ARTICLE

Deep Learning-Assisted Single-Molecule Detection of Protein Post-translational Modifications with a Biological Nanopore

Chan Cao,* Pedro Magalhães,# Lucien F. Krapp,# Juan F. Bada Juarez, Simon Finn Mayer, Verena Rukes, Anass Chiki, Hilal A. Lashuel,* and Matteo Dal Peraro*
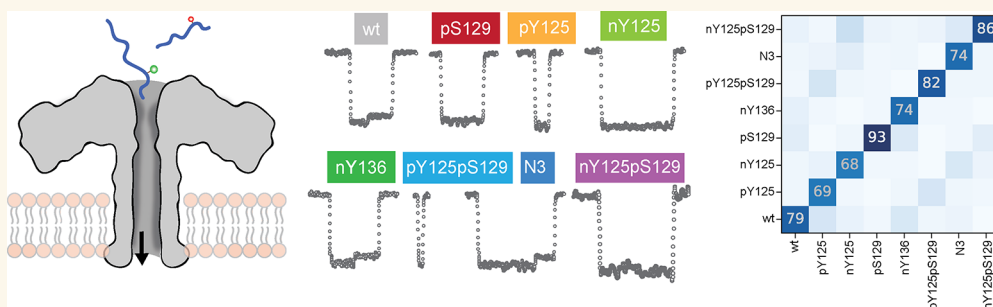
Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Protein post-translational modifications (PTMs) play a crucial role in countless biological processes, profoundly modulating protein properties on both spatial and temporal scales. Protein PTMs have also emerged as reliable biomarkers for several diseases. However, only a handful of techniques are available to accurately measure their levels, capture their complexity at a single molecule level, and characterize their multifaceted roles in health and disease. Nanopore sensing provides high sensitivity for the detection of low-abundance proteins, holding the potential to impact single-molecule proteomics and PTM detection, in particular. Here, we demonstrate the ability of a biological nanopore, the pore-forming toxin aerolysin, to detect and distinguish α-synuclein-derived peptides bearing single or multiple PTMs, namely, phosphorylation, nitration, and oxidation occurring at different positions and in various combinations. The characteristic current signatures of the α-synuclein peptide and its PTM variants could be confidently identified by using a deep learning model for signal processing. We further demonstrate that this framework can quantify α-synuclein peptides at picomolar concentrations and detect the C-terminal peptides generated by digestion of full-length α-synuclein. Collectively, our work highlights the advantage of using nanopores as a tool for simultaneous detection of multiple PTMs and facilitates their use in biomarker discovery and diagnostics.

KEYWORDS: biological nanopores, protein post-translational modifications, deep-learning, single-molecule sensing, α-synuclein

P roteins are the major molecular building blocks of life, dictating the spatial and temporal occurrence of most biological functions.[1] Their functional properties are mainly determined by their native three-dimensional structure, subcellular localization, and interactome. Post-translational modifications (PTMs) play an important role in regulating all these properties of proteins and thus represent molecular switches for coordinating, diversifying, and regulating cellular function.[2] Moreover, the majority of protein-based biopharmaceuticals approved or in clinical trials bear some form of PTMs in order to enhance their efficacy for therapeutic applications.[3] Despite the impressive recent advances in biochemistry and structural biology, the development of

sensitive methods to detect PTMs, generate site-specifically modified proteins, and decipher the PTM code of proteins still lags behind.

The standard methods currently available for the detection of PTMs are mass spectrometry (MS), and enzyme-linked

antibody-based assays,[4,5] while immunoprecipitation coupled to MS (IP/MS) has also been increasingly used to detect and quantify low abundant protein species. Recently, great progress has been made toward developing and improving MS methods and immunoassays to map, detect, and quantify modified proteins in biological samples. For instance, proximity extension assays[6] and single molecule array (Simoa)[7] have shown high sensitivity for detecting PTMs. One of the major drawbacks of these immunoassays is that they are based on antibodies that are developed against single PTMs. Single-molecule techniques such as electron tunneling have also achieved the detection of single PTM on peptides (i.e., phosphorylation) with high sensitivity.[5,8,9] However, the development and optimization of these methods did not account for the presence of multiple and/or neighboring PTMs.[5] One example is demonstrated in our recent work, which shows that neighboring PTMs interfere with the detection of phosphorylated α-synuclein at S129 by most pS129 antibodies.[10] Therefore, a method to accurately identify protein PTMs is essential as increasing evidence suggests the PTM code is combinatorial in nature and involves complex interplay and cross-talk among PTMs.[11,12]

PTMs have been implicated in a wide variety of diseases, such as characterizing the changes in protein PTM types and levels is highly relevant for early diagnosis and monitoring the progression of diseases, as well as for evaluating the efficacy of emerging therapies.[4,13] For example, several neurodegenerative diseases including Alzheimer's disease (AD) and Parkinson's disease (PD) are caused by the accumulation of misfolded and aggregated proteins in the brain regions that are affected by the disease (e.g., amyloid-β in amyloid plaques, Tau in neurofibrillary tangles, and α-synuclein in Lewy bodies and Lewy neurites). One shared characteristic among these proteins is that their aggregated forms are heavily modified. PTMs have emerged as key signatures of disease pathologies and are commonly used as the primary (bio)markers of disease progression and pathology formation, spreading and clearance in response to therapies.[14,15]

Nanopore sensing is an approach based on ionic current readout that can detect a single molecule as it is passing through a nanometer scale pore.[16] Such pores are either biological assemblies of proteins embedded in a lipid membrane or are fabricated by a solid-state material.[17] When a molecule of interest passes through such a pore, the electric current signal is modulated and exquisitely sensitive to the molecule of interest and thus can provide information about its size, mass, charge, composition, structure and conformation in real-time.[18,19] Nanopore sensing has achieved sequencing of ultralong DNA,[20,21] and this success has inspired its application for peptide and protein analyses.[22−25] Compared with MS, nanopore sensing is faster, cheaper, free of ionization, and has an ultrasensitive detection of biomolecules in solution at the single-molecule level. These distinguishing features, in particular its sensitivity, make nanopores perfectly suited for protein analysis since there is no biochemical method available for protein amplification, like polymerase chain reaction, that has been widely used for DNA amplification. Recently, several groups have explored the potential of using nanopore approaches to detect protein PTMs, including phosphorylation,[26−28] acetylation,[29] propionylation,[30] glycosylation,[31,32] and ubiquitination.[33] However, these studies focus on identifying either different types of modifications or the same modification at different positions but do not consider the

simultaneous occurrence of different types and combinations of PTM species that are of clear clinical relevance. Moreover, the detection of low-abundance protein PTMs from clinical samples is still poorly investigated.

Here, we demonstrate that an engineered form of aerolysin, a founding number of a major class of β-pore-forming toxins which has shown an excellent sensing capability for DNA and peptide analysis,[34−37] can be used to detect peptides containing single or multiple types of PTMs. As a model system, we used peptides derived from the C-terminal domain of the presynaptic protein α-synuclein, the primary constituent of the pathological hallmarks of PD (i.e., Lewy bodies and Lewy neurites). This region is characterized by the clustering of different types of PTMs (e.g., phosphorylation, nitration and oxidation) that are associated with pathology formation and disease progression in PD and other neurodegenerative diseases.[38] Recent studies also suggest that specific α-synuclein PTMs (i.e., pS129) are elevated in the cerebral spinal fluid (CSF) of PD patients and could be developed into diagnostic biomarkers for PD and potentially for other neurodegenerative diseases.[39,40]

We observe that phosphorylation, nitration, and oxidation significantly modulate the current amplitude and dwell time of the signals in comparison with the wild-type (wt) α-synuclein. Additionally, we show that the α-synuclein peptides bearing multiple PTMs and different PTMs combinations induce characteristic ionic current signatures, which provides an opportunity to investigate PTMs crosstalk and monitor the PTMs dynamics in the future. A deep learning approach was specifically developed to process the signals, providing an automatic and fast way to classify the current readouts associated with each PTM combination. We also demonstrate the ability of aerolysin pores to detect unmodified α-synuclein peptides when mixed with red blood cells (RBCs) at picomolar concentration as well as the possibility of using cathepsin D to generate the specific C-terminal peptides that bear most PTMs, providing promising ground for actual clinical applications. In summary, our work presents a single-molecule approach able to detect clinically relevant low-abundance α-synuclein PTMs. Compared to other analytical methods like MS and ELISA,[41,42] this nanopore-based approach offers detection at single-molecule resolution and provides promising ground to decipher PTMs patterns relevant for disease diagnostics.

RESULTS AND DISCUSSION

Characterization of the C-Terminal Region of α-Synuclein with a Nanopore. The ability of aerolysin to detect PTMs was examined by means of single-channel recording experiments, as shown in Figure 1a. A lipid membrane separates the chamber into two compartments, *cis* and *trans*, and holds an aerolysin pore that connects them. An engineered variant of aerolysin is used here, namely K238A, where lysine residues at position 238 are substituted by alanine residues, providing a significantly enhanced resolution for biomolecular sensing compared to wt aerolysin.[43] As mentioned above, we focused on α-synuclein as a model system because of the strong links between its PTM profile and the brain pathology of several neurodegenerative diseases, most notably in PD.[38,44]

The C-terminal domain of α-synuclein, encompassing the residues 124−140 (α-syn$_{124−140}$) (Figure 1b) harbors several PTMs that are found in pathological α-synuclein aggregates in
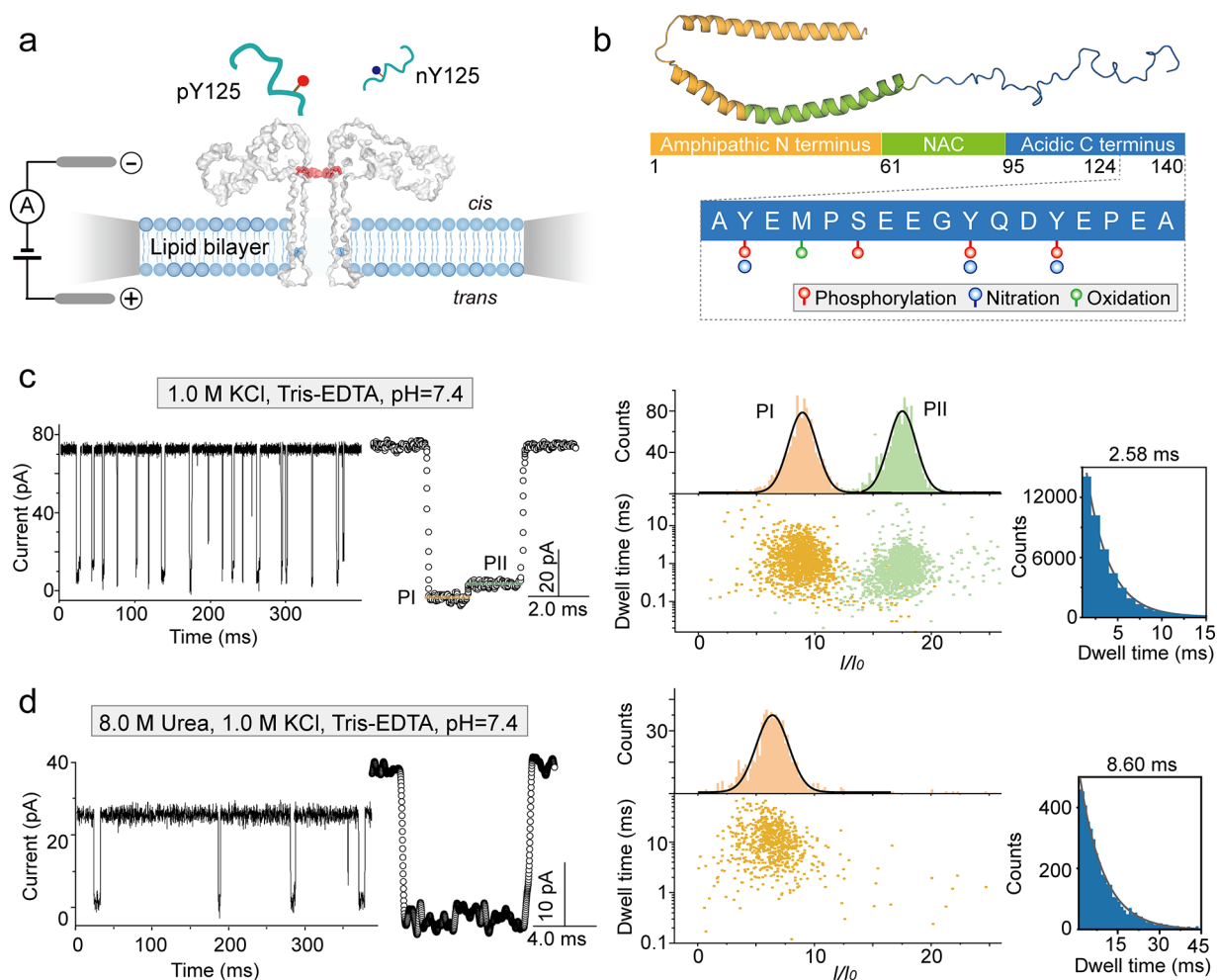
Figure 1. Characterization of wild-type α-synuclein C-terminal peptides. (a) Illustration of the single-channel recording setup composed of two chambers separated by a lipid membrane that is formed across an orifice; the chambers are named *cis* and *trans*. Voltage is applied across the pore using two Ag/AgCl electrodes. **(b)** The structure of α-synuclein in full-length (PDB: 1xq8), including the amphipathic N-terminus (yellow), nonamyloid-component (NAC) domain (green), and the disordered C-terminal fragment (blue). Phosphorylation (red circle), nitration (blue circle), and oxidation (green circle) of residues 124−140 are highlighted. Nanopore single channel recording of wt α-syn$_{124-140}$ peptide in 1.0 M KCl **(c)** and in 1.0 M KCl, 8.0 M urea **(d)** solutions, respectively, both buffered with 10 mM Tris, and 1.0 mM EDTA at pH 7.4. Left panels of (c) and (d) report raw current traces, middle panels typical events, while the right panels show scatter plots, I/I_0 percentage, and dwell time histograms. All data were obtained by applying a voltage of +100 mV.

the brain of patients with PD and other neurodegenerative diseases.[44] Many of these PTMs, including phosphorylation and nitration, have emerged as reliable markers of α-synuclein pathology formation in human brains and animal models of PD and related synucleinopathies.[38] The C-terminal domain of α-synuclein is rich in proline residues, it is highly negatively charged and does not adopt a stable secondary structure in the monomeric state of the protein,[45] which in this context is particularly convenient as these peptides can be easily driven to translocate through the nanopore by the potential applied across the lipid membrane. All peptides were prepared using Fmoc-based solid-phase peptide synthesis and purified as previously reported[46−48] (see Methods).

α-Syn$_{124-140}$ was added into the *cis* compartment and when a positive voltage was applied to the *trans* compartment, clear and reproducible blockades of ionic current were obtained (Figure 1c,d). For wt α-syn$_{124-140}$, a peculiar and well-recognizable 2-level blockade event was observed and the signal often contained a higher residual current at the last fraction (Figure 1c, middle). As a result, two populations were observed (Figure 1c, right): one exhibited a relative current percentage (I/I_0, see Methods) of 9.0 ± 2.0, while the other 17.1 ± 2.0 (named hereafter *PI* and *PII*, respectively). These values correspond to the mean and standard deviation derived from a Gaussian fit of the relative current histogram.

To gain further insights into this peculiar two-level signal, we considered the structure of this α-synuclein fragment. Based on previous studies,[49] there is a highly hydrophobic segment between residues 125−129 that can produce a structurally compact cluster likely responsible for the initial lower residual current. To test this hypothesis, we reasoned that destabilizing this conformation by the addition of urea could produce signals expected for a more extended peptide.[49] We thus used 8.0 M urea (along with 1.0 M KCl, 10 mM Tris, 1.0 mM EDTA, and pH 7.4) to induce complete unfolding of the peptide, while aerolysin remained folded and functional (Supplementary Figure S1). Under these conditions, the open pore current was lower (25 ± 1.8 pA at +100 mV) compared to the same salt concentration without urea (72 ± 1.5 pA at +100 mV). This is because the additional urea

significantly decreased the mobility of the ions.[50,51] However, in these conditions, wt α-syn$_{124-140}$ translocated with a simple one-level signal as evidenced by one population in the scatter plot and the I/I_0 histograms (Figure 1d). These one-level current signals suggest that the addition of urea results in a more extended conformation of the polypeptide chain, which facilitates its translocation through the pore. In addition, due to the decrease of ionic mobility, the dwell time became much longer, 8.6 ± 0.24 ms, approximately 4 times longer than the condition without urea. These results demonstrate that the conformation of largely disordered peptides such as those obtained from α-synuclein can be monitored at the single-molecule level using a nanopore.

Deep Learning-Assisted Detection of Different PTMs and Their Combinations. Next, we assessed the feasibility of using nanopores for detecting PTMs of α-syn$_{124-140}$ peptide. We investigated different types of modifications occurring at the same amino acid (i.e., phosphorylation and nitration of Y125, pY125 and nY125), or the same type of PTM located at different positions (i.e., pS129 and pY125, nY125 and nY136). Moreover, we considered peptides bearing multiple PTMs of the same type (i.e., pY125pS129 and nY125nY133nY136) and of different types (i.e., nY125pS129).

As shown in Figure 2a, the current signatures of the different PTM combinations appeared to all have distinct current characteristics. Compared to the unmodified peptide, the pY125 and nY125 α-syn$_{124-140}$ peptide presented only one level and their relative current were lower than the wt peptide (Supplementary Figure S2). This decrease in the relative current value is consistent with a deeper blockade induced by the additional volume due to the presence of the PTMs. Moreover, the dwell time varied significantly with different types of PTMs, the fitted values being 2.58 ± 0.4 ms for wt, 0.55 ± 0.08 ms for pY125, and 4.51 ± 0.5 ms for nY125, indicating that phosphorylation speeds up the translocation process, while nitration significantly slows it down. The faster translocation of pY125 could be induced by the additional negatively charged phosphate group. In addition, we investigated the effect of oxidation of methionine at position 127 of α-syn$_{124-140}$ (oM127, Supplementary Figure S3). Similar to pY125 and nY125, oM127 also showed only one current level.

When the same PTM occurs at different positions in the peptide sequence, the ionic current and dwell time were modulated differently. For example, the dwell time of pS129 (3.62 ± 0.2 ms) is 6.5 times longer than that of pY125 (0.55 ± 0.08 ms). We hypothesize that phosphorylation at Y125 (pY125) disrupts the hydrophobic cluster of α-syn$_{124-140}$,[49] all amino acids become more easily exposed and translocate in a linear fashion. In this case, the additional phosphate group of pY125 contributes to faster translocation under the applied voltages. Similar results were obtained when we compared nY125 and nY136 α-syn$_{124-140}$ (Supplementary Figures S2b and S4b). Unlike nY125, which only showed one population, the relative current of nY136 was less pronounced compared to that of wt and showed two populations like wt α-syn$_{124-140}$. This suggests that disrupting the local structure facilitates the detection of PTMs in proteins by the aerolysin pore. As reported recently, this could be also achieved by using chemical denaturants such as urea or guanidinium chloride.[52]

To explore the possibility of using the nanopore for the detection of multiple PTMs occurring simultaneously on the same peptide, we measured the ionic current response of

double-phosphorylated α-syn$_{124-140}$, pY125pS129, and triple nitrated α-syn$_{124-140}$, nY125nY133nY136 (Supplementary Figure S4c,d). Here, the nanopore results showed that compared to the unmodified peptide (wt) and the singly phosphorylated peptides (i.e., pY125 and pS129), the translocation speed of pY125pS129 (0.45 ± 0.02 ms) was even faster, likely due to its increased negative charge. While the dwell time of nY125nY133nY136 was around 2-fold longer (5.18 ± 0.23 ms) compared to wt and nY136, and slightly longer than nY125: its relative current was the lowest among all peptides (7.2 ± 1.0) since the three modifications contribute to increasing the overall volume of the peptide and therefore induce a deeper blockade of ionic current. In the case of the peptides containing two types of PTMs at different positions, i.e., nitration at Y125 and phosphorylation at S129 (nY125pS129, Supplementary Figure S4e), only one population was observed. The relative current of nY125pS129 was between the values of the single modifications nY125 and pS129, as observed also for the width of the relative current distribution. The dwell time of nY125pS129 was slightly longer than that of pS129, but identical with nY125. For all of these peptides, the dwell time decreases as the voltage increases, indicating that the collected signals are indeed induced by peptides translocating through the nanopore. Altogether, these observations demonstrate that the engineered K238A aerolysin can capture the diversity of PTMs.

To classify these PTMs in a more precise, unbiased, and automatic way that would allow a translation to clinical applications in the future, we developed a tailored deep learning approach for processing the nanopore current readouts. A long short-term memory (LSTM) recurrent neural network was integrated to read the local extrema of events followed by a multilayer perceptron (MLP) to classify the peptides (Figure 2b), a pipeline which is similar to our previous analysis of informational polymers[53] (see details in Methods). We first detected the events with a cutoff threshold at 3σ from the open pore current and then computed statistical features (i.e., mean residual current, standard deviation, and dwell time) and extracted the local extrema of the events (Supplementary Figure S5 and Methods). For each single peptide, we randomly used 75% of the recorded data to feed the deep learning algorithm and train the model; the remaining 25% was used as validation set to compute the accuracy of the model for each single peptide. The events are filtered using a threshold on the predictions confidence of the model which is translated into a selection percentage.

As shown in Figure 2c, this approach allows for differentiating wt α-syn$_{124-140}$ from peptides that contain different PTMs at the same position Y125 (i.e., pY125 and nY125) with an accuracy of 94% at a 50% selection percentage. Control experiments with a mixture of wt and pY125 at different concentration ratios were performed to further test the deep learning model (Figure 2d). First, a mixture of an equimolar ratio (1:1) was measured in nanopore experiments and the percentage of assignment of wt and pY125 was 52.8 ± 3.5% and 43.2 ± 3.1%, respectively. This is in line with the theoretical predictions. When the ratio of wt:pY125 was changed to 1:3, the percentage of wt assignment decreased to 19.0 ± 0.3% while pY125 increased to 78.2 ± 0.7%. Finally, a mixture ratio of 3:2 (wt:pY125) was tested. As shown in Figure 2d, the assignments of wt and pY125 were as expected 61.5 ± 0.1% and 35.0 ± 0.2%, respectively. Therefore, the expectation is that any sequence in the library of peptides (wt, pY125, or
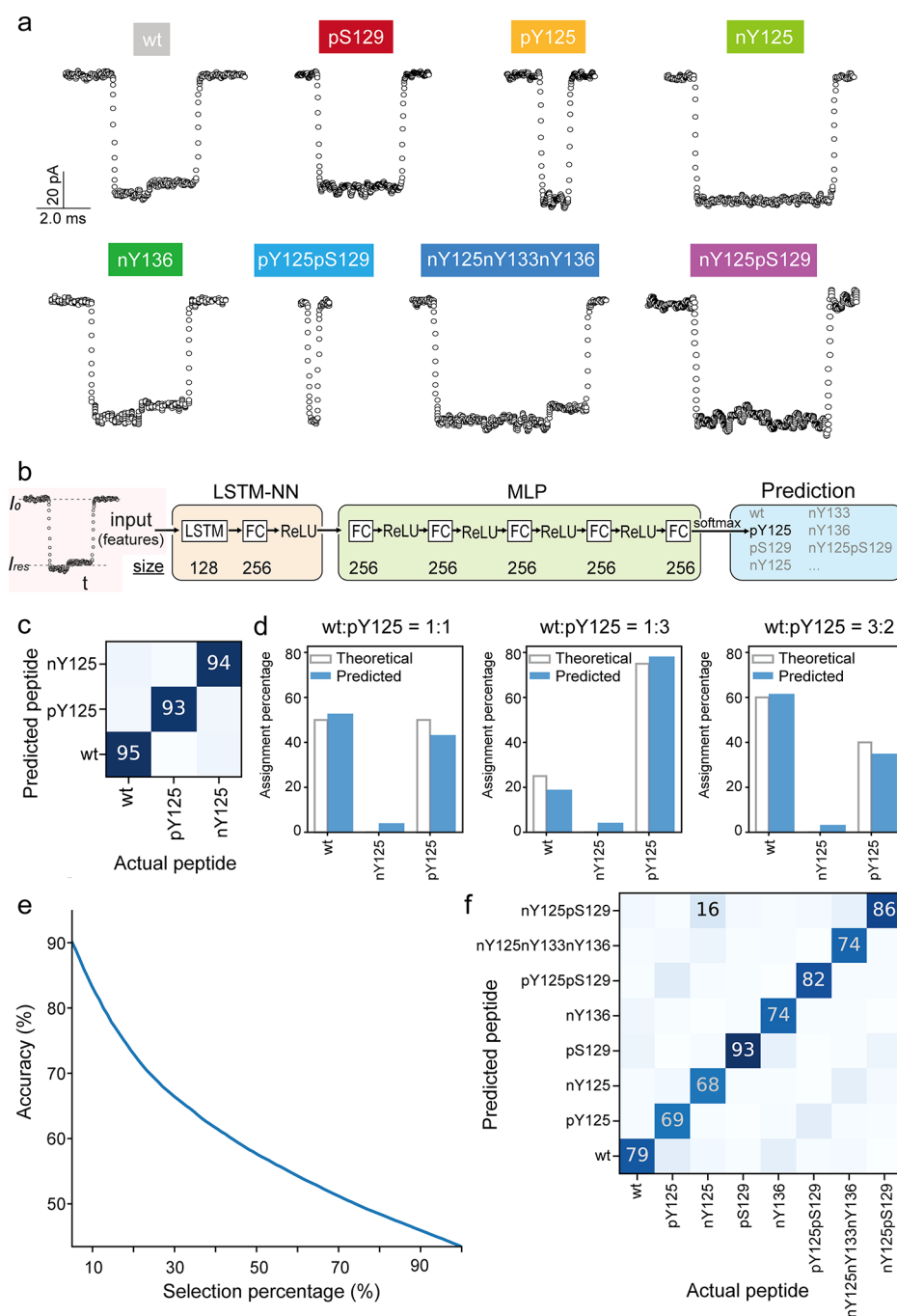
Figure 2. Discrimination of α-synuclein peptide PTMs by deep learning. (a) Typical ionic current signals of different PTM types and their spatial combinations. The scale shown in the left top panel applies to all of the events. (b) Deep learning approach to signal processing: an LSTM recurrent neural network was used to read the events, followed by an MLP to predict the peptides. (c) The normalized confusion matrix over the predictions shown as percentage of wt, pY125, and nY125 peptides classification using a deep learning approach. (d) Assignment percentage of different mixture samples of wt and pY125 at a concentration ratio of 1:1 (left), 1:3 (middle), and 3:2 (right). The theoretical accuracy is shown by the white columns while the predicted accuracy is represented by blue columns. (e) Selection percentage versus averaged accuracy obtained from deep learning approach of peptides, including wt, pY125, nY125, pS129, pY125pS129, nY125nY133nY136, and nY125pS129. (f) The normalized confusion matrix over the predictions shown as percentage of wt, pY125, nY125, pS129, pY125pS129, nY125nY133nY136, and nY125pS129 classification. Columns represent actual peptides from the test set, while rows represent the peptides that the deep learning algorithm assigned them to. All data were obtained using 1.0 M KCl, 10 mM Tris, and 1.0 mM EDTA buffer at pH 7.4 by applying a voltage of +100 mV.

nY125) can be identified directly with a probability of 94%. Furthermore, mixture experiments of wt and pS129 were performed (Supplementary Figure S6), for which the deep learning model could predict the different mixture ratios with a probability of 96%.

Data from wt and the 7 tested α-syn$_{124-140}$ PTM variants (i.e., pY125, nY125, pS129, nY136, pY125pS129, nY125-nY133nY136, and nY125pS129) have been used as a training set for the same deep learning model, with a resulting accuracy of 78.2% using a selection percentage of 25% (i.e., accuracy
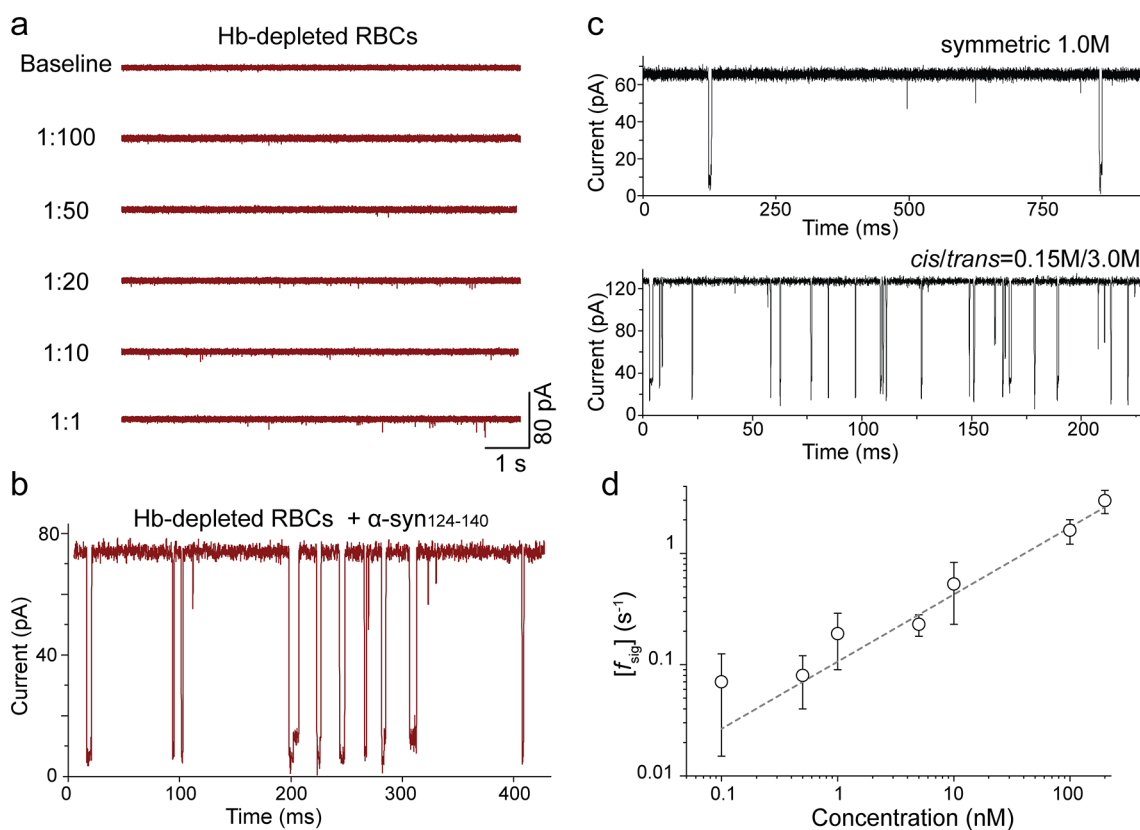
Figure 3. Detection of α-synuclein peptides at low abundance concentration. (a) Raw current traces after addition of serial dilutions of Hb-depleted RBC samples. For 1:1 condition, 3 μL of red blood cell lysates devoid of hemoglobin were added to the 150 μL nanopore chamber (the total protein content is 0.4−0.5 μg/μL). (b) Raw current traces of Hb-depleted RBCs sample in the presence of 6 μM wt α-syn$_{124−140}$ peptides. The RBCs used here is 1:10 dilution. (c) Raw current traces showing signals of α-syn$_{124−140}$ in symmetric salt (top, 1.0 M KCl) and in a gradient salt concentration (bottom, 0.15 M/3.0 M (*cis*/*trans*) KCl). (d) The correlation between signal frequency f_{sig} and the concentration of wt α-syn$_{124−140}$ in gradient salt conditions ranging from 100 pM to 200 nM peptide in the *cis* chamber under +100 mV applied voltage. The error bars represent standard deviations from at least three independent experiments under the same conditions.

dependence on selection percentage as shown in Figure 2e). If the event is too noisy or does not contain enough information as in the case of short events, the prediction confidence would be lower. The selection percentage allows for a trade-off between accuracy and filtering without requiring heavy filtering during the preprocessing of the events. As illustrated in Figure 2f, 13% of pY125 was confused with pY25pS129, because the signals of both pY125 and pY25pS129 showed very short dwell times. This could be further improved if higher-bandwidth instruments have been employed. Additionally, 6% of nY125 was confused with nY125nY133nY136, while the other 16% of this peptide was confused with nY125pS129, 10% of nY136 was confused with pS129 and the other 9% was confused with wt. We think this prediction confusion was caused by the similar dwell times between these modified peptides. Nonetheless, the deep learning approach we developed here provides high reading accuracy for all of the tested PTMs peptides, particularly for pS129 which is one of the most relevant biomarkers for synucleinopathies. This is encouraging also because a similar deep learning model can be trained for other systems, thus finding general applicability for signal processing in nanopore sensing experiments.

Detection of α-Synuclein Peptides at Low Abundance Concentration. To date, most nanopore sensing applications are performed using purified peptides or protein samples (as done hereinabove). There is however the need to develop methodologies whereby detection can be performed on biological samples.[54] Nevertheless, even though some nanopores have been used for the detection of biomolecules in biological fluids samples (e.g., plasma,[55,56] serum,[54,57−61] sweat,[62] urine,[62] and saliva[62]), the task of measuring biomarker peptides is often challenging due to the intrinsic low sensitivity and selectivity of the techniques applied to bodily fluids.

We previously showed that aerolysin pores possess high sensitivity for biomolecular sensing, especially its engineered variants.[43,63] Therefore, we explored if K238A aerolysin was able to detect α-synuclein peptides in clinically relevant conditions, red blood cells (RBCs) devoid of hemoglobin (Hb) proteins. We first determined the dilution necessary to ensure a stable baseline of the ionic current through serial dilutions of Hb-depleted RBC samples. In the literature, no particular instabilities of the lipid bilayer were observed when testing RBCs except when dilutions were lower than 1:50.[62] Based on these previous reports, we scanned a range of dilutions up to 1:100 (Figure 3a) in our setting and observed that after the addition of 3 μL of Hb-depleted RBCs extracts into the *cis* chamber, no obvious effects were detected for the membrane stability. Interestingly, during the experiments with Hb-depleted RBCs, we could continually collect data for up to 3 h, indicating that the aerolysin pore remained invariantly intact likely due to the ultrastability provided by its double β-barrel structure.[35] Moreover, when the concentration of Hb-depleted RBCs samples increased in the chamber, occasional signals were also observed (Supplementary Figure S7).
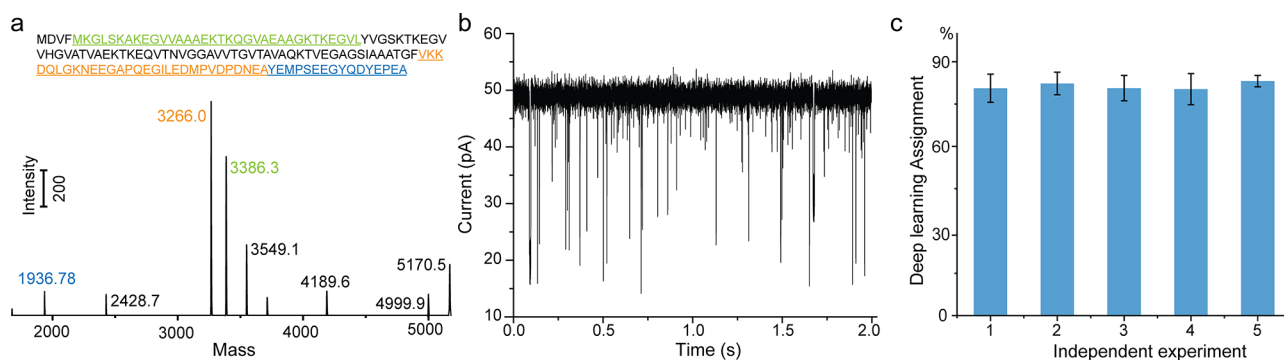
Figure 4. Detection of C-terminal peptides of α-synuclein after CtsD digestion. (a) Possible peptide fragments from CtsD digestion based on previous studies[66] and MS characterization of CtsD digestion of full-length α-synuclein. (b) Raw current trace from nanopore measurement of CtsD-digested samples. (c) Deep learning result from 5 independent experiments aimed at classifying α-syn$_{125-140}$ from the mixture sample.

Notably, all raw current traces shown in Figure 3 were collected 30 min after samples were added.

Next, we sought to quantify the presence of α-synuclein wt peptide within the Hb-depleted RBCs, with the aim of mimicking the real complexity of pathological conditions within a clinical setting. Based on the previous tests, dilutions of 1:10 were chosen as the background for Hb-depleted RBCs extracts. As illustrated in Figure 3b, after the addition of wt α-syn$_{124-140}$, typical two-level signals were recorded.

While these results represent the potential of aerolysin pores to detect C-terminal α-synuclein peptide, another prerequisite toward the development of aerolysin pores into an efficient single-molecule proteomic device for biomarker detection is the ability to detect molecules in low abundance, as often found in clinical samples. In previous studies, the frequency of signals has been used to quantify the detected molecule. Therefore, we used the equation $f_{sig} = k_{on} [\text{PTMs}]_0$ to explore the theoretical detection limit of aerolysin nanopores for sensing α-synuclein peptides. As illustrated in Supplementary Figure S8, we verified that the frequency of blockade events, f_{sig}, was proportional to the concentration of α-synuclein ranging from 120 nM to 24 μM, when +100 mV were applied using a symmetric buffer of 1.0 M KCl. The physiological concentration of α-synuclein in RBCs is however much lower than these tested conditions, namely 26.2 $\pm$ 3.0 μg/mL,[64,65] which corresponds to 35 nM if 3 μL of the original sample was added into the nanopore system. Therefore, to enhance the detection limit of our setting, we measured peptides in a gradient of salt concentration, which was shown to be an efficient way to increase the capture rate for ssDNA.[56,66] An asymmetrical KCl buffer solution consisting of 0.15 M in the *cis* chamber and 3.0 M in the *trans* was thus used (Figure 3c), which enabled the detection of wt α-synuclein peptides at far lower concentrations (i.e., ~100 pM, Figure 3d) than when using canonical symmetric buffer conditions.

Altogether, these results demonstrate that K238A aerolysin nanopores can detect C-terminal peptides of α-synuclein at ~nM concentration (factoring in the dilution conditions), a condition which is overlapping with the concentration at which they are found in a RBC clinical sample. Considering that an even lower detection limit could be achieved by applying higher voltages, minimizing the volume of the chambers, or optimizing the pore variants, the possibility of developing an effective nanopore-based tool for the diagnosis of synucleino-pathies appears within reach.

Detection of α-Synuclein Peptides in Complex Mixtures. Next, we sought to determine if the presence of multiple fragments, produced for instance by proteolytic cleavage of α-synuclein, could influence the nanopore detection of C-terminal peptides bearing the desired PTMs. This aims at mimicking and describing the actual α-synuclein mixtures that one could find in an actual clinical sample. For this purpose, we choose to work with cathepsin D (CtsD) protease as previous studies showed that CtsD-mediated digestion of α-synuclein results in the generation of a C-terminal fragment spanning residues 125−140 (α-syn$_{125-140}$), in addition to four other α-synuclein fragments of various length[67,68] (Figure 4a). The C-terminal fragment contains the great majority of known disease-associated PTM sites (p/nY125, pS129, nY133, nY136), including all the ones we investigate above for α-syn$_{124-140}$. To the best of our knowledge, CtsD is the only enzyme that cleaves α-synuclein and results in the generation of a peptide fragment that contains all the most relevant PTM sites, thus making it the ideal choice for future studies to detect α-synuclein C-terminal peptides bearing single or multiple PTMs. Other enzymes which cleave within the C-terminal domain, e.g., Glu-C, give rise to much shorter peptides, thus precluding the detection of single peptides containing multiple PTMs.[44]

We incubated CtsD with full-length α-synuclein (see details in Methods) and detected the resulting products by nanopore experiments afterward. In parallel, the same sample was characterized by MS and, as shown in Figure 4a, CtsD digestion produced three main peptide fragments, α-syn$_{5-38}$, α-syn$_{95-124}$, and α-syn$_{125-140}$. After the CtsD-digested α-synuclein sample was added into the nanopore system, diverse current signals were obtained (Figure 4b). Under these conditions, the deep learning model was able to identify the C-terminal peptides α-syn$_{125-140}$ with an average accuracy of 81.4 $\pm$ 1.1% from five independent experiments (Figure 4c), implying that the other proteolytic fragments, while having an effect of the current readout do not compromise the ability to detect the α-syn$_{125-140}$ fragment. It should be noted that there is one amino acid difference between peptides used to train the deep learning model and this α-syn$_{125-140}$ fragment; a difference which however does not seem to affect significantly the accuracy of predictions, likely because Ala124 does not contribute in a significant way to modulate the current signal during translocation, while not holding any PTM. Moreover, as α-syn$_{5-38}$ is positively charged, we do not expect it to be

captured by the nanopore when using +100 mV voltage. Therefore, the only possible fragment that may interfere with the measurement is α-syn$_{95-124}$ since it is negatively charged. However, it should be noted that the length of α-syn$_{95-124}$ is two times longer than α-syn$_{125-140}$ which may lead to a low capture rate in the conditions used here. As shown for DNA, the capture of long ssDNA is dramatically reduced in the aerolysin nanopore.[69]

CONCLUSIONS

In this work, we have shown that an engineered aerolysin nanopore can detect and distinguish peptides carrying different types and numbers of PTMs occurring at different residues. In addition, the measurements can capture subtle structural features that are challenging to be characterized by other biophysical methods without fluorescently labeling or modifying the peptides/proteins. Using a deep learning model for signal processing, all investigated PTMs could be automatically identified in a supervised context, which means that this approach can be scaled up to identify more PTMs or scaled down to fit a specific application. Importantly, this nanopore approach can reach a detection limit as low as 100 picomolar concentration and is amenable for high-throughput applications, which are challenging for other techniques such as MS. Finally, one major advantage of this nanopore-based approach is that it enables simultaneous detection of several protein PTMs which is difficult to achieve using immunoassay/ antibody-based methods. This is because the presence of multiple PTMs alters the biochemical properties of the antibody-targeting epitopes.

Altogether, our findings advance the development of biological nanopores into efficient single-molecule proteomic devices and diagnostic tools. We envision that this could be achieved by detecting the circulating peptides directly, or by isolating target proteins from various biological fluids using specific antibodies[70] and then digesting them into smaller peptides and subsequently detecting them through a nanopore as reported in the recent studies.[67,71] In the future, clinical samples could be treated directly with the specific protease, rather than extracting the target proteins, and subsequently detected and classified using nanopores. Several very recent studies have shown that nanopore sensors can be used identify different types of proteins by assessing characteristic blockage event features linked to peptide fragments produced by digestion of the proteins with proteases.[72,73] Such nanopore-based molecular diagnosis platform would hold also promise to detect multiple biomarkers simultaneously, broadly extending the potential field of application, for instance for the detection of biomarker signatures for cancer diagnosis.[74]

The ability of the aerolysin pore to detect and distinguish between long peptides bearing multiple PTMs also provides a more precise remapping of PTMs in proteins at the single-molecule level, which remains a challenge for MS and other methods. In conclusion, the nanopore-based technology presented here, besides the natural advantages of being fast, cheap, label-free, and high-throughput, provides the possibility of being developed into a portable diagnostic device with medical and commercial potential.

METHODS

Synthesis of C-Terminal α-Synuclein Peptides and Their PTMs. The majority of peptides used in nanopore experiments were produced as described in previous works.[44−46] A few peptides

(oM127 and wt shown in Supplementary Figure S2) were provided by GenicBio Limited. All peptides were characterized by Liquid chromatography—mass spectrometry (LC-MS) as previously described.[70] Their purity was also assessed by UPLC analysis, on a Waters Acquity H-Class system using a C18 column (with UV detection at 214 nm and run time of 4 min (gradient 10% to 90% acetonitrile) with 0.6 mL/min flow rate.

RBCs were prepared similarly as previously described.[75] Briefly, RBCs were lysed and further treated using the Hemovoid kit (Biotech Support Group), aiming to remove hemoglobin (the most abundant protein in RBCs) and to enrich low-abundance proteins such as α-synuclein. Importantly, all biological fluids used in this study are derived from healthy controls.

CtsD digestion experiments were performed in a total volume of 100 μL. Specifically, 2 μL of 500 μM CtsD (Sigma-Aldrich Chemie GmbH, Buchs, SG Switzerland), 3 μL of 100 μM full-length α-synuclein and 95 μL of buffer (40 mM sodium acetate, 50 mM NaCl and 5 mM DTT, pH 5) incubated at 37 °C for 20 h at 300 rpm in a ThermoMixer C.

Aerolysin Productions. The recombinant K238A aerolysin proteins were generated from the aerolysin gene in the pET22b vector with a C-terminal hexa-histidine tag as described in our previous work,[43,53] and then expressed and purified from BL21 DE3 pLys *E. coli* cells. Cells were grown to an optical density of 0.6−0.7 in Luria−Bertani (LB) media. Protein expression was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and subsequent growth overnight at 20 °C. Cell pellets were resuspended in lysis buffer (20 mM Sodium phosphate pH 7.4, 500 mM NaCl) mixed with cOmplete Protease Inhibitor Cocktail (Roche) and then lysed by sonication. The resulting suspensions were centrifuged (12.000 rpm for 35 min at 4 °C), and the supernatants were purified through a HisTrap HP column (GE Healthcare) previously equilibrated with lysis buffer. The protein was eluted with a gradient over 40 column volumes of elution buffer (20 mM Sodium phosphate pH 7.4, 500 mM NaCl, 500 mM Imidazole), and buffer was exchanged into a final buffer (20 mM Tris, pH 7.4, 500 mM NaCl) using a HiPrep Desalting column (GE Healthcare). The purified protein was flash-frozen in liquid nitrogen and stored at −20 °C.

Single-Channel Recording Experiments. Phospholipids of 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC) powder (Avanti Polar Lipids, Alabaster, USA) were dissolved in octane (Sigma-Aldrich Chemie GmbH, Buchs, Switzerland) for a final concentration of 1.0 mg per 100 μL. Purified protein was diluted to the concentration of 0.2 μg/mL and then incubated with Trypsin-agarose (Sigma-Aldrich Chemie GmbH, Buchs, SG Switzerland) for 2 h under 4 °C temperature. The solution was finally centrifuged to remove trypsin.

Nanopore single-channel recording experiments were performed on Orbit Mini equipment (Nanion, Munich, Germany) and an Axopatch 200B Amplifier system (Molecular Devices, San Jose, USA). In the Orbit Mini setup, DPhPC membranes were formed across a MECA 4 recording chip that contains a 2 × 2 array of cylindrical 50 μm diameter in a highly inert polymer. Each of the four cavities contains an individual integrated Ag/AgCl-microelectrode and sustains one DPhPC bilayer. If not indicated otherwise, the measurement chamber temperature was set to 20 °C. Data was collected at 10 kHz sampling rate with a 5 kHz low-pass filter.

Ionic strength gradient experiments were carried out as follows. Teflon films with 50 μm apertures were mounted in Teflon chambers using high-vacuum grease (Dow Corning Corporation, Midland, MI, USA). The films separated two compartments (*cis/trans*) only connected through the Teflon film aperture with one Ag/AgCl electrode in each compartment. Apertures were pretreated with 1 μL of 2% (v/v) hexadecane in hexane on both sides using a standard pipet, and the chamber was mounted in the recording setup. DPhPC bilayers were formed by folding as described previously.[76,77] Briefly, electrolyte solution was added to both sides taking care that the level stayed below the aperture, and lipids (10 mg/mL in pentane) were added onto the electrolyte surface in both compartments. After the pentane evaporated, the electrolyte level was raised above the

aperture, and a lipid bilayer was formed. The quality of the lipid bilayer was monitored through its capacitance, and its stability was verified through the application of 150 mV over the course of at least 5 min. After peptide addition, the *cis* chamber was carefully mixed by pipetting up and down. Currents were sampled at 200 kHz and low-pass filtered at 100 kHz with an Axopatch 200B (Molecular Devices, LLC., San Jose, CA, USA).

Peptides (lyophilized powder) were prediluted in 10 mM Tris and 1.0 mM EDTA solution (pH 7.4) to a stock concentration of 500 μM and added to the *cis* side of the chamber in 1.0 M KCl solution buffered with 10 mM Tris and 1.0 mM EDTA (pH 7.4) to the final concentration indicated in the figure caption. All experiments shown here were repeated with at least 4 different pores.

Signals Processing and Classifications Using Deep Learning. The signal processing was done in the same way as previously reported.[53] The open pore current distribution is measured by fitting a Gaussian function on the peak distribution of the current with the highest mean current. The events are extracted using a current threshold at 3σ from the open pore current distribution. The relative current percentage (I/I_0) is computed from the mean open pore current (I_0) and the mean residual current (I). The dwell time, average relative current, relative standard deviation of the current $\sigma_{rel} = \frac{\sigma}{\sigma_0}$ (σ_0 is the value of the open pore current standard deviation and σ is the residual current standard deviation) and local extrema are computed. The events are selected on the basis of the dwell time (0.2 to 100.0 ms) and the average relative current (0 to 40%) discarding the events that are too short, too long, or that do not block the current sufficiently.

The machine learning pipeline is composed of two steps. The first one is the classification of every event, and the second is the assessment of the quality of the prediction of the classifier. The neural network architecture for both the classification and the assessment is a long, short-term memory (LSTM) neural network followed by a multilayer perceptron (MLP) using the position in time and relative current of the local extrema for each event as input features. The features are rescaled by a fixed factor to decrease the training time. The classifier is composed of an LSTM with state size 128 without any activation function followed by 6 fully connected hidden layers of size 256 with rectified linear unit (ReLU) as activation functions and finally an output layer of size 8 with softmax activation function. The assessment is done with a scaled down version of the classifier with a LSTM with a state of size 32, 3 fully connected hidden layers of size 64 with hyperbolic tangent activation functions, and an output layer of size 1 with sigmoid activation function. The neural networks for the classification and assessment are trained together using a 3-parts loss function. The first part is the full classification cross-entropy loss of the predictions from the classifier and the peptides label. The second part is the assessment of the cross-entropy loss between the predicted and actual prediction validity from the classifier. The third part is the reinforcement classification loss, which is the full classification cross-entropy loss scaled by the assessment prediction.

The deep learning classifier is benchmarked using cross-validation, where the first (in time) 75% of the events of a measurement is used for training and the last 25% for validation. In the case of the mixture experiments, we use the holdout method as they are not present in the training data set (or test/validation data sets). The data set composed of nY125, nY125nY133nY136, nY125pS129, nY136, pS129, pY125, pY125pS129, and wt contains 519,000 events. The data set for oM127 and wt contains 83,000 events. The mixture experiments (WT:pY125) data set contains 88,999 events.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.3c08623.

> $I-V$ curve for the aerolysin K238A mutant in urea condition, the ionic current signals of the different PTMs on the peptide, the contour plot and confusion matrix for wt and oM127 peptides, the mixture experiments of wt and pS129 peptide, the red blood cells lysate recording and the correlation between signal frequency and analyte concentration (Figures S1−S8) (PDF)

AUTHOR INFORMATION

Corresponding Authors

Chan Cao − *Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland; Department of Inorganic and Analytical Chemistry, Chemistry and Biochemistry, University of Geneva, 1211 Geneva, Switzerland;* ⊙ orcid.org/0000-0003-2592-0690; Email: chan.cao@unige.ch

Hilal A. Lashuel − *Laboratory of Molecular and Chemical Biology of Neurodegeneration, Brain Mind Institute, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland;* ⊙ orcid.org/0000-0001-7682-8320; Email: hilal.lashuel@eplf.ch

Matteo Dal Peraro − *Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland;* Email: matteo.dalperaro@epfl.ch

Authors

Pedro Magalhães − *Laboratory of Molecular and Chemical Biology of Neurodegeneration, Brain Mind Institute, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland*

Lucien F. Krapp − *Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland*

Juan F. Bada Juarez − *Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland*

Simon Finn Mayer − *Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland*

Verena Rukes − *Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland;* ⊙ orcid.org/0009-0008-2704-179X

Anass Chiki − *Laboratory of Molecular and Chemical Biology of Neurodegeneration, Brain Mind Institute, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, EPFL, Lausanne 1015, Switzerland*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsnano.3c08623

Author Contributions

C.C., M.D.P., and H.L. conceived the project and designed the study. C.C. performed the experiments on Orbit Mini, supervised all single-channel experiments and did data analysis. P.M. characterized all C-terminal α-synuclein peptides and prepared all clinical samples. L.F.K. performed signal processing and developed the deep learning pipeline. J.F.B.J. conducted CtsD digestion and nanopore experiments. S.F.M. and V.R. carried out experiments on the asymmetric buffer. A.C. synthesized C-terminal α-synuclein peptides with different PTMs forms. C.C., L.F.K., P.M., S.F.M., H.L., and M.D.P. interpreted the data. C.C., H.L., and M.D.P. wrote the manuscript with input from all authors.

Author Contributions

#P.M. and L.F.K. equally contributed.

Notes

ACKNOWLEDGMENTS

REFERENCES

(1) Marth, J. D. A Unified Vision of the Building Blocks of Life. *Nat. Cell Biol.* **2008**, *10*, 1015−1015.

(2) Wang, Y.-C.; Peterson, S. E.; Loring, J. F. Protein Post-Translational Modifications and Regulation of Pluripotency in Human Stem Cells. *Cell Res.* **2014**, *24*, 143−160.

(3) Walsh, G.; Jefferis, R. Post-Translational Modifications in the Context of Therapeutic Proteins. *Nat. Biotechnol.* **2006**, *24*, 1241−1252.

(4) Pagel, O.; Loroch, S.; Sickmann, A.; Zahedi, R. P. Current Strategies and Findings in Clinically Relevant Post-Translational Modification-Specific Proteomics. *Expert Rev. Proteomics* **2015**, *12*, 235−253.

(5) Restrepo-Pérez, L.; Joo, C.; Dekker, C. Paving the Way to Single-Molecule Protein Sequencing. *Nat. Nanotechnol.* **2018**, *13*, 786−796.

(6) Petrera, A.; von Toerne, C.; Behler, J.; Huth, C.; Thorand, B.; Hilgendorff, A.; Hauck, S. M. Multiplatform Approach for Plasma Proteomics: Complementarity of Olink Proximity Extension Assay Technology to Mass Spectrometry-Based Protein Profiling. *J. Proteome Res.* **2021**, *20*, 751−762.

(7) Wu, D.; Milutinovic, M. D.; Walt, D. R. Single Molecule Array (Simoa) Assay with Optimal Antibody Pairs for Cytokine Detection in Human Serum Samples. *Analyst* **2015**, *140*, 6277−6282.

(8) Banerjee, A.; Mastrangelo, C. H. Resurgence of Electron Quantum Tunneling Sensors. *Micro* **2022**, *2*, 679−698.

(9) Ohshiro, T.; Tsutsui, M.; Yokota, K.; Furuhashi, M.; Taniguchi, M.; Kawai, T. Detection of Post-Translational Modifications in Single Peptides Using Electron Tunnelling Currents. *Nat. Nanotechnol.* **2014**, *9*, 835−840.

(10) Lashuel, H. A.; Mahul-Mellier, A.-L.; Novello, S.; Hegde, R. N.; Jasiqi, Y.; Altay, M. F.; Donzelli, S.; DeGuire, S. M.; Burai, R.; Magalhães, P.; Chiki, A.; Ricci, J.; Boussouf, M.; Sadek, A.; Stoops, E.; Iseli, C.; Guex, N. Revisiting the Specificity and Ability of Phospho-S129 Antibodies to Capture Alpha-Synuclein Biochemical and Pathological Diversity. *NPJ Parkinsons Dis.* **2022**, *8*, 136.

(11) Chiki, A.; DeGuire, S. M.; Ruggeri, F. S.; Sanfelice, D.; Ansaloni, A.; Wang, Z.-M.; Cendrowska, U.; Burai, R.; Vieweg, S.; Pastore, A.; Dietler, G.; Lashuel, H. A. Mutant Exon1 Huntingtin Aggregation Is Regulated by T3 Phosphorylation-Induced Structural Changes and Crosstalk between T3 Phosphorylation and Acetylation at K6. *Angew. Chem., Int. Ed.* **2017**, *56*, 5202−5207.

(12) Chiki, A.; Zhang, Z.; Rajasekhar, K.; Abriata, L. A.; Rostami, I.; Krapp, L. F.; Boudeffa, D.; Dal Peraro, M.; Lashuel, H. A. Investigating Crosstalk Among PTMs Provides Novel Insight Into the Structural Basis Underlying the Differential Effects of Nt17 PTMs on Mutant Httex1 Aggregation. *Front. Mol. Biosci.* **2021**, *8*, 686086.

(13) Mnatsakanyan, R.; Shema, G.; Basik, M.; Batist, G.; Borchers, C. H.; Sickmann, A.; Zahedi, R. P. Detecting Post-Translational Modification Signatures as Potential Biomarkers in Clinical Mass Spectrometry. *Expert Rev. Proteomics* **2018**, *15*, 515−535.

(14) Dubois, B.; Feldman, H. H.; Jacova, C.; Hampel, H.; Molinuevo, J. L.; Blennow, K.; DeKosky, S. T.; Gauthier, S.; Selkoe, D.; Bateman, R.; Cappa, S.; Crutch, S.; Engelborghs, S.; Frisoni, G. B.; Fox, N. C.; Galasko, D.; Habert, M.-O.; Jicha, G. A.; Nordberg, A.; Pasquier, F.; et al. Advancing Research Diagnostic Criteria for Alzheimer's Disease: The IWG-2 Criteria. *Lancet Neurol.* **2014**, *13*, 614−629.

(15) Janelidze, S.; Mattsson, N.; Palmqvist, S.; Smith, R.; Beach, T. G.; Serrano, G. E.; Chai, X.; Proctor, N. K.; Eichenlaub, U.; Zetterberg, H.; Blennow, K.; Reiman, E. M.; Stomrud, E.; Dage, J. L.; Hansson, O. Plasma P-Tau181 in Alzheimer's Disease: Relationship to Other Biomarkers, Differential Diagnosis, Neuropathology and Longitudinal Progression to Alzheimer's Dementia. *Nat. Med.* **2020**, *26*, 379−386.

(16) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13770−13773.

(17) Venkatesan, B. M.; Bashir, R. Nanopore Sensors for Nucleic Acid Analysis. *Nat. Nanotechnol.* **2011**, *6*, 615−624.

(18) Ying, Y.-L.; Cao, C.; Long, Y.-T. Single Molecule Analysis by Biological Nanopore Sensors. *Analyst* **2014**, *139*, 3826−3835.

(19) Cao, C.; Long, Y.-T. Biological Nanopores: Confined Spaces for Electrochemical Single-Molecule Analysis. *Acc. Chem. Res.* **2018**, *51*, 331−341.

(20) Jain, M.; Koren, S.; Miga, K. H.; Quick, J.; Rand, A. C.; Sasani, T. A.; Tyson, J. R.; Beggs, A. D.; Dilthey, A. T.; Fiddes, I. T.; Malla, S.; Marriott, H.; Nieto, T.; O'Grady, J.; Olsen, H. E.; Pedersen, B. S.; Rhie, A.; Richardson, H.; Quinlan, A. R.; Snutch, T. P.; et al. Nanopore Sequencing and Assembly of a Human Genome with Ultra-Long Reads. *Nat. Biotechnol.* **2018**, *36*, 338−345.

(21) Laszlo, A. H; Derrington, I. M; Ross, B. C; Brinkerhoff, H.; Adey, A.; Nova, I. C; Craig, J. M; Langford, K. W; Samson, J. M.; Daza, R.; Doering, K.; Shendure, J.; Gundlach, J. H Decoding Long Nanopore Sequencing Reads of Natural DNA. *Nat. Biotechnol.* **2014**, *32*, 829−833.

(22) Houghtaling, J.; List, J.; Mayer, M. Nanopore-Based, Rapid Characterization of Individual Amyloid Particles in Solution: Concepts, Challenges, and Prospects. *Small* **2018**, *14*, 1802412.

(23) Brinkerhoff, H.; Kang, A. S. W.; Liu, J.; Aksimentiev, A.; Dekker, C. Infinite Re-Reading of Single Proteins at Single-Amino-Acid Resolution Using Nanopore Sequencing. *bioRxiv*, July 14, 2021. DOI: 10.1101/2021.07.13.452225.

(24) Yan, S.; Zhang, J.; Wang, Y.; Guo, W.; Zhang, S.; Liu, Y.; Cao, J.; Wang, Y.; Wang, L.; Ma, F.; Zhang, P.; Chen, H.-Y.; Huang, S. Single Molecule Ratcheting Motion of Peptides in a *Mycobacterium Smegmatis* Porin A (MspA) Nanopore. *Nano Lett.* **2021**, *21*, 6703−6710.

(25) Alfaro, J. A.; Bohländer, P.; Dai, M.; Filius, M.; Howard, C. J.; van Kooten, X. F.; Ohayon, S.; Pomorski, A.; Schmid, S.; Aksimentiev, A.; Anslyn, E. V.; Bedran, G.; Cao, C.; Chinappi, M.; Coyaud, E.; Dekker, C.; Dittmar, G.; Drachman, N.; Eelkema, R.; Goodlett, D.; et al. The Emerging Landscape of Single-Molecule Protein Sequencing Technologies. *Nat. Methods* **2021**, *18*, 604−617.

(26) Rosen, C. B.; Rodriguez-Larrea, D.; Bayley, H. Single-Molecule Site-Specific Detection of Protein Phosphorylation with a Nanopore. *Nat. Biotechnol.* **2014**, *32*, 179−181.

(27) Ying, Y.-L.; Yang, J.; Meng, F.-N.; Li, S.; Li, M.-Y.; Long, Y.-T. A Nanopore Phosphorylation Sensor for Single Oligonucleotides and Peptides. *Research* **2019**, *2019*, 1−8.

(28) Li, S.; Wu, X.; Li, M.; Liu, S.; Ying, Y.; Long, Y. T232K/K238Q Aerolysin Nanopore for Mapping Adjacent Phosphorylation Sites of a Single Tau Peptide. *Small Methods* **2020**, *4*, 2000014.

(29) Ensslen, T.; Sarthak, K.; Aksimentiev, A.; Behrends, J. C. Resolving Isomeric Posttranslational Modifications Using a Biological

Nanopore as a Sensor of Molecular Shape. *J. Am. Chem. Soc.* **2022**, *144*, 16060−16068.

(30) Zhang, L.; Gardner, M. L.; Jayasinghe, L.; Jordan, M.; Aldana, J.; Burns, N.; Freitas, M. A.; Guo, P. Detection of Single Peptide with Only One Amino Acid Modification via Electronic Fingerprinting Using Reengineered Durable Channel of Phi29 DNA Packaging Motor. *Biomaterials* **2021**, *276*, 121022.

(31) Restrepo-Pérez, L.; Wong, C. H.; Maglia, G.; Dekker, C.; Joo, C. Label-Free Detection of Post-Translational Modifications with a Nanopore. *Nano Lett.* **2019**, *19*, 7957−7964.

(32) Fahie, M. A.; Chen, M. Electrostatic Interactions between OmpG Nanopore and Analyte Protein Surface Can Distinguish between Glycosylated Isoforms. *J. Phys. Chem. B* **2015**, *119*, 10198−10206.

(33) Wloka, C.; Van Meervelt, V.; van Gelder, D.; Danda, N.; Jager, N.; Williams, C. P.; Maglia, G. Label-Free and Real-Time Detection of Protein Ubiquitination with a Biological Nanopore. *ACS Nano* **2017**, *11*, 4387−4394.

(34) Degiacomi, M. T.; Iacovache, I.; Pernot, L.; Chami, M.; Kudryashev, M.; Stahlberg, H.; van der Goot, F. G.; Dal Peraro, M. Molecular Assembly of the Aerolysin Pore Reveals a Swirling Membrane-Insertion Mechanism. *Nat. Chem. Biol.* **2013**, *9*, 623−629.

(35) Cirauqui, N.; Abriata, L. A.; van der Goot, F. G.; Dal Peraro, M. Structural, Physicochemical and Dynamic Features Conserved within the Aerolysin Pore-Forming Toxin Family. *Sci. Rep.* **2017**, *7*, 13932.

(36) Cao, C.; Ying, Y.-L.; Hu, Z.-L.; Liao, D.-F.; Tian, H.; Long, Y.-T. Discrimination of Oligonucleotides of Different Lengths with a Wild-Type Aerolysin Nanopore. *Nat. Nanotechnol.* **2016**, *11*, 713−718.

(37) Ouldali, H.; Sarthak, K.; Ensslen, T.; Piguet, F.; Manivet, P.; Pelta, J.; Behrends, J. C.; Aksimentiev, A.; Oukhaled, A. Electrical Recognition of the Twenty Proteinogenic Amino Acids Using an Aerolysin Nanopore. *Nat. Biotechnol.* **2020**, *38*, 176−181.

(38) Oliveira, L. M. A.; Gasser, T.; Edwards, R.; Zweckstetter, M.; Melki, R.; Stefanis, L.; Lashuel, H. A.; Sulzer, D.; Vekrellis, K.; Halliday, G. M.; Tomlinson, J. J.; Schlossmacher, M.; Jensen, P. H.; Schulze-Hentrich, J.; Riess, O.; Hirst, W. D.; El-Agnaf, O.; Mollenhauer, B.; Lansbury, P.; Outeiro, T. F. Alpha-Synuclein Research: Defining Strategic Moves in the Battle against Parkinson's Disease. *NPJ Parkinsons Dis.* **2021**, *7*, 65.

(39) Eusebi, P.; Giannandrea, D.; Biscetti, L.; Abraha, I.; Chiasserini, D.; Orso, M.; Calabresi, P.; Parnetti, L. Diagnostic Utility of Cerebrospinal Fluid *α*-Synuclein in Parkinson's Disease: A Systematic Review and Meta-Analysis: CSF *α*-Synuclein for PD Diagnosis. *Mov. Disord.* **2017**, *32*, 1389−1400.

(40) Majbour, N. K.; Vaikath, N. N.; van Dijk, K. D.; Ardah, M. T.; Varghese, S.; Vesterager, L. B.; Montezinho, L. P.; Poole, S.; Safieh-Garabedian, B.; Tokuda, T.; Teunissen, C. E.; Berendse, H. W.; van de Berg, W. D. J.; El-Agnaf, O. M. A. Oligomeric and Phosphorylated Alpha-Synuclein as Potential CSF Biomarkers for Parkinson's Disease. *Mol. Neurodegener.* **2016**, *11*, 7.

(41) Conibear, A. C. Deciphering Protein Post-Translational Modifications Using Chemical Biology Tools. *Nat. Rev. Chem.* **2020**, *4*, 674−695.

(42) Nakayasu, E. S.; Gritsenko, M.; Piehowski, P. D.; Gao, Y.; Orton, D. J.; Schepmoes, A. A.; Fillmore, T. L.; Frohnert, B. I.; Rewers, M.; Krischer, J. P.; Ansong, C.; Suchy-Dicey, A. M.; Evans-Molina, C.; Qian, W.-J.; Webb-Robertson, B.-J. M.; Metz, T. O. Tutorial: Best Practices and Considerations for Mass-Spectrometry-Based Protein Biomarker Discovery and Validation. *Nat. Protoc.* **2021**, *16*, 3737−3760.

(43) Cao, C.; Cirauqui, N.; Marcaida, M. J.; Buglakova, E.; Duperrex, A.; Radenovic, A.; Dal Peraro, M. Single-Molecule Sensing of Peptides and Nucleic Acids by Engineered Aerolysin Nanopores. *Nat. Commun.* **2019**, *10*, 4918.

(44) Schmid, A. W.; Fauvet, B.; Moniatte, M.; Lashuel, H. A. Alpha-Synuclein Post-Translational Modifications as Potential Biomarkers for Parkinson Disease and Other Synucleinopathies. *Mol. Cell. Proteomics* **2013**, *12*, 3543−3558.

(45) Meuvis, J.; Gerard, M.; Desender, L.; Baekelandt, V.; Engelborghs, Y. The Conformation and the Aggregation Kinetics of *α*-Synuclein Depend on the Proline Residues in Its C-Terminal Region. *Biochemistry* **2010**, *49*, 9345−9352.

(46) Hejjaoui, M.; Butterfield, S.; Fauvet, B.; Vercruysse, F.; Cui, J.; Dikiy, I.; Prudent, M.; Olschewski, D.; Zhang, Y.; Eliezer, D.; Lashuel, H. A. Elucidating the Role of C-Terminal Post-Translational Modifications Using Protein Semisynthesis Strategies: *α*-Synuclein Phosphorylation at Tyrosine 125. *J. Am. Chem. Soc.* **2012**, *134*, 5196−5210.

(47) Fauvet, B.; Butterfield, S. M.; Fuks, J.; Brik, A.; Lashuel, H. A. One-Pot Total Chemical Synthesis of Human [Small Alpha]-Synuclein. *Chem. Commun.* **2013**, *49*, 9254−9256.

(48) Haj-Yahya, M.; Fauvet, B.; Herman-Bachinsky, Y.; Hejjaoui, M.; Bavikar, S. N.; Karthikeyan, S. V.; Ciechanover, A.; Lashuel, H. A.; Brik, A. Synthetic Polyubiquitinated *α*-Synuclein Reveals Important Insights into the Roles of the Ubiquitin Chain in Regulating Its Pathophysiology. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 17726−17731.

(49) Bertoncini, C. W.; Jung, Y.-S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. Release of Long-Range Tertiary Interactions Potentiates Aggregation of Natively Unstructured *α*-Synuclein. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 1430−1435.

(50) Kawahara, K.; Tanford, C. Viscosity and Density of Aqueous Solutions of Urea and Guanidine Hydrochloride. *J. Biol. Chem.* **1966**, *241*, 3228−3232.

(51) Fologea, D.; Uplinger, J.; Thomas, B.; McNabb, D. S.; Li, J. Slowing DNA Translocation in a Solid-State Nanopore. *Nano Lett.* **2005**, *5*, 1734−1737.

(52) Yu, L.; Kang, X.; Li, F.; Mehrafrooz, B.; Makhamreh, A.; Fallahi, A.; Foster, J. C.; Aksimentiev, A.; Chen, M.; Wanunu, M. Unidirectional Single-File Transport of Full-Length Proteins through a Nanopore. *Nat. Biotechnol.* **2023**, *41*, 1130.

(53) Cao, C.; Krapp, L. F.; Al Ouahabi, A.; König, N. F.; Cirauqui, N.; Radenovic, A.; Lutz, J.-F.; Peraro, M. D. Aerolysin Nanopores Decode Digital Information Stored in Tailored Macromolecular Analytes. *Sci. Adv.* **2020**, *6*, No. eabc2661.

(54) Sze, J. Y. Y.; Ivanov, A. P.; Cass, A. E. G.; Edel, J. B. Single Molecule Multiplexed Nanopore Protein Screening in Human Serum Using Aptamer Modified DNA Carriers. *Nat. Commun.* **2017**, *8*, 1552.

(55) Sauvage, V.; Boizeau, L.; Candotti, D.; Vandenbogaert, M.; Servant-Delmas, A.; Caro, V.; Laperche, S. Early MinION™ Nanopore Single-Molecule Sequencing Technology Enables the Characterization of Hepatitis B Virus Genetic Complexity in Clinical Samples. *PLoS One* **2018**, *13*, No. e0194366.

(56) Wang, Y.; Zheng, D.; Tan, Q.; Wang, M. X.; Gu, L.-Q. Nanopore-Based Detection of Circulating microRNAs in Lung Cancer Patients. *Nat. Nanotechnol* **2011**, *6*, 668−674.

(57) Wang, S.; Haque, F.; Rychahou, P. G.; Evers, B. M.; Guo, P. Engineered Nanopore of Phi29 DNA-Packaging Motor for Real-Time Detection of Single Colon Cancer Specific Antibody in Serum. *ACS Nano* **2013**, *7*, 9814−9822.

(58) Kukwikila, M.; Howorka, S. Nanopore-Based Electrical and Label-Free Sensing of Enzyme Activity in Blood Serum. *Anal. Chem.* **2015**, *87*, 9149−9154.

(59) Yu, J.; Cao, C.; Long, Y.-T. Selective and Sensitive Detection of Methylcytosine by Aerolysin Nanopore under Serum Condition. *Anal. Chem.* **2017**, *89*, 11685−11689.

(60) Fahie, M. A.; Yang, B.; Mullis, M.; Holden, M. A.; Chen, M. Selective Detection of Protein Homologues in Serum Using an OmpG Nanopore. *Anal. Chem.* **2015**, *87*, 11143−11149.

(61) Chuah, K.; Wu, Y.; Vivekchand, S. R. C.; Gaus, K.; Reece, P. J.; Micolich, A. P.; Gooding, J. J. Nanopore Blockade Sensors for Ultrasensitive Detection of Proteins in Complex Biological Samples. *Nat. Commun.* **2019**, *10*, 2109.

(62) Galenkamp, N. S.; Soskine, M.; Hermans, J.; Wloka, C.; Maglia, G. Direct Electrical Quantification of Glucose and Asparagine from Bodily Fluids Using Nanopores. *Nat. Commun.* **2018**, *9*, 4085.

(63) Cao, C.; Yu, J.; Wang, Y.-Q.; Ying, Y.-L.; Long, Y.-T. Driven Translocation of Polynucleotides Through an Aerolysin Nanopore. *Anal. Chem.* **2016**, *88*, 5046−5049.

(64) Barbour, R.; Kling, K.; Anderson, J. P.; Banducci, K.; Cole, T.; Diep, L.; Fox, M.; Goldstein, J. M.; Soriano, F.; Seubert, P.; Chilcote, T. J. Red Blood Cells Are the Major Source of Alpha-Synuclein in Blood. *Neurodegener. Dis.* **2008**, *5*, 55−59.

(65) Klatt, S.; Roberts, A.; Lothian, A.; Masters, C. L.; Cappai, R.; Fowler, C.; Roberts, B. R. Optimizing Red Blood Cell Protein Extraction for Biomarker Quantitation with Mass Spectrometry. *Anal. Bioanal. Chem.* **2020**, *412*, 1879−1892.

(66) Wanunu, M.; Morrison, W.; Rabin, Y.; Grosberg, A. Y.; Meller, A. Electrostatic Focusing of Unlabelled DNA into Nanoscale Pores Using a Salt Gradient. *Nat. Nanotechnol.* **2010**, *5*, 160−165.

(67) Sorrentino, Z. A.; Giasson, B. I. The Emerging Role of α-Synuclein Truncation in Aggregation and Disease. *J. Biol. Chem.* **2020**, *295*, 10224−10244.

(68) McGlinchey, R. P.; Lee, J. C. Cysteine Cathepsins Are Essential in Lysosomal Degradation of α-Synuclein. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 9322−9327.

(69) Hu, Z.-L.; Li, M.-Y.; Liu, S.-C.; Ying, Y.-L.; Long, Y.-T. A Lithium-Ion-Active Aerolysin Nanopore for Effectively Trapping Long Single-Stranded DNA. *Chem. Sci.* **2019**, *10*, 354−358.

(70) Kumar, S. T.; Jagannath, S.; Francois, C.; Vanderstichele, H.; Stoops, E.; Lashuel, H. A. How Specific Are the Conformation-Specific α-Synuclein Antibodies? Characterization and Validation of 16 α-Synuclein Conformation-Specific Antibodies Using Well-Characterized Preparations of α-Synuclein Monomers, Fibrils and Oligomers with Distinct Structures and Morphology. *Neurobiol. Dis.* **2020**, *146*, 105086.

(71) McGlinchey, R. P.; Lacy, S. M.; Huffer, K. E.; Tayebi, N.; Sidransky, E.; Lee, J. C. C-Terminal α-Synuclein Truncations Are Linked to Cysteine Cathepsin Activity in Parkinson's Disease. *J. Biol. Chem.* **2019**, *294*, 9973−9984.

(72) Afshar Bakshloo, M.; Kasianowicz, J. J.; Pastoriza-Gallego, M.; Mathé, J.; Daniel, R.; Piguet, F.; Oukhaled, A. Nanopore-Based Protein Identification. *J. Am. Chem. Soc.* **2022**, *144*, 2716−2725.

(73) Lucas, F. L. R.; Versloot, R. C. A.; Yakovlieva, L.; Walvoort, M. T. C.; Maglia, G. Protein Identification by Nanopore Peptide Profiling. *Nat. Commun.* **2021**, *12*, 5795.

(74) Borrebaeck, C. A. K. Precision Diagnostics: Moving towards Protein Biomarker Signatures of Clinical Utility in Cancer. *Nat. Rev. Cancer* **2017**, *17*, 199−204.

(75) Vicente Miranda, H.; Cássio, R.; Correia-Guedes, L.; Gomes, M. A.; Chegão, A.; Miranda, E.; Soares, T.; Coelho, M.; Rosa, M. M.; Ferreira, J. J.; Outeiro, T. F. Posttranslational Modifications of Blood-Derived Alpha-Synuclein as Biochemical Markers for Parkinson's Disease. *Sci. Rep.* **2017**, *7*, 13713.

(76) Montal, M.; Mueller, P. Formation of Bimolecular Membranes from Lipid Monolayers and a Study of Their Electrical Properties. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 3561−3566.

(77) Mayer, M.; Kriebel, J. K.; Tosteson, M. T.; Whitesides, G. M. Microfabricated Teflon Membranes for Low-Noise Recordings of Ion Channels in Planar Lipid Bilayers. *Biophys. J.* **2003**, *85*, 2684−2695.

(78) Cao, C.; Magalhães, P.; Krapp, L. F.; Bada Juarez, J. F.; Mayer, S.; Rukes, V.; Chiki, A.; Lashuel, H. A.; Dal Peraro, M. Deep Learning-Assisted Single-Molecule Detection of Protein Post-Translational Modifications with a Biological Nanopore. *bioRxiv*, September 8, 2023. DOI: 10.1101/2023.09.08.555904 (accessed on November 13, 2023).