In memory of Professor Marek Cieplak

REVIEW

Protein Dynamics in Tight Tunnels

M. Wojciechowski and M. Chwastyk*

Institute of Physics, Polish Academy of Sciences, al. Lotników 32/46, PL-02668 Warsaw, Poland

Doi: 10.12693/APhysPolA.145.S61 *e-mail: chwastyk@ifpan.edu.pl

We investigate the impact of narrow tunnels, such as the ribosomal exit tunnel and the entrance of the proteasome channel, on the dynamics of proteins with and without knots. Our exploration delves into the potential driving forces behind protein chain movement and their individual significance. Furthermore, within the framework of protein degradation facilitated by the proteasome, we analyze how the presence of knots influences the protein's entry into the proteasome chamber through diverse approaches. This discussion illustrates how molecular dynamics simulations within a coarse-grained structure-based model provide valuable insights into these intricate molecular processes.

topics: molecular dynamics simulations, knots, ribosome, proteasome

1. Introduction

Proteins serve as the workhorses of biology, participating in virtually every aspect of life, from fundamental cellular processes to complex physiological functions. Their versatility and diversity make them integral to the functioning of living organisms. Within a cell, the creation and degradation of proteins are crucial processes. The former involves the extrusion of proteins from the ribosomal exit tunnel, while the latter is related to the translocation of proteins into the chamber within the proteasome, a complex structure containing ATPases associated with diverse cellular activities (AAA+ proteases). Protein degradation plays a pivotal role in maintaining cellular homeostasis by breaking down damaged or unwanted conformations [1–5]. The structure and function of these proteins are tightly regulated to ensure proper turnover and overall cellular function. The crystallographic structures of the ribosome (PDB: 5XY3) [6] and the proteasome (PDB: 7QO3) [7], shown in Fig. 1 (see [8–10]), are provided by the Protein Data Bank (PDB).

Among all proteins, there is a special category that contains knotted conformations in their native state. This topology is not very common, as less than 2% of known structures from the human proteome are knotted [11]. It is extremely important to conduct research on them because knots in proteins have functional significance, and understanding the knotting process can provide insights into various biological processes [12]. Moreover, knots can influence the stability and folding kinetics of proteins, so knowledge about them can be applied to design more stable structures or predict the folding behavior of novel protein sequences [13]. Finally, a very important aspect of examining knotted proteins is their role in neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's diseases, where misfolded proteins aggregate and form insoluble deposits in the brain [14]. Knots in proteins may affect their propensity to aggregate and the structure of the resulting aggregates, potentially influencing disease progression. Therefore, part of our research on protein degradation is devoted to this not very common but super important group of structures.

2. CG Model of ribosome and proteasome

Computer simulations of the processes mentioned above can be carried out using either all-atom or coarse-grained (CG) approaches. Due to the significant conformational changes involved in protein creation and degradation, simulating them at the all-atom level can be challenging. Therefore, we employ a model developed by Professor Marek Cieplak's group over many years [15–17]. In this model, the protein is represented using a structure-based approach, as described in references [15, 18–21], with a chirality potential responsible for maintaining backbone stiffness. Each amino acid is depicted by a single bead positioned at the C_α position, and interactions between beads are based on whether the residues belong to the contact map or not. If two residues are on the contact map, the Lennard-Jones potential is applied to them, with the well depth denoted as ϵ. The contact map is determined using the overlap criterion between all atoms of residues, as observed in the fully folded native state. The characteristic length of these interactions corresponds to distances from the same state. Interactions between the remaining residues consist only of the repulsive

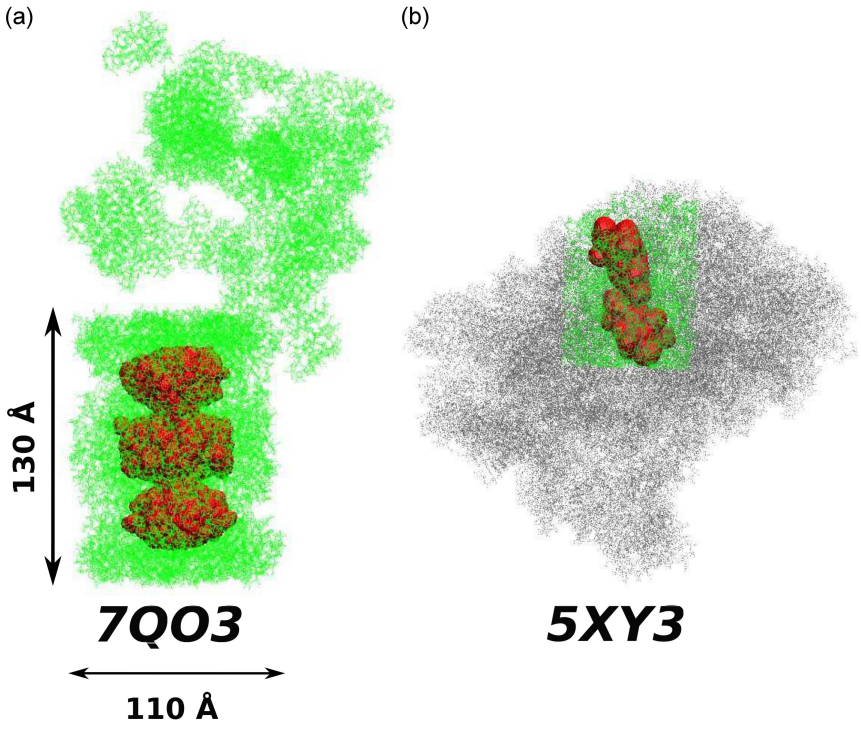

Fig. 1. Representation of the proteasome (panel (a)) and the ribosome (panel (b)) based on PDB crystallographic structures 7QO3 and 5XY3, respectively. In the proteasome, the green color denotes all heavy atoms of its structure. For the ribosome, heavy atoms are colored gray or green, with green representing only the atoms considered in our simulation. Red indicates cavities within both structures detected using the SPACEBALL server [8–10]. Both structures are presented at the same scale.

part of the Lennard-Jones potential, with a characteristic length of 4 Å, beyond which the potential is turned off. Temperature control and the influence of the solvent are introduced by the Langevin thermostat, with the room temperature set at $\approx 0.35\epsilon/k_{\mathrm{B}}$.

The ribosome is a macromolecular machine within a cell responsible for protein synthesis during the process of messenger RNA (mRNA) translation. Ribosomes facilitate the linkage of amino acids in a specific order as directed by the codons present in mRNA molecules. In eukaryotic cells, such as those found in humans, ribosomes are composed of two primary subunits: the small ribosomal subunit (40S) and the large ribosomal subunit (60S). These subunits consist of 2–6 RNA chains and approximately 50 proteins, totaling between 100 000 and 220 000 atoms [22–24]. Within both subunits, ribosomal RNA (rRNA) molecules provide the framework for the ribosome structure and play essential roles in catalyzing the chemical reactions involved in protein synthesis. Protein synthesis occurs at the peptidyl transferase center (PTC), and the newly created chain is directed toward the ribosomal exit tunnel. The size and geometry of this tunnel significantly depend on different domains of life, including bacteria, archaea, and eukarya [25], as well as the specific organism within each domain. The inner walls of the tunnel are rough and highly irregular, featuring several constriction sites. The narrowest constriction, with a radius of approximately 8 Å, is observed in eukaryotes, followed by a slightly wider one, of around 11 Å in radius, in archaeal ribosomes, and even wider in the bacterial case, with an average radius of approximately 15 Å. Due to the high complexity of the ribosome, we simplify its structure to only the cylinder with a radius of 70 Å containing the ribosomal exit tunnel aligned with its longitudinal axis, as illustrated in Fig. 1b, where all heavy atoms of the ribosome are colored gray. The bottom of the cylinder is positioned at PTC, and its length extends to encompass the farthest atoms from PTC. The tunnel under consideration is marked in green. Since the created cylinder is composed of 11 680 atoms, it is still a very complex system, and simulations of all its atoms would be highly inefficient. In this research, we focused on the impact of confinement inside the cylinder, making the interactions between its atoms less critical. Thus, we simplified our system by considering the ribosome structure as rigid. The interactions between the protein and the atoms of the tunnel are purely repulsive, as each ribosomal atom contributes to a soft repulsive potential, truncated at 4 Å, with an amplitude of ϵ. The bottom of the confinement cylinder is modeled as a repulsive wall characterized by the potential $\frac{3\sqrt{3}}{2}\epsilon\,(\sigma_0/z)^9$, where z signifies the distance from the plate and $\sigma_0 = 4 \times 2^{-1/6}$. This wall prevents any backward steps.

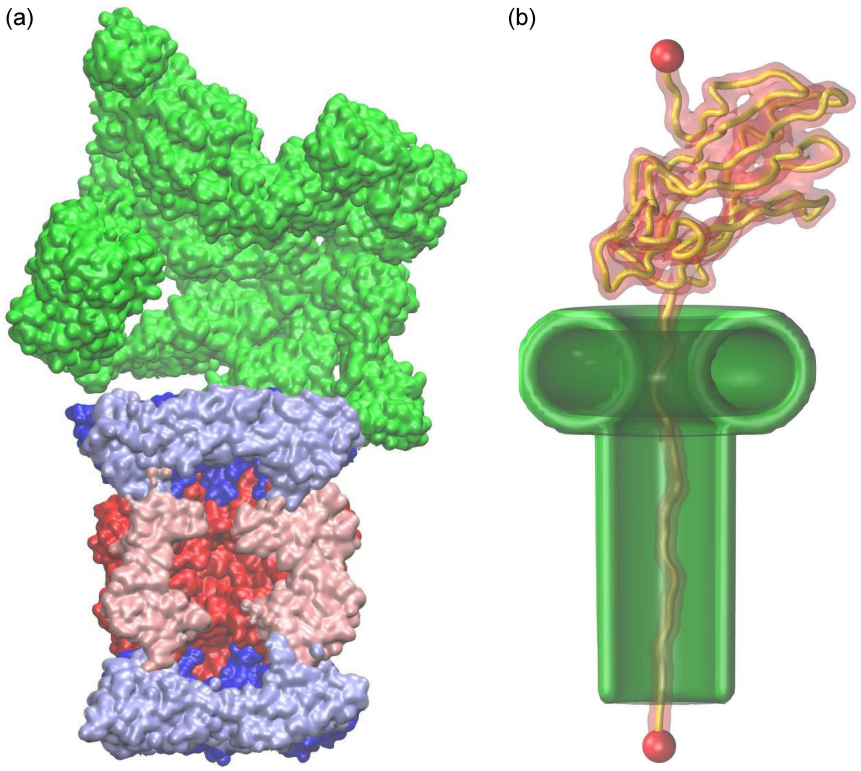

Fig. 2. The cross-section of the crystal structure of the 26S proteasome (PDB: 7QO3) and its theoretical model. The structures' parts closer to the reader are omitted to reveal inner chambers in both cases. In panel (a), the β-rings are colored red or pink, and the α-rings are blue or cyan, with the 19S cap in green. Panel (b) depicts a schematic representation of the entrance to the proteasome. A dragged protein must fit into a hole with a diameter of approximately 14 Å, represented by the torus. Below the torus, there is a straight pipe that keeps the chain unfolded. The entrance and pipe are depicted in green, while the protein is shown in yellow and red tube representation. The N- and C-ends of the chain are represented as red spheres. The pulling force is parallel to the pipe and is attached to the protein terminus (at the bottom of the picture).

On the other hand, we have the proteasome, a highly complicated and well-organized protein complex, as illustrated in Fig. 1a. It consists of multiple subunits arranged into a cylindrical structure resembling a barrel with an open entrance and exit. Protein degradation occurs within the inner chamber of this barrel-like structure. In eukaryotic cells, such as those in humans, the proteasome typically has a 26S structure, as depicted in Fig. 2. This 26S proteasome consists of two main components: the 20S core particle (CP) and the 19S regulatory particle (RP). The 19S regulatory particle recognizes proteins tagged for degradation with ubiquitin molecules, unfolds them, and translocates these target proteins into the central channel of the 20S core particle. Within the complex chamber, the target protein encounters proteolytic active sites that cleave it into smaller peptide fragments. These remaining peptide fragments are subsequently processed and released from the proteasome for recycling or presentation as antigens in immune responses. In our simulations, the presence of the proteasome structure is not directly modeled by the potentials of its individual atoms. Instead, we employ a highly simplified model consisting of a torus as the source of a continuous repulsive part of the Lennard-Jones potential, defined on its surface with a characteristic length of 6 Å. The major radius of the torus, $R_t = 13$ Å, and the minor radius, $r_t = 6$ Å, result in an entrance diameter of ≈ 14 Å, while the diameter of the entrance in a proteasomal crystal structure is 13 Å [26–28]. We enlarged it to accommodate the flexibility of the hole. Below the torus, we introduce a narrow tunnel, but its presence does not originate from the proteasome shape. The tunnel is used solely for technical reasons and prevents the elongated chain from refolding since we do not simulate the degradation process but only protein pulling through the entrance. In our simulations, the chain inside the tunnel is considered degraded, and it no longer impacts the simulation.

Despite the fundamentally different biophysics of protein creation and degradation, one aspect of these processes is shared — the protein must pass through a narrow tunnel or entrance with rugged walls. In the case of the ribosome, this occurs as the protein is pushed out through the ribosomal exit tunnel, while with the proteasome, the chain is pulled inside with the assistance of adenosine triphosphate (ATP) [1, 5].

In computer simulations, the creation of a protein in a ribosome is typically implemented by placing a fully synthesized chain in proximity to the peptidyl transferase center (PTC) and then monitoring the folding process [29–32]. Simulations often model the protein's exit from the ribosome using steered molecular dynamics with a constant pulling speed applied to the N-terminus [29]. Alternatively, some simulations apply a constant force to more accurately replicate the natural process [30]. In our simulations, we demonstrate that the protein chain can exit the ribosomal tunnel without external manipulation. We employ a sequential growth method, where each amino acid emerges at specific time intervals after the previous one is synthesized. As mRNA translation proceeds from the 5' to 3' ends, proteins are synthesized from the N-terminus to the C-terminus, causing the N-terminus to emerge first. Our approach resembles the one described by P.T. Bui and T.X. Hoang [33], with the distinction that our repulsive potential accounts for all heavy atoms of the ribosome part considered in simulations, and any backward motion is prevented by repulsion from the lower tunnel wall. Additionally, in our case, the growth process is implemented in a quasi-continuous manner.

Protein degradation, on the other hand, involves pulling the chain into the proteasome, and our focus is exclusively on this pulling process. Unlike the protein dynamics in the ribosomal exit tunnel, achieving this requires the application of a dragging force. The simplest approach is to drag the chain into the proteasome and through the torus hole at a constant speed. While this approach is not realistic, it serves to illustrate the difference between AFM-like protein stretching (AFM — atomic force microscopy) [18, 19] and proteasome-assisted unfolding. To implement a more realistic scheme, we start pulling with a constant force, which can be implemented either continuously or periodically. In the latter case, the force is applied for a specified time, approximately 4.5 μs, to either the C- or N-terminal chain ends, depending on the pulling scenario, and then turned off for an equal amount of time, allowing the protein to retract. This reflects the generation of force upon delivery of ATP (non-continuous). The final scheme represents a "ratchet-like mechanism" involving pulling for a maximum of 4.5 μs (or pulling 1 nm of chain), followed by blocking the protein's retraction during the absence of force. This mechanism is supported by biological proteasome action, where certain parts of the proteasome undergo bending upon ATP delivery, generating the force capable of dragging a small portion of the chain [34]. Subsequently, the protein chain is prevented from retracting and awaits the next ATP cycle, making periodic force application closely mimic biological evidence.

The aforementioned simplified description of on-ribosome protein creation and protein degradation by the proteasome could be improved by using more complex and detailed approaches. However, in our research, we deal with large conformational changes in proteins that occur within seconds, making simulations of such processes very expensive and time-consuming. The advantage of simulations within a coarse-grained model is that, thanks to their efficiency, a large number of different trajectories for a given process can be conducted, and the final result is obtained by averaging the considered quantities.

3. On-ribosome protein folding

We initiate simulations of protein chain creation by the ribosome with a single residue (the N-terminus) placed at the PTC. The next residue appears within a specified time, denoted as t_w, referred to as the waiting time. Determining its value is not straightforward, but in our previous work [35], we demonstrated that a waiting time of 5 000 τ is sufficient, and times longer than this value do not significantly affect the outcomes. Here, τ is equal to 1 ns, which is the characteristic time in our model. This means that in our simulations, protein synthesis occurs $\approx$ 4–5 orders of magnitude faster compared to biological systems. Based on the previous results [35], we assume that this waiting time value is sufficient, and thanks to the acceleration of adding new residues, our simulations are much more efficient, allowing us to examine large conformational changes in the newly created protein. As mentioned in Sect. 2, the process of the protein leaving the ribosomal exit tunnel can be implemented in various ways. In our model and simulations, our intention was to closely simulate the biological process. This means that there is no external force aiding the squeezing of the protein N-terminus through the rough channel. The rigid walls are also unhelpful in this process. The only forces acting on the N-terminus of the newly created protein facilitating its travel to the ribosome exit are the repulsive forces from neighboring protein residues, the repulsive forces from the bottom wall, and the atoms comprising the ribosome structure. In the case of smooth walls, the protein chain would easily slide toward the ribosomal exit tunnel. The walls of the real ribosomal exit tunnel are rough, and the movement of the first residue is not easy because it can become jammed in the alleys or nooks of the walls. Such a situation is presented in panel (b) in Fig. 3, which shows the distance of the N- and C-terminus from the PTC in function of time. The graph illustrates that the N-terminus became jammed $\approx$ 60 Å from the PTC, and all of the protein's residues became crowded into the space between this point and the PTC. This scenario is unrealistic because proteins do not typically jam during their creation in real situations, and it can be considered an artifact of the model. Nevertheless, this observation highlights the significant impact of irregularities in the exit tunnel walls on protein dynamics.

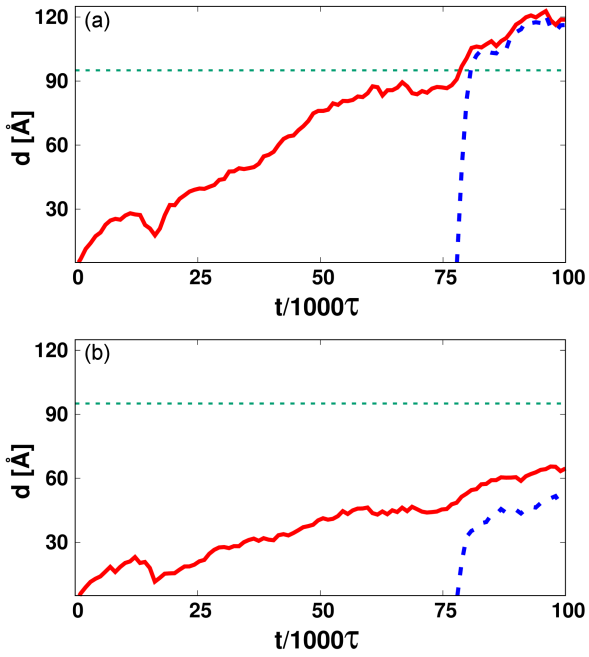

Fig. 3. The distances between the PTC and the N-terminus of the 1J85 protein chain (solid red line) or its C-terminus (dotted blue line) plotted as a function of time during the simulation within the ribosomal exit tunnel. Panel (a) illustrates the scenario where the N-terminus has reached the exit from the ribosomal structure, denoted by the green dotted line, while dragging the C-terminus outward. Panel (b) depicts the situation where the N-terminus was jammed within the tunnel. In both cases, the simulations were conducted at room temperature with a waiting time of $t_w = 100$ ns.

Most often, when the t_w is long enough, the N-terminus can find its way to reach the exit from the ribosome, as illustrated in Fig. 3a. Here arises a question regarding the mechanisms that play a crucial role in protein dynamics within the ribosomal exit tunnel. In our previous research [25], we proposed that the protein's movement toward the outside of the tunnel starts at the PTC and is primarily influenced by diffusion, interactions with the tunnel walls, and the increase in entropy associated with the escape. However, we did not investigate which of these mechanisms is crucial. By looking at Fig. 3a, we see that there is a rapid increase in the N-terminus' distance from the PTC when it reaches the top of the tunnel, denoted by the green dotted line. Above the outer surface of the ribosome, the protein began to fold, native contacts began to form, and there was a decrease in the protein's potential energy. As a consequence, the C-terminus of the created protein started to be pulled by the folded protein and traveled very quickly, without any jamming, outside the ribosome. This observation suggests that the change in potential energy is crucial for the protein's exit from the tunnel.

4. Protein degradation by the proteasome

Simulating protein degradation within the proteasome presents several challenges. As the proteasome employs ATP to pull proteins periodically and restricts backward motions, the most realistic simulation scenario involves periodic force pulling with a "ratchet-like mechanism" [34]. Here, we examine several pulling scenarios, ranging from the simplest to the most advanced, to assess pulling efficiency in each case.

In our study of protein behavior within the proteasome, we have considered three primary methods of protein stretching. The first method, referred to as "type AFM," involves standard AFM stretching and serves as a template and reference for proteasome-aided unfolding. In the second method, we pull the protein by its ends in the presence of a proteasome model, either at the C-end or N-end, while the other end remains free. This method of pulling is called "type I." Considering stretching from both termini is justified by the fact that the degradation mechanism is preceded by the protein marking with special tags, which can be attached to either the C- or N-terminal chain ends. The last type of pulling, referred to as "type II," involves pulling by one terminus while the other remains attached to its original position. All of these pulling types can be realized by different scenarios. The simplest one is constant speed pulling, in which we pull the protein end at a constant speed. This type of protein pulling generates a characteristic AFM-like force graph with peaks, as depicted in the top parts of each panel in Fig. 4. Each peak on the force-displacement graph corresponds to the breaking of contacts within a specific part of the protein structure and represents the protein's resistance to mechanical stress. The most resilient part of the structure yields the highest force peak, and its height, denoted as $F_{\max}$, can be used to characterize the protein mechanostability. Both types of proteasome-assisted pulling with constant speed are comparable to standard AFM stretching experiments and, as demonstrated in our previous research [36], they facilitate the pulling process. Lower forces were measured in most of the proteins considered. However, in the case of barnase (PDB: 1BNR) and titin (PDB: 1TIT), slightly higher force values were observed when pulling from the C- or N-termini compared to standard AFM stretching. This suggests that the presence of the torus facilitates the breaking of contacts and the unfolding of the protein chain, but in our simulations, we also observed the impact of the torus presence on knotted protein dynamics and their degradation.

Since the entrance to the proteasome is narrower than the average width of a knot in a knotted protein, the knot can block the entrance, making the proteasome useless [36]. However, in our simulations, we have demonstrated that the presence of the proteasome can untie deep knots, which are

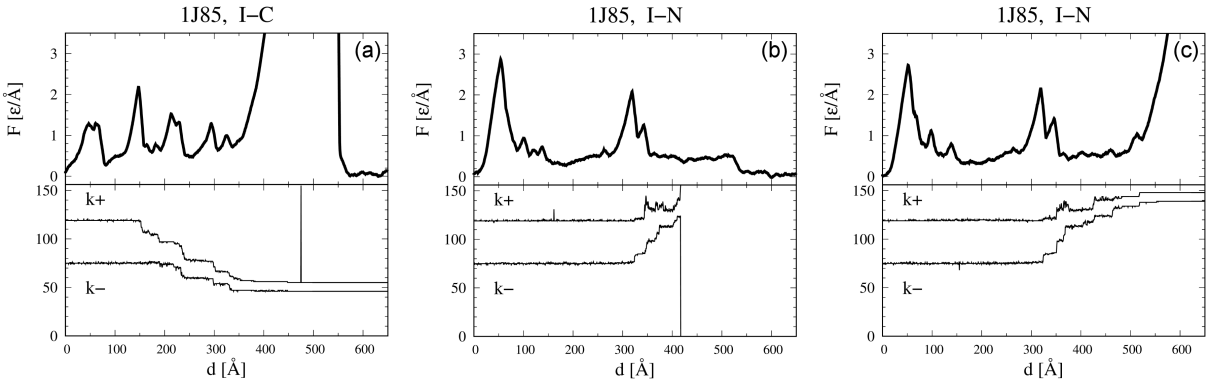

Fig. 4. The simulation results of the deeply knotted protein 1J85 stretching. In both AFM-like stretching and "type II" pulling, the knot almost always tightens. Panel (b) illustrates "type I" pulling, where the knot may slide along the chain and become untightened. Alternatively, the knot may slide down and tighten, similar to AFM stretching, as depicted in panels (a) and (c). The top part of each panel displays force versus distance, while the bottom ones show the position of the knot in the sequence. The knot is considered untied when its position along the protein chain is one or it equals the length of the chain.

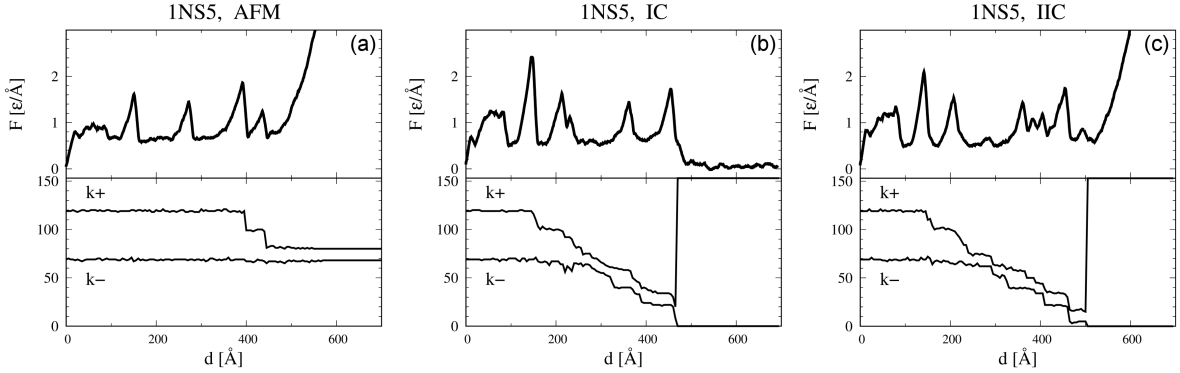

Fig. 5. This figure is similar to Fig. 4. All panels depict the simulation results of the deeply knotted protein 1NS5. Here, we observe knot untying only for C-end pulling (panels (b) and (c)). The N-end simulations only show knot tightening (panel (a)).

defined as knots where the termini are distant from the protein ends [37]. In our research, we considered two deeply knotted proteins: YibK methyltransferase from *Haemophilus influenzae* (PDB: 1J85) and YBEA from *E. coli* (PDB: 1NS5). In the case of the 1J85 protein, the ends of the knot, which consists of 156 residues, are located at residues 75 and 120 in the protein's native state. For 1NS5 (153 residues), the positions of the knot are at residues 69 and 119.

Figure 4 displays the results of our simulations of the degradation of the 1J85 protein with a knot. In this case, we observe that the AFM experiment leads to the tightening of the knot ends, as expected. Knot untying was observed only for "type I-N" simulations, and one trajectory resulting in this behavior is presented in Fig. 4b. Since the protein chain is not smooth, geometric constraints can arise, allowing knot tightening. We observed such trajectories as well, as shown in panels (a) and (c) in Fig. 4. Panel (a) illustrates how the knot blocks the entrance, but if the dragging force is very large, the knot may be pushed through the entrance. However, this is non-physical, as the value of the force is unrealistic ($F \sim 17\epsilon$ Å^{-1}). This occurrence was not observed in every trajectory; in some trajectories, the knot remained at the entrance even at larger forces ($F > 20\epsilon$ Å^{-1}). For other types of pulling, we observe only jamming.

The same observations were made in the case of another deeply knotted protein, 1NS5. The results of the AFM stretching of this protein are presented in Fig. 5a. The outcomes of the stretching using the method called "type I" are shown in Fig. 5b. Interestingly, when stretching the 1NS5 protein using the "type II" method, we also observed knot untying. In this method, we grasp the opposite end of the protein, making knot untying less straightforward. In the simplest scheme, the opposite end of the protein fluctuates during pulling and may accidentally pass out from the loop, thus untying the knot. However, in "type II," such a situation is not possible. Therefore, the only method to untie the knot is to move the entire loop to pass around the immobilized end.

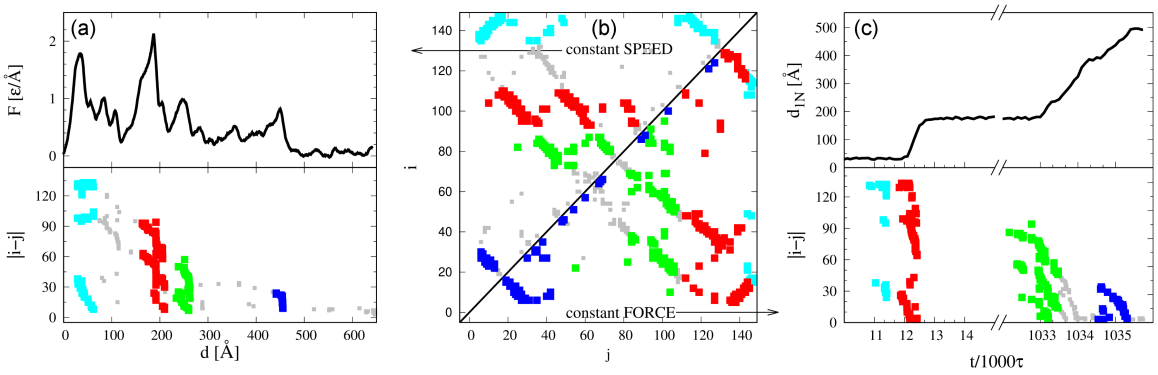

Fig. 6. The simulation results of protein 1AOH. Panel (a) presents an example dataset of constant speed pulling ("type I-C"). The protein is pulled into the proteasome by one end at a constant speed, resisting dragging with the depicted force. The largest force peak reaches approximately $\sim 2.3\epsilon$ Å^{-1}. The top part of panel (a) displays the force graph as a function of distance. Panel (c) illustrates an example of constant force pulling ($F = 1.6\epsilon$ Å^{-1}) ("type I-C"). The protein is pulled into the proteasome with a constant force while monitoring its distance from the proteasome entrance. The top part of this panel displays the distance graph as a function of time. Panel (b) presents the contact map (amino acid i versus j) for both simulation types. The breakings of specific contacts during the simulation of protein stretching with constant speed (panel (a)) and constant force (panel (c)) are indicated at the bottom of both panels by different colors corresponding to the colors of the contacts marked on the contact map presented in panel (b).

Another interesting aspect is the role of the pulling direction (C- versus N-end). When dragging from the C-end ("type I-C" and "type II-C"), we observed untying in 32% and 29% of cases, respectively. However, in the case of N-end dragging, untying is almost impossible (4% for "type I-N" and 0% for "type II-N"). Therefore, the direction of pulling (or knot sliding) may play an important role. Untying a knot in "type II" is more complicated since the other end is restrained, but it is still possible.

Interestingly, we observe in our simulations that the presence of the proteasome does not affect the measured mechanostability of knotted proteins, in contrast to similar simulations of proteins without any knots. This shows that the presence of knots on the protein chain affects the functionality of the proteasome, even leading to the proteasome becoming inactive when the entrance is blocked by the knotted protein.

According to our understanding of the mechanism of protein degradation by the proteasome, pulling the protein at a constant speed does not mimic the behavior of a protein in the presence of the proteasome. As explained in Sect. 2, a more realistic approach is pulling with a constant force. In this case, the protein chain can also be stretched using pulling types I or II. The comparison between the methods of protein pulling with constant speed and constant force is presented in Fig. 6.

Pulling with a constant force, apart from better mimicking the real mechanism, allows for the use of a lower force and, consequently, reduces the work needed to unravel the protein. This is possible thanks to the support from random forces provided by the Langevin thermostat. These forces facilitate protein unfolding or the untying of knots because they can resolve small steric clashes that may occur during protein pulling. In this case, the constant force only determines the direction of the chain. What is particularly interesting is how to compare the results obtained by pulling with a constant force and constant speed. One possible approach is well described in the literature [38] — a number of stretching simulations were conducted with different values of force, and the average pulling speed was calculated for each case. Protein mechanostability is defined as the extrapolation of the force values to a pulling speed of 80 residues/s, which is the pulling speed in the real proteasome [34]. Figure 6 shows different shapes of the graphs obtained during pulling with a constant force (panel (c)) or constant speed (panel (a)). The graph obtained during constant force pulling presents the distance L from the entrance to the proteasome as a function of time. The graph resembles multiple stair steps with different widths. Each step is associated with the breaking of a particular part of the structure. The contacts broken during the appearance of the particular steps are marked at the bottom part in panel (c). The time, $\tau_k \sim e^{\Delta E_k}$, necessary to break a particular group of contacts is related to the energy barrier, ΔE_k, that must be crossed [39]. The longest time is connected to the highest energy barrier.

Protein pulling with a constant force is also challenging. During such simulations, it is possible that steric clashes are so strong that thermal fluctuations cannot remove them. Additionally, knots in the protein chain also pose problems. In our simulations with continuous, constant force, we observed the proteasome becoming jammed by knots in proteins. For low pulling forces ($F \leq 1.6\epsilon$ Å^{-1}), only

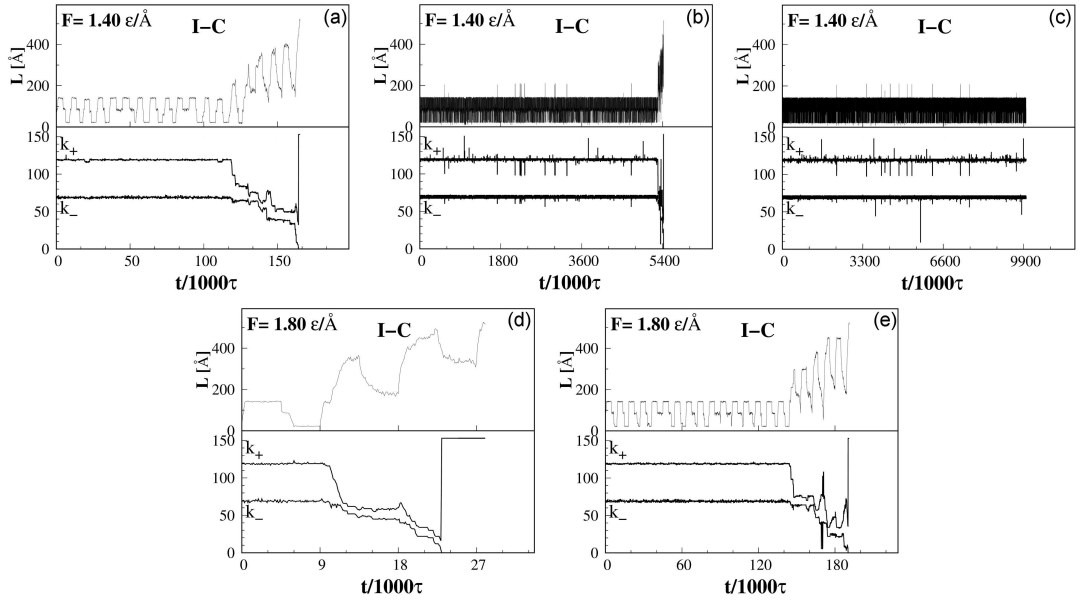

Fig. 7. The simulation results for protein 1NS5 within the periodic force model. Panels (a–c) present results obtained with a pulling force of $F = 1.40\epsilon$ Å^{-1}, while panels (d–e) present results obtained for $F = 1.80\epsilon$ $^{-1}$. Panels (a), (b), and (c) depict fast protein untying, a long simulation completed with knot untying, and a long unsuccessful simulation, respectively. Successful unknotting simulations are shown by a short simulation (panel (d)) and a long simulation (panel (e)).

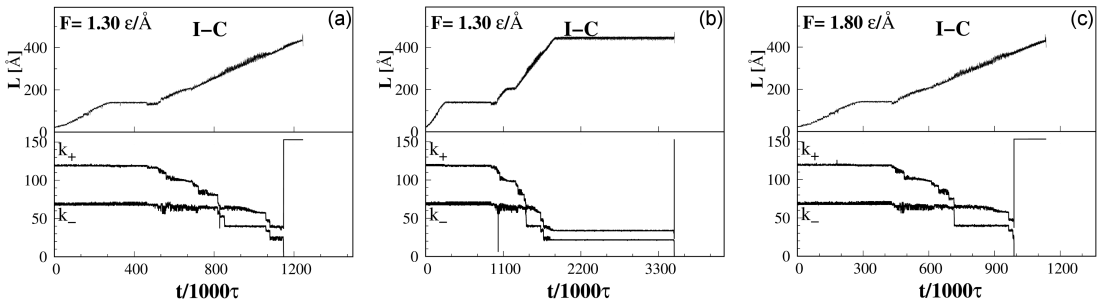

Fig. 8. The successful simulation results for 1NS5 within the ratchet force model. Panels (a), (b), and (c) depict fast pulling with a force of $F = 1.30$ ϵ/Å, slow pulling with the same force, and fast pulling again with a force of $F = 1.80$ ϵ $^{-1}$, respectively.

a portion of the protein without any knots was pulled into the chamber, and the pulled protein was halted at the position where the knot began. This indicates that the knot remained at its initial position. For higher forces, we observed the tightening of knots, but they did not move along the chain. This observation suggests that knotted proteins can block the proteasome, which cannot be true because in living cells, knotted proteins are usually degraded. To avoid this problem and achieve an even more realistic situation, we switch from the continuous application of constant force to a simulation with a constant force applied at given time intervals. Figure 7 shows five trajectories of protein pulling with different forces.

Applying the force in time intervals is also challenging because, in the absence of the force, we observe a backward movement of the chain, similar to the situation when part of the protein folds outside of the ribosome. However, this backward movement can be helpful in removing potential geometrical constraints. Repeatedly applying the force for a specific duration and then allowing the chain to relax can facilitate chain reorientation, increasing the possibility of the chain passing through the entrance. The number of times the force needs to be applied depends strongly on the protein, pulling force, and force application duration. This approach is also interesting because, as we showed in Fig. 7, the periodic force application can even enable the degradation of a knotted protein, in contrast to continuous force pulling. Figure 7 also depicts the varying times needed to untie knots in the knotted protein 1NS5, depending on the mentioned conditions. This process is very inefficient due to the return movements of the protein chain.

To prevent this behavior, we have implemented a "ratchet-like mechanism" that prevents the protein from moving backward when the force is switched off, mimicking the waiting for a new ATP portion in a real situation. Moreover, we do not specify the pulling period based on time, but rather consider the length of the pulled chain. We stop the pulling process when $d_{step} = 3.8$ Åof the chain is drawn into the proteasome. The size d_{step} was chosen to be comparable to the average size of a single amino acid. It is worth mentioning that if a geometrical clash occurs and no chain movement is possible, we pause the pulling process for 4.5 μs, allowing the Langevin thermostat to address the issue for another 4.5 μs before applying the force again. Importantly, no protein backward movement is allowed during this time. The advantage of this mechanism, compared to simple pulling with a constant force, is that we prevent pulling the same part of the chain many times, which can occur when it escapes during the absence of force. Avoiding multiple pulls of the same part of the protein chain using the "ratchet-like mechanism" is also favorable because periodically applying a small force has a lower likelihood of causing strong steric clashes. The trajectories from the simulation of the 1NS5 knotted protein, as presented in Fig. 8, exhibit this behavior. Moreover, we can see that this approach is much more stable, as the times necessary for protein unfolding are comparable regardless of the simulation parameters. The "ratchet-like mechanism" is considered the most realistic approach for pulling proteins into the proteasome.

5. Conclusions

In our simulations, we have demonstrated that the reduction in the protein's potential energy is crucial for the process of a protein exiting the ribosomal tunnel. As the partially folded protein emerges from the ribosome, it generates a pulling force on the remaining portion within the tunnel. While interactions between the protein chain and the ribosomal tunnel walls, along with diffusion, are important, their significance for the protein's movement outside the tunnel is secondary. This was confirmed by simulations of pulling the protein into the proteasome chamber without the "ratchet-like mechanism," resulting in the protein being retracted from the proteasome. This backward movement also correlates with a decrease in the protein's potential energy. It is worth noting that further investigation is needed, as these simplified models may not encompass all aspects of the processes under consideration.

Furthermore, our simulations lend support to the hypothesis regarding the directionality of knot tightening and untying. Given the critical role of the ribosome structure in simulations of knotting the 1J85 protein by pulling the C-terminus across a specially created loop, it suggests that knot degradation should be more feasible from the C-end. This implies that pulling the 1J85 protein from the N-terminus enables knot sliding and untying. This finding aligns with our simulation results using the constant-speed scheme for 1J85. The untying of the knot and successful engulfing of the protein were achievable only with the "type I-N" scheme.

Acknowledgments

The authors acknowledge Professor Marek Cieplak for encouraging the investigation of protein dynamics, explored using coarse-grained models. This research has received support from the National Science Centre (NCN), Poland, under grant No. 2018/31/B/NZ1/00047 and the European H2020 FETOPEN-RIA-2019-01 grant PathoGel-Trap No. 899616. The computer resources were supported by the PL-GRID infrastructure.

References

[1] F. Türker, E.K. Cook, S.S. Margolis, *Cell Chem. Biol.* **28**, 903 (2021).

[2] G.A. Collins, A.L. Goldberg, *Cell* **169**, 792 (2017).

[3] M. Bochtler, L. Ditzel, M. Groll, C. Hartmann, R. Huber, *Annu. Rev. Biophys. Biomol. Struct.* **28**, 295 (1999).

[4] S. Gottesman, *Annu. Rev. Genet.* **30**, 465 (1996).

[5] J.A.M. Bard, E.A. Goodall, E.R. Greene, E. Jonsson, K.C. Dong, A. Martin, *Annu. Rev. Biochem.* **87**, 697 (2018).

[6] Z. Li, Q. Guo, L. Zheng, Y. Ji, Y.T. Xie, D.H. Lai, Z.R. Lun, X. Suo, N. Gao, *Cell Res.* **27**, 1275 (2017).

[7] K.Y.S. Hung, S. Klumpe, M.R. Eisele et al., *Nat. Commun.* **13**, 838 (2022).

[8] M. Chwastyk, M. Jaskolski, M. Cieplak, *FEBS J.* **281**, 416 (2014).

[9] M. Chwastyk, M. Jaskolski, M. Cieplak, *Proteins* **84**, 1275 (2016).

[10] M. Chwastyk, E.A. Panek, J. Malinowski, M. Jaskólski, M. Cieplak, *Front. Mol. Biosci.* **7**, 591381 (2020).

[11] A.P. Perlinska, W.H. Niemyska, B.A. Gren, M. Bukowicki, S. Nowakowski, P. Rubach, J.I. Sulkowska, *Prot. Sci.* **32**, e4631 (2023).

[12] P. Virnau, L.A. Mirny, M. Kardar, *PLoS Comput. Biol.* **15**, e122 (2006).

[13] J.I. Sułkowska, P. Sułkowski, P. Szymczak, M. Cieplak, *Proc. Natl. Acad. Sci.* **105**, 19714 (2008).

[14] M. Cieplak, M. Chwastyk, Ł. Mioduszewski, B.R.H. de Aquino, *Prog. Mol. Biol. Transl. Sci.* **174**, 79 (2020).

[15] M. Cieplak, T.X. Hoang, M.O. Robbins, *Proteins* **49**, 114 (2002).

[16] M. Sikora, J.I. Sułkowska, M. Cieplak, *Plos Comput. Biol.* **5**, e1000547 (2009).

[17] J.I. Sulkowska, M. Cieplak, *Biophys. J.* **95**, 3174 (2008).

[18] M. Chwastyk, A. Galera-Prat, M. Sikora, Á. Gómez-Sicilia, M. Carrión-Vázquez, M. Cieplak, *Proteins* **82**, 717 (2014).

[19] M. Gunnoo, P.-A. Cazade, A. Orlowski, M. Chwastyk, H. Liu, D.T. Ta, M. Cieplak, M. Nashde, D. Thompson, *Phys. Chem. Chem. Phys.* **20**, 22674 (2018).

[20] Y.N. Zhao, M. Chwastyk, M. Cieplak, *Sci. Rep.* **7**, 39851 (2017).

[21] M. Chwastyk, A.P. Bernaola, M. Cieplak, *Phys. Biol.* **12**, 046002 (2015).

[22] R. Young, H. Bremer, *Biochem. J.* **160**, 185 (1976).

[23] K. Boström, M. Wettesten, J. Borén, G. Bondjers, O. Wiklund, S.O. Olofsson, *J. Biol. Chem.* **261**, 13800 (1986).

[24] N.T. Ingolia, L.F. Lareau, J.S. Weissman, *Cell* **147**, 789 (2011).

[25] M. Chwastyk, M. Cieplak, *Front Mol. Biosci.* **8**, 692230 (2021).

[26] A.M. Ruschak, T.L. Religa, S. Breuer, S. Witt, L.E. Kay, *Nature* **467**, 868 (2010).

[27] F. Zhang, M. Hu, G. Tian, P. Zhang, D. Finley, P.D. Jeffrey, Y. Shi, *Mol. Cell* **34**, 473 (2009).

[28] M. Groll, L. Ditzel, J. Löwe, D. Stock, M. Bochtler, H.D. Bartunik, R. Huber, *Nature* **386**, 463 (1997).

[29] D.A. Nissley, Q.V. Vu, F. Trovato, N. Ahmed, Y. Jiang, M.S. Li, E.P. O'Brien, *J. Am. Chem. Soc.* **142**, 6103 (2020).

[30] P. Dabrowski-Tumanski, M. Piejko, S. Niewieczerzal, A. Stasiak, J.I. Sulkowska, *J. Phys. Chem. B* **122**, 11616 (2018).

[31] J. Frank, R.L. Gonzalez Jr., *Annu. Rev. Biochem.* **79**, 381 (2010).

[32] A.H. Elcock, *Plos Comput. Biol.* **2**, e98 (2006).

[33] P.T. Bui, T.X. Hoang, *J. Chem. Phys.* **153**, 045105 (2020).

[34] R.A. Maillard, G. Chistol, M. Sen, M. Righini, J. Tan, C.M. Kaiser, C. Hodges, A. Martin, C. Bustamante, *Cell* **145**, 459 (2011).

[35] M. Chwastyk, M. Cieplak, *J. Phys. Chem. B* **124**, 11 (2020).

[36] M. Wojciechowski, P. Szymczak, M. Carrión-Vázquez, M. Cieplak, *Biophys. J.* **107**, 1661 (2014).

[37] Y.I. Zhao, M. Chwastyk, M. Cieplak, *J. Chem. Phys.* **146**, 225102 (2017).

[38] M. Wojciechowski, Á. Gómez-Sicilia, M. Carrión-Vázquez, M. Cieplak, *Mol. BioSyst.* **12**, 2700 (1016).

[39] J.I. Steinfeld, J.S. Francisco, W.L. Hase, *Chemical Kinetics and Dynamics*, 2nd ed., Prentice Hall, 1999.