Nanoscale

PAPER

View Article Online
View Journal | View Issue

Cite this: *Nanoscale*, 2019, **11**, 9920

Translocation intermediates of ubiquitin through an α-hemolysin nanopore: implications for detection of post-translational modifications†

Emma Letizia Bonome,[a] Fabio Cecconi [b] and Mauro Chinappi *[c]

Received 28th December 2018,
Accepted 15th April 2019

DOI: 10.1039/c8nr10492a

rsc.li/nanoscale

Nanopore based sensors constitute a promising approach to single molecule protein characterization being able, in principle, to detect sequences, structural elements and folding states of proteins and polypeptide chains. In narrow nanopores, one of the open issues concerns the coupling between unfolding and translocation. Here, we studied the ubiquitin translocation in an α-hemolysin nanopore, the most widely used pore for nanopore sensing, *via* all-atom molecular dynamics simulations. We completely characterize the co-translocational unfolding pathway finding that robust translocation intermediates are associated with the rearrangement of secondary structural elements, as also confirmed by coarse grained simulations. An interesting recurrent pattern is the clogging of the α-hemolysin constriction by an N-terminal β-hairpin. This region of ubiquitin is the target of several post-translational modifications. We propose a strategy to detect post-translational modifications at the N-terminal using the α-hemolysin nanopore based on the comparison of the co-translocational unfolding signals associated with modified and unmodified proteins.

Introduction

Nanopore devices are emerging as enabling technologies to achieve reliable, fast and point-of-care protocols for fundamental biology and biomedical applications. One of the main features of nanopore-based sensing devices is their capability of operating at a single molecule level, *i.e.* the signal is due to the interaction of a single molecule with the nanopore. In the proteomic field, nanopores have been proposed for a number of different tasks[51] such as analysis of proteins and peptide folding,[40,43,46,58,68] determination of conformational changes,[62,65] interaction with DNA,[52] aggregation of proteins and amyloids[25,26] and discrimination between different peptide chains.[18,29,34] Moreover, the capability of specific molecules to alter the translocation probabilities of peptides has been used to propose a nanopore based sensor for the detection of uranil ions.[55] Interestingly, recent studies show that even modifications of few amino acids can be distinguished[5,48] and that graphene nanogap devices can potentially sense single peptide bonds, suggesting possible applications for protein sequencing.[19]

Beside the larger variety of pores employed (solid-state[14,26,27,36,50,65] or biological[3,48,52,62]) and the number of different sensing strategies, there are three fundamental issues in nanopore based devices:[19] (i) capture, the molecule in the bulk of the reservoir has to move in the pore region and engage the pore, (ii) residence, the molecule has to reside in the pore for a time interval long enough to record a stable signal and (iii) distinguishability, different molecules (or different conformations of the same molecule) have to leave different signals.

In the context of protein and peptide sensing, all these three issues are challenging. A combination of electrophoresis and electroosmotic contribution is potentially able to control the capture and translocation speed,[4,10,11,13,29,42] while, concerning distinguishability, the univocal correspondence between signals and protein conformations inside the pore is not straightforward.

Few years ago, Rodriguez-Larrea and Bayley[53,54] proposed a generic strategy to import and translocate folded proteins independent of their charge. They covalently added an oligonucleotide to one of the two protein terminals. This enables a voltage-driven unidirectional movement through a nanopore. The narrower section of the αHL pore is smaller than the typical size of folded globular proteins, hence, protein translocation is necessarily associated with its unfolding. This co-

[a]*Dipartimento di Ingegneria Meccanica e Aerospaziale Sapienza Università di Roma, Roma, 00185, Italy*
[b]*CNR-Istituto dei Sistemi Complessi UoS Sapienza, Via dei Taurini 19, Roma, 00185, Italy*
[c]*Dipartimento di Ingegneria Industriale, Università di Roma Tor Vergata, Roma, 00133, Italy. E-mail: mauro.chinappi@uniroma2.it*

†Electronic supplementary information (ESI) available. See DOI: 10.1039/C8NR10492A

translocational unfolding process resulted in a multilevel current signal that can be interpreted as a multistep translocation similar to the one previously observed in coarse grained simulations.[6,30,61] All-atom molecular dynamics simulations in αHL and graphene nanopores showed that the multistep translocation is mainly ruled by the native state features of the translocating proteins.[9,23] Interestingly, Rosen *et al.*[56] showed that the presence of phosphorylation sites alters the current level associated with the co-translocational unfolding pathway of Thioredoxin. Another promising generic strategy for controlling translocation of the folded protein was proposed by Nivala and Akeson[44,45] where the translocation was carried out by the action of the AAA + unfoldase ClpX protein that acts on a histidine tag added to the protein terminal.

These promising results leave open several interesting questions. From the biophysical point of view, the link between the protein native state and the multistep co-translocational unfolding pathway needs to be clarified, supporting the current picture with additional cases concerning different proteins. This should allow assessing which features are generic and which are specific of protein–pore interactions. On the biomedical application side, it is not clear how to exploit multistep translocation for specific biochemical analysis. Indeed, although multistep translocation does not provide a full control of the translocation speed, its high reproducibility suggests that the differences in the translocation pathways can be used to detect specific protein features. In this framework, coarse grained[2,16,17,20,30,49,61] and all-atom[9,23,38,67] simulations have been shown to be reliable tools to investigate the translocation unfolding and translocation pathways.

In this work, we present a computational study on the co-translocation unfolding of ubiquitin (Ubq) through the α-hemolysin nanopore (αHL). We explored both pulling directions, *i.e.* Ubq is pulled through the αHL nanopore from either its N- or C-terminus. We find that, in both cases, the co-translocational unfolding is a multistep process where Ubq gets stuck along the pore in long-lived states, termed translocation intermediates (or stalls). The more evident ones are associated with specific structural rearrangements of Ubq secondary and tertiary structures. These translocation intermediates are highly reproducible among the different simulation replicas. Other translocation intermediates are associated with contingent interactions between the unfolded region of the Ubq and the pore. These stalls are shorter and less reproducible among the different replicas. Interestingly, a recurrent pattern in the co-translocational unfolding is the clogging of the αHL constriction by an N-terminal β-hairpin. This region can be the target of several post-translational modifications (PTMs). An approach to exploit the co-translocational unfolding pathway to characterize the PTMs is discussed.

Methodology

Simulations were performed using the NAMD software.[47] The CHARMM36 force field[12,63] was employed to model lipids, proteins and TIP3P water molecules.[35] NBFIX corrections were applied for ions.[39] The particle mesh Ewald (PME) summation method is used for the electrostatics.[8]

System set-up

The system consists of two main elements: (i) the αHL channel embedded in a lipid membrane and (ii) the ubiquitin protein (Ubq). The αHL structure (pdb id: 7AHL)[59] was downloaded from the OPM database.[37] The membrane is constituted of a phosphatidylcholine (POPC) lipid bilayer. The system is built and equilibrated following protocols already reported in the literature[1,22] and described in our previous work.[10] The final configuration of the αHL-membrane system, including water and ions (2 M KCl), is a periodic box of dimensions L_{x1} = 126.49 Å, L_{y1} = 124.61 Å and L_{z1} = 217.5 Å and the number of atoms is ~370 000. Ubiquitin (pdb_id: 1UBQ[64]) is solvated and ionized (2 M KCl) using the VMD software[31] and independently pre-equilibrated. The Ubq box dimensions after a constant area NPT equilibration are $L_{x2} = L_{x1}$, $L_{y2} = L_{y1}$ and L_{z2} = 233 Å. Then, the two systems were merged and equilibrated for additional 2 ns (flexible cell, constant area NPT, T = 310 K, P = 1 atm). The resulting periodic box has dimensions L_{x3} = 126.49 Å, L_{y3} = 124.61 Å and L_{z3} = 435 Å. The relatively large size in z allows the fully stretched Ubq to be accommodated in the z-direction. The full system is constituted of 681 519 atoms.

Dedicated steered molecular dynamics simulations were employed to bring the Ubq at the pore vestibule entrance (*cis* side). In particular, the protein terminus (C or N) is placed at ~10 Å above the αHL *cis* entrance before translocation runs. Fig. 1B shows the system configuration at the end of the equilibration for the C-pulling simulation.

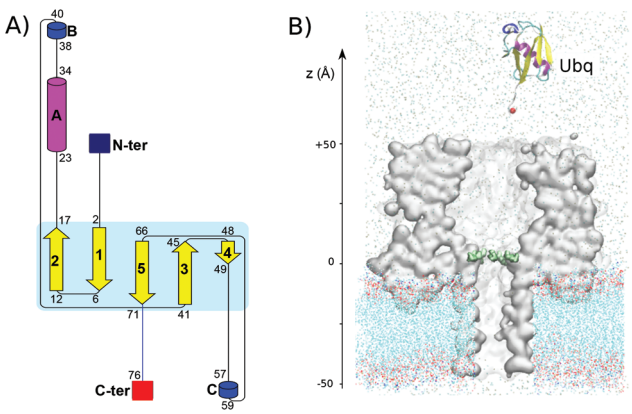

Fig. 1 System components and simulation set-up. Panel A reports the ubiquitin (Ubq) topological diagram.[60] Yellow arrows correspond to beta strands while helices are represented as cylinders (magenta for standard α-helices and blue for 3–10 helices). Panel B reports a sketch of the pore region before the translocation. The αHL nanopore is embedded in a lipid bilayer and the Ubq is placed at ~10 Å above the *cis* entrance. Water is not reported, while Cl^- and K^+ atoms are represented as red and blue dots in the electrolyte solution. The panel refers to the pulling from the C-terminus. The origin of the reference system is in the center of the pore constriction with the z-axis laying along the pore axis and oriented from *trans* to *cis*. Consequently, αHL roughly corresponds to the region -50 Å < z < 50 Å.

Non-equilibrium simulations

The protein translocation pathway is explored by constant velocity (cvSMD) and constant force (cfSMD) steered molecular dynamics.[33] For cfSMD runs, a force $\mathbf{F} = (0,0,F)$ is applied to the α-carbon of the pulled terminus. In the cvSMD, the α-carbon of the pulled terminus is attached to a dummy atom *via* a virtual spring. This dummy atom is moved at a constant velocity $\mathbf{v}$ parallel to the z-direction and the force $\mathbf{F}_\mathbf{v}$ acting on the pulled terminus is measured.[33] Fig. S1† reports the time evolution of the z-component of $\mathbf{F}_\mathbf{v}$ for a pulling velocity $|\mathbf{v}| = 0.025$ Å ps^{-1}. It is apparent that the force needed to unfold the Ubq is ≃1 nN. This information provides an order of magnitude of tentative values of the constant force F to be applied in the cfSMD protocol in order to get a complete multistep translocation in a time window that can be explored by all-atom MD simulations.[9]

Results

The ubiquitin (Ubq) is a small protein constituted of 76 amino acids. Fig. 1A shows its topological diagram.[60] The structure of Ubq is characterized by one β-sheet, formed by five β-strands (β1–β5), an α-helix HA and two 3_{10}-helices, HB and HC. To investigate the Ubq translocation pathway, we performed a set of constant force steered molecular dynamics simulations (cfSMDs),[33] where a force F parallel to the z-direction is applied to the α-carbon of the C or N-terminus. In the next two sections, we discussed the N and C-terminus cases.

C-Terminus pulling

We have selected three different force values: $F_1 = 0.75$ nN, $F_2 = 0.65$ nN and $F_3 = 0.6$ nN. We performed several simulations, each one of them is labelled by a code TFS_x, with T being the pulling terminus (C or N), F the code for the force (*e.g.* 1 for F_1, 2 for F_2, 3 for F_3), and $S = \Omega$ means that the run starts from the native state, whereas $S = \Upsilon$ means a restart from a specific translocation stage, see below. Since we performed several replicas for each simulation, we used the subscript x to indicate the replica index. Table S1† reports the list of all the simulations for C- and N-pulling.

For C-pulling, a common overall co-translocational unfolding pathway emerges from our simulations. As a first step, the strand β5 unfolds and enters the pore. This step splits the native conformation into two structural clusters packing against each other in the Ubq fold. The first cluster includes β1, β2 and HA, while the second one includes β3 and β4 and HC. This split into two halves has been conjectured in the free pulling experiments of poly-Ubq by Schlierf *et al.*,[57] and then seen by Irback *et al.*[32] in the simulations of single Ubq pulling. The co-translocational unfolding of the second cluster is associated with stall I and II, while stall III, IV and V are associated with elements of the first cluster, Fig. 2. In particular,

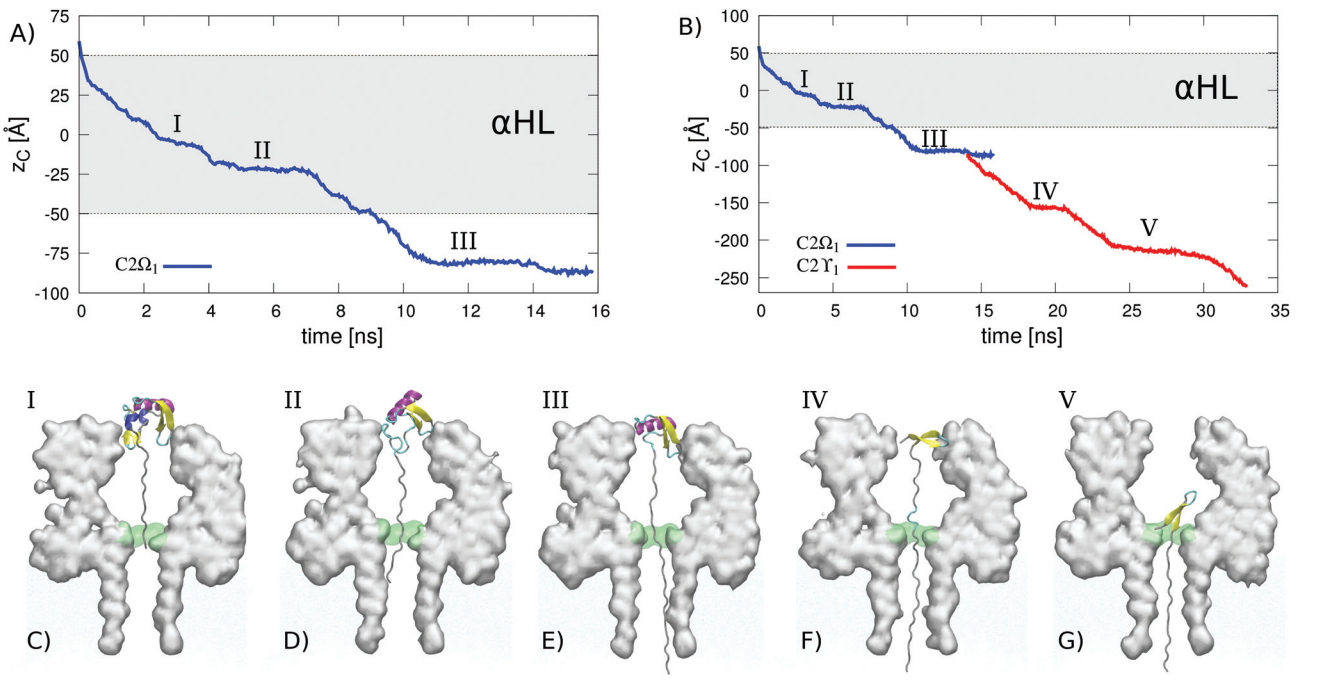

Fig. 2 C-Terminus pulling. (A) Time evolution of the position z_C of the α-carbon of the C-terminus for the simulation starting from the native state pulled with a force $F_2 = 0.65$ nN, simulation code $C2\Omega_1$. The grey rectangle represents the αHL region, –50 Å < z < 50 Å. The translocation starts with the C-terminus coil (residues 71–76) entering in the *cis* side of the pore. In this initial stage, indicated as 0, all the secondary structure elements are still folded. Then, three different stalls (I, II, III) are apparent, with the corresponding translocation intermediates being reported in panels C, D and E, respectively. For simulation $C2\Omega_1$, the Ubq does not complete the translocation in a time window of 50 ns (not shown) and gets stuck in the translocation intermediate III. (B) Complete protein unfolding. The blue line refers to the translocation pathway already shown in panel A, while the red line corresponds to the simulation $C2\Upsilon_1$ obtained using an initial conformation just after stall III (see text). In the red curve, two additional stalls, IV and V corresponding to the conformations reported in panels F and G, are apparent.

stall IV and V are both due to the hairpin formed by β1 and β2. In stall IV, the hairpin remains blocked outside the αHL vestibule, while in stall V, it clogs the pore constriction. As discussed in the last section, this occurrence could be employed to sense specific post-translational modifications. The next paragraphs describe in detail the overall pathway and the computational results.

C-Pulling stalls I, II and III. Fig. 2A reports the time evolution of the position z_C of the α-carbon of the C-terminus for the simulation starting from the native conformation Ω, Fig. 1B, pulled with force F_2 (simulation code $C2\Omega_1$). The origin of the reference system is at the pore constriction and the z-axis points toward the vestibule, so that the translocation corresponds to a decrease of z_C. The simulation starts with Ubq placed ∼10 Å above the pore *cis* entrance, $z_C \simeq 60$ Å. In the first part of the translocation, the Ubq reaches the *cis* entrance in the native state. The C-terminus coiled structure (residues 71–76) easily enters the pore and the translocation slows down only when β5 reaches the pore entrance ($z_C = 30$ Å). After this, the native secondary structure elements begin to unfold, starting from the strand β5. The translocation smoothly proceeds until stall I is observed at times in the range 3 ns < t < 4 ns (at which $z_C \simeq -5$ Å). The corresponding translocation intermediate is represented in Fig. 2C. The native β-sheet, β5–β1 is no more present, but all the other secondary structure elements are still folded and are stuck at the pore entrance. Then, β4 and β3 unfold and partially enter the pore. The subsequent stall, corresponding to $z_C = -20$ Å and times 5 ns < t < 7 ns, is characterized by staying outside the pore of the remaining secondary elements, HB, HA, β2 and β1, see the translocation intermediate II in Fig. 2D. At the same time, the residues 41–49, natively belonging to β3 and β4, occupy the pore vestibule in a random coil conformation, possibly forming isolated bridges as shown in Fig. S2† (frame 102). The translocation of this random coil is characterized by a sequence of short stalls corresponding to small rearrangements of this unfolded region, −50 Å < z_C < −20 Å, in the time interval 7 ns < t < 10 ns, that end when the C-terminus moiety (residues 40–76) is completely stretched, while the N-terminus elements (HA, β2 and β1) are still folded and clog the pore entrance, translocation intermediate III, Fig. 2E, $z_C = -80$ Å, 10 ns < t < 15 ns. This conformation is not unfolded by the force F_2 in a time window of 50 ns (data not shown).

To overcome this stall and explore the complete unfolding pathway, we have performed a set of 5 independent cfSMDs using the higher force F_1 starting from the native conformation Ω (code = $C1\Omega_x$, with x = 1, ⋯, 5). Four of them result in a complete translocation in a time window T_w = 20 ns, see Fig. S3,† while for one of them, the protein stalls in a translocation intermediate corresponding to $z_C = -180$ Å, see Fig. S3(D).† In all the $C1\Omega_x$ simulations, stall I and III are always present (although, in several cases, they are less evident when compared to the $C2\Omega_1$ simulation discussed above). Stall II, instead, is not always present. We have attributed this occurrence to the lack of secondary structural elements associated with stall II. In other words, stalls associated with translocation intermediates linked to specific secondary structure rearrangements are reproducible, while stalls due to contingent interaction between random coils and the pore walls, are not strictly reproducible among the different realizations of the translocation process.

C-Pulling stalls IV and V. To complete the translocation pathway, we have selected the first conformation after stall III in the $C1\Omega_2$ run (Fig. S4†) and have used it as the initial condition, dubbed ϒ, for simulations at force F_2. Fig. 2B shows the complete time evolution of z_C where the first part (blue) is the already discussed translocation pathway of $C2\Omega_1$, and the second part (red) corresponds to the simulation $C2\Upsilon_1$. After stall III, the translocation is characterized by the unfolding of HA while the β2 and β1 are folded. In particular, β2 and β1 are parallel to the pore entrance, Fig. 2F (stall IV, $z_C = -155$ Å and 17 ns < t < 20 ns). Then, the two β-strands rotate, enter the vestibule, and get stuck at the pore constriction (stall V, z_C = −205 Å at times 23 ns < t < 28 ns, Fig. 2G). After stall V, β2 and β1 unfold and the translocation curve rapidly drops. The described phenomenology also occurs in the other replicas, see Fig. S5† (code $C2\Upsilon_x$). As a further confirmation of the described scenario, we performed a set of five independent simulations with a smaller force F_3 = 0.6 nN, (code $C3\Upsilon_x$). In almost all simulations, stalls IV and V are present, see Fig. S6.†

Our results can be compared with those of Xu *et al.*[67] where Ubq unfolding through the carbon nanotube is investigated *via* a cvSMD protocol. Xu *et al.*[67] observed distinct structural arrangements (motifs) that remain intact as the protein is pulled through the nanopore. The motifs they identified are very similar to translocation intermediates I, II, III and IV, although, in our case, stall II is characterized by the unfolding of residues 41–49, natively belonging to β3 and β4.

In summary, our C-pulling simulations indicate that, during the translocation, the Ubq alternates phases where it smoothly unfolds to stalls where portions of the protein are blocked in translocation intermediates that appear to be associated with specific secondary structural elements. A similar scenario, characterized by a wealth of intermediates, was f also ound in the co-translocational folding–unfolding of RNA structures. Precisely, a nanopore pulling assay has been used for dissecting the sequential unfolding mechanism of pseudoknots that constitute longlived intermediates in RNA molecules.[69,70] Our simulations show that intermediates associated with specific secondary structural elements are highly reproducible among the replicas. Interestingly, the difference between the structures of the same intermediate in different replicas, measured as RMSD, is just a few Angstrom (see the structural analysis discussed in section S3 of the ESI†) indicating that actually the translocation intermediates involving folded structures, are very similar. Instead, intermediates due to contingent interaction with the pore wall, do not occur in all the replicas.

N-Pulling

Fig. 3A shows the typical time evolution of the position z_N for the simulation starting from the native conformation Ω using

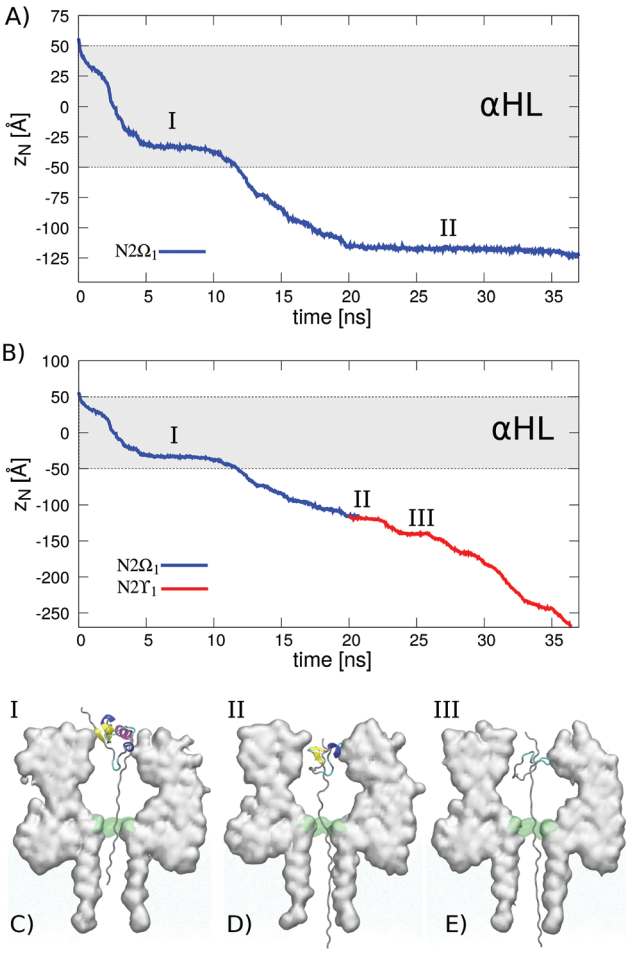

Fig. 3 N-Terminus pulling. (A) Time evolution of the position z_N of the α-carbon of the N-terminus for the simulation starting from the native state pulled into the nanopore by a force F_2 = 0.65 nN, simulation code $N2\Omega_1$. The gray rectangle represents the region corresponding to αHL, −50 Å < z < 50 Å. The translocation starts with the N-terminus coil (residues 1–2) entering in the *cis* entrance pore. In this initial stage, indicated as 0, all the secondary structure elements are still folded. Then, two different stalls, I and II, occur, with the corresponding translocation intermediates being reported in panels C and D, respectively. For simulation $N2\Omega_1$, the protein does not complete the translocation and gets stuck in the intermediate II. (B) Complete protein unfolding. The blue line refers to the translocation pathway already shown in panel A, while the red line corresponds to the simulation $N2\Upsilon_1$ obtained using an initial conformation just after stall II (see the text). In the red curve, one additional stall, III, corresponding to the conformation reported in panels E, is apparent. (C–E) Snapshots of the translocation intermediates I–III.

a pulling force F_2 = 0.65 nN (code = $N2\Omega_1$). During the translocation, two stalls (I and II) occur. As soon as the Ubq touches the pore, it gets stuck at the *cis* entrance and a short stall, 0, is observed at z_N = 30 Å. All the secondary structure elements are still folded although the tertiary structure is only slightly deformed. Then, β1, β2, and a part of HA unfold and translocate into the nanopore. Stall I, occurring at z_N = −35 Å in the time interval 5 ns < t < 12 ns, is characterized by the presence in the pore vestibule of the secondary structural elements HB (partially folded), HA, β3, β4, HC and β5, as shown in Fig. 3C. After this stall, the translocation is characterized by several short stalls, corresponding to the progressive unfolding of the HA (−100 Å < z_N < −75 Å and 13 ns < t < 15 ns), and β5, while β3, β4 and HC are folded and still blocked in the vestibule, Fig. 3D, stall II (z_N = −120 Å and 20 ns < t < 35 ns). The residues 66–71, natively belonging to β5, occupy the pore vestibule in a random coil conformation; this conformation is not unfolded by the force F_2 in a time window of ∼40 ns. To overcome this stall and explore the complete unfolding pathway, the same approach presented for the C-pulling has been used, *i.e.* we selected a conformation Υ just after stall II, see Fig. S8.† Fig. 3B shows the complete time evolution of z_N. After stall II, the beta strands β3 and β4 unfold and the proteins get stuck into the vestibule in a random coil, stall III, Fig. 3E. As this unstructured coil disentangles, the translocation proceeds with only short and not reproducible stalls associated with interactions between Ubq and the pore. The presented co-translocational unfolding pathway is confirmed by four additional independent replicas at F_2, code $N2\Upsilon_x$, Fig. S9,† and five replicas at F_1 = 0.75 nN, Fig. S7.† With the exception of a single replica, all the runs share the same sequence of translocation intermediates. Also in this case, the other stalls that appear in the z_N plots are not associated with secondary structure rearrangements but to contingent interactions between Ubq and the pore walls.

The main difference between C-pulling and N-pulling cases is that for C-pulling, the unfolding mainly occurs at the *cis* entrance (stalls I–IV) while, for N-pulling, the Ubq enters the vestibule after stall I. From the structural point of view, this difference is associated with the different unfolding of the helix HA. Indeed, HA is the largest secondary structure element and Ubq can enter the vestibule only after HA unfolds. In N-pulling, HA unfolds in the first stages of translocation (HA is close to the N-term, see the topological diagram in Fig. 1) allowing the protein to soon enter the vestibule just after stall I. In C-pulling, instead, HA unfolds only in the last stages of the translocation, after stall III.

Native structure signature in co-translocational unfolding

In order to better clarify the role of native structure elements on the co-translocational unfolding pathway, we employed the backward burial analysis.[7] The backward burial analysis is a simple method to attempt to predict the translocation pathway from the native structure, Bacci *et al.*,[7] that was shown to provide useful insight into the analysis of Thioredoxin co-translocational unfolding.[23] The backward burial of a residue i is the number of long-range native contacts that are formed with the untranslocated portion of the chain. For C-pulling and N-pulling, the expressions for the backward burial of the i-th residue are

$$B_C(i) = \sum_{j=1}^{i-\delta} \Delta_{ij}, \quad B_N(i) = \sum_{j=i+\delta}^{m} \Delta_{ij}, \qquad (1)$$

where Δ_{ij} is the contact matrix, in our case, Δ_{ij} = 1 if the Cα of residues i and j in the native structure are within 7.5 Å and 0 otherwise, m is the total number of residues and δ is set in

order to only include interactions that are far enough in the sequence, here $|i-j| > \delta = 5$. In essence, backward burial is a measure of the interaction of a residue with the untranslocated portion of the proteins, assumed to be in a folded state.

Fig. 4A reports $B_C(i)$ and $B_N(i)$. For C-pulling (red, right vertical axis), the backward burial plot has to be read from right to left, *i.e.* from the residue 76 (C-terminal) to residue 1. The first peak is located in the portion of the chain corresponding to β5, *i.e.* $i \sim 66$–71 and it is due to the interaction of β5 with β1 and β3, see also the contact map in Fig. 4B. This peak is the one that causes stall 0, where the protein is completely folded outside the pore. The second peak arises in the region $i \sim 56$–68 and refers to the interaction between HC and the loop of β2 and HA and roughly corresponds to the rearrangement involved in stalls I and II. These first two peaks in B_C show a good correspondence with the co-translocational unfolding pathway. Instead, stall III is not associated with a peak in the B_C since it is mainly due to the unfolding of HA that is transversally stuck at the pore entrance, see Fig. 2E. Finally, the B_C profile in Fig. 4A predicts stalls IV and V involving the residues 11–18; the latter being associated with the unzipping of the β1–β2 hairpin.

A similar analysis can be performed for N-pulling, blue areas in Fig. 4A to be read from left to right, *i.e.* from residue 1 (N-terminal), to residue 76. The first peak, residues ~1–8, corresponds to the interaction of β1 with β2 and β5 and it is associated with the first stall 0 when the entire protein is outside the pore and almost completely folded. Stall I is associated with the contacts between the loop connecting β2 to HA, residues ~17–23, with the loop connecting β4 to HC, residues 50–57. Similarly, the third peak, residues ~40–45, corresponds to the interaction among β3, β5 and β4 and it is associated with stall II. In stall III, instead, all the secondary structures are unfolded, so, no information can be extracted from the native structure analysis.

As a further check of the crucial influence of the native structure on the co-translocational unfolding pathway, we performed coarse grained simulations of the Ubq translocation. The protein is modelled using a weighted version of the Gō model,[15,21] while the αHL pore is modelled as two coaxial confining cylinders: a larger one for the vestibule and a narrower one for the barrel, see Fig. 4E, with the details being reported in section S1 of the ESI.† Such a coarse grained model allows to collect a large number of translocation events using limited computational resources. To describe the co-translocational

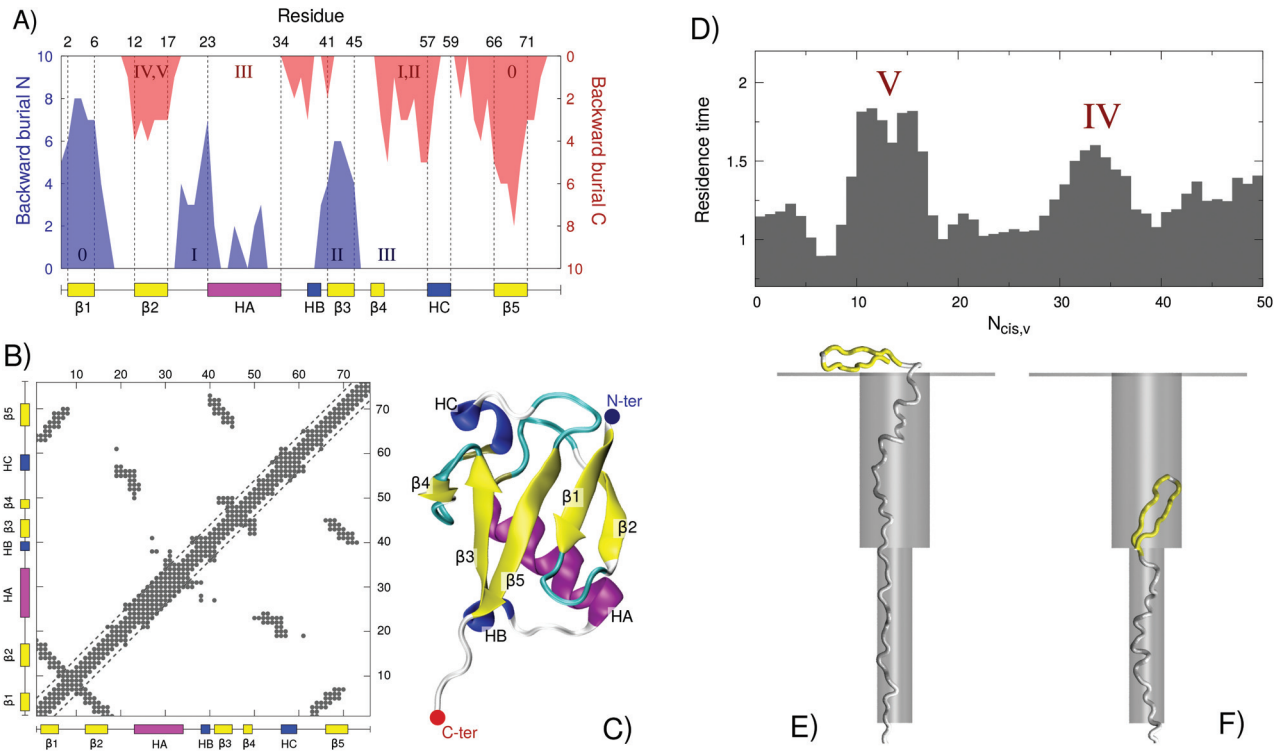

Fig. 4 Native structure signature in co-translocational unfolding. (A) Backward burial. The horizontal axis reports the index of the residue. A sketch of the secondary structure is reported in the bottom of the graph (yellow rectangles refer to β-strands, while magenta and blue rectangles refer to α and 3–10 helices). Blue areas refer to N-pulling. Red areas refer to C-pulling and the backward burial axis for B_C is reversed; see the red left *y*-label. The roman digits refer to the correspondence between some of the stalls observed in the co-translocational unfolding simulations and the backward burial graph. (B) Contact map Δ_{ij}. Points refer to $\Delta_{ij} = 1$, *i.e.* the Cα of residues *i* and *j* in the native structure are within 7.5 Å. (C) Cartoon view of the Ubq native structure (pdb_id: 1UBQ[64]). Image drawn using the VMD software.[31] (D) Coarse grained simulations. Histogram of residence time. $N_{cis,v}$ is the number of residues that have not yet crossed the constriction. The residence time is rescaled by the average residence time between stall IV and V. The peaks at $30 < N_{cis,v} < 35$ and $11 < N_{cis,v} < 16$ correspond to stalls IV and V, respectively. Panel (E) and (F) display typical snapshots of stalled conformations referred to Stall IV and V respectively, obtained from the coarse-grained computational model.

unfolding, we selected two collective variables, the number $N_{cis,v}$ of residues upward the barrel (*i.e.* residues that still did not engaged the barrel) and the number N_{cis} of residues at the *cis* side. The histograms of the residence time into N_{cis}-states in Fig. S11† display a large peak at the beginning of the translocation stage, $55 < N_{cis} < 70$. As in the all-atom simulations, this main bottleneck is associated with the unfolding of the β5. To analyze in more detail, in the last part of the translocational pathway, we employed the same strategy used for all-atom simulations, *i.e.* we restarted the run after the main bottleneck. The histograms of the residence time for $N_{cis,v}$ are reported in Fig. 4D, where the two peaks correspond to stalls IV and V, the typical conformations of which are reported in Fig. 4E and F. For completeness, we calculated the RMSD between Stall V conformations from all-atom MD and coarse grained simulations. The average RMSD is ≃1.5 Å while the maximum is ≃4 Å. This larger deviation is mainly due to the partial unfolding of the first two residues of the N-term occurring during the last frames of stall V in the all-atom simulations. This minor unfolding happens when the N-term interacts with the constriction, see, *e.g.* Fig. 5F and it is not present in the coarse grained simulation.

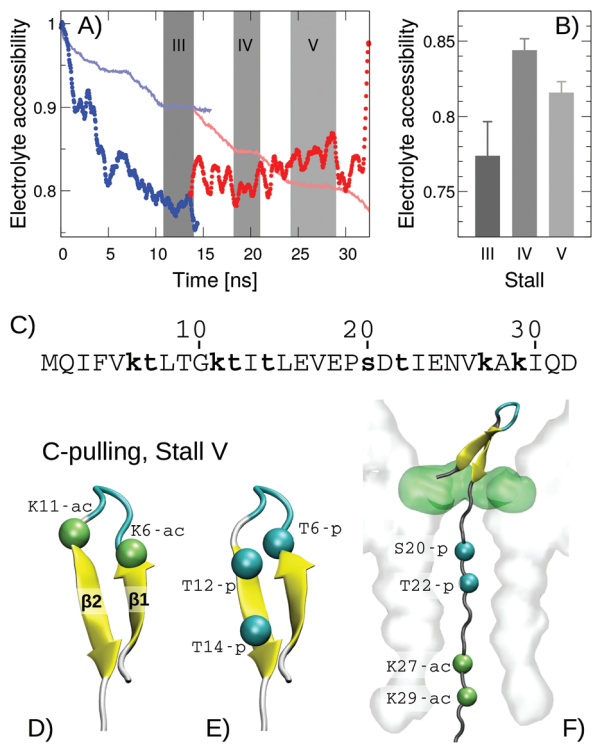

Fig. 5 Application to post-translational modification (PTM) sensing. (A) Circles refer to the electrolyte accessibility, eqn (2), for the translocation already shown in Fig. 2, with z_C being reported as a solid line. (B) Electrolyte accessibility for translocational intermediates corresponding to stalls III, IV and V. (C–F) PTMs located close to the secondary structure elements involved in stall V. Panel C reports the Ubq sequence for residues 1–32. Bold lower case letters refer to amino acids for which PTMs were reported in the literature as annotated in the PhosphoSitePlus database.[28] Panels D and E report acetylation (green, suffix – ac) and phosphorylation (cyan, suffix – p) sites for β1 and β2 while panel F shows the PTMs that would be located in the barrel in stall V.

Discussion

A crucial question arising in nanopore sensing protocols is whether the pore occlusion due to the translocating molecule can leave an unambiguous signature in the ionic current trace. It is, hence, interesting to briefly speculate on possible applications of the observed co-translocational unfolding pathway for specific sensing protocols. It is well known that also small variations of the volume available for the electrolyte passage can be detected by αHL nanopore sensing systems.[5,40,53,56] In our case, the stalls observed are due to different conformations of the protein in the pore, so, in principle, they can be associated with different current levels as seen, for instance, in Thioredoxin co-translocational unfolding.[53,54] However, our simulations indicated that, although the majority of the translocation bottlenecks are caused by reproducible rearrangements of specific portions of the ubiquitin secondary structure, the overall co-translocational unfolding pathway always shows also (i) stalls due to partially unfolded structures (see *e.g.* stall II from C-pulling and stall III from N-pulling), (ii) stalls due to pore–protein interactions (see *e.g.* stalls "*" in the ESI Fig. S3, S6 and S7†). Type (ii) stalls occur only in some simulations, hence, will presumably add additional unreproducible current levels in an experimental current trace. Instead, type (i) stalls, being quite reproducible (*i.e.* they occur in almost all the runs), are expected to give rise to well defined current levels. These mixture of reproducible and unreproducible stalls seem to exclude the possibility of using a co-translocational unfolding pathway as an unambiguous signature for single protein identification.

Application to the detection of post-translational modifications

Some stable features of the co-translocational unfolding pathway can be exploited for nanopore sensing. Stall V (C-pulling), corresponding to the β1–β2 hairpin stuck at the pore constriction, is expected to give rise to a well defined current signal. Indeed, after stall V, no other bottlenecks are present, so the identification of stall V on an experimental current blockade trace should be unambiguous because stall V is always the last one.

To make this argument more quantitative, we try to estimate the trend of the current levels associated with different translocation intermediates from our structural data. The current mainly depends on the portion of the pore that is accessible to the electrolyte. We define the electrolyte accessibility as

$$c(t) = \frac{R_0}{R(t)}, \qquad (2)$$

where R_0 and $R(t)$ are estimations of the pore clogging for the empty pore and for the pore occupied by the translocating protein at time t. Eqn (2) and the estimation of the pore clogging R are inspired by a continuum quasi-1D model for pore resistance, see the ESI S2† and ref. 24. In brief, we

divided the pore region in N_z slabs of size $\Delta z = 1$ Å and calculated R as

$$R = \sum_{i=1}^{N_z} \frac{\rho}{A_i} \Delta z, \quad (3)$$

with A_i being the area available for the electrolyte passage across the i-th section and ρ the electrolyte resistivity. Note that the value of ρ is not relevant for the electrolyte accessibility because it cancels out in eqn (2). Details on the calculation of A_i can be found in the ESI S2.†

Fig. 5A reports the electrolyte accessibility $c(t)$ for the co-translocational unfolding already discussed for C-pulling, Fig. 2. At $t = 0$, the pore is empty, so, $R(t) = R_0$ and $c(t) = 1$. In the first 10 ns, the Ubq C-terminal enters the pore and, consequently, the electrolyte accessibility $c(t)$ reduces. It is worth noting, that in an experiment, the current trace associated with this stage of the translocation process will be dominated by the signal due to the tail added to the C-terminal for importing the proteins, see *e.g.* ref. 44 and 53. The first stall due to a conformation where the pore is completely filled by the Ubq is stall III, Fig. 2E, that gives rise to the minimum electrolyte accessibility. Then, Ubq continues to unfold and $c(t)$ decreases, stall IV. A further decrease occurs when the β1–β2 hairpin moves from the vestibule entrance (stall IV) to the constriction (stall V). After the complete passage of Ubq, we recover $c(t) = 1$. Fig. 5B, reports the electrolyte accessibility for stall III, IV and V, with the data being obtained from all the replicas where these stalls are present.

The unambiguous trace associated with stall IV and V can be exploited for the detection of post-translational modifications (PTMs) at the N-term. Indeed, PTMs on β1 and β2 are expected to produce detectable signals both in terms of blockade duration (less stable conformations should reduce the dwell time in a stall) and in the current level.[56] Fig. 5C shows the PTMs on the Ubq moiety involved in stall V, as annotated on the PhosphoSitePlus database.[28] Some PTMs directly affect residues belonging to β1 or β2, see Fig. 5D (acetylation sites) and 5E (phosphorylation sites). Other PTM sites, instead, are located on residues that, in stall V, occupy the αHL barrel (Fig. 5F) and, hence, can alter the corresponding current blockade level.

Conclusion

We studied the co-translocational unfolding of ubiquitin through an α-hemolysin pore *via* extensive all-atom molecular dynamics simulations. Our data indicates that the main ingredient controlling the co-translocation unfolding pathway of ubiquitin in the αHL nanopore is the native protein structure and not the specific pore–protein interaction. Several stalls are associated with the unfolding of a specific region of the protein as also confirmed *via* coarse grained simulations. Beyond these structural translocational intermediates, we clearly show that also other stalls due to interaction between the unfolded ubiquitin moiety and the αHL surface occur along the translocation. These stalls are less reproducible than the ones involving secondary structure rearrangements, although, in principle, mutations of the pore (as did, for instance in ref. 66 and 41) can induce specific protein–pore interactions making some stalls stable and reproducible. The complexity of the co-traslocational unfolding pathway may raise doubts on the actual possibility of using the co-translocational unfolding signal for protein sensing. Nevertheless, the last part of the C-pulling translocation is characterized by a high reproducible pattern where the β-hairpin constituted by β1 and β2, gets stuck first at the pore entrance and, then, at pore constriction. This very stable pattern can be potentially exploited to detect post-translational modifications (PTMs) on this region. Indeed, PTMs can alter the stability of this structure (potentially changing the dwell time for a stall) and affect the current blockade level as well.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This research used computational resources from CINECA (NATWE project), TGCC (PRACE project 2014112673) and the Swiss National Super-computing Centre (CSCS), project ID sm11. We would like to thank Dr Edoardo Milanetti for the precious discussions on structural analysis.

References

1 A. Aksimentiev and K. Schulten, Imaging α-hemolysin with molecular dynamics: ionic conductance, osmotic permeability, and the electrostatic potential map, *Biophys. J.*, 2005, **88**(6), 3745–3761.
2 A. Ammenti, F. Cecconi, U. Marini Bettolo Marconi and A. Vulpiani, A statistical model for translocation of structured polypeptide chains through nanopores, *J. Phys. Chem. B*, 2009, **113**(30), 10348.
3 A. Asandei, M. Chinappi, H. K. Kang, C. H. Seo, L. Mereuta, Y. Park and T. Luchian, Acidity-mediated, electrostatic tuning of asymmetrically charged peptides interactions with protein nanopores, *ACS Appl. Mater. Interfaces*, 2015, **7**(30), 16706–16714.
4 A. Asandei, I. Schiopu, M. Chinappi, C. H. Seo, Y. Park and T. Luchian, Electroosmotic trap against the electrophoretic force near a protein nanopore reveals peptide dynamics during capture and translocation, *ACS Appl. Mater. Interfaces*, 2016, **8**(20), 13166–13179.
5 A. Asandei, A. E. Rossini, M. Chinappi, Y. Park and T. Luchian, Protein nanopore-based discrimination between selected neutral amino acids from polypeptides, *Langmuir*, 2017, **33**(50), 14451–14459.

6 M. Bacci, M. Chinappi, C. Casciola and F. Cecconi, Role of denaturation in maltose binding protein translocation dynamics, *J. Phys. Chem. B*, 2012, **116**(14), 4255–4262.
7 M. Bacci, M. Chinappi, C. M. Casciola and F. Cecconi, Protein translocation in narrow pores: Inferring bottlenecks from native structure topology, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2013, **88**(2), 022712.
8 P. F. Batcho, D. A. Case and T. Schlick, Optimized particle-mesh Ewald/multiple-time step integration for molecular dynamics simulations, *J. Chem. Phys.*, 2001, **115**(9), 4003–4018.
9 E. L. Bonome, R. Lepore, D. Raimondo, F. Cecconi, A. Tramontano and M. Chinappi, Multistep current signal in protein translocation through graphene nanopores, *J. Phys. Chem. B*, 2015, **119**(18), 5815–5823.
10 E. L. Bonome, F. Cecconi and M. Chinappi, Electroosmotic flow through an α-hemolysin nanopore, *Microfluid. Nanofluid.*, 2017, **21**(5), 96.
11 M. Boukhet, F. Piguet, H. Ouldali, M. Pastoriza-Gallego, J. Pelta and A. Oukhaled, Probing driving forces in aerolysin and α-hemolysin biological nanopores: electrophoresis versus electroosmosis, *Nanoscale*, 2016, **8**(43), 18352–18359.
12 B. R. Brooks, C. L. Brooks, A. D. MacKerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, *et al.*, Charmm: the biomolecular simulation program, *J. Comput. Chem.*, 2009, **30**(10), 1545–1614.
13 S. Buyukdagli, Facilitated polymer capture by charge inverted electroosmotic flow in voltage-driven polymer translocation, *Soft Matter*, 2018, **14**(18), 3541–3549.
14 P. Cadinu, G. Campolo, S. Pud, W. Yang, J. B. Edel, C. Dekker and A. P. Ivanov, Double barrel nanopores as a new tool for controlling single-molecule transport, *Nano Lett.*, 2018, **18**(4), 2738–2745.
15 F. Cecconi, C. Guardiani and R. Livi, Testing simplified proteins models of the hpin1 ww domain, *Biophys. J.*, 2006, **91**(2), 694–704.
16 F. Cecconi, M. Bacci and M. Chinappi, Protein transport across nanopores: A statistical mechanical perspective from coarse-grained modeling and approaches, *Protein Pept. Lett.*, 2014, **21**(3), 227–234.
17 F. Cecconi, M. A. Shahzad, U. M. B. Marconi and A. Vulpiani, Frequency-control of protein translocation across an oscillating nanopore, *Phys. Chem. Chem. Phys.*, 2017, **19**(18), 11260–11272.
18 A. E. Chavis, K. T. Brady, G. A. Hatmaker, C. E. Angevine, N. Kothalawala, A. Dass, J. W. Robertson and J. E. Reiner, Single molecule nanopore spectrometry for peptide detection, *ACS sens.*, 2017, **2**(9), 1319–1328.
19 M. Chinappi and F. Cecconi, Protein sequencing via nanopore based devices: a nanofluidics perspective, *J. Phys.: Condens. Matter*, 2018, **30**(20), 204002.
20 M. Chinappi, F. Cecconi and C. M. Casciola, Computational analysis of maltose binding protein translocation, *Philos. Mag.*, 2011, **91**(13–15), 2034–2048.
21 C. Clementi, H. Nymeyer and J. N. Onuchic, Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? an investigation for small globular proteins, *J. Mol. Biol.*, 2000, **298**(5), 937–953.
22 J. R. Comer, D. B. Wells and A. Aksimentiev, *Modeling nanopores for sequencing DNA*, 2012, http://bionano.physics.illinois.edu/tutorials/modeling-nanopores-sequencing-dna.
23 D. Di Marino, E. L. Bonome, A. Tramontano and M. Chinappi, All-atom molecular dynamics simulation of protein translocation through an α-hemolysin nanopore, *J. Phys. Chem. Lett.*, 2015, **6**(15), 2963–2968.
24 G. Di Muccio, A. E. Rossini, D. Di Marino, G. Zollo and M. Chinappi, Insights into protein sequencing with an α-hemolysin nanopore by atomistic simulations, *Sci. Rep.*, 2019, **91**(1), 6440.
25 N. Giamblanco, D. Coglitore, A. Gubbiotti, T. Ma, E. Balanzat, J. M. Janot, M. Chinappi and S. Balme, Amyloid growth, inhibition and real-time enzymatic degradation revealed with single conical nanopore, *Anal. Chem.*, 2018, **90**(21), 12900–12908.
26 N. Giamblanco, D. Coglitore, J. M. Janot, P. E. Coulon, B. Charlot and S. Balme, Detection of protein aggregate morphology through single antifouling nanopore, *Sens. Actuators, B*, 2018b, **260**, 736–745.
27 N. Haridasan, S. K. Kannam, S. Mogurampelly and S. P. Sathian, Translational mobilities of proteins in nanochannels: A coarse-grained molecular dynamics study, *Phys. Rev. E*, 2018, **97**(6), 062415.
28 P. V. Hornbeck, B. Zhang, B. Murray, J. M. Kornhauser, V. Latham and E. Skrzypek, Phosphositeplus, 2014: mutations, ptms and recalibrations, *Nucleic Acids Res.*, 2014, **43**(D1), D512–D520.
29 G. Huang, K. Willems, M. Soskine, C. Wloka and G. Maglia, Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with Frac nanopores, *Nat. Commun.*, 2017, **8**(1), 935.
30 L. Huang, S. Kirmizialtin and D. Makarov, Computer simulations of the translocation and unfolding of a protein pulled mechanically through a pore, *J. Chem. Phys.*, 2005, **123**, 124903.
31 W. Humphrey, A. Dalke and K. Schulten, *Vmd: visual molecular dynamics*, 1996, vol. 14(1), pp. 33–38.
32 A. Irbäck, S. Mitternacht and S. Mohanty, Dissecting the mechanical unfolding of ubiquitin, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**(38), 13427–13432.
33 B. Isralewitz, M. Gao and K. Schulten, Steered molecular dynamics and mechanical functions of proteins, *Curr. Opin. Struct. Biol.*, 2001, **11**(2), 224–230.
34 Z. Ji, S. Wang, Z. Zhao, Z. Zhou, F. Haque and P. Guo, Fingerprinting of peptides with a large channel of bacteriophage phi29 dna packaging motor, *Small*, 2016, **12**(33), 4572–4578.
35 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, **79**(2), 926–935.

36 J. S. Lee, B. Peng, A. C. Sabuncu, S. Nam, C. Ahn, M. J. Kim and M. Kim, Multiple consecutive recapture of rigid nanoparticles using a solid-state nanopore sensor, *Electrophoresis*, 2018, **39**(5–6), 833–843.

37 M. A. Lomize, A. L. Lomize, I. D. Pogozheva and H. I. Mosberg, Opm: orientations of proteins in membranes database, *Bioinformatics*, 2006, **22**(5), 623–625.

38 B. Luan, T. Huynh, J. Li and R. Zhou, Nanomechanics of protein unfolding outside a generic nanopore, *ACS Nano*, 2015, **10**(1), 317–323.

39 Y. Luo and B. Roux, Simulation of osmotic pressure in concentrated aqueous salt solutions, *J. Phys. Chem. Lett.*, 2009, **1**(1), 183–189.

40 C. Madampage, O. Tavassoly, C. Christensen, M. Kumari and J. Lee, Nanopore analysis: An emerging technique for studying the folding and misfolding of proteins, *Prion*, 2012, **6**(2), 110–118.

41 G. Maglia, M. R. Restrepo, E. Mikhailova and H. Bayley, Enhanced translocation of single dna molecules through α-hemolysin nanopores by manipulation of internal charge, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**(50), 19720–19725.

42 P. Malgaretti, I. Pagonabarraga and J. Rubi, Entropic electrokinetics: Recirculation, particle separation, and negative mobility, *Phys. Rev. Lett.*, 2014, **113**(12), 128301.

43 C. Merstorf, B. Cressiot, M. Pastoriza-Gallego, A. Oukhaled, J. M. Betton, L. Auvray and J. Pelta, Wild type, mutant protein unfolding and phase transition detected by single-nanopore recording, *ACS Chem. Biol.*, 2012, **7**(4), 652–658.

44 J. Nivala, D. B. Marks and M. Akeson, Unfoldase-mediated protein translocation through an α-hemolysin nanopore, *Nat. Biotechnol.*, 2013, **31**(3), 247–250.

45 J. Nivala, L. Mulroney, G. Li, J. Schreiber and M. Akeson, Discrimination among protein variants using an unfoldase-coupled nanopore, *ACS Nano*, 2014, **8**(12), 12365–12375.

46 G. Oukhaled, J. Mathe, A. Biance, L. Bacri, J. Betton, D. Lairez, J. Pelta and L. Auvray, Unfolding of proteins and long transient conformations detected by single nanopore recording, *Phys. Rev. Lett.*, 2007, **98**(15), 158101(4).

47 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, Scalable molecular dynamics with namd, *J. Comput. Chem.*, 2005, **26**(16), 1781–1802.

48 F. Piguet, H. Ouldali, M. Pastoriza-Gallego, P. Manivet, J. Pelta and A. Oukhaled, Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore, *Nat. Commun.*, 2018, **9**(1), 966.

49 C. Plata, Z. N. Scholl, P. E. Marszalek and A. Prados, Relevance of the speed and direction of pulling in simple modular proteins, *J. Chem. Theory Comput.*, 2018, **14**(6), 2910–2918.

50 L. Restrepo-Pérez, S. John, A. Aksimentiev, C. Joo and C. Dekker, Sds-assisted protein transport through solid-state nanopores, *Nanoscale*, 2017, **9**(32), 11685–11693.

51 J. W. Robertson and J. E. Reiner, The utility of nanopore technology for protein and peptide sensing, *Proteomics*, 2018, 1800026.

52 J. Rodríguez, S. Learte-Aymamí, J. Mosquera, G. Celaya, D. Rodríguez-Larrea, M. E. Vázquez and J. L. Mascareñas, Dna-binding miniproteins based on zinc fingers, assessment of the interaction using nanopores, *Chem. Sci.*, 2018, **9**(17), 4118–4123.

53 D. Rodriguez-Larrea and H. Bayley, Multistep protein unfolding during nanopore translocation, *Nat. Nanotechnol.*, 2013, **8**(4), 288–295.

54 D. Rodriguez-Larrea and H. Bayley, Protein co-translocational unfolding depends on the direction of pulling, *Nat. Commun.*, 2014, **5**, 4841.

55 G. M. Roozbahani, X. Chen, Y. Zhang, R. Xie, R. Ma, D. Li, H. Li and X. Guan, Peptide-mediated nanopore detection of uranyl ions in aqueous media, *ACS sens.*, 2017, **2**(5), 703–709.

56 C. B. Rosen, D. Rodriguez-Larrea and H. Bayley, Single-molecule site-specific detection of protein phosphorylation with a nanopore, *Nat. Biotechnol.*, 2014, **32**, 179–181.

57 M. Schlierf, H. Li and J. M. Fernandez, The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**(19), 7299–7304.

58 W. Si and A. Aksimentiev, Nanopore sensing of protein folding, *ACS Nano*, 2017, **11**(7), 7091–7100.

59 L. Song, M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley and J. E. Gouaux, Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore, *Science*, 1996, **274**(5294), 1859–1865.

60 A. Stivala, M. Wybrow, A. Wirth, J. C. Whisstock and P. J. Stuckey, Automatic generation of protein structure cartoons with pro-origami, *Bioinformatics*, 2011, **27**(23), 3315–3316.

61 P. Tian and I. Andricioaei, Repetitive pulling catalyzes co-translocational unfolding of barnase during import through a mitochondrial pore, *J. Mol. Biol.*, 2005, **350**(5), 1017–1034.

62 V. Van Meervelt, M. Soskine, S. Singh, G. K. Schuurman-Wolters, H. J. Wijma, B. Poolman and G. Maglia, Real-time conformational changes and controlled orientation of native proteins inside a protein nanoreactor, *J. Am. Chem. Soc.*, 2017, **139**(51), 18640–18646.

63 K. Vanommeslaeghe and A. D. MacKerell Jr., Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing, *J. Chem. Inf. Model.*, 2012, **52**(12), 3144–3154.

64 S. Vijay-Kumar, C. E. Bugg and W. J. Cook, Structure of ubiquitin refined at 1.8 Å resolution, *J. Mol. Biol.*, 1987, **194**(3), 531–544.

65 P. Waduge, R. Hu, P. Bandarkar, H. Yamazaki, B. Cressiot, Q. Zhao, P. C. Whitford and M. Wanunu, Nanopore-based measurements of protein size, fluctuations, and conformational changes, *ACS Nano*, 2017, **11**(6), 5706–5716.

66 Y. Q. Wang, C. Cao, Y. L. Ying, S. Li, M. B. Wang, J. Huang and Y. T. Long, Rationally designed sensing selectivity and sensitivity of an aerolysin nanopore via site-directed mutagenesis, *ACS Sens.*, 2018, **3**(4), 779–783.

67 Z. Xu, S. Zhang, J. K. Weber, B. Luan, R. Zhou and J. Li, Sequential protein unfolding through a carbon nanotube pore, *Nanoscale*, 2016, **8**(24), 12143–12151.

68 Y. L. Ying, S. C. Liu, X. Shi, L. Wh, W. Yj and Y. T. Long, The hidden transition paths during the unfolding of individual peptides with a confined nanopore, ChemRxiv6394925v1, 2018. DOI:DOI: 10.26434/chemrxiv.6394925.v1.

69 X. Zhang, X. Xu, Z. Yang, A. J. Burcke, K. S. Gates, S. J. Chen and L. Q. Gu, Mimicking ribosomal unfolding of RNA pseudoknot in a protein channel, *J. Am. Chem. Soc.*, 2015, **137**(50), 15742, DOI: 10.1021/jacs.5b07910.

70 X. Zhang, D. Zhang, C. Zhao, K. Tian, R. Shi, X. Du, A. J. Burcke, J. Wang, S. J. Chen and L. Q. Gu, Nanopore electric snapshots of an RNA tertiary folding pathway, *Nat. Commun.*, 2017, **8**, 1458.