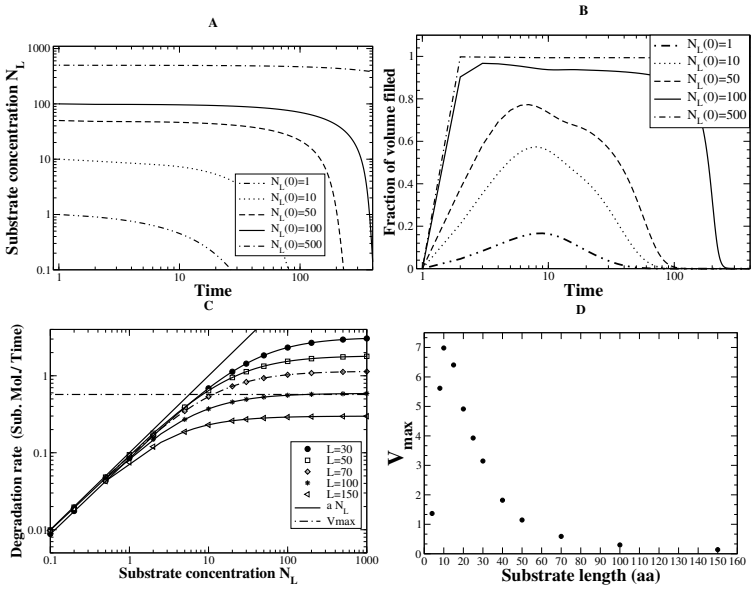

Fig. 10. A. Substrate consumption for a different initial concentration, $N_L(0)$. **B.** Filling of the proteasome core particle (in amino acids). **C.** Michaelis-Menten log-log plot; each symbol-curve denotes a substrate of a different length obtained from the model presented in Eqs. (26, 27) in conditions where the proteasome is in a bath of substrate in order to prevent substrate limiting effects. The solid line is given by the initial slope $\hat{a}N_L$ and the dashed-dot line is the V_{max}, both predicted with the simplified model (see [27]) for the standard parameter values (see Table 2) and an average efflux $\bar{e} = 0.2$. **D.** V_{max} as a function of the substrate length. It increases for substrate shorter than 10 aa and decreases for longer substrates. The initial substrate concentration is $N_L(0) = 6000$.

To study the Michaelis-Menten kinetics, we fix the substrate concentration by fixing $N(t) = N(0)$ and let the model approach the corresponding steady state. At the steady state we measure the degradation rate as the number of substrate molecules in solution which are lost per unit time, and depict that as a function of the substrate concentration and the length of the substrate, L (see 10C). This reveals a family of Michaelis-Menten curves for the various lengths of the substrate. The longer the substrate, the smaller the maximum degradation rate, V_{max}, and the smaller the Michaelis-Menten constant, K_m. The degradation rate at low substrate concentrations is fairly independent on the length of the substrate (see 10C).

Fig. 10D shows the maximum degradation rate V_{max}, calculated numerically from the full model Eqs. (26, 27), as a non-linear function of the substrate length. It increases for small substrates with a maximum at ca. 10 aa and decreases with the inverse of the length for longer substrates. Very small fragments are weakly degraded because of the low cleavage rate for fragments shorter than 9 aa. This

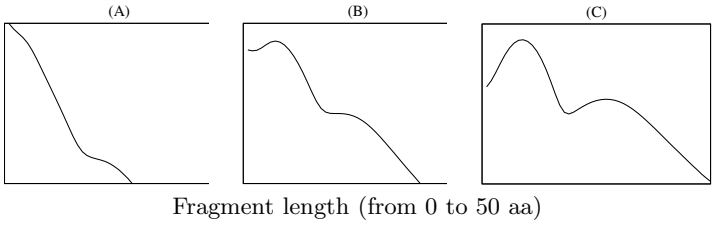

Fig. 11. Length distributions of the fragments outside the proteasome. From left to right the efflux rate $\hat{e}$ increases $\hat{e} = (0.1, 1, 10)$. Each distribution has been taken at the time 170 (A), 86 (B), 76 (C) when 20 % of substrate degraded. The vertical on each plot is the log frequency (from 0 to 0.25). The influx rate $\hat{a} = 0.01$. Note that the distributions are insensitive to the variation in the influx rate $\hat{a}$ [27].

explains why the degradation rate decreases for very short fragments. For longer fragments, V_{max} was found to decrease with the inverse of the length. This is due to the fact that the CP volume is finite and thus the influx decreases with the increase of the substrate length. Therefore, the degradation rate decreses accordingly. This compares well with experimental results (see [27] and refs therein). Interestingly, the ratio of V_{max}/K_m increases with the increase of the substrate length [8] and saturates for long substrates reaching 50% of its maximum for substrate of 23 aa which indicating that both constants decrease with the length of the substrate in the same ratio.

Length distribution of the fragments. *In vitro* experiments generate cleavage products that range from 2 to 35 aa with an average length of $7 - 8$ aa (see [27] and refs therein). Using size exclusion chromatography and on-line fluorescence detection, [23] showed that the products generated by the wild type (WT) proteasome have a length distribution with three broad peaks corresponding to lengths of $2 - 3$, $8 - 10$, and $20 - 30$ aa, respectively.

In 11 we show how the fragment length distribution depends on the size of the gate, i.e., on the influx and efflux rates $\hat{a}$ and $\hat{e}$, as calculated with the model Eqs. (26, 27). For an intermediate efflux rate, we obtain three-peaked distributions similar to those observed in experiments [23] for a wide range of influx rates. Note that the first peak has its maximum at 1 aa but we call this decreasing slope "peak" for simplicity. In our model, the three-peaked distributions are the result of the cleavage machinery, which tends to cut fragments of $8 - 10$ aa, and the efflux of products, which favors the short fragments.

In the Panel A the efflux is slow compared to the cleavage ($c/\hat{e} = 10$). As a consequence, most substrate molecules are fragmented extensively before they are exported, and one observes short fragments in the solution. Increasing the efflux rate 10-fold (the panel B) gives a similar time scale to the efflux and to the cleavage, and allows for a three-peak distribution. Another 10-fold increase of the efflux rate (see Panel C) makes cleavage the limiting factor. The ratio of

long to short fragments increases. Because the residence time of fragments in the CP is short, there is less fragmentation, and the first peak at 1-3 aa decreases.

When the efflux rate and the cleavage rate have a similar time scale we observe three peaks in the distribution of fragments (see 11B). Similar to what is observed experimentally [23], the third peak is much smaller than the other two, and the second peak is larger than the first peak. In our model, the first peak corresponding to the small fragments reflects an efficient cleavage mechanism where fragments are repeatedly cleaved before they are released from the CP. These "rest" products do not collapse to single amino acids because cleavage of very short fragments is improbable in our model (see 9B). The second peak corresponding to fragments with a length of $8-10$ aa, is the result of the preference to cut at $\mu = 9$ aa. Fragments are found in a broad peak around 9 aa, because of the variation in the cleavage (i.e., standard deviation of the Gaussian function). The third peak around 25 aa found in the WT distribution is due to the efflux function. It results from the high probability of a fraction of intermediate 25-35 aa fragments to exit the proteasome. As they would have been a source for fragments of length 15-25 aa, the production of fragments of this length drops. This intuitive explanation elucidates the presence of the third peak.

6 Discussion

6.1 Development of modelling

This work provides insights on mathematical modelling of the kinetic and translocation properties of the proteasome degradation. A realistic choice for the models parameters is hard to obtain. Both model are phenomenological and designed to qualitatively capture the main features of the kinetics and transport. It is important to note that there is no contradiction between these two models, the kinetic model operates on the mesoscopic scale whereas the transport model on the microscopic one. Hence, the transport model can be included inside the parameters of the kinetic model. At the present stage the models discussed cannot be used for quantitative predictions and serves as an explanation of the proteasome product size distribution. We identify two possible directions which can be taken in order to develop these models to be able to perform quantitative predictions. A first possibility is to implement a specific protein sequence, which certainly influences a cleavage pattern and the proteasome-protein interaction potential, and description of in- and outfluxes, taking into account the possibility of gate opening and closing [23].

It is important to note that both, the transport and kinetic model of the proteasome allows the inclusion of the sequence specific cleavage strengths. In the transport model the cleavage specificity can be included if we set the cleavage probability dependent on the position in the peptide chain. Let us simulate the degradation of the Casein as in [23] with constant cleavage rates and with sequence specific cleavage rates computed with Netchop algorithms. Casein has a length of $L_p = 188$ aa and the cleavage pattern as in in Fig. 12 right. All other parameters are the same as in the case of the three peak length distribution,

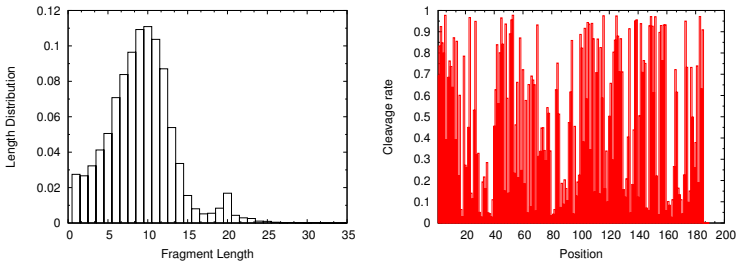

Fig. 12. Length distribution as a result of Casein degradation, with non sequence specific cleavage rates (left). Casein cleavage strength pattern computed with Netchop. (right).

see Fig. 12 left. We have rescaled the cleavage strength to have the same mean value of 0.001. To our surprise the length distributions of not sequence specific case and sequence specific case are practically identical (see Fig. 12 left). Hence to model the length distribution as a result of different translocation rates, it is not so important to consider the cleavage pattern of the substrate. However, for shorter substrates the cleavage pattern can significantly change the length distribution.

Cleavage specificity can be also included in the kinetic model. For this instead of the cleavage matrix which depends only on the distance from the protein end one should include the cleavage coefficients which depend on the protein sequence. This would be possible if one consider separately the concentration equations for all fragments generated from the initial substrate. Then taking the cleavage strength and the substrate concentration as an input of the mathematical model one can predict the dynamics of the fragment concentration. The results which illustrate this approach will be published by us elsewhere.

Finally for effects where both the translocation and kinetic properties are important one can merge both models. This can be done via including the transport corrections in the kinetic model described here. In this case different transport rate functions will result in different influx coefficients and, hence, different kinetic properties of the proteasome. This modelling approach is especially important for the study of the proteasome role in the neurodegenerative diseases, as we discuss it in the next subsection.

6.2 Kinetics models and neurodegenerative associated proteasome degradation

A broad array of human neurodegenerative diseases share strikingly similar histopathological features, such as the presence of insoluble protein deposits, such as the neurofibrillary tangles and neuritic plaques of Alzheimers disease, the Lewy bodies of Parkinsons disease, and the intranuclear inclusions of Huntingtons disease [42].

As the 26S proteasome is the cellular proteolytic machinery involved in the clearance of uniquitinated proteins, this has led to the suggestion that a chronic imbalance between their generation and processing may be the primary cause for the formation of protein deposits. Recent experimental data revealed an impaired 26S proteasome activity in neurodegnerative diseases [39,48]. Using our kinetics model several possible scenarios to explain proteasome dysfunction in neurodegenerative diseases can be envisaged. For instance the transport of elongated molecules, such as the substrates found in neurodegenerate brains, may require a longer time. Therefore, the delayed transport can be responsible for keeping proteasome sequestrated for a long time lowering the degradation rate.

Kinetics models also would predict and provide quantitative results of proteasome inhibition through a reduced catalytic activity of the 20S proteasome chamber. A reduced cleavage rate c can result in a filling up of the proteasome chambers and therefore in an impairment of the degradation process. Moreover, a non-trivial scenario arise in the kinetics model presented here when a regime of reduced cleavage activity is considered. In this condition the degradation rate decreases with the increase of the influx rate of substrate within the proteasome (data not shown). This unexpected result is a direct consequence of the filling of the proteasome and of the the internal dynamics of undegraded substrates and intermediate long fragments. In fact, when both the influx and the cleavage rate are very low, the proteasome is almost empty and the substrates entering the chamber are immediately degraded and fragments are ejected. Increasing the influx rate the proteasome fills up much faster. Consequently, becauseof the low cleavage rate, many long intermediate fragments will be generated. The proteasome fills up at a high rate, reducing the number of substrate molecules entering the proteasome chamber per unit time. A quasi steady state regime will be reached, where proteasomes are filled up with long intermediate fragments. This unexpected result might be experimentally explored: experiments with different substrate concentrations in a low cleavage regime could prove this theoretical expectation.

Modeling the ATP-dependent activity of the 19S base and the mechanical transport of substrate molecules within the 20S core are suggestive and intriguing phenomena that can help the understanding of proteasome blockage and of the impaired activity observed in neurodegenerative associated protein degradation. Unfortunately to date, no such model has been proposed, the main reason being the very little information about these mechanisms. One model-candidate for analysing this problem is the ratchet-kinetic model, intent to describe the function of the transport through the 19S and the kinetics of cleavage within the 20S core particle. This complete model can unravel the dynamics of transport of a given substrate through the 19S subunits and the successive kinetics within the core particle, where substrate specific cleavage machinery is implemented.

Therapeutic application of these future results can be envisaged. For instance, the deficiency of ATP as major cause of deficient transport for elongated and/or modified molecules can be tested quantitatively. Providing ATP source

in this impaired system, such as Alzheimer disease or Parkinson disease, to the proteasome system can result in a reduction of substrate accumulation. Predictions with mathematical models can be tested by manipulating the activity of the proteasome in vivo. This is a challenging task but will undoubtedly provide more insight into the pathogenesis of neurodegenerative diseases and in new therapeutic options.

Acknowledgments. AZ acknowledges a financial support from the VW-Stiftung. AZ and JK from the European Union through the Network of Excellence BioSim, Contract No. LSHB-CT-2004-005137. FL from the Australian Research Council through the Discovery Grant scheme (DP0556732 and DP0664970).

References

1. J. Adams, V.J. Palombella, and P.J. Elliott. Proteasome inhibition: a new strategy in cancer treatment. *Investigational New Drugs*, 18:109–121, 2000.
2. T. N. Akopian, A. F. Kisselev, and A. L. Goldberg. Processive degradation of proteins and other catalytic properties of the proteasome from *thermoplasma acidophilum*. *The Journal of Biological Chemistry*, 272:1791–1798, 1997.
3. M. Bier. Processive motor protein as an overdamped brownian stepper. *Phys. Rev. Lett.*, 91:148104, 2003.
4. C.J. Brokaw. protein-protein ratchets: stochastic simulation and application to processive enzymes. *Biophys. J.*, 81:1333–1344, 2001.
5. P. Cascio, C. Hilton, A. F. Kisselev, K. L. Rock, and A. L. Goldberg. 26S proteasomes and immunoproteasomes produce mainly n-extended versions of an antigenic peptide. *EMBO J.*, 20:2357–2366, 2001.
6. A. Ciechanover. The ubiquitin-proteasome proteolytic pathway. *Cell*, 79:13–21, 1994.
7. O. Coux, K. Tanaka, and A. L. Goldberg. Structure and functions of the 20S and 26S proteasomes. *Annu. Rev. Biochem.*, 65:801–847, 1996.
8. I. Dolenc, E. Seemuller, and W. Baumeister. Decelerated degradation of short peptides by the 20S proteasome. *FEBS. Lett.*, 434:357–361, 1998.
9. Q. P. Dou, D.M. Smith, K.G. Daniel, and A. Kazi. Interruption of tumor cell cycle progression through proteasome inhibition: implications for cancer therapy. *Progress in Cell Cycle Research*, 5:441–446, 2003.
10. D.T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.
11. M. Glickman and A. Ciechanover. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.*, 82:373–428, 2002.
12. A. L. Goldberg, P. Cascio, T. Saric, and K. L. Rock. The importance of the proteasome and subsequent proteolytic steps in the generation of antigenic peptides. *Mol. Immunology*, 39:147–164, 2002.
13. M. Groll and R. Huber. Substrate access and processing by the 20S proteasome core particle. *Int. J. Biochem. Cell. Biol.*, 35:606–616, 2003.
14. W. Hilt and D. H. Wolf. Proteasomes of the yeast s. cerevisiae: genes, structure and functions. *Mol. Biol. Rep.*, 21:3–10, 1995.
15. H.G. Holzhuetter, C. Frommel, and P.M. Kloetzel. A theoretical approach towards the identification of cleavage-determining amino acid motifs of the 20s proteasome. *J. Mol. Biol.*, 286:1251–1265, 1999.

16. H. G. Holzhütter and P. M. Kloetzel. A kinetic model of vertebrate 20S proteasome accounting for the generation of major proteolytic fragments from oligomeric peptide substrates. *Biophysical Journal*, 79:1196–1205, 2000.
17. P. Jung and P. Hänggi. Resonantly driven brownian motion: Basic concepts and exact result. *Phys. Rev. A*, 41:2977, 1990.
18. C. Kesmir, A.K. Nussbaum, H. Schild, V. Detours, and S. Brunak. Prediction of proteasome cleavage motifs by neural networks. *Protein Engineering*, 15:287–296, 2002.
19. A. F. Kisselev, T. N. Akopian, and A. L. Goldberg. Range of sizes of peptide products generated during degradation of different proteins by archaeal proteasomes. *J. Biol. Chem.*, 273:1982–1989, 1998.
20. P. M. Kloetzel. Antigen processing by the proteasome. *Nat. Rev. Mol. Cell. Biol.*, 2:179–187, 2001.
21. P. M. Kloetzel. Generation of major histocompatibility complex class I antigens: functional interplay between proteasomes and TPPII. *Nature Immunology*, 5:661–669, 2004.
22. P. M. Kloetzel. The proteasome and MHC class I antigen processing. *Biochem. Biophys. Acta.*, 1695:225–233, 2004.
23. A. Köhler, P. Cascio, D. S. Leggett, K. M. Woo, A. L. Goldberg, and D. Finley. The axial channel of the proteasome core particle is gated by the rpt2 ATPase and controls both substrate entry and product release. *Molecular Cell*, (7):1143–1152, 2001.
24. C. Kuttler, A. K. Nussbaum, T. P. Dick, H. G. Rammensee, H. Schild, and K. P. Hadeler. An algorithm for the prediction of proteasomal cleavages. *J. Mol. Biol.*, 298:417–429, 2000.
25. P.S. Landa, A.A. Zaikin, and L. Schimansky-Geier. Effect of the potential shape and of a brownian particle mass on noise-induced transport. *Chaos, Solitons & Fractals*, 12:1459–1471, 2001.
26. B. Lankat-Buttgereit and R. Tampe. The transporter associated with antigen processing: function and implications in human diseases. *Physiol Rev*, 82:187–204, 2002.
27. F. Luciani, C. Kesmir, M. Mishto, M. Or-Guil, and R. J. de Boer. A mathematical model of protein degradation by the proteasome. *Biophysical Journal*, 88:2422–2432, 2005.
28. F. Luciani and A. Zaikin. Mathematical models of the proteasome product size distribution. 2006. (submitted).
29. O. Lund, S. Brunak, and et. al. Creating a virtual immune system. *Journal of Biological Physics*, 2006. (in press).
30. M. Nielsen, C. Lundegaard, O. Lund, and C. Kesmir. The role of the proteasome in generating cytotoxic t-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics*, 57:33–41, 2005.
31. A. Nussbaum. *From the test tube to the World Wide Web: The cleavage specificity of the proteasome*. PhD thesis, Eberhard-Karls-Universitaet Tuebingen, 2001.
32. A. K. Nussbaum, T. P. Dick, W. Kielholz, M. Schirle, S. Stevanovic, K. Dietz, W. Heinemeyer, M. Groll, D. H. Wolf, R. Huber, H. G. Rammensee, and H. Schild. Cleavage motifs of the yeast 20S proteasome β subunits deduced form digests of enolase 1. *Proc. Natl. Acad. Sci. USA*, 95:12504–12509, 1998.
33. A.K. Nussbaum, C. Kuttler, K.P. Hadeler, H.G. Rammensee, and H. Schild. PA-ProC: a prediction algorithm for proteasomal cleavages available on the WWW. *Immunogenetics*, 53:87–94, 2001.

34. R. Z. Orlowski. The role of the ubiquitin-proteasome pathway in apoptosis. *Cell Death Differ.*, 6:303–313, 1999.
35. B. Peters, K. Janek, U. Kuckelkorn, and H. G. Holzhütter. Assessment of proteasomal cleavage probabilities from kinetic analysis of time-dependent product formation. *J. Mol. Biol.*, 318:847–862, 2002.
36. B. Peters, K. Janek, U. Kuckelkorn, and H. G. Holzhutter. Assessment of proteasomal cleavage probabilities from kinetic analysis of time-dependent product formation. *J. Mol. Biol.*, 318:847–862, 2002.
37. J. M. Peters, Z. Cejka, J. R. Harris, J. A. Kleinschmidt, and W. Baumeister. Structural features of the 26S proteasome complex. *J. Mol. Biol.*, 234:932–937, 1993.
38. M. Rechsteiner, C. Realini, and V. Ustrell. The proteasome activator 11S REG (PA28) and class I antigen presentation. *Biochem. J.*, 345:1–15, 2000.
39. C.A. Ross and M.A. Poirier. What is the role of protein aggregation in neurodegeneration? *Nat. Rev. Mol. Cell. Biol.*, 6:891–898, 2005.
40. K. Sakamoto. Ubiquitin-dependent proteolysis: its role in human diseases and the design of therapeutic strategies. *Mol. Genet. Metab.*, 77:44–56, 2002.
41. G. Schmidtke, S. Emch, M. Groettrup, and H. G. Holzhuetter. Evidence for the existence of a non-catalytic modifier site of peptide hydrolysis by the 20S proteasome. *J. Biol. Chem.*, 275:22056–22063, 2000.
42. A. L. Schwartz and A. Ciechanover. The ubiquitin-proteasome pathway and pathogenesis of human diseases. *Annu. Rev. Med.*, 50:57–74, 1999.
43. N. Shastri and S. Schwab. Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. *Annu. Rev. Immunol.*, 20:463–493, 2002.
44. R. L. Stein, F. Melandri, and L. Dick. Kinetic characterization of the chymotrypic activity of the 20S proteasome. *Biochemistry*, 35:3899–3908, 1996.
45. R. Stohwasser, U. Salzmann, J. Giesebrecht, P. M. Kloetzel, and H. G. Holzhuetter. Kinetic evidence for facilitation of peptide channelling by the proteasome activator pa28. *Eur. J. Biochem.*, 267:6221–6230, 2000.
46. T. Tamura, I. Nagy, A. Lupas, F. Lottspeich, Z. Cejka, G. Schoofs, K. Tanaka, R. De Mot, and W. Baumeister. The first characterization of a eubacterial proteasome: the 20S complex of rhodococcus. *Curr. Biol.*, 5:766–774, 1995.
47. S. Tenzer, B. Peters, S. Bulik, O. Schoor, C. Lemmel, M. M. Schatz, P. M. Kloetzel, H. G. Rammensee, H. Schild, and H. G. Holzhuetter. Modeling the MHC class I pathway by combining predictions of proteasomal clavage, TAP transport and MHC class I binding. *CMLS, Cell. Mol. Life Sci.*, 62:1025–1037, 2005.
48. L.G. Verhoef, K. Lindsten, M.G. Masucci, and N.P. Dantuma. Aggregate formation inhibits proteasomal degradation of polyglutamine proteins. *Hum. Mol. Genet.*, 11:2689–2700, 2002.
49. T. Wenzel, C. Eckerskorn, F. Lottspeich, and W. Baumeister. Existence of a molecular ruler in proteasomes suggested by analysis of degradation products. *FEBS Letters*, 349:205–209, 1994.
50. A. Zaikin and J. Kurths. Optimal length transportation hypothesis to model proteasome product size distribution. *Journal of Biological Physics*, 2006. (in press, DOI: 10.1007/s10867-006-9014-z).
51. A. Zaikin, A. Mitra, D. Goldobin, and J. Kurths. Influence of transport rates on the protein degradation by the proteasome. *Biophysical Reviews and Letters*, 2006. (in press).
52. A. Zaikin and T. Pöschel. Peptide-size-dependent active transport in the proteasome. *Europhysics Letters*, 69:725–731, 2005.

53. B. Y. Zeng, A. D. Medhurst, M. Jackson, S. Rose, and P. Jenner. Proteasomal activity in brain differs between species and brain regions and changes with age. *Mech Ageing Dev*, 126:760–766, 2005.