Blobs form during the single-file transport of proteins across nanopores

Sauciuc, Adina; Whittaker, Jacob; Tadema, Matthijs; Tych, Katarzyna; Guskov, Albert; Maglia, Giovanni

Published in:
Proceedings of the National Academy of Sciences

DOI:
10.1073/pnas.2405018121

Publication date:
2024

PNAS RESEARCH ARTICLE | CHEMISTRY BIOCHEMISTRY OPEN ACCESS

Blobs form during the single-file transport of proteins across nanopores

Adina Sauciuc[a], Jacob Whittaker[a], Matthijs Tadema[a], Katarzyna Tych[a], Albert Guskov[a], and Giovanni Maglia[a,1]

Edited by William DeGrado, University of California San Francisco, San Francisco, CA; received March 12, 2024; accepted August 5, 2024

The transport of biopolymers across nanopores is an important biological process currently under investigation for the rapid analysis of DNA and proteins. While the transport of DNA is generally understood, methods to induce unfolded protein translocation have only recently been discovered (Yu et al., 2023, Sauciuc et al., 2023). Here, we found that during electroosmotically driven translocation of polypeptides, blob-like structures typically form inside nanopores, often obstructing their transport and preventing addressing individual amino acids. This is in contrast with the electrophoretic transport of DNA, where the formation of such structures has not been reported. Comparisons between different nanopore sizes and shapes and modifications by different surface chemistries allowed formulating a mechanism for blob formation. We also show that single-file transport can be achieved by using 1) nanopores that have an entry and an internal diameter smaller than the persistence length of the polymer, 2) nanopores with a nonsticky (i.e., nonaromatic) inner surface, and 3) moderate translocation velocities. These experiments provide a basis for understanding polypeptide transport under confinement and for improving the design and engineering of nanopores for protein analysis.

nanopore | protein sequencing | translocation | electroosmosis

Significance

Biopolymers may be sequenced as they linearly translocate across nanopores. In the case of highly charged polymers like DNA, single-file translocation is obtained by applying a strong electric field across the nanopore. Proteins are weakly charged, and they translocate across nanopores under the effect of the electroosmotic flow. We found that under this regime, unraveled proteins typically form blob-like structures, which often stop their transport, clogging the nanopore. We propose a mechanism for blob formation and identify the properties of nanopores that allow single-file transport of polypeptides.

Polymer transport across nanopores has been extensively studied using uniformly charged polymers. For example, the electrophoretic translocation of single-stranded DNA (ssDNA) has been reported for a range of nanopores including alpha-hemolysin (αHL) (1), MspA (2), and FraC (3). Despite these nanopores having different entry diameters (of 2 nm, 4 nm, and 6 nm, respectively) and constriction sizes (1.5 nm, 1 nm, and 1.5 nm, respectively), electrophoretically driven ssDNA transport has been always reported as single-file in all nanopores.

Proteins, however, often fold into well-defined structures and do not have a uniform charge. In the presence of guanidinium chloride, their transport across a nanopore has been observed using aerolysin and αHL nanopores (4–6). However, with a possible few exceptions (7), proteins typically required a ssDNA (8) or a poly-10 aspartate (9) electrophoretic tag at the N- or C-termini to initiate the electrophoretic driven transport. In urea or using unstructured polypeptides, transport was observed with engineered CytK (homologue to αHL) nanopores, whereas acidic residues were introduced to induce a strong electro-osmotic flow (EOF) [$P(K^+)/P(Cl^-)$] ratio >~3.5 (10).

Here, exploring the electroosmotically driven transport of urea-destabilized proteins across CytK, MspA, aerolysin, and lysenin nanopores, to our surprise, we found that polypeptides typically form blob-like coiled structures inside the nanopore during translocation. Single-file transport of unstructured proteins, which is crucial in protein sequencing applications, is only observed using a nanopore with a specific size, shape, and chemical composition and under moderate translocation speeds.

Results

CytK Structure and Protein Transport. It was recently shown that polypeptide translocation can be observed using CytK-2E-4D (shortened as CytK-4D) (10) where 4 Asp residues were introduced in the β-barrel region of the nanopore to introduce a strong electro-osmotic flow (EOF) [$P(K^+)/P(Cl^-)$ ratio of 4.0 ± 0.1]. The latter induced the translocation of unstructured protein substrates even against relatively strong electrophoretic forces (EPF). Similar results were observed using WT-αHL nanopores, a homologue of CytK, during translocation induced by both EP and EOF forces (9).

To identify the characteristics in nanopores that allow for single-file translocation, we resolved the structure of CytK-K128D-T147D-K155D, a CytK mutant with enhanced solubility which allows the transport of unstructured or urea-destabilized proteins. The structure was resolved at a resolution of 4.4 Å (*SI Appendix*, Table S1) and, as previously predicted by homology modeling (10, 11), the nanopore was

Author affiliations: [a]Chemical Biology I, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen 9747 AG, The Netherlands

Author contributions: A.S., J.W., and G.M. designed research; A.S., J.W., and M.T. performed research; A.G. contributed new reagents/analytic tools; A.S., K.T., and A.G. analyzed data; and A.S. and G.M. wrote the paper.

Competing interest statement: Giovanni Maglia has equity in and is a founder, director, and shareholder of Portal Biotech Limited, a company engaged in the development of nanopore technologies. This work was not supported by Portal Biotech Limited.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: giovanni.maglia@rug.nl.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2405018121/-/DCSupplemental.

Published September 12, 2024.

heptameric and very similar to α-hemolysin. The most striking difference between the two nanopores is a narrower *cis* entry in CytK (1.1 nm) compared to αHL (2.3 nm, Fig. 1*A* and *SI Appendix*, Fig. S1).

Protein Translocation Across CytK, MspA, lysenin, and Aerolysin Nanopores. In this work, we compared the translocation of malE219a (a maltose-binding protein variant containing the G220D and E221P destabilizing mutations) in 2 M urea (1 M KCl and 15 mM HEPES buffer pH 7.5) (10) using four different nanopores: CytK-4D (Fig. 1*A*), MspA (Fig. 1*B*), lysenin (Fig. 1*C*), and an engineered aerolysin nanopore (Fig. 1*D*). Under these conditions, malE219 is expected to be mainly unfolded in solution (12, 13). The CytK-4D nanopore comprises a relatively narrow entry of 2.4 nm immediately followed by a constriction of 1.1 nm, a vestibule of ~2.6 nm at its widest point, and a ~1.2 nm β-barrel (1.1 nm at its narrowest point). MspA-M2 mutant (2), which has been employed in DNA (14) and peptide (15–18) sequencing, has an asymmetric hour-glass shape

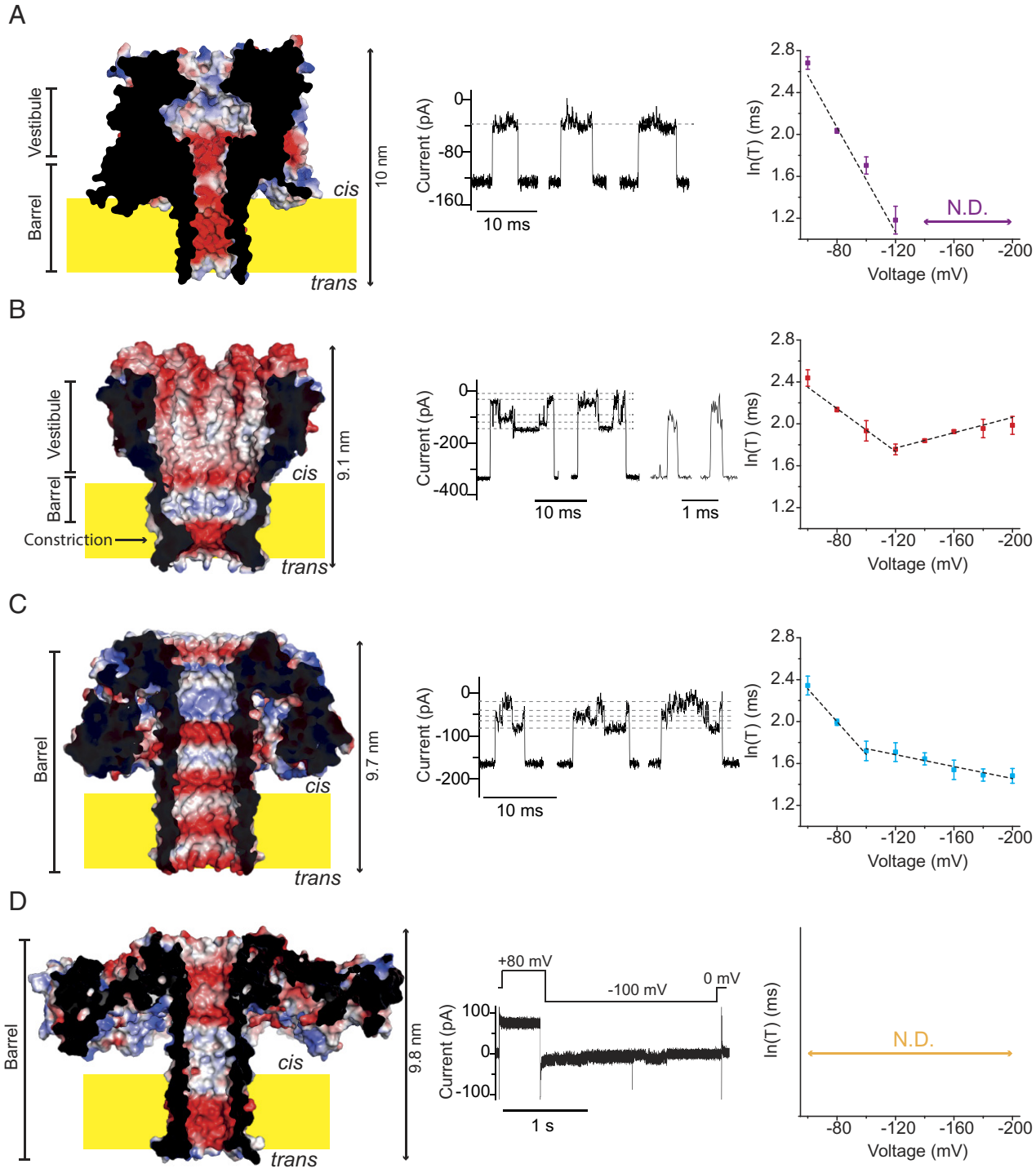

Fig. 1. Translocation of the malE219a protein across highly cation-selective nanopores: CytK-4D (*A*), MspA WT (*B*), lysenin WT (*C*), and aerolysin 1D-3N (*D*), colored by calculated electrostatic potential (Pymol). The middle traces show representative events obtained upon addition of malE219a on the cis side of the nanopores. The right graphs indicate the dwell time (τ) dependence on the applied potential corresponding to the translocation through the nanopores. Data were collected at −100 mV in 1 M KCl, 2 M urea, and 15 mM HEPES buffer pH 7.5. Traces were sampled at 50 kHz and filtered at 10 kHz.

and comprises a broad vestibule (~4 nm at the *cis* entry), followed by a ~3 nm β-barrel region and a 1.0 nm constriction at the *trans* exit (Fig. 1*B*). To test for electroosmotically driven transport of polypeptides, here, we employed the highly cation selective [$P(K^+)/P(Cl^-)$ of 4.1 ± 0.2] wild-type nanopore. The lysenin nanopore has a cylindrical shape, where the top half contains two 1.5 nm constrictions defined by K37 and K45, while the bottom half of the nanopore is ~2 nm wide (Fig. 1*C*). Lysenin is cation selective, as indicated by the reported $P(K^+)/P(Cl^-)$ ratio of 7.5 ± 0.5 (19). Finally, we used the aerolysin nanopore, which has been employed in peptide identification (11) and amino acid recognition (20), and has been used to translocate proteins destabilized by guanidinium chloride (4, 7). This nanopore has a 1.2 to 1.5 nm wide cylindrical lumen (defined as the water facing region of the nanopore, Fig. 1*D*). Aer-WT, however, is weakly anion selective [$P(K^+)/P(Cl^-)$ of 0.76 ± 0.07 (11)], and extensive engineering was required to promote a strong EOF for the translocation of unstructured proteins in the absence of denaturant or in urea. R220, K242, and R282 were substituted with Asn and K238 with Asp, which created a highly cation selective nanopore [aer-1D-3N, [$P(K^+)/P(Cl^-)$ 3.95 ± 0.08, Fig. 1*D*]. The introduction of an additional Asp at position 272 further increased the cation selectivity to 4.68 ± 0.19 (aer-2D-3N). In the presence of 2 M urea, all nanopores showed stable currents (*SI Appendix*, Fig. S2).

As also reported earlier (10), in the −60 to −120 mV range, the translocation of urea-destabilized malE219a across CytK-4D induced well-defined blockades with an average excluded current of 71.9 ± 0.8 % at −80 mV. The dwell time associated with the events decreased sharply with the applied potential from 24.4 ± 6.7 ms to 3.3 ± 0.5 ms (Fig. 1*A*), demonstrating the translocation of the polypeptide across the nanopore. Long-lasting current events were occasionally observed, which were characterized by the following pattern: an initial full block level followed by a subsequent ~70 % $I_{ex\%}$ level [defined as $(I_o - I_B)/ I_o$ percent, with I_o the open pore current and I_B the blocked pore current] and eventually reaching full block or the open pore current (*SI Appendix*, Fig. S3). These long-lasting events were observed with other urea-destabilized proteins (*SI Appendix*, Fig. S3), increased with the applied bias, and became prominent at voltages above −120 mV (*SI Appendix*, Fig. S4), preventing the use of these voltage regimes for polypeptide characterization.

Urea-destabilized malE219a also induced blockades to WT-MspA nanopore starting at −60 mV. Two types of events occurred in the −60 to −120 mV range, which we term type I and type II events (Fig. 1*B* and *SI Appendix*, Fig. S5). Type I events were characterized by frequent shifts between levels. The shallowest and the deepest levels were rather well-defined and predominant and the transitions between these levels consisted of various intermediate levels (*SI Appendix*, Fig. S5). Under this voltage regime, the dwell time decreased from 11.5 ± 0.9 ms to 5.8 ± 0.3 ms (Fig. 1*B*) which is consistent with protein translocation. Type II events were shorter than type I and associated with a smaller excluded current. Their dwell time increased with the potential (*SI Appendix*, Fig. S5), suggesting these events may correspond to the capture and release back in the *cis* compartment. At higher applied potentials, only type I events were observed, whose dwell time gradually increased to reach 7.3 ± 0.6 ms at −200 mV. This behavior is consistent with the protein not translocating under high voltage regimes. The average $I_{ex\%}$ increased with the applied potential across the whole range: from 70.2 ± 1.7 % at −60 mV to 87.2 ± 1.7 % at −200 mV (*SI Appendix*, Fig. S5).

Urea-destabilized malE219a induced multilevel events to WT-lysenin nanopores throughout the entire range of tested potentials (Fig. 1*C*). Events appeared at −60 mV, and their dwell time decreased during the entire range of the applied bias (up to −200 mV), suggesting that the protein translocated through the nanopore across the entire voltage range. Nonetheless, two regimes were still observed with different trends in translocation times with a turning point at −100 mV (Fig. 1*C* and *SI Appendix*, Fig. S6). The dwell time in the first voltage range decreased from 10.5 ± 0.9 ms at −60 mV to 5.6 ± 0.6 ms at −100 mV. After the turning point at −100 mV, the dwell time weakly decreased from 5.5 ± 0.5 ms at −120 mV to 4.4 ± 0.3 ms at −200 mV. The $I_{ex\%}$ increased over the whole range of sampled potentials: from 69.5 ± 0.3 % at −60 mV to 73.9 ± 3.1 % at −200 mV.

The addition of the same urea-destabilized protein to aerolysin 1D-3N or 2D-3N nanopores resulted in long full block events at any of the sampled potentials (Fig. 1*D* and *SI Appendix*, Fig. S7) that most often required reverting the applied potential to obtain the open pore current. This is surprising given that unfolded protein translocation was observed with WT-Aer in the presence of molar concentrations of guanidinium-chloride (a charged chaotropic agent). Increasing the salt concentration to 1.8 M KCl led to the appearance of translocation events, characterized by a reduced dwell time as the applied potential increased (622 ± 25.5 ms at −80 mV up to 181 ± 3.9 ms at −140 mV, aerolysin 1D-3N, *SI Appendix*, Fig. S8), suggesting that there is an electrostatic interaction between the nanopore ~14 nm cap domain and the unraveled polypeptide during translocation.

The importance of the surface interaction prior to threading was tested by translocating the urea-destabilized protein substrate from the *trans* side, where the lipid bilayer is in close proximity with the aperture of the nanopore. This is in contrast with the *cis* side, where the bilayer is ~5 nm away from the entry of the nanopores. Therefore, polypeptides are expected to interact with the lipid bilayer when threading is initiated from the *trans* side and not from the *cis* side (Fig. 1). Blockades initiated from the *trans* side were either long and often irreversible (i.e., the applied potential had to be reversed to reopen the nanopore), and often associated with multiple levels (*SI Appendix*, Figs. S9–S12), suggesting that the polypeptides likely interacted with the lipid bilayer prior entering the nanopore.

Stretched and Coiled Protein Transport. Proteins unfolded by denaturant exhibit polymer-like behavior, whose flexibility can be quantified through their persistence length (L_p). In polypeptides, the L_p is influenced by multiple factors such as the amino acid composition and geometry along the chain (21, 22), the ionic strength of the solution (21), and nanoscale confinement (7, 23, 24). Reported values for the persistence length of an unstructured peptide in solution vary between a single residue (25, 26) and five residues (27), equivalent to a contour length of 0.34 to 1.8 nm [based on an average length gain per amino acid of 0.34 (28) to 0.36 (26, 29, 30) nm]. In many reports, the L_p values are determined under applied force, which may not correspond to the L_p of a polypeptide in aqueous solution (21, 31). In dilute conditions (consistent with the 0.3 μM concentration used here), polypeptides are in a predominantly coiled state characterized by a radius of gyration (Rg) ranging from ~3 nm for a 124 amino acid (AA) protein to ~8 nm for a 549 AA protein (32). Interpolation from previous data (32) suggests that unfolded malE219a (412 amino acids) would have a Rg of ~7 nm (*SI Appendix*, Fig. S13*A*).

Under confinement, i.e., inside a nanopore, if the pore diameter (D) is much smaller than the persistence length ($D \ll L_p$), the polymer follows the Odijk regime, adopting a stretched configuration. For $D \gg L_p$, the de Gennes regime is observed, where the

polymer forms structures referred to as blobs (7, 33) (*SI Appendix*, Fig. S13*B*). In intermediate regimes, nonuniform blobs or back-folding may occur (34).

Inside a nanopore, a polymer can also adopt several different conformations. However, the shape of the nanopore, the solvent flow and the friction of the polymer during translocation are expected to strongly influence the formation of such structures (35, 36). Here, we refer to as "coil" the ensemble of weakly interacting and partially folded conformations that unstructured polypeptides can adopt either in solution or inside a nanopore. For "blobs" we refer to the structure that a polypeptide forms inside a nanopore by filling its entire volume. Although the exact structure of blobs is difficult to predict, blobs are likely to include multiple wounded strands, which are expected to retard or block the translocation of the polypeptide.

MspA has a near conical shape with a large ~4 nm *cis* entry and a narrow 1.0 nm constriction, suggesting that urea-destabilized proteins are likely to enter the nanopore in a partially coiled state (Rg ~ 7 nm). The multilevel events can then be explained by the unraveling or translocation of coils/blobs of different sizes (Fig. 2*A*). The shallow current blockade not associated with translocation (type II events) most likely reflects coils protruding and then retracting to the *cis* side. On the other hand, the electroosmotic flow and electrophoretic forces can also compact the coil further inside the nanopore lumen driving the formation of blob-like structures (Fig. 2*A*). Upon reaching the constriction, the blobs will then unravel or translocate as blobs of different sizes, in turn explaining multiple-level events. The longer dwell times and increased $I_{ex\%}$ at V > −120 mV likely resulted from compression of the blobs against the constriction, which in turn reduces the ionic current and weakens the EOF, hindering translocation under these voltages.

Lysenin has a cylindrical shape, whose entry diameter (1.5 nm) is similar to the polymer's persistence length (~1.8 nm if considering the upper limit), suggesting the polypeptide undergoes a certain degree of unfolding before threading (35, 36). Blob-like structures or back-folding might then occasionally occur in the second half of the nanopore where an enlargement is observed (Fig. 2*B*), explaining the multilevel current events. Polypeptide translocation appears to occur across the entire voltage range tested, suggesting that blob-like structures that may form can be pushed out of the nanopore.

Aerolysin 1D-3N consists of a 1.2 to 1.5 nm wide barrel, which should be ideal for linear transport. Urea-destabilized proteins, however, can only translocate at high ionic strength. This is likely because of the presence of a large cap domain (14 nm wide) which induces electrostatic interactions with the polypeptide prior translocation (Fig. 2*C*).

CytK-4D has a 1.1 nm *cis* entry constriction and a consistently ~1.2 nm wide barrel (Fig. 2*D*). In the −60 mV/−120 mV regime, relatively uniform current blockades suggest that blob-like structures are not formed (Figs. 2*D* and 3). Above −120 mV, many

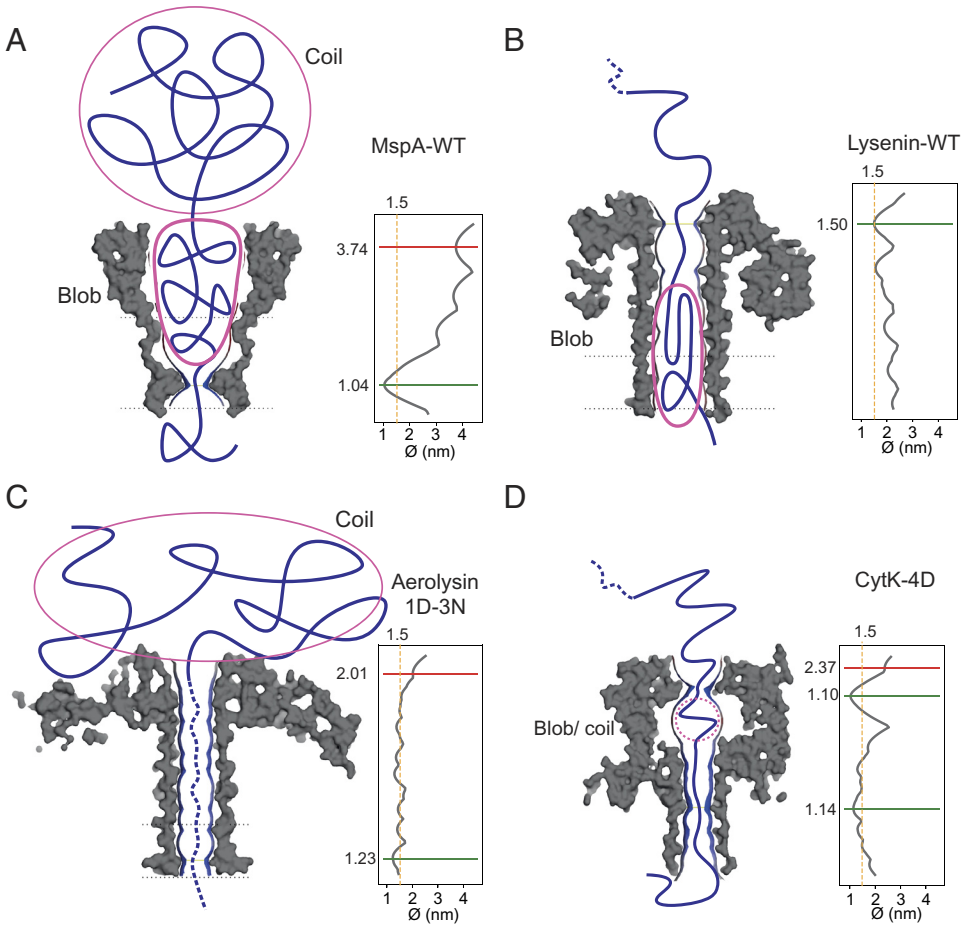

Fig. 2. Blob formation during protein translocation across nanopores. In MspA (*A*), the entry of the pore is ~3-fold larger than the persistence length of a polypeptide, suggesting that blobs can be formed in the lumen of the nanopore. In lysenin (*B*), the nanopore entry is similar to the persistence length of the polypeptide, suggesting that the polymer enters the nanopore as a single file. Then, the lysenin barrel opens to about 2 nm and coils can form. In aerolysin 1D-3N (*C*), the polypeptide may coil at the nanopore entry due to the size and composition of the cap domain. In CytK-4D (*D*), the nanopore has two narrow points with diameters smaller than the L_p of the polypeptide, possibly preventing the formation of blobs.

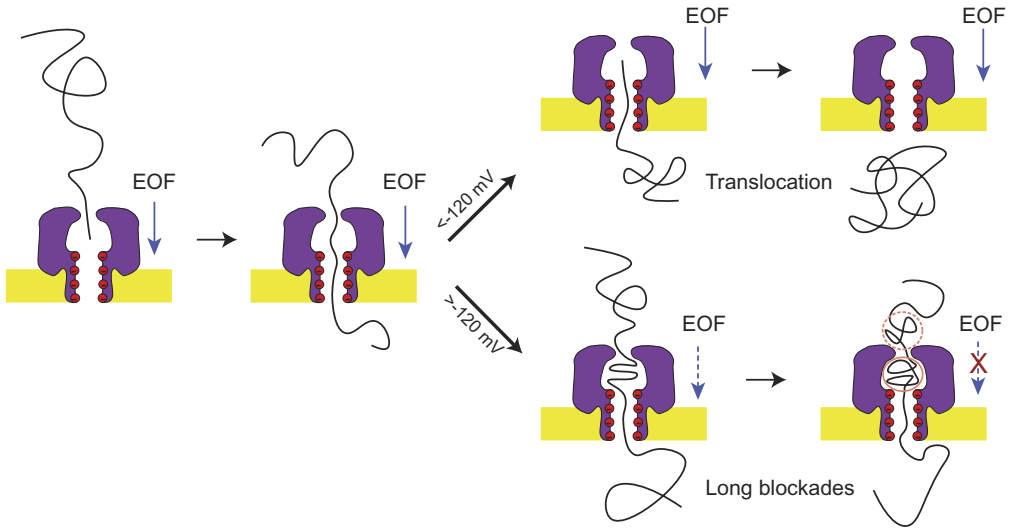

Fig. 3. Possible mechanism for polypeptide translocation across CytK-4D at low potentials (<−120 mV) and high potentials (>−120 mV). In the higher regime, coils and blobs may start forming in the vestibule of the nanopore as the velocity of transport is increased.

long-lasting events permanently blocked the nanopore. A likely explanation is that a higher transport speed increases the likelihood of sudden changes in the speed of transport, which could favor the formation of local blob/coiled structures, possibly in the vestibule of the nanopore (see below). Such structures would then reduce the ionic current and the corresponding EOF preventing their unraveling. This reduction may further hamper the unraveling of coils from solution as they approach the nanopore and prevent the transport across the nanopore (Fig. 3).

Role of the Nanopore Size, Vestibule, and Aromatic Interactions in Blob-like Structure Formation. The relevance of a narrow entry was probed by stepwise truncation of the first 8 N-terminal residues within CytK-4D, which define the entry of the nanopore (Fig. 4*A* and *SI Appendix*, Fig. S1). ΔN2, ΔN4, and ΔN6 truncations did not affect the translocation substantially (Fig. 4*B* and *SI Appendix*, Figs. S14–S16). On the other hand, the ΔN8 truncation, which broadened the diameter at the *cis* constriction site to ~2.7 nm without affecting the stability of the nanopore (*SI Appendix*, Fig. S17), promoted the occurrence of long-lived multilevel events (Fig. 4*B* and *SI Appendix*, Fig. S18), compatible with the formation of blobs. It is worth noting that the initial methionine amino acid (M_{-1}) may still be present.

If blobs/coils are formed inside the vestibule, amino acid substitutions are expected to affect the threading of substrates. The introduction of aromatic residues (A109W and T157F/W) or substitution of an aliphatic residue for a polar one (V110T) in the vestibule near the β-barrel entry (Fig. 4 *A* and *C* and *SI Appendix*, Figs. S19–S22) resulted in longer events compared to CytK-4D (*SI Appendix*, Table S3), showing that the polypeptide is likely to interact with the vestibule of CytK during translocation. This indicates that transient structures, such as coils or blobs, may form during translocation as the polypeptide samples the vestibule space of CytK prior to threading to the β-barrel region.

The role of polymer–pore interactions was further investigated by introducing an aromatic residue within the β-barrel of CytK-4D (CytK-4D-S126F, *SI Appendix*, Fig. S23). This substitution reduced the translocation velocity by ~10-fold across all potentials (*SI Appendix*, Table S4), and the $I_{ex\%}$ increased (e.g., to 89.2 ± 0.6 % compared to 71.3 ± 0.8 % for CytK-4D at −100 mV). Furthermore, current signatures similar to those obtained from MspA and lysenin nanopores were observed (Fig. 4*D*), suggesting the formation of blob-like structures. A likely interpretation is that S126F leads to the coiling of the substrate within the vestibule by contributing to sudden changes in the local velocity once the polypeptide interacts with the introduced aromatic residue.

To confirm the role of aromatic interactions on blob formation, we tested tzatziki, a model substrate devoid of aromatic residues (except for a single Tyr at the Cterminus) and rich in disorder-promoting residues such as Ser, Gly, Asn, and Gln, and charged residues (Fig. 4*D*). Tzatziki was associated with a smooth translocation through CytK-4D (Fig. 4*C*, $I_{ex\%}$ of ~90%, not affected by urea in solution). The presence of S126F did not alter the current signature of this substrate (*SI Appendix*, Fig. S24), despite a slight decrease in $I_{ex\%}$ (2 to 3 %, *SI Appendix*, Table S5) and a ~2-fold decreased velocity compared to CytK-4D. These observations indicate that polypeptide–nanopore interactions (e.g., through hydrophobic interactions) modulate the velocity and the formation of blob-like structures within the nanopore.

Electrophoretic Versus Electroosmotic Polymer Translocation. The observation that linearized translocation can only be obtained under a narrow range of conditions is surprising. This is because the transport of ssDNA, which has a similar persistence length as polypeptides, is not associated with the formation of blob-like structures through a range of different conditions (1, 3). A likely interpretation lies in the fact that the DNA is uniformly charged, and its translocation is dominated by the electrophoretic force. Under electrophoretic transport, the electroosmotic component acts in the opposite direction of the polymer translocation and it contributes to stretching the polymer most likely preventing the formation of blob-like structures.

The translocation of two model unstructured polypeptides helps to further shed light on the transport of polypeptides across nanopores (10). Highly positively charged S1, *SI Appendix*, Fig. S25), allowed to test the electrophoretic-driven transport of a polypeptide. During the translocation of S1 across CytK-WT (10) (EPF-driven transport) the long-lived events were absent, indicating that as for ssDNA, an EOF opposing the direction of

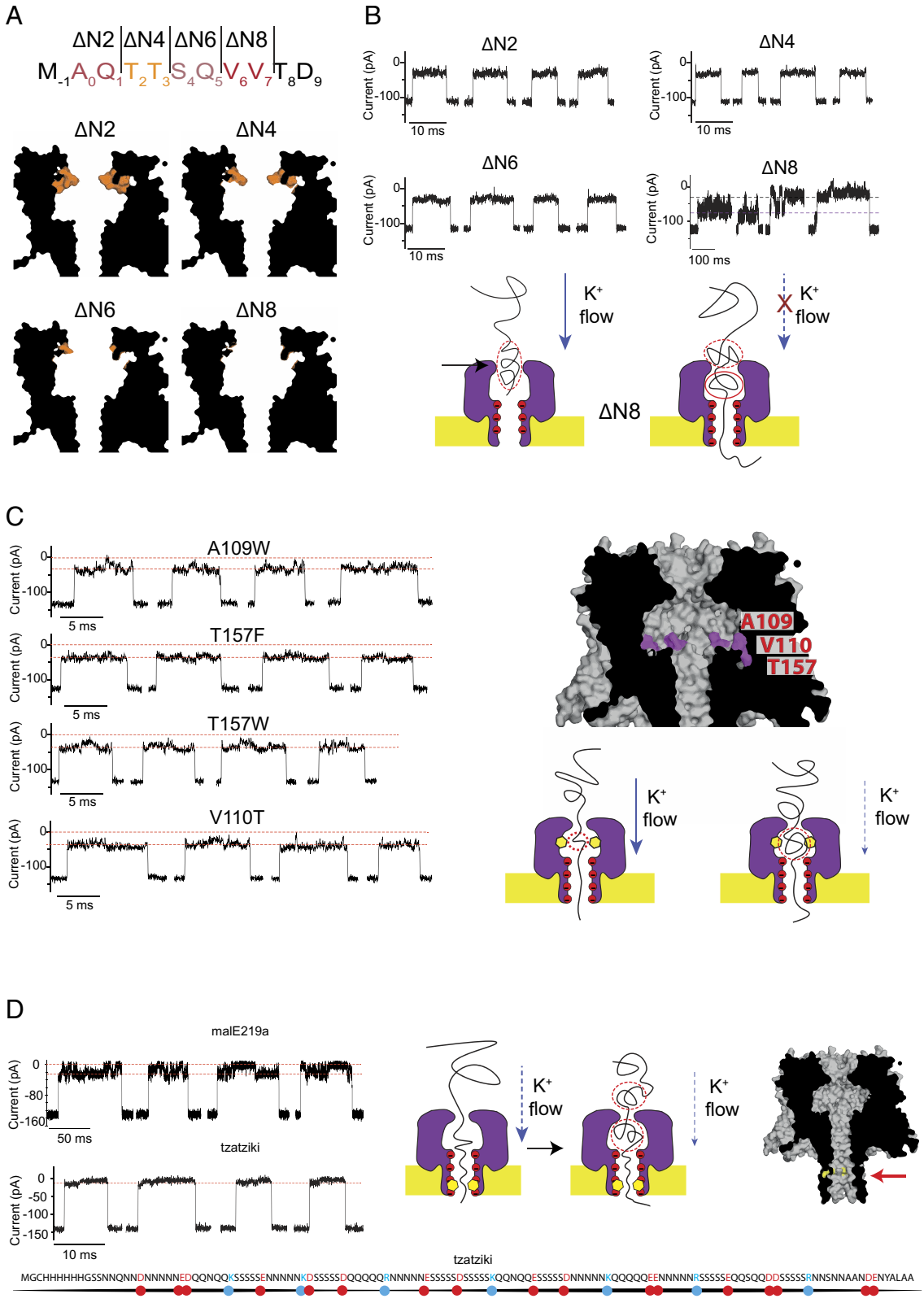

Fig. 4. Mutagenesis in the CytK-4D nanopore addressing the formation of coil and blob-like structures within CytK-4D. (*A-B*) Truncation at the *cis* entry of the nanopore. (*A*) Top: Sequence of the first eight residues in CytK. Nanopore structures showing the truncations where the first eight residues of CytK are colored in orange. (*B*) Traces (−80 mV) from the N-terminal truncation mutants and potential mechanism in the ΔN8 truncation. (*C*) Role of the vestibule. *Left*: Traces from the vestibule mutants (−100 mV). *Right*: Representation of the mutations in the vestibule of CytK and potential translocation mechanism. (*D*) Role of the barrel. On the left the traces from malE219a and tzatziki translocation through CytK-4D-S126F (−100 mV) and on the right potential explanation for this behavior. Data were collected in 1 M KCl, 2 M urea, and 15 mM HEPES buffer pH 7.5. Traces were sampled at 50 kHz and filtered at 10 kHz.

translocation prevents the formation of blob-like structures also with polypeptides. Weakly negatively charged tzatziki devoid of hydrophobic amino acids did not induce long-lived events, suggesting that under electroosmotic-driven transport blob-like structures in proteins are facilitated by aromatic/aliphatic interactions with the nanopore.

Conclusions

The translocation of polypeptides across membranes is an important biological process that is not easily monitored by biophysical techniques. In biotechnology, the transport of unfolded polypeptides across nanopores under an applied potential is under investigation for identification and sequencing of proteins at the single-molecule level.

In this work, we found that during the translocation across nanopores of weakly charged, unstructured polymers, blob-like structures form. This process is augmented by interactions between the polypeptide and the nanopore. The formation of blob-like structures might be prevented by using nanopores having both an entry and a lumen smaller than the persistence length of the polymer (<~1.2 nm), and by removing residues in the nanopore that might interact with the hydrophobic residues in the translocating polypeptide.

Materials and Methods

Chemicals and Reagents. We list the chemicals and their corresponding suppliers as follows: Ampicillin sodium salt was purchased from Fisher Bio Reagents; chloramphenicol (≥98.0) from Sigma Life Science); urea (≥99.5%), guanidinium chloride (≥99.5%, biochemistry), Isopropyl β-D-thiogalactopyranoside (≥99.0%, dioxin-free, animal-free), LB medium, 2xYT medium, NaCl (≥99.5%), HEPES (PUFFERAN® CELLPURE® (≥99.5%), imidazole (≥99%), KCl (≥99.5%), Tris(2-carboxyethyl)phosphine hydrochloride (≥98.0%), Dodecyl-β-D-maltoside (≥99%) from Roth; and n-hexadecane (99% from Acros Organics; protease inhibitors (Pierce™ Protease inhibitor Mini tablets, EDTA-free); GeneJET gel extraction kit, GeneJET PCR purification kit, GeneJET Plasmid Miniprep kit were purchased from (Thermo Scientific); Ni-NTA agarose from Qiagen; Strep Tactin® Sepharose® and D-desthiobiotin from IBA Lifesciences; DPhPC from Avanti polar lipids, n-pentane from Sigma-Aldrich; Quick Start Bradford 1× Dye Reagent (Bio-Rad); DNA primers and gBlock™ from IDT. BL21(DE3) strain harboring the pET-PfuX7 plasmid was kindly provided by Prof. Dr. Oscar Kuipers.

Protein sequences of the nanopores.

- CytK WT:

MAQTTSQVVTDIGQNAKTHTSYNTFNNEQADNMTMSLKVTFIDDPSADKQIAVIN
TTGSFMKANPTLSDAPVDGYPIPGASVTLRYPSQYDIAMNLQDNTSRFFHVAPTNAVEET
TVTSSVSYQLGGSIKASVTPSGPSGESGATGQVTWSDSVSYKQTSYKTNLIDQTNKHVKW
NVFFNGYNNQNWGIYTRDSYHALYGNQLFMYSRTYPHETDARGNLVPMNDLPTLTNS
GFSPGMIAVVISEKDTEQSSIQVAYTKHADDYTLRPGFTFGTGNWVGNNIKDVDQKTF
NKSFVLDWKNKKLVEKKGSAHHHHHH

- MspA WT:

M G L D N E L S L V D G Q D R T L T V Q Q W D T F L N G V F P L D R N R L T R E W F H S G R
A K Y I V A G P G A D E F E G T L E L G Y Q I G F P W S L G V G I N F S Y T T P N I L I D D G D I T A P P F
G L N S V I T P N L F P G V S I S A D L G N G P G I Q E V A T F S V D V S G A E G G V A V S N A H G T
V T G A A G G V L L R P F A R L I A S T G D S V T T Y G E P W N M N G S A G S A W S H P Q F E K

- lysenin WT:

MSAKAAEGYEQIEVDVVAVWKEGYVYENRGSTSVDQKITITKGMKNVNSETRTVTATH
SIGSTISTGDAFEIGSVEVSYSHSHEESQVSMTETEVYESKVIEHTITIPPTSKFTRWQLNAD
VGGADIEYMYLIDEVTPIGGTQSIPQVITSRAKIIVGRQIILGKTEIRIKHAERKEYMTVVSR
KSWPAATLGHSKLFKFVLYEDWGGFRIKTLNTMYSGYEYAYSSDQGGIYFDQGTDNPKQRW
AINKSLPLRHGDVVTFMNKYFTRSGLCYDDGPATNVYCLDKREDKWILEVVGSGHHHHHH

- Aer-WT:

MAEPVYPDQLRLFSLGQGVCGDKYRPVNREEAQSVKSNIVGMMGQWQISGLA
NGWVIMGPGYNGEIKPGTASNTWCYPTNPVTGEIPTLSALDIPDGDEVDVQWRLVHD
SANFIKPTSYLAHYLGYAWVGGNHSQYVGEDMDVTRDGDGWVIRGNNDGGCDGY
RCGDKTAIKVSNFAYNLDPDSFKHGDVTQSDRQLVKTVVGWAVNDSDTPQSGYDVTL
RYDTATNWSKTNTYGLSEKVTTKNKFKWPLVGETELSIEIAANQSWASQNGGSTTTSLSQ
SVRPTVPARSKIPVKIELYKADISYPYEFKADVSYDLTLSGFLRWGGNAWYTHPDNRPNW
NHTFVIGPYKDKASSIRYQWDKRYIPGEVKWWDWNWTIQQNGLSTMQNNLARVLR
PVRAGITGDFSAESQFAGNIEIGAPVPLAADSKVRRARSVDGAGQGLRLEIPLDAQELSG
LGFNNVSLSVTPAANQGSSHHHHHH

Cloning of the Nanopores. Plasmids containing the mutant CytK nanopores were constructed by means of USER cloning (37, 38). PfuX7 DNA polymerase was prepared as previously described (39) and used to generate the fragments used in the USER reaction. Homology regions of 8 to 13 bps were defined in the vicinity of the ATG codon, the mutation site, and downstream of the stop codon to mediate the joining of the empty vector with the gene fragments. The full insert was divided into upstream and downstream fragments at the mutation site and amplified with uracil-containing primers. The empty pT7-SC1 backbone (AmpR) was separately linearized using uracil-containing primers. The PCRs were performed as in ref. 39, with the exception of the extension time, which was decreased from 1 min/kb to 30 s/kb. The PCR products were purified either by gel extraction (if by-products were present) or by cleanup from the PCR mix. The upstream (F1) and downstream (F2) fragments were added to the empty vector (V) in a molar ratio F1:F2:V of 3:3:1 in the USER reaction, which was carried out as in ref. 38: 25 min at 37 °C, followed by 10 min incubation at 60 °C. Finally, the mixture was cooled down to room temperature (22 or 20 °C) for 15 min and subsequently stored on ice/ at 4 °C until the transformation step. The annealed fragments resulting in circularized plasmids were transformed into chemically (RbCl protocol) competent *E. coli* cells using the heat shock procedure (42 °C for 70 s followed by incubation on ice for 10 min) and the cells were selected on a LB agar plate supplemented with 100 μg/mL ampicillin and 1% glucose. Plasmids from individual colonies were isolated and the introduction of the mutations was confirmed by Sanger sequencing (Macrogen/ Eurofins).

pT7-SC1 bearing MspA-WT was generated from pT7-SC1 by means of fusion PCR followed by USER cloning. First, three gene fragments were generated, with the overlap regions for fusion PCR being in the vicinity of the 90 to 93 residues and the 134 to 139 residues. In the second PCR round, the three fragments were joined together in the presence of U-containing primers. This insert and the empty pT7-SC1 vector were used for USER cloning.

pT7-SC1 bearing proaerolysin K238D, previously prepared in our laboratory (11), was used to generate the 1D-3N and 2D-3N aerolysin mutants. The mutations, namely R220N, R282N, K242N, and S272D were introduced as described for the CytK mutants.

pT7-SC1 bearing lysenin-WT was previously prepared in our laboratory from the plasmid kindly provided by Gregor Anderluh.

Expression and Purification of the Nanopores. The plasmids encoding for the nanopores were electroporated into BL21(DE3) electrocompetent cells using a Bio-Rad Micro Pulser (bacterial setting). The cells were selected on plates supplemented with 100 μg/mL ampicillin and 1% glucose and grown at 37 °C overnight. Next, several transformants were resuspended in 200 mL (proaerolysin, lysenin, and MspA) or 100 mL (CytK) LB medium containing 100 μg/mL ampicillin, such that the starting optical density at 600 nm (OD_{600}) was 0.05 to 0.1. Cells were grown at 37 °C, 180 RPM until an OD_{600} of 0.6 to 0.8. At this point, protein expression was induced with 0.5 mM IPTG after the cultures were chilled on ice for 5 to 10 min. The cells were harvested (7,500 rpm, 5 min) following an incubation of 19 to 21 h at 25 °C, 180 RPM. The resulting cell pellets were incubated for at least 30 min at −80/−70 °C, prior to purification. Pellets from 100 mL cultures were resuspended in 20 to 25 mL ice-cold lysis buffer (Table 1), while pellets from 200 mL cultures were resuspended in 30 to 40 mL lysis buffer. Subsequent steps were performed at 4 to 6 °C, unless stated otherwise. The cell suspension was sonicated with a Branson sonifier 450 at 25% duty cycle, 2.5 output control for 2 to 3 min in the case of the CytK mutants, or with a Branson sonifier 550 using pulses at 15% amplitude, 5 s on, 7 s off, total 5 min on time in the case of MspA, lysenin, and aerolysin mutants. Following sonication, the cellular debris was removed (8,000 RPM, 20 min). The resulting supernatant was incubated with 200 μL Ni^{2+}-NTA slurry (50% suspension) in the case of CytK, Aer, and lysenin or StrepTactin slurry in the case of MspA, pre-equilibrated, and prewashed with 1 mL lysis buffer (Table 1) for 20 to 40 min with shaking. The beads were briefly pelleted (1,500 RPM, 1 min) and transferred to the column (2 mL bed volume biospin chromatography, BioRad) while allowing the flow through pass, at RT. The column was washed in steps with a total of 10 mL wash buffer (Table 1). The protein was eluted with 150 μL elution buffer (Table 1) in three elution fractions. The presence of the SDS-stable CytK mutant oligomers and MspA oligomers was confirmed by SDS-PAGE, omitting the heating step in the sample preparation.

Oligomerization Lysenin. The concentration of the elution fractions was determined by the Bradford assay. Sphingomyelin/DPhPC liposomes (50/50) were added to 0.25 mg/mL lysenin sample in a 10:1 w/w ratio and then incubated at 37 °C for 30 min. The lysenin oligomers were extracted from the liposomes by adding 6%

Table 1. Buffers used in the purification of the nanopores

Nanopore	Lysis buffer	Wash buffer	Elution buffer
CytK	50 mM HEPES, 150 mM NaCl, 10 mM imidazole, pH 7.4 + 0.02% DDM	50 mM HEPES, 150 mM NaCl, 30 mM imidazole, pH 7.4 + 0.02% DDM	50 mM HEPES, 150 mM NaCl, 250 mM imidazole, pH 7.4 + 0.02% DDM
MspA	50 mM HEPES, 150 mM NaCl, pH 7.4 + 0.02% DDM	50 mM HEPES, 150 mM NaCl, pH 7.4 + 0.02% DDM	50 mM HEPES, 150 mM NaCl, pH 7.4 + 0.02% DDM + 3 mM desthiobiotin
Lysenin	50 mM HEPES, 150 mM NaCl, 10 mM imidazole, pH 7.4	50 mM HEPES, 150 mM NaCl, 30 mM imidazole, pH 7.4	50 mM HEPES, 150 mM NaCl, 300 mM imidazole, pH 7.4
Aerolysin	50 mM HEPES, 500 mM NaCl, 10 mM imidazole, pH 7.4	50 mM HEPES, 500 mM NaCl, 30 mM imidazole, pH 7.4	50 mM HEPES, 500 mM NaCl, 250 mM imidazole, pH 7.4

LDAO (in MQ) in a 1:10 v/v ratio to the protein–liposome solution. Next, the mixture was further diluted ~25-fold in 150 mM NaCl, 15 mM Tris + 0.02% DDM (total final volume 6 mL). Then, 50 μL pre-equilibrated Ni^{2+}-NTA beads were added, and the solution was incubated at room temperature for 10 min with shaking. The beads were transferred to the column, washed with a total of 10 mL 150 mM NaCl and 15 mM Tris + 0.02% DDM to remove the LDAO, and finally eluted with 100 μL EBO buffer (15 mM TrisHCl pH 7.5, 150 mM NaCl, 200 mM Na_2EDTA pH 8.0, and 0.02 % DDM). The oligomers were stored at 4 °C for several months.

Oligomerization aerolysin. The proaerolysin elution fractions were combined and subsequently desalted by using an amicon filter 10 K MWCO and 50 mM HEPES and 150 mM NaCl, pH 7.4 buffer. This sample was used for oligomerization. To promote its oligomerization, the proaerolysin protein (~0.5 mg/mL) was cleaved by adding 0.01 mg/mL trypsin in a proaerolysin: trypsin volume ratio 10:1. Additionally, 0.02% DDM was added during cleavage, in order to improve the stability of the oligomers. This mix was incubated at room temperature for 20 min and then stored at 4 °C for 2 to 3 wk.

Substrates. The protein substrates were prepared as previously described (10).

Crystallization of the CytK Nanopore. For crystallization and structure determination, we selected the CytK 2E-2D-T147D mutant due to its improved solubility compared to any other mutant. pT7-SC1 harboring this CytK mutant was electroporated into electrocompetent *E. coli* BL21(DE3) and transformants were selected on LB agar plates supplemented with 100 μg/mL ampicillin and 1% glucose (37 °C, overnight). Then, 800 mL culture was grown, and the mutants were purified as described in the previous section.

Following size exclusion chromatography, the protein was kept in a solution of 50 mM HEPES, pH 7.4, 150 mM NaCl, and 0.02% DDM which was subsequently concentrated to 8 mg/mL using a 30,000 MWCO concentrator. CytK was mixed with the crystallization buffer in a ratio of 2:1 and all crystals were obtained via sitting drop vapor diffusion using Molecular Dimensions screening kits (Calibre Scientific). Crystal hits were observed in numerous conditions using the MemGold I and Morpheus II screening kits, the crystals which appeared the most optimal were further optimized with the custom-made screens made with the DragonFly (SPT Labtech). The optimized conditions produced yellow, rectangular prismatic crystals of approximately 500 microns in length along the longest axis in the following condition: 0.005 M Manganese(II) chloride tetrahydrate, 0.005 M Cobalt(II) chloride hexahydrate, 0.005 M Nickel(II) chloride hexahydrate, 0.005 M Zinc acetate dihydrate, 0.1 M MOPSO, Bis-Tris, pH 6.5, 35%(w/v) glycerol, and 30%(w/v) PEG 4000. Crystals were briefly soaked in 50%(w/v) PEG 4000 for cryoprotection and subsequently flash-frozen in liquid nitrogen in preparation for diffraction experiments at a synchrotron. Data were collected at beamline ID23-1 (ESRF, Grenoble).

The most optimal crystal of CytK diffracted at 4.3 Å resolution (*SI Appendix*, Table S1). Data were processed with XDS (40) and structures were solved by Molecular Replacement with Phaser (41) using a previously published homologous model of alpha-hemolysin (PDB ID: 7AHL). Manual rebuilding was performed with COOT (42) and refinement with PHENIX (43). In all chains, minor truncation of small amino acid strings was necessary as there was no assignable electron density nearby. Data collection and refinement statistics are given in *SI Appendix*, Table S1. The refined model was deposited into the PDB repository with accession code 8RJ8. Images were prepared using UCSF ChimeraX (44).

Calculation of Inner Diameter Profiles. The HOLE2 program was used as is ref. 45, according to the MDAnalysis Python ("G. van Rossum, Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, May 1995.") module (46) to probe the inner radius of each pore along their long axes.

As the program yields radii in Å, the profiles were first converted to diameter in nm using the formula Ø = r*2/10 before being plotted.

The global minimum and maximum of these profiles were computed using numpy (47) and plotted along with the profiles using matplotlib (48).

The molecular structures used in the calculation were taken from the Orientations of Proteins in Membranes (OPM) database (49) unless otherwise specified.

The PDB IDs corresponding to each nanopore's wild-type structure are as follows:

Lysenin: 5GAQ (50)
MspA: 1UUN (51)
Aerolysin: 5JZT (52)

Since the crystal structure created in this work (PDB ID 8RJ8) is incomplete, we built a homology model for the purpose of calculating the radius profile. The model was built using MODELLER software (53) version 10.5. We separated the monomers from the crystal structure and then duplicated and redistributed them to complete the heptameric structures. This way we have seven incomplete structures to use as templates for the homology model, each filling in the gaps left by other structures.

The triangulated surface was converted to a representation suitable for use in PyMOL and that software was used to render the graphics [1. Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. (2015).]

Electrophysiology Measurements. Recordings in planar lipid bilayers were carried out using a chamber comprising two compartments, separated by a 25 μm thick Teflon membrane containing an aperture of approximately 100 μm as described earlier (54). A droplet (half the quantity contained in a 10 μL glass capillary) consisting of n-hexadecane dissolved in n-pentane (6.25%) was applied on the Teflon membrane. When the n-pentane evaporated, 500 μL buffer, followed by two droplets of DPhPC lipids in n-pentane (5 mg/mL) were added in each compartment. Ag/AgCl electrodes were connected to the two compartments via agarose bridges (2.5% agarose, 3 M KCl solution), grounding the *cis* compartment. Measurements were performed using an Axon™ Digidata® 1550B digitizer and an Axopatch 200B amplifier (Molecular Devices) and recorded with the Clampex 11.1 software.

Ion selectivity. Buffers: Buffer A–2 M KCl, 15 mM HEPES, pH 7.5, Buffer B–0.5 M KCl, 15 mM HEPES, pH 7.5, and Buffer C–0 M KCl, 15 mM HEPES, pH 7.5,

The ion selectivity of the aerolysin mutants and MspA WT was determined using the buffers described above. First, both the *cis* and *trans* compartments were filled with 500 μL buffer A, and a single nanopore was isolated. Next, the pipet offset was adjusted to 0 pA at 0 mV bias. The I/V curve was determined between −30 and +40 mV, in 5 mV steps, using a 10 kHz sampling rate coupled with a 2 kHz Bessel filter. Next, the KCl concentration in the *trans* compartment was decreased to approximately 0.5 M by perfusing with buffer C (5 × 100 μL). Subsequent perfusion with buffer B (7 × 100 μL) ensured that the final concentration of KCl in the *trans* compartment was correctly fine-tuned to 0.5 M. A similar protocol was used to record the I/V curve of the nanopore in these conditions. This second I/V curve was used to determine the reversal potential

from the linear function fitting the data points between −20 and +20 mV. Finally, the ion selectivity, expressed as the fraction p_{K+}/p_{Cl-} was calculated using the formula below and each ion selectivity was established from triplicate experiments.

$$\frac{p_{K+}}{p_{Cl-}} = \frac{[a_{Cl^-}]_{trans} - [a_{Cl^-}]_{cis} \times e^{V_r F/RT}}{[a_{K+}]_{trans} \times e^{V_r F/RT} - [a_{K+}]_{cis}},$$

where [a] is the activity of the K^+ or Cl^- in the *cis* or trans compartment, V_r is the reversal potential, which is obtained from the experiments, F corresponds to the Faraday constant (96 485 C/mol), R the gas constant (8.3145 J mol^{-1} K^{-1}), and T the temperature (298 K).

Translocation experiments. Single pores were isolated and the pore orientation was determined based on the I/V curve (1 M KCl and 15 mM HEPES, pH 7.5). Then, 2 M urea was introduced into the system by exchanging 125 μL buffer with 125 μL 1 M KCl, 8 M urea, and 15 mM HEPES, pH 7.5, in both compartments. Next, 0.3 μM malE219a was added in the *cis* compartment and translocation was induced by applying negative potential. Alternatively, the protein was added in the *trans* compartment and events were induced by applying positive potential. Each nanopore (substrate in either cis or trans) was tested in triplicate.

Data Analysis Translocation Experiments. The generated files were analyzed using the Clampfit 11.1 software. First, the open pore current (level 0, L0) and the corresponding noise, σ, were obtained from a full point histogram. Second, L1, corresponding to the detection limit, was set at 10σ, (note that Clampfit sets its half-way detection algorithm at 5σ). Events were recorded using voltage protocols in which the potential was ramped through positive and negative applied potentials. In the case of the *cis*-to-trans translocation, events were detected at negative potentials from approximately 2 s of recording times at negative bias. The resulting L1 data points were used to construct the log(dwell time) vs amplitude scatter plot, from which the amplitude boundaries of the event cluster were defined. Events faster than 100 μs were ignored. Next, using the amplitude boundaries, the logarithmic histogram of the dwell time and the conventional histogram of either the amplitude or the $I_{ex\%}$ were constructed. In both cases, the bin value was set such that the distribution within the histogram would resemble a Gaussian distribution as much as possible. The values for the log(dwell time) and either the amplitude or the I_{ex} were established by fitting a Gaussian function to the histogram, whose μ is either log(dwell time) or the amplitude/$I_{ex\%}$. The dwell time and $I_{ex\%}$ are reported from triplicate measurements, with the values representing the average obtained from three individual repeats and the corresponding SD.

Data, Materials, and Software Availability. Electrophysiology data have been deposited in Zenodo (DOI: 10.5281/zenodo.13356734) (55). The crystal structure of the CytK mutant was deposited in the PDB (56) database (8RJ8). All study data are included in the article and/or Supporting Information.

ACKNOWLEDGMENTS. We thank Gregor Anderluh for kindly providing a plasmid containing lysenin.

1. D. Stoddart, A. J. Heron, E. Mikhailova, G. Maglia, H. Bayley, Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 7702–7707 (2009).
2. T. Z. Butler, M. Pavlenok, I. M. Derrington, M. Niederweis, J. H. Gundlach, Single-molecule DNA detection with an engineered MspA protein nanopore. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 20647–20652 (2008).
3. C. Wloka, N. L. Mutter, M. Soskine, G. Maglia, Alpha-helical fragaceatoxin C nanopore engineered for double-stranded and single-stranded nucleic acid analysis. *Angewandte Chemie Int. Edition* **55**, 12494–12498 (2016).
4. M. Pastoriza-Gallego *et al.*, Evidence of unfolded protein translocation through a protein nanopore. *ACS Nano* **8**, 11350–11360 (2014).
5. G. Oukhaled *et al.*, Unfolding of proteins and long transient conformations detected by single nanopore recording. *Phys. Rev. Lett.* **98**, 158101 (2007).
6. M. Pastoriza-Gallego *et al.*, Dynamics of unfolded protein transport through an aerolysin pore. *J. Am. Chem. Soc.* **133**, 2923–2931 (2011).
7. B. Cressiot *et al.*, Dynamics and energy contributions for transport of unfolded pertactin through a protein nanopore. *ACS Nano* **9**, 9050–9061 (2015).
8. D. Rodriguez-Larrea, H. Bayley, Protein co-translocational unfolding depends on the direction of pulling. *Nat. Commun.* **5**, 4841 (2014).
9. L. Yu *et al.*, Unidirectional single-file transport of full-length proteins through a nanopore. *Nat. Biotechnol.* **41**, 1130–1139 (2023), 10.1038/s41587-022-01598-3.
10. A. Sauciuc, B. Morozzo della Rocca, M. J. Tadema, M. Chinappi, G. Maglia, Translocation of linearized full-length proteins through an engineered nanopore under opposing electrophoretic force. *Nat. Biotechnol.* **42**, 1275–1281 (2023), 10.1038/s41587-023-01954-x.
11. R. C. A. Versloot, S. A. P. Straathof, G. Stouwie, M. J. Tadema, G. Maglia, **B**-barrel nanopores with an acidic-aromatic sensing region identify proteinogenic peptides at low pH. *ACS Nano* **16**, 7258–7268 (2022).
12. S. Raffy, N. Sassoon, M. Hofnung, J. Betton, Tertiary structure-dependence of misfolding substitutions in loops of the maltose-binding protein. *Protein Sci.* **7**, 2136–2142 (1998).
13. S. Y. Chun, S. Strobel, P. Bassford, L. L. Randall, Folding of maltose-binding protein. Evidence for the identity of the rate-determining step in vivo and in vitro. *J. Biol. Chem.* **268**, 20855–20862 (1993).
14. E. A. Manrao *et al.*, Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat. Biotechnol.* **30**, 349–353 (2012).
15. H. Brinkerhoff, A. S. W. Kang, J. Liu, A. Aksimentiev, C. Dekker, Multiple rereads of single proteins at single–amino acid resolution using nanopores. *Science* **1979**, 1509–1513 (2021).
16. Z. Chen *et al.*, Controlled movement of ssDNA conjugated peptide through Mycobacterium smegmatis porin A (MspA) nanopore by a helicase motor for peptide sequencing application. *Chem. Sci.* **12**, 15750–15756 (2021).
17. S. Yan *et al.*, Single molecule ratcheting motion of peptides in a mycobacterium smegmatis Porin A (MspA) nanopore. *Nano Lett.* **21**, 6703–6710 (2021).
18. I. C. Nova *et al.*, Detection of phosphorylation post-translational modifications along single peptides with nanopores. *Nat. Biotechnol.* **42**, 710–714 (2023), 10.1038/s41587-023-01839-z.
19. A. Bogard *et al.*, The ionic selectivity of Lysenin channels in open and sub-conducting states. *Membranes (Basel)* **11**, 897 (2021).
20. H. Ouldali *et al.*, Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore. *Nat. Biotechnol.* **38**, 176–181 (2020).
21. G. Stirnemann, D. Giganti, J. M. Fernandez, B. J. Berne, Elasticity, structure, and relaxation of extended proteins under force. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3847–3852 (2013).
22. F. Gräter, P. Heider, R. Zangi, B. J. Berne, Dissecting entropic coiling and poor solvent effects in protein collapse. *J. Am. Chem. Soc.* **130**, 11578–11579 (2008).
23. P. Cifra, Z. Benková, T. Bleha, Persistence lengths and structure factors of wormlike polymers under confinement. *J. Phys. Chem. B* **112**, 1367–1375 (2008).
24. S. Vaitheeswaran, D. Thirumalai, Interactions between amino acid side chains in cylindrical hydrophobic nanopores with applications to peptide stability. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17636–17641 (2008).
25. M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, H. E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by afm. *Science* **1979**, 1109–1112 (1997).
26. M. Carrion-Vazquez *et al.*, The mechanical stability of ubiquitin is linkage dependent. *Nat. Struct. Mol. Biol.* **10**, 738–743 (2003).
27. J. N. Bright, T. B. Woolf, J. H. Hoh, Predicting properties of intrinsically unstructured proteins. *Prog Biophys. Mol. Biol.* **76**, 131–173 (2001).
28. G. Yang *et al.*, Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 139–144 (2000).
29. H. Dietz, M. Rief, Protein structure by mechanical triangulation. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1244–1247 (2006).
30. F. Oesterhelt *et al.*, Unfolding pathways of individual bacteriorhodopsins. *Science* **1979**, 143–146 (2000).
31. B. S. Hanson, D. Head, L. Dougan, The hierarchical emergence of worm-like chain behaviour from globular domain polymer chains. *Soft Matter* **15**, 8778–8789 (2019).
32. J. E. Kohn *et al.*, Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12491–12496 (2004).
33. M. Daoud, P. G. De Gennes, Statistics of macromolecular solutions trapped in small pores. *J. de Phys.* **38**, 85–93 (1977).
34. L. Dai, C. B. Renner, P. S. Doyle, The polymer physics of single DNA confined in nanochannels. *Adv. Colloid Interface Sci.* **232**, 80–100 (2016).
35. S. Daoudi, F. Brochard, Flows of flexible polymer solutions in pores. *Macromolecules* **11**, 751–758 (1978).
36. C. T. A. Wong, M. Muthukumar, Polymer capture by electro-osmotic flow of oppositely charged nanopores. *J. Chem. Phys.* **126**, 164903 (2007).
37. J. Bitinaite *et al.*, USERTM friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Res.* **35**, 1992–2002 (2007).
38. A. M. Cavaleiro, S. H. Kim, S. Seppälä, M. T. Nielsen, M. H. H. Nørholm, Accurate DNA assembly and genome engineering with optimized uracil excision cloning. *ACS Synth Biol.* **4**, 1042–1046 (2015).
39. M. H. Nørholm, A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol.* **10**, 21 (2010).
40. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
41. A. J. McCoy *et al.*, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
42. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
43. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
44. E. F. Pettersen *et al.*, UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
45. O. S. Smart, J. M. Goodfellow, B. A. Wallace, The pore dimensions of gramicidin A. *Biophys. J.* **65**, 2455–2460 (1993).
46. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).
47. C. R. Harris *et al.*, Array programming with NumPy. *Nature* **585**, 357–362 (2020).
48. J. D. Hunter, Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
49. M. A. Lomize, I. D. Pogozheva, H. Joo, H. I. Mosberg, A. L. Lomize, OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).
50. M. Bokori-Brown *et al.*, Cryo-EM structure of lysenin pore elucidates membrane insertion by an aerolysin family protein. *Nat. Commun.* **7**, 11293 (2016).

51. M. Faller, M. Niederweis, G. E. Schulz, The structure of a mycobacterial outer-membrane channel. *Science* **1979**, 1189–1192 (2004).
52. I. Iacovache *et al.*, Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process. *Nat. Commun.* **7**, 12062 (2016).
53. B. Webb, A. Sali, Comparative protein structure modeling using modeller. *Curr. Protoc Bioinformatics* **54**, 5.6.1–5.6.37 (2016).
54. G. Maglia, A. J. Heron, D. Stoddart, D. Japrung, H. Bayley, Analysis of single nucleic acid molecules with protein nanopores. *Methods Enzymol.* **475**, 591–623 (2010).
55. A. Sauciuc, Data from "Blobs form during the single-file transport of proteins across nanopores" Zenodo. https://doi.org/10.5281/zenodo.13356734. Deposited 21 August 2024.
56. J. J. Whittaker, A. Sauciuc, A. Guskov, PDB structure "CytK nanopore mutant", Protein Data Bank. https://www.rcsb.org/structure/unreleased/8RJ8. Deposited 20 December 2023.