Engineering Biological Nanopore Approaches towards Protein Sequencing

Xiaojun Wei,[1,2] Tadas Penkauskas,[3,4] Joseph E. Reiner,[5] Celeste Kennard,[1] Mark J. Uline,[1,2] Qian Wang,[6] Sheng Li,[7] Aleksei Aksimentiev,[8] Joseph W.F. Robertson,[3] and Chang Liu[1,2*]

[1] Biomedical Engineering Program, University of South Carolina, Columbia, SC 29208, United States
[2] Department of Chemical Engineering, University of South Carolina, Columbia, SC 29208, United States
[3] Biophysics and Biomedical Measurement Group, Microsystems and Nanotechnology Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, United States
[4] School of Engineering, Brown University, Providence, RI 02912, United States
[5] Department of Physics, Virginia Commonwealth University, Richmond, VA 23284, United States
[6] Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208, United States
[7] School of Data Science, University of Virginia, Charlottesville, VA 22903, United States
[8] Department of Physics and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

*Address correspondence to: changliu@cec.sc.edu

Abstract
Biotechnological innovations have vastly improved the capacity to perform large-scale protein studies, while the methods we have for identifying and quantifying individual proteins are still inadequate to perform protein sequencing at the single-molecule level. Nanopore-inspired systems devoted to understanding how single molecules behave have been extensively developed for applications in genome sequencing. These nanopore systems are emerging as prominent tools for protein identification, detection, and analysis, suggesting realistic prospects in novel protein sequencing. This review summarizes recent advances in biological nanopore sensors toward protein sequencing, from the identification of individual amino acids to the controlled translocation of peptides and proteins, with attention on device and algorithm development and the delineation of molecular mechanisms with the aid of simulations. Specifically, the review aims to offer recommendations for the advancement of nanopore-based protein sequencing from an engineering perspective, highlighting the need for collaborative efforts across multiple disciplines. These efforts should include chemical conjugation, protein engineering, molecular simulation, machine learning-assisted identification, and electronic device fabrication to enable practical implementation in real-world scenarios.

Keywords: nanopore, amino acid, peptide, protein sequencing, engineering, molecular simulation, machine learning, instrumentation

Introduction

Proteins are the structural elements and machinery in all living cells that are responsible for functioning biological architecture and homeostasis.[1] Function or dysfunction of proteins largely depends on the sequence of amino acids (AAs) making up their primary structure, which is naturally synthesized from information encoded in genes. Proteins have natural variability due to differences in the start site of translation, resulting in similar proteins with variable lengths. Additional complexity can be found in the degree and type of posttranslational modifications (PTMs).[2] Thus, sequencing of single proteins or small populations of proteins is expected to fundamentally enhance the understanding of all living systems. Core to this advancement are new and emerging tools to measure the sequence of intact proteins, along with their PTMs.[3] The primary structure of proteins consists of a linear sequence drawn from 20 proteinogenic AAs with an average volume of about 0.1 nm^3, linked by peptide bonds separated by only 0.38 nm.[4] These fundamental building blocks lead to macromolecules that fold into secondary and tertiary three-dimensional structures to produce biochemically active entities. The advances in nanopore DNA sequencing inform many of the methods more recently developed for protein sequencing, but many new challenges emerge. DNA and RNA require identification of only four nucleobases. Furthermore, low-abundance DNA/RNA molecules can be enzymatically amplified outside the cell, whereas such an amplification mechanism is absent for proteins. Thus, the development of highly sensitive, high-throughput protein sequencing is a required technological advancement to transform many areas of the biotechnology field.[5]

Current technologies for protein sequencing.

Currently, the only principal methods available for protein sequencing and identification that have been put into large-scale practice are Edman degradation and mass spectrometry (MS), or their combinations.[4,6] Edman degradation consists of cyclic chemical reactions that label, cleave, and identify the AA at the terminus of a protein one at a time, allowing the ordered identification of the protein's AA sequence from the *N*- to the *C*-terminus using liquid chromatography. However, it is limited to the analysis of homogenous protein samples and read lengths typically <50 AAs, which falls far below the median protein length in eukarya (419 AAs), bacteria (306 AAs), and archaea (288 AAs).[7,8] Edman degradation also relies on the presence of a free α-amino group at the peptide's *N*-terminus, making accurate detection of PTMs not always possible.[9] In addition, each ~45 minute degradation cycle makes the total process quite time-consuming.

MS is a sophisticated tool for identifying proteins based on the mass-to-charge (m/z) ratio of fragmented peptides and is the current gold standard for protein sequencing.[10] MS utilizes two approaches, namely bottom-up and top-down, to infer the primary structure of a protein.[4,11] The bottom-up approach prevalently used for protein sequencing involves protein extraction from cells or tissue, protein digestion with the protease trypsin, ionization of the resulting peptides, separation and detection of these ionized peptides, and subsequent mass analysis. The measured masses of the peptides function as fingerprints that are later matched with known proteins found in databases using search engines such as Mascot or Sequest.[12] Recently, Li *et al.* developed a *de novo* machine learning algorithm for data-independent acquisition MS, which involves restructuring graph-based neural networks with features derived from fragment ion peaks, providing an alternative to traditional databases or spectral libraries.[13] These MS-based proteomics workflows with specialized bioinformatics tools can undoubtedly benefit the future framework of single molecule techniques for protein sequencing. However, MS has its limitations,[14] primarily in regard to its narrow dynamic

range (up to 10^4 or 10^5) of protein concentrations, which is significantly lower than the desired dynamic range for comprehensive proteome analysis (up to 10^8 or 10^9).[15] Other drawbacks include limitations on detection limit, low throughput, insufficient readable peptide length, and low sequence coverage. These shortcomings are further magnified when differentiating protein isoforms in which identical AAs are rearranged or when detecting protein heterogeneity for unambiguous identification of protein variations, such as PTMs (*e.g.*, phosphorylation, acetylation, methylation, glycosylation, and ubiquitination, *etc*.). These slight differences in protein composition provide vital information for understanding biological processes and diseases.[16]

Emerging single-molecule protein sequencing technologies.

To address the above challenges, some highly parallel, single-molecule techniques have been envisioned in the past decade for next-generation protein analysis and sequencing.[17,18] These emerging methods use tunneling currents,[19,20] fluorescence,[21] resistive-pulse nanopores,[3,22,23] and other nanotechnologies to sequence or identify individual proteins, down to the single-molecule level or in single cells.[24,25] The precision of these advanced measurement tools have the capacity to create many opportunities in biomedical research, with applications ranging from proteomics of single cells and bodily fluids to sensing and classifying low-abundance protein biomarkers for disease screening and precision diagnostics.[21] Among the methods driving this era of advanced nanodevices for single-molecule recognition was the tunneling current method, as reported by Zwolak and Di Ventra for DNA sequencing.[26-28] This method was expanded to detect AAs using two metal electrodes separated by a nanogap (0.7-2 nm), comparable to the size of typical AA molecules. In 2014, Kawai and Taniguchi *et al*. demonstrated that 12 different AAs and a PTM (phosphotyrosine) could be identified by trapping electron tunnelling currents measured as the individual molecules pass through the nanoscale gap between electrodes.[29] Furthermore, by coating electrodes with a layer of recognition molecules (4(5)-(2-mercaptoethyl)-1H-imidazole-2-carboxamide, ICA), Lindsay *et al*. demonstrated the potential of nanogap-based recognition tunnelling for identifying both individual AAs and short peptides. This method exhibits the capability of distinguishing between molecular classes, such as enantiomers and isobaric isomers.[30]

Subsequently, several fluorescence-based methods were proposed and have since shown promising proof-of-concepts for single-molecule protein fingerprinting.[21,25,31] By combining Edman chemistry, single-molecule microscopy, and stable synthetic fluorophore chemistry, Marcotte and Anslyn *et al*. developed a high-throughput fluorosequencing technique to identify protein fragments by the millions in parallel.[32] This approach involves selective fluorescence labeling of cysteine and lysine residues in fragmented proteins, immobilization of the labeled peptides in a single-molecule microscope stage perfusion chamber, and subsequent quantification of peptide fluorescence through total internal reflection fluorescence (TIRF). Repeated cycles of Edman degradation and imaging can reveal the decrease in the intensity of the fluorescent dyes on each molecule in each cycle to determine the presence of cysteine and lysine and their positions. A variation on this theme was proposed by Joo and Meyer *et al*., who combined the protease-motor complex, ClpXP, with fluorescence resonance energy transfer (FRET) to detect low-abundance proteins.[33] These fluorescence-based fingerprinting methods are highly attractive for protein sequencing due to their scalability and utilization of established techniques, such as Edman chemistry or extensively researched proteases. Additionally, these methods incorporate workflows familiar to MS database searching. Recently, a dynamic approach for protein sequencing that differs markedly from the above approaches has been developed using a semiconductor chip to measure fluorescence intensity,

lifetime, and binding kinetics in real-time.[25] Peptides are immobilized in nanoscale reaction chambers and *N*-terminal AAs are instantaneously detected with dye-labeled *N*-terminal AA recognizers. Aminopeptidases are used to remove individual *N*-terminal AAs, exposing subsequent AAs for recognition. This direct sequencing technique eliminates the need for the complex chemistry and fluidics involved in other methods and allows for the identification of multiple AAs in an information-rich manner.

These innovative approaches have contributed to revolutionizing conventional proteomic technologies. However, there are still obstacles in terms of the potential for achieving *de novo*, full-length protein sequencing. Tunneling current methods are limited to the analysis of AAs or short peptides and lack the experimental basis to control the translocation of polypeptides through the nanogap. Fluorescence-based protein identification is impeded by the lack of distinct organic fluorophores for detecting the 20 different AAs without substantial signal crosstalk and constraint in the chemical repertoire to orthogonally label each of the AAs.[34]

Nanopore approaches aimed at protein sequencing.

The nascent field of nanopore sensing offers promises to revolutionize proteomics with single-molecule sensitivity and long reads beyond the tunneling current- and fluorescence-based approaches.[35-40] In a typical nanopore system, two electrolyte-filled flow-cells are separated by a thin membrane that hosts a nanoscale pore.[41] Under the applied voltage bias served by immersed electrodes in the flow-cells, the single molecules are driven through the narrow channel of the nanopore and induce characteristic ionic current modulations, which carry information about the composition, charge distribution, structure, and sequence of the translocating molecule.[42-45]

Nanopores are primarily composed of solid-state materials (*i.e.*, semiconductor-based), nanopipettes (*i.e.*, pulled glass capillaries), or biological molecules (*e.g.*, proteinaceous).[46,47] Solid-state nanopores and nanopipettes are microfabricated with inorganic materials that instill a chemical robustness, allowing them to be used across a wide range of experimental conditions. To date, many approaches based on these devices have been developed for discriminating residue substitutions in a single protein molecule,[48] decoding proteoforms,[49] calling the AA sequence of a protein by using a sub-nanometer diameter nanopore,[50] sensing folded proteins,[51,52] extracting the generic properties (*e.g.*, volume, dipole, and shape, *etc*.) of proteins,[53-55] or profiling protein conformations.[56] For protein sequencing, many simulations and computational assessments with solid-state nanopores have been expolored.[57-59] The use of sub-nanometer solid-state pores to detect AAs in proteins at high speed and throughput is intriguing.[60,61] The low capacitance of these pores allows for rapid detection, and semiconductor manufacturing enables high throughput. High voltage is crucial for precise control over peptide translocation and to minimize diffusion and back-stepping, albeit at the cost of decreasing analyte dwell time. Recently, Feng *et al*. reported the direct experimental identification of 16 natural AA by using MoS_2 nanopores, which can provide a sub-1 Dalton resolution.[60] Despite the potential benefits, in terms of genome and proteome sequencing, the large sizes of solid-state pores are not particularly beneficial, as sequencing requires a narrow band of smaller pores (*i.e.*, diameters < 3 nm), with a degree of precision challenging for even the most advanced semiconductor manufacturing techniques that can achieve 2 nm line-rules.[62,63] Additionally, current solid-state pores lack scalable manufacturing processes and sufficient surface wetting control. Although controlled breakdown (CBD) tools are currently under rapid development and active study to increase scalability, overcoming all abovementioned challenges remains a critical, yet unfulfilled, requirement for the practical implementation of solid-state pore-based platforms in protein analysis.[64]

In contrast, biological nanopores spontaneously assemble and insert into lipid bilayers with a precision dictated by the protein structure, which guarantees the small size and the reproducibility of the pore.[42,65-67] This inherent reproducibility allows precise engineering of the internal pore structure to achieve constriction diameters of 1.0 nm to 4.0 nm.[68] Both constriction geometry and local charge can be modified through established protein engineering tools by introducing functional groups or mutations in the primary structure of the protein.[69] This fine control has made biological nanopores arguably the most promising candidates to drive the next developments in single-molecule proteomics.[70-72] To date, biological nanopore-based sequencing technology has been established for nucleic acids by combining a high-resolution nanopore to obtain electrical signals, a molecular motor that controls polymer motion, and robust computer algorithms for deciphering and "calling" bases as they transit the sensing region of the pore.[73,74] The commercial success of portable nanopore sequencers[75,76] for single DNA molecules has inspired extensive research and development of single-molecule techniques for protein sequencing by extending the concepts applied to the relatively simple nucleic acids (four unique bases) to proteins (20 unique AAs and countless PTMs).[77-80]

Despite some exciting recent advancements, nanopore-based protein sensing is still in its infancy, facing tremendous obstacles unique to proteins and proteomics: *(1) Discrimination of the 20 proteinogenic AAs as building blocks of proteins.* Firstly, certain AAs display negligible differences in volume, measuring as little as 0.001 nm^3, and have a narrow maximum volume range of only 0.06-0.23 nm^3.[81] Hence, either conferring more specific features on the AAs or improving the spatial resolution of biological nanopores is required for accurate detection. Secondly, PTMs alter the properties of proteins through the cleavage of peptide bonds and/or the addition of a modifying group to one or more AAs, making the identification of AAs more complex.[82,83] *(2) Identification of peptides/proteins.* Compared to DNA sequencing through biological nanopores, the analysis and sequencing of peptides/proteins pose a greater challenge due to their intricate secondary and tertiary structures (folded configurations) and non-uniform charge distributions. These features can complicate or prevent the capture and translocation of peptides and proteins across nanopores.[84] In addition, distinguishing peptides by length or mass can be difficult, given the abundance of peptides that share comparable physical properties. *(3) Controlling the motion of peptides and proteins.* Overly fast and uneven translocation of a peptide strand impedes sufficient interaction with the nanopore. This makes identification problematic, especially for long polypeptides and rare proteins. Enhancing the interaction strength and duration between single molecules and the constricted region of the nanopore could greatly improve the signal-to-noise ratio of sensing and reduce the reading errors of the nanopore. *(4) Signal processing and understanding.* Signals from whole proteins or protein fragments with the full set of 20 proteinogenic AAs are complex and quite difficult to interpret. This necessitates the development of molecular dynamics (MD) simulations that facilitate the understanding of these signals. Furthermore, challenges posed by the abundant ionic current data call for machine learning techniques to assist the data analysis.

The last five years have seen significant advances toward addressing the above challenges with the development of specifically engineered nanopores to identify individual AAs and peptides, advances in signal processing algorithms, and refinement of MD simulations for understanding complex signals. Undoubtedly, these important breakthrough studies are reshaping the nanopore field from third-generation gene sequencing towards single-molecule proteomics, and perhaps a future of multi-omics. Here, we review nanopore protein sequencing from an engineering perspective, with emphasis on supporting research in computation, hardware, bioinformatics, and proteomics as a

system (**Figure 1**). Combined, these tools provide a roadmap for future routes to efficient sequencing and scaling-up for commercial applications.

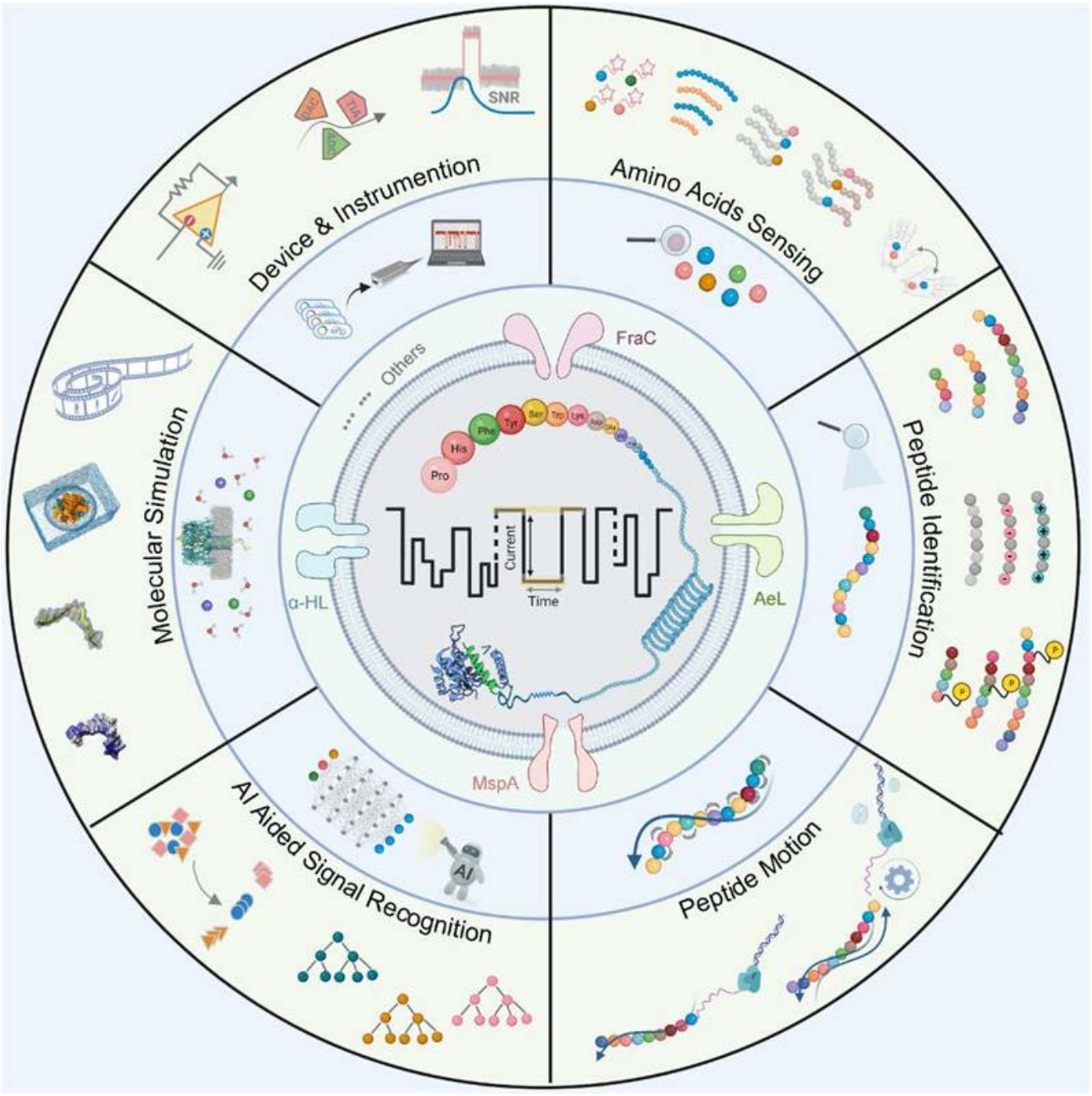


Figure 1. Engineering methods surrounding biological nanopore technologies for protein sequencing. Four most widely used biological nanopores (inner ring) and six peripheral engineering methods for nanopore sensing (outer ring) with great potential to contribute to protein sequencing are shown. (Created with Biorender.com)[§]

[§]Certain commercial materials, equipment, and instruments may be identified in this work to describe the experiments as completely as possible. In no case does such an identification imply a recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials, equipment, or instrument identified are necessarily the best available for the purpose. The authors declare no other competing interest.

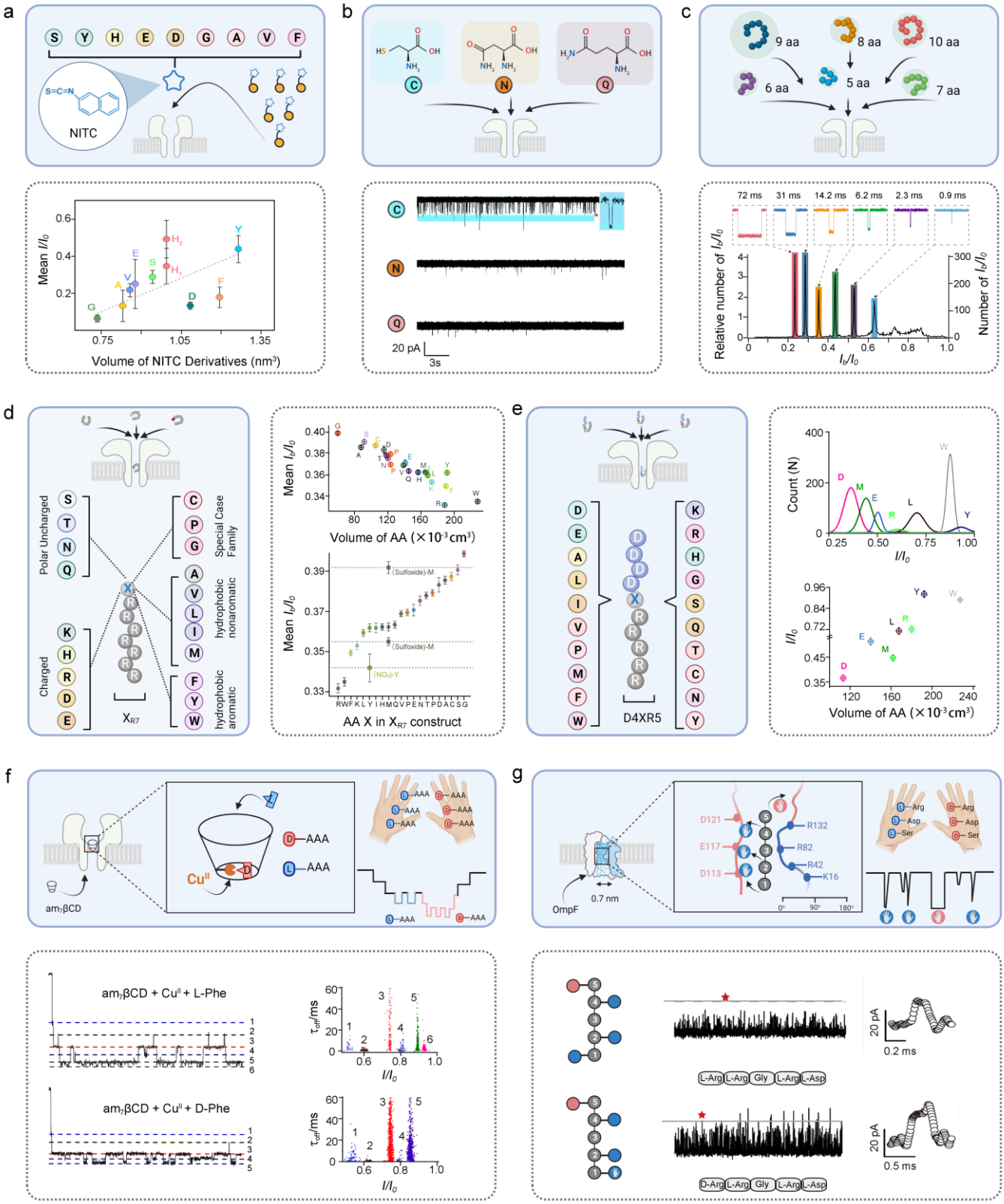

a
b
c
NITC
9 aa
8 aa
10 aa
6 aa
5 aa
7 aa
Mean I/I0
Volume of NITC Derivatives (nm³)
20 pA
3s
72 ms
31 ms
14.2 ms
6.2 ms
2.3 ms
0.9 ms
Relative number of Ib/I0
Number of Ib/I0
Ib/I0
d
e
Polar Uncharged
Special Case Family
Charged
hydrophobic nonaromatic
hydrophobic aromatic
XR7
Mean Ib/I0
Volume of AA (×10⁻³cm³)
AA X in XR7 construct
D4XR5
Count (N)
I/I0
f
g
amβCD
Cu
OmpF
0.7 nm
D121
E117
D113
R132
R82
R42
K16
amβCD + Cu + L-Phe
amβCD + Cu + D-Phe
τoff/ms
L-Arg L-Arg Gly L-Arg L-Asp
D-Arg L-Arg Gly L-Arg L-Asp
20 pA
0.2 ms
0.5 ms

Figure 2. Discrimination of amino acids with a biological nanopore. Schematic illustration of the identification strategies and the corresponding representative results in the dashed box. **(a)** Nine AAs were derivatized with NITC at the *N*-terminus and then translocated through the α-HL nanopore. Box: mean relative current blockade produced by each NITC derivative versus its spatial volume. **(b)** Single AAs (C, N, and Q) traversing across an AeL nanopore. Box: raw current traces. The red star denotes a typical current blockade event of C in blue shadow. **(c)** Detection of Arg (R) peptides with different lengths in an equimolar mixture using the AeL nanopore. Box: typical current event of six distinct populations (top) and the corresponding histogram (bottom). **(d)** Recognition of 20 AAs (X) in a cationic carrier of seven Args (R7) by the AeL nanopore. Box: mean relative residual current produced by the X_{R7} probes versus volume of AA (top); experimentally determined mean I/I_0 value for all 20 X_{R7} peptides (bottom). **(e)** Recognition of an AA (X) with a bipolar D_4XR_5 peptide carrier, in which five Arg residues and four Asp (D) residues were chemically linked to the target AA. Box: identification of D, M, E, R, L, W, and Y based on the I/I_0 values (top); relationship of current blockade against volume (bottom). **(f)** Recognition of am7βCD-Cu^{II} complex-functionalized α-HL nanopore for AA enantiomers. Box: current traces (left) and the corresponding scatter plots (right) showing the interaction of pores with am7βCD, Cu^{II}, and either L-Phe (top) or D-Phe (bottom). **(g)** Depiction of an OmpF trimeric protein sensing a single peptide. The enlarged part represents the zig-zag alignment of each AA sidechain of N-Arg-Arg-Gly-Arg-Asp in bulk. Box: typical nanopore-based readouts for Mol-1 and Mol-5. Red star points denote enlarged events in the original trace. Data in the boxes are extracted with permission from these references: (a): 94, (b): 107, (c): 109,[§] (d): 110, (e): 117, (f): 122, (g): 123 and all corresponding schematic diagrams are created with Biorender.com.

[§]https://creativecommons.org/licenses/by/4.0/

1. Sensing individual amino acids

The primary structure of a protein is defined by the sequence of AAs, which is important for determining the protein's spatial structure and function. Current biological nanopore techniques lack sufficient spatiotemporal sensitivity to identify all 20 proteinogenic AAs and diverse PTMs. Over the past decade, numerous research efforts have investigated AA sensing and reading using biological nanopores with the aid of chemical modifications to nanopore proteins and analytes.

Derivatization-assisted amino acid identification with an α-hemolysin nanopore.

The α-hemolysin (α-HL) nanopore has been the most well-studied biological nanopore for sensing small molecules, nucleotides, and proteins since it was reported by Menestrina in 1987.[85] Early works by Kasianowicz and Bezrukov demonstrated α-HL to be sensitive to single molecule-interactions through H^+ and D^+ association to AA residues inside the pore.[86,87] In 1996, single-stranded DNA was measured while translocating through the pore, sparking a race to realize nanopore sequencing.[88,89] α-HL is a self-assembling nanopore that has a heptameric, mushroom-like pore-forming structure. The narrowest constriction region of the α-HL nanopore is about 1.4 nm, located at the connection between the vestibule lumen (2.6-4.6 nm) and a β-barrel transmembrane stem (2.0 nm).[90] The distinct size, structure, and base pairing interactions of DNA bases contribute to the relative ease of sequencing DNA using the α-HL pore. In contrast, structurally complex peptides composed of many similar-sized AAs present greater difficulty in achieving high-resolution differentiation by α-HL, calling for methods to enhance the pore's sensing capability. Based on the Edman peptide degradation reaction,[91,92] a series of *N*-terminal derivatization were developed to assist the identification of individual AAs using the α-HL nanopore.[93,94] Liu and Wang *et al.*

employed four derivatization reagents, including ortho-phthalaldehyde (OPA), 2,3-naphthalenedicarboxaldehyde (NDA), phenyl isothiocyanate (PITC), and 2-naphthylisothiocyanate (NITC), to derivatize the *N*-termini of five representative AAs (Ala, Phe, Tyr, Asp, and His).[93] Compared to bare AAs, which have minimal or no detectable current blockade events through α-HL nanopores, the derivatized AAs exhibited a significant increase in the event frequency, particularly those modified with NDA and NITC, under the same experimental conditions. NITC proved to be the superior derivatization reagent, enabling precise identification of all five of the tested AAs. Using derivatized AAs as detection surrogates not only prolonged interactions of analytes with the sensing region of the pore,[95] but also limited their conformational variations. In a subsequent study using the NITC-derivatization method, the authors expanded the analysis to discriminate nine AAs of three classifications: nonpolar (G, A, V, and F), polar (S, Y), and charged (H, E, and D) (**Figure 2a**).[94] They also revealed that the current blockade reading is largely proportional to the spatial excluded volume of each derivative.

While identifying all 20 natural AAs can be considered the first step toward nanopore protein sequencing, their ultrasmall volume and molecular weight differences prohibit their full discrimination using traditional analyses of ionic current traces (*i.e.,* blockade, dwell time, *etc*.).[96-98] By revisiting the origin of nanopore current blockades from a theoretical perspective, Long *et al*. have demonstrated that the noncovalent interaction between nanopore and analyte significantly influences ion mobility within the nanopore. Understanding this mechanism offers an additional strategy for guiding the design of engineered nanopores to achieve high temporal and spatial resolution, thereby enhancing the ability to discriminate AAs of similar molecular weight and volume.[99]

From direct readout to peptide-assisted amino acid discrimination.

Several protein channels, other than α-HL, with diverse physical and chemical properties have been investigated for the use as nanopore sensors of small molecules. Analyte-pore interaction plays one of the most important roles in sensitivity and specificity, and different analytes require unique nanopore characteristics (pore diameter, morphology, charge distribution, hydrophobicity, *etc*.) for optimal detection and identification. One prominent nanopore that has been explored is aerolysin (AeL), a toxin secreted by *Aeromonas hydrophila*.[100] AeL has many interesting features as a nanopore sensor, including a more highly charged lumen and smaller diameter (~1.0 nm) than α-HL.[101-103] The AeL pore was determined to have superior sensitivity for some analytes, established by its ability to differentiate poly(ethylene glycol) species, which was speculatively linked to its highly charged interior surface.[104] Further evidence that the presence of these charges enhance the AeL pore's sensing capability can be drawn from studies using the pore to discriminate oligomers 2 to 20 bases in length[105]and distinguish methylated cytosine from cytosine.[106] Initial attempts at the direct identification of single AAs, without modifications or labelling, involved the analysis of single cysteine molecules (0.11 nm^3), taking advantage of AeL's reported sensing capabilities.[107] In this work, cysteine produced current blockades distinct from both asparagine (Asn, 0.12 nm^3) and glutamine (Gln, 0.15 nm^3), highlighting the importance of considering factors other than volume exclusion (*e.g.*, spatial conformation, specific functional groups) for analyzing the characteristic interaction between a single AA and the nanopore sensing interface (**Figure 2b**). While these results are encouraging, reading bare AAs directly using a biological nanopore cannot currently achieve the recognition of all 20 proteinogenic AAs.

Recent advancements indicate that expanding single AAs into polypeptides and using

polypeptides as carriers have significantly enhanced the identification of individual AAs. The sensing zone of the AeL pore is ~2 nm long, about three-four times longer than the reading heads used for DNA sequencing, which can slow down the movement of the peptide to aid AA detection.[108] In the 2018 study of Oukhaled *et al.*,[109] arginine (R) peptides of six different lengths, differing by a single AA in length (5, 6, 7, 8, 9, and 10 AAs), were discriminated independently or in an equimolar mixture with single AA resolution by using a recombinant AeL nanopore, without any physical or chemical modifications (**Figure 2c**). Building upon this work, Oukhaled and Aksimentiev *et al.*[110] reported an attempt at identifying all 20 proteinogenic AAs by linking each AA residue to a carrier peptide comprising seven arginines (R7) to facilitate capture by an AeL nanopore (**Figure 2d**). The net positive charge conferred by the arginine heptapeptide ensured exclusive electrophoretic transport of the peptides through the pore in a unidirectional manner. Based on the correlation between the molecular volume of each residue and its blockade magnitude, 13 AAs (R, W, F, K, L, N, T, P, D, A, C, S, and G) were distinguished in superimposed histograms of the relative residual current values, obtained individually. The study also demonstrated the method's ability to distinguish K, H, D, E, and R within an equimolar mixture of the R7-AA peptides, in addition to identifying mixtures of species that differed by only one hydroxyl group (F and Y) and of structural isomers (L and I). The authors note that overlap in relative residual current distributions prevented identification of seven AAs, a problem that is compounded in analyte samples containing mixtures of many species, and emphasize the need for further refinement of analytical techniques in order to improve the accuracy and reliability of protein analysis. They claim that with the implementation of chemical modifications, instrumentation advances, and nanopore engineering, their methodology may achieve enhanced identification of AAs, individually and in mixtures, and perhaps further ambitions such as parallel protein sequencing.

In other studies adopting peptide-assisted approaches for AA identification, a channel of bacteriophage T7 DNA packaging motor was used as a single pore sensing instrument, in combination with enzymatic digestion assays. The T7 motor channel has a relatively large constriction region of ~3.9 nm, yet has been purported to distinguish peptides at lower concentrations than AeL and to be more flexible in regard to modification with different recognition elements.[111] Clear mappings of five peptides with 8, 9, 10, 11, and 12 arginine (R) residues were achieved[112] and peptides composed of a mixture of acidic and basic AAs were differentiated[113] after digestion with protease. These studies demonstrate the possibility of improving the facility of nanopore sensing for the discrimination and recognition of AAs with peptide-assisted methods.

Computational and simulation-based studies in combination with preliminary experiments have shown that using macromolecules engineered with oppositely charged groups at either end, together with an applied potential difference across the nanopore, creates a means of enhancing the capture rate and increasing the residence time of the analyte. These features are attributed to the resulting electrostatic "tug-of-war" effect between the charged extrema of the macromolecule and the oppositely biased voltage of the pore setup.[114,115] Dekker and Joo *et al.* applied this concept of a bipolar peptide as a carrier to resolve various AA chemical modifications in a fingerprinting scheme. The chosen peptide, with a negative (10 glutamate) *N*-terminus and positive (10 arginine) *C*-terminus, was stretched and stalled inside a pore with an applied negative (-90 mV) *trans* bias. By attaching different chemical groups to a single cysteine in the central part of the peptide and measuring their effect on the nanopore signals, the research showed clear differentiation between labeled and unlabeled peptides and achieved sensitive detection for labels with various physicochemical

properties like mass and geometry. These properties were found to correlate with the relative current blockade, while label charge and size showed a direct relationship with translocation time.[116] Most recently, by linking four aspartic acid residues (D) to the *N*-terminal (negatively charged) and five arginine residues (R) to the *C*-terminal (positively charged) of individual AAs, a bipolar peptide probe was formed to enable electrical recognition of nine proteinogenic AAs (D, M, E, R, L, W, Y, N and A) using an AeL nanopore, based on their distinct current blockades and dwell times (**Figure 2e**).[117] Beyond AA recognition, the above studies hold significant implications for various biotechnology applications, including the detection of PTMs of peptides/proteins, as elaborated upon in the following section.

Amino acid isomerization sensing.

Chiral molecules exhibit differences in physiological behavior that directly affect chemical and biological activity and make them pharmacologically distinct.[95] Chirality, along with other forms of stereoisomerism induced by PTMs or asymmetric synthesis, introduces substantial complexity to the identification of isomeric AAs and peptides, due to their similar physiochemical properties. For example, L- and D-AA isomers possess nearly identical chemical structures, differing only in their stereochemistry, making it difficult to differentiate them using conventional analytical techniques. Several studies have taken advantage of the high sensing resolution of nanopores to develop ways of advancing the ability to discriminate stereoisomers, specifically enantiomers, at the single-molecule level.[118-121] By taking advantage of a chiral Cu^{II}-phenanthroline complex within an α-HL nanopore, Bayley *et al.* were able to differentiate L- and D-AAs (Tyr, Phe, Cys, and Asp) directly with the nanopore.[119] Kang *et al*. later introduced a modified strategy to discriminate all enantiomers of natural aromatic AAs in 2017, using a metal-organic (Cu^{II}-am7βCD) complex-functionalized protein nanopore. In this sensing system, the Cu^{II} plugging valve played the crucial role of holding the chiral AAs in the nanocavity for a sufficient registering time. The interactions between enantiomers of the aromatic AAs and the am7βCD-infused nanopore cavity caused distinctive current signatures (**Figure 2f**).[122] Although these nanopore-based approaches have demonstrated the possibility of sensing chiral molecules by designing a specific chiral environment, they have limitations when it comes to identifying chiral AAs in peptides with identical volumes. Chiral isomers differ only by the opposite orientation of their sidechains. Consequently, the chirality of an analyte plays a pivotal role in determining the orientation of its sidechain during interaction with a nanopore, thereby influencing the distinctive single-molecule current pattern utilized for analyte identification. Taking advantage of this relationship, Winterhalter and Ying *et al*. recently employed the bacterial outer-membrane channel, OmpF, as a chiral biosensor to track sidechain orientation and identify chiral and positional AA isomers within model peptides.[123] The narrow, 0.7 nm constriction zone of OmpF is composed of a positively charged ladder (K16, R42, R82, and R132) and a negatively charged pocket (D113, E117, and D121) on opposite sides (**Figure 2g**), forming a lateral electrostatic field that provokes sidechain reorientation during peptide translocation and consequent, distinct ionic current fluctuations and residence time changes. Analysis of these current patterns enables identification of peptides containing chiral AAs and positional isomers.

These studies, among others, have made progress in the recognition of AA isomers, either with individual AA molecules or within peptides, through the persistent development of specialized pore environments or modifications. Such momentum highlights that continuation of this trend will see the types of advancements needed to apply these methodologies to large-scale translational and commercial operations.

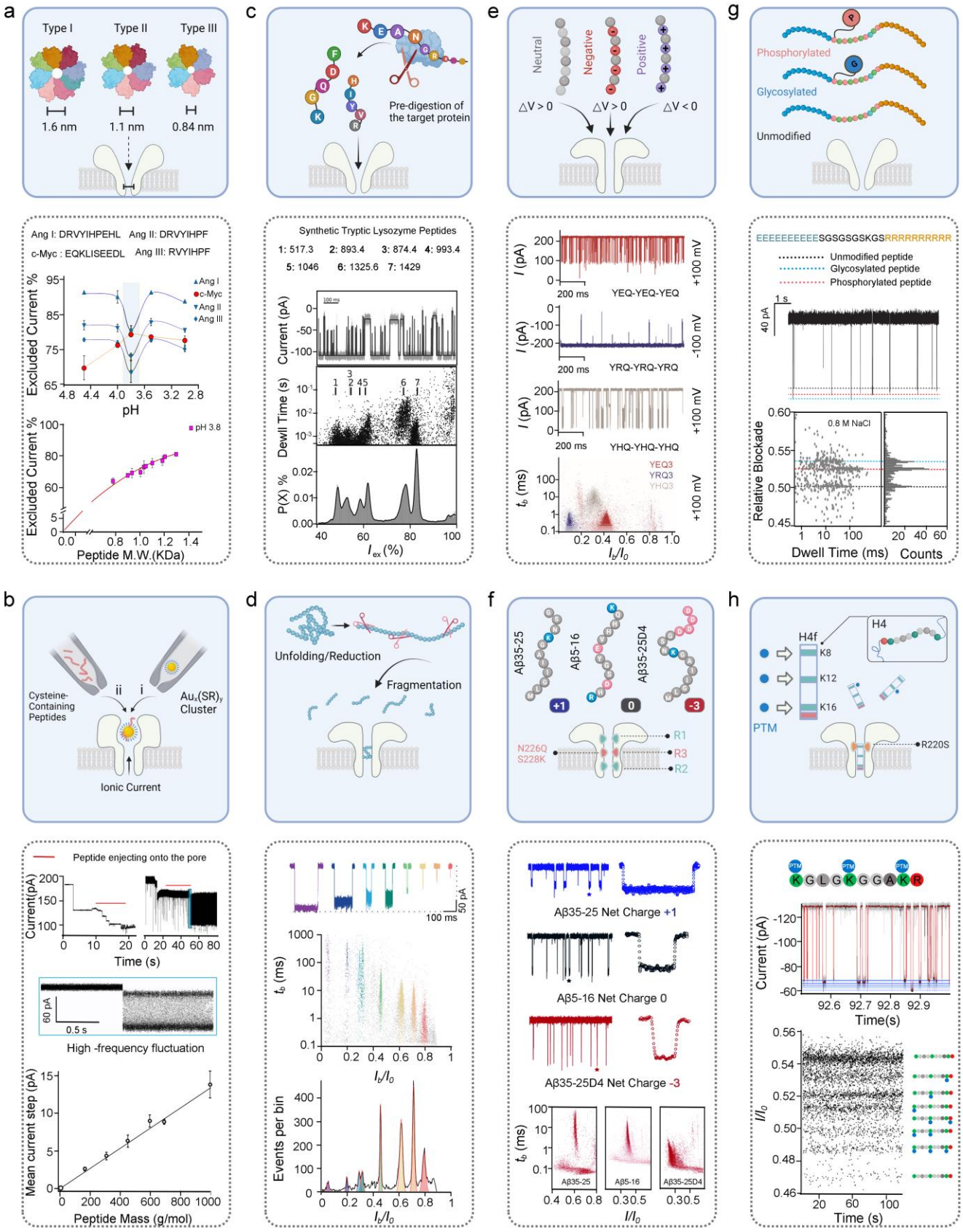

a
Type I
Type II
Type III
1.6 nm
1.1 nm
0.84 nm
Ang I: DRVYIHPEHL Ang II: DRVYIHPF
c-Myc : EQKLISEEDL Ang III: RVYIHPF
Excluded Current %
pH
Peptide M.W.(KDa)
pH 3.8
b
Cysteine-Containing Peptides
Au$_x$(SR)$_y$ Cluster
Ionic Current
Peptide enjecting onto the pore
Current(pA)
Time (s)
High -frequency fluctuation
Mean current step (pA)
Peptide Mass (g/mol)
c
Pre-digestion of the target protein
Synthetic Tryptic Lysozyme Peptides
1: 517.3 2: 893.4 3: 874.4 4: 993.4
5: 1046 6: 1325.6 7: 1429
Current (pA)
Dewll Time (s)
P(X) %
I_{ex} (%)
d
Unfolding/Reduction
Fragmentation
100 ms
50 pA
t_b (ms)
I_b/I_0
Events per bin
e
Neutral
Negative
Positive
△V > 0
△V > 0
△V < 0
I (pA)
200 ms
YEQ-YEQ-YEQ
YRQ-YRQ-YRQ
YHQ-YHQ-YHQ
+100 mV
-100 mV
YEQ3
YRQ3
YHQ3
t_b (ms)
I_b/I_0
f
Aβ35-25
Aβ5-16
Aβ35-25D4
+1
0
-3
N226Q
S228K
R1
R3
R2
Aβ35-25 Net Charge +1
Aβ5-16 Net Charge 0
Aβ35-25D4 Net Charge -3
t_b (ms)
I/I_0
g
Phosphorylated
Glycosylated
Unmodified
EEEEEEEEEESGSGSGSKGSRRRRRRRRRR
Unmodified peptide
Glycosylated peptide
Phosphorylated peptide
1 s
40 pA
0.8 M NaCl
Relative Blockade
Dwell Time (ms)
Counts
h
H4f
H4
K8
K12
K16
PTM
R220S
KGLGKGGAKR
Current (pA)
Time(s)
I/I_0
Time (s)

Figure 3. Identification of peptides with biological nanopores. Schematic diagrams of the sensing strategies and the corresponding representative results in the dashed box. **(a)** Three types of FraC nanopores with different diameters. Box: pH dependence of the Iex% for four peptides using Frac-T2 (top) and the relationship between the Iex% and the mass of peptides (bottom). **(b)** Peptide attachment methodology with nanopore-based cluster analysis. Box: low frequency fluctuation (top-left) and high-frequency fluctuation (top-right) resulting from peptides and cluster-peptides, respectively; the resulting ligand dynamics exhibits two-state fluctuations that can be analyzed to identify the target peptide (middle); linear dependence between the mean current step size and the ligand mass (bottom). **(c)** Peptides are pre-hydrolyzed by protease and measured as they translocate the FraC nanopore. Box: lysozyme fingerprinting using a nanopore. **(d)** Unfolding and cleaving the protein into multiple polypeptide fragment types, analyzed with an AeL nanopore. Box: typical current blockade events of polypeptide fragments (top); scatter plot of t_D versus I/I_0 (middle); and histogram of the I/I_0 values (bottom). **(e)** Analyzing different polypeptides that have identical length, but different net charges and different charge distributions with an AeL nanopore. Box: typical current traces of polypeptides and discrimination of polypeptides through event scatter plots. **(f)** Three electrostatic constricted regions of N226Q/S228K AeL for heterogeneously charged peptide sensing. Box: typical current traces of heterogeneously charged peptides obtained with an N226Q/S228K AeL nanopore (top); scatter plots between current blockade and duration (bottom). **(g)** Label-free detection of both phosphorylation and O-glycosylation and their discrimination from unmodified peptides using a FraC nanopore. Box: typical current trace obtained for a measurement on a mixture of three peptides (top); scatter plot and blockade histogram of the mixture (bottom). **(h)** Discrimination of acetylation-derived positional isomers with the R220S variant AeL nanopore. The fragment (H4f.) of the full-length human H4 protein was modified at three different positions on lysin AAs. Box: current trace of an experiment recorded using the R220S pore in the simultaneous presence of eight different H4f. variants (top); together with the scatter plot (bottom). The error bars in a and b represent the standard deviations calculated from at least three independent repeats. Data in the boxes are extracted with permission from the references: (a): 132,[§] (b): 140, (c): 143,[§] (d): 144, (e): 146, (f): 147, (g): 156, (h): 160, and all corresponding schematic diagrams are created with Biorender.com. [§]https://creativecommons.org/licenses/by/4.0/

2. Profiling and identifying peptides

In proteomics research, the objective of accurate and efficient sensing of peptides and proteins is a tremendous undertaking, due to the remarkable heterogeneity of these molecules and the abundance of dynamic PTMs that underlie their functions. This has sparked the inspiration to develop high-throughput single-molecule techniques to identify and sequence proteins that excel in the areas where traditional MS methods fall short. Compared to the identification of AAs, which focuses primarily on detecting the unique electrical signal generated by each AA passing through the pore, investigating peptides or proteins with nanopore technology requires greater consideration of additional properties. The mass and length of the analyte, which are much more variable among peptides/proteins than simple AAs, may affect how it interacts with the nanopore; larger molecules, in particular, may require altered nanopore configurations. Other factors include heterogeneity in the distribution of charge in the target peptides or proteins, the occurrence of PTMs, and the techniques used for folded structures (such as engineering nanopores to better accommodate the folded conformations, employing unfoldases to obtain more linear analyte structures, or using multiple pores simultaneously for a more complete representation of analyte morphology and chemical nature). Among the early demonstrations of protein sensing using nanopores presented in the early 2000's were studies investigating the capability of the α-HL nanopore in transporting and identifying

a repeating peptide sequence, (Gly-Pro-Pro)n.[124,125] During this time, the interaction kinetics of the α-HL nanopore with various peptides were also detailed[126] and hydrophobic interaction-induced peptide translocation velocities were disclosed.[127,128] These works have helped lay a solid foundation for the ongoing progress toward reliable differentiation and quantification of proteins, peptides, and their PTMs using nanopores over the past two decades.

Mass and length profiling of peptides.

The transmembrane region of the Fragaceatoxin C (FraC) nanopore is unique compared to other biological nanopores, with eight α-helices that funnel down to a sharp and narrow constriction of ~1.5 nm at the *trans* exit of the pore, well-suited for protein-sequencing applications.[129-131] Maglia *et al*. modulated several experimental parameters to investigate their effects on the capacity to identify peptides and proteins of varied mass and length using the FraC nanopore. By substituting a pair of AA residues of the pore for serine at the membrane-FraC interface and adjusting the electrolyte environment and several purification conditions, the authors were able to observe the spontaneous assembly of FraC nanopores into groups of three dominant pore sizes (1.6 nm, 1.1 nm, and 0.84 nm). These engineered FraC nanopores demonstrated the ability to detect peptides with lengths ranging from 4 to 22 AAs and to differentiate alanine and glutamate, which differ in molecular weight by only 44 Da (**Figure 3a**). The authors also found that at a specific system pH (3.8), the depth of the peptide current blockade scaled with peptide mass, irrespective of the chemical composition of the analyte. This provides an efficient single-molecule identification strategy that could potentially be useful for sequencing peptides/proteins in real-time, without prior knowledge of the analyte identity.[132]

Inspired by studies that applied a single molecule nanopore spectrometry (SMNS) method to achieve highly resolved mass spectrums of polydisperse polyethylene glycol (PEG) mixtures,[133-135] Reiner *et al*. sought to expand the SMNS technique to biologically significant peptides. By introducing a 2 nm $Au_{25}(SG)_{18}$ negatively charged cluster[136] to the *cis* side of an α-HL nanopore, the resulting peptide analyte-cluster interactions increased both the frequency and residence time of peptide current blockades.[137] Generally, this effect is caused by the local electrostatic attraction between the anionic metallic nanocluster(s) at the *cis* end of the nanopore and cationic analyte entering the pore from the *trans* end, which slows the analyte's complete translocation through the pore and/or reduces the occurrence of its back-diffusion through its point of entry. This improves the accuracy of blockade depth estimates, consequently enhancing the resolution of blockade distributions. Several works have reported on using nanopore sensors to characterize metallic nanoclusters, taking advantage of the tendency of these clusters to become trapped inside the pore for extended periods. This provides the opportunity to, for example, delineate the kinetic activity and reordering of ligands on a nanocluster's surface in real time[138] and observe and quantify the efficiency of real-time ligand exchange on isolated nanoclusters,[139] allowing for the optimization of cluster surfaces for varied medical applications. The latter of these studies also illustrates a way of adapting the ability to characterize ligand exchange on individual metallic clusters within a nanopore for the purpose of sensing small peptides, where the cluster is used to convert the signals from heterogeneous peptide exchange to distinctive fluctuations in electrical current. This was later expanded upon to better understand what physical and chemical properties of peptides gives rise to different types of fluctuations upon peptide-cluster linkage. This research found that the presence of a cysteine residue on the peptide was required for its attachment to the nanocluster and that the peptide-cluster interaction resulted in either stepwise current transitions or high-frequency one- or

two-state fluctuations, depending on the given peptide (**Figure 3b**).[140] The nature of these current patters was governed by various factors, such as the peptide's mass, charge, sequence, length, and configuration. Both types of peptide-induced signals could be analyzed to identify peptides in mixtures. These studies have laid the investigative groundwork for identifying small, cysteine-containing peptides by integrating metallic nanoclusters inside a nanopore sensor.

Some studies have implemented a fingerprinting method whereby nanopore current blockade data of enzymatically cleaved protein fragments were converted into spectra. These spectra were used to identify the analytes of interest and validate those results with databases of known peptide and protein signatures (as are used in MS).[141,142] In a work using a G13F-FraC-T1 nanopore to directly sample protease-digested proteins, it is shown that the resulting current spectra compare well with those obtained by MS. The method is hence deemed viable for reliable protein recognition by spectral matching, akin to peptide fingerprinting. The study involves the tryptic digest of 10 different proteins, with molecular weights between 12.4 and 66.5 kDa, which generates clustered current events from the individual trypsinated peptides as they translocate the nanopore (**Figure 3c**).[143] Recently, a similar proof-of-concept study was reported by Oukhaled's group,[144,145] demonstrating that hydrolyzed polypeptide fragments can be detected and classified at the single-molecule limit using an AeL nanopore (**Figure 3d**). In this work, three native proteins (myoglobin, lysozyme, and cytochrome c) with similar molecular mass were treated with the trypsin protease, producing polypeptide fragments that were defined and discriminated based on the nanopore's ionic current and then compared to a signal database to successfully identify the original protein. Though protein sequencing using nanopore systems remains an ongoing effort, these advancements in nanopore-based single-molecule fingerprinting represent a meaningful step toward this goal, capitalizing on the existing protein databases used with mass spectrometry while offering a comparatively much simpler methodology.

Transport behavior of charged peptides.

Charge is one of the most important parameters to consider in peptide recognition with nanopores.[115] Variability in charge distribution along a peptide's sequence can hinder the ability of nanopores to generate a clear and distinct signal. Compared with the consistent and predictable current patterns given by uniformly, negatively charged polynucleotides, the presence of charge heterogeneity in peptides can lead to irregular interactions with the nanopore, resulting in inconsistent or ambiguous current disruptions that complicate signal interpretation. Recent efforts have been devoted to controlling the capture and translocation behavior of heterogeneously charged peptides and polypeptides with mutant nanopores.[146-148] One such work studied the translocation of 11 identical-length polypeptides with different net charges and charge distributions to better understand the mechanism by which trypsin-cleaved polypeptide fragments are captured, and how they engender different ionic current drops in the pore.[146] The results show that, under relatively high ionic strength and high applied voltages, negative, positive, and neutral polypeptides can be driven into an AeL pore by the same applied voltage polarity (**Figure 3e**). Specifically, the relatively high ionic strength (4 M *vs.* 1 M KCl) increases the capture rate of polypeptides, independent of their net charge; high voltages (+100 mV *vs.* +50 mV) favor relatively long residence times of positive and neutral polypeptides, but drastically decrease those of negative polypeptides, with a larger reduction in residence time for peptides that are more negatively charged. In addition to the experimental conditions, the ion selectivity and sensing ability of AeL are mostly controlled by electrostatics and

the narrow diameter of the double β-barrel cap.[149] Thus, accurate molecular detection of peptides may be achieved by engineering single/multi-site mutations of the nanopore. Taking this approach, Ying *et al*. designed a N226Q/S228K mutant AeL nanopore to control the capture and translocation of heterogeneously charged peptides by inserting an additional electrostatic constriction, with positive charge, between two natural sensing regions of the pore. The presence of positive charges at the constriction induced a high electroosmotic flow (EOF) velocity, resulting in a significant increase in frequency (up to 8-fold) for heterogeneously charged peptides, with a net charge ranging from +1 to -3 (**Figure 3f**).[147]

Detection of post-translational modifications (PTMs) of peptides.

Post-translational modifications, or PTMs, are one of the mechanisms by which the protein pool of a cell can be expanded by orders of magnitude for realizing diverse sets of functionalities.[150] This entails the chemical modification of protein structures by the addition of a functional group to one or more AA residues to finely tune the protein's properties and activity.[151] PTMs of proteins constitute an essential regulatory mechanism involved in almost all cellular events. Thus, technologies used for their detection at the single-molecule level are in high damand.[152,153] For example, two major PTMs, phosphorylation and glycosylation, are important regulatory processes, but have also been implicated in pathogenic pathways of many diseases, such as diabetes, cancer, and Alzheimer's.[154,155] Joo and Dekker *et al*.[156] demonstrated a proof-of-principle for the label-free differentiation of both phosphorylation and O-glycosylation, and their mutual discrimination from unmodified peptides, using a FraC nanopore and a model peptide, modified at a serine site with either of these PTMs (**Figure 3g**). The experiments in this work employ a model system with bespoke peptides, designed to be simple, but present an important initial step in demonstrating the feasibility of PTM recognition at the single-molecule level using biological nanopore sensors.

PTMs can alter the translocation speed of peptides, which could affect the ability of the nanopore to detect them. For example, phosphorylation can accelerate the traversing speed of a negatively charged substrate while significantly enhancing the translocation frequency of a positively charged substrate.[157] Maglia and Walvoort *et al.*[158] sought to address the issue of rapid translocation speeds of proteins with PTMs that prevented their sufficient observation under a range of typical experimental conditions. Focusing on natural, hydrophilic (glyco-) peptides generated from the proteolytic cleavage of proteins, they found that by using a mutated FraC nanopore (G13F-FraC) with an aromatic constriction region and combining high electrolyte concentrations (3 M LiCl) with low pH (pH 3), the dwell time of the glycopeptides increased, allowing for their accurate, selective detection. To test if these conditions could be generally applied to detect and quantify other PTMs, they repeated the experiment in the same environment with rhamnosylated cyclic peptides and proteins, again with favorable results.

Another work achieving prolonged residence time of peptides with PTMs inside a nanopore, conducted by Long and Ying *et al.*,[159] focused on tau phosphorylation and attributed its success to the design of a T232K/K238Q mutant AeL nanopore. The strengthened electrostatic interaction at the T232K site and the high repulsion barrier at the K238Q site worked synergistically to reduce the translocation speed of the phosphorylated tau peptide by tens to hundreds of milliseconds, down to about 5-70 ms per AA. This was the slowest reported translocation velocity within any nanopore studied at the time. The technique was able to overcome the difficulty in studying tau phosphorylation, caused by its multiple and adjacent phosphorylation sites within the tau sequence, boasting a nearly 100% accuracy in identifying distinct distributions within a mixture of unlabeled

pS262-, pT263-, pS262/pT263-, and unphosphorylated tau peptides. The effective nanopore design featured in this study offers an encouraging engineering concept that can be implemented in several other applications. The strategy is, however, limited to charge-conferring modifications and is incapable of differentiating true positional isomers. To address these issues, Behrends and Aksimentiev *et al.* offered a more generalized approach, showing that whole-molecule sensing by an engineered AeL nanopore (entrapment of an entire peptide inside the pore) is capable of differentiating PTMs solely on the basis of the positions of acetylated and methylated lysine residues. To isolate position as the factor underlying peptide differentiation ability, the work analyzed peptide sequences, derived from human histone H4 protein, of identical mass (**Figure 3h**).[160] Unlike sequencing by stepwise threading, this method detected PTMs and their positions by sensing the shape of a fully entrapped peptide, thus eliminating the need for controlled translocation.

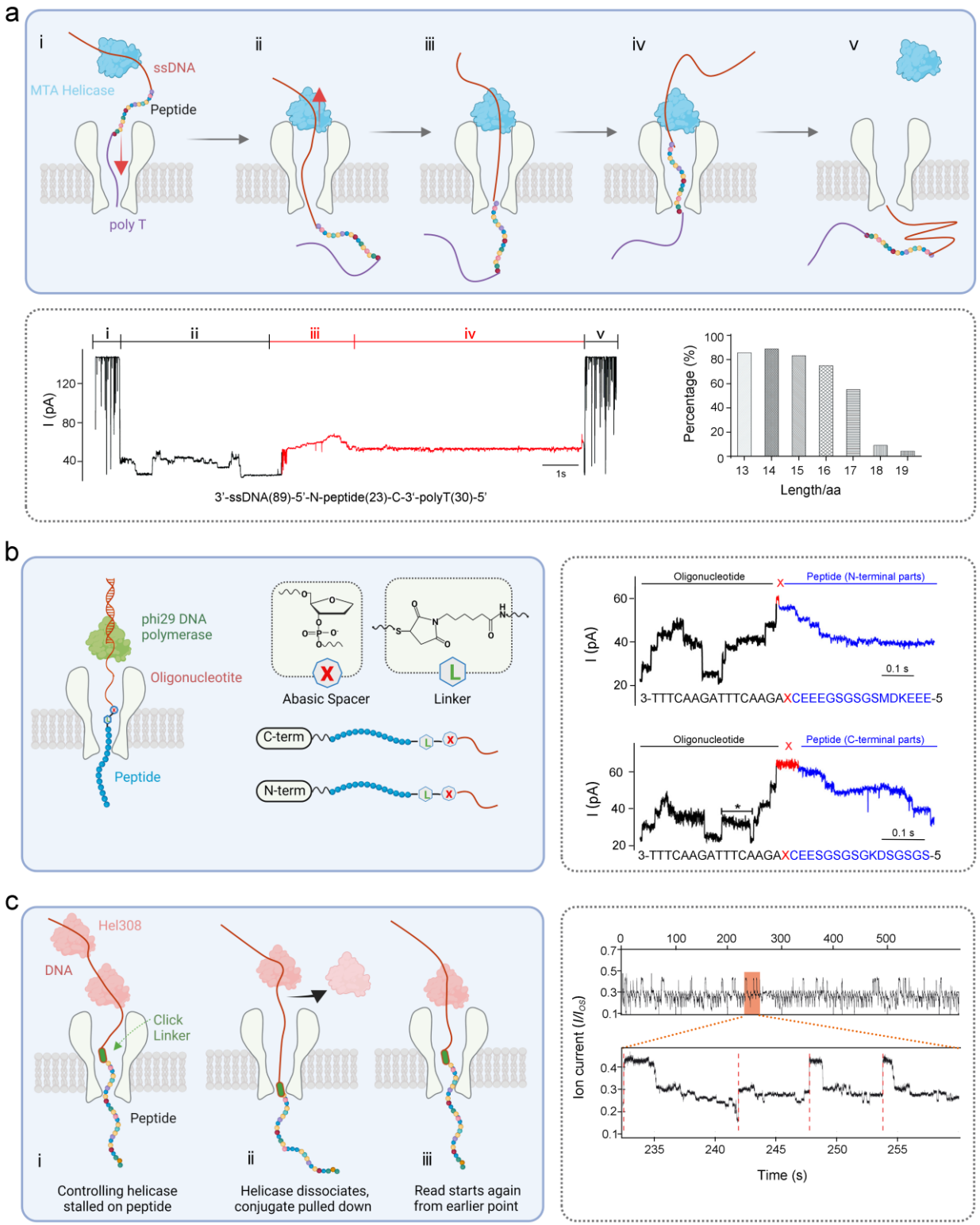

a
i
ii
iii
iv
v
ssDNA
MTA Helicase
Peptide
poly T
I (pA)
120
80
40
1s
3'-ssDNA(89)-5'-N-peptide(23)-C-3'-polyT(30)-5'
Percentage (%)
100
80
60
40
20
0
13
14
15
16
17
18
19
Length/aa
b
phi29 DNA polymerase
Oligonucleotite
Peptide
Abasic Spacer
Linker
C-term
N-term
Oligonucleotide
Peptide (N-terminal parts)
I (pA)
60
40
20
0.1 s
3-TTTCAAGATTTCAAGAXCEEEGSGSGSMDKEEE-5
Oligonucleotide
Peptide (C-terminal parts)
I (pA)
60
40
20
0.1 s
3-TTTCAAGATTTCAAGAXCEESGSGSGKDSGSGS-5
c
Hel308
DNA
Click Linker
Peptide
i
ii
iii
Controlling helicase stalled on peptide
Helicase dissociates, conjugate pulled down
Read starts again from earlier point
Ion current (I/I0)
0
100
200
300
400
500
0.7
0.5
0.3
0.1
0.4
0.3
0.2
0.1
235
240
245
250
255
Time (s)

Figure 4. Controlling peptide translocation through a biological nanopore. Schematic diagrams of the sensing strategies and the corresponding representative results in the dashed boxes. **(a)** Schematic view of peptide sequencing achieved by helicase-driven translocation of DNA-peptide conjugates through the MspA-M2 nanopore. Box: the sequencing signal of the DNA-peptide conjugate with the changes in the ionic current profile corresponding to the stages of the translocation marked in the schematic diagram (left); percentage of translocation events in which the polyT signal was detected for peptides of each length (right). **(b)** Nanopore-induced phase-shift sequencing strategy to observe the ratcheting motion of peptide-oligonucleotide conjugate (POC) with an MspA nanopore. The abasic spacer (X) serves as a signal marker separating the oligonucleotide and the peptide. The linker (L) conjugates the two parts. Box: representative trace of *N*-termini conjugated POC (top) and *C*-termini conjugated POC (bottom). **(c)** Rereading is facilitated by helicase queueing with a MspA nanopore. Box: highly repetitive ion current signal corresponding to numerous rereads of the same section of an individual peptide. The expanded plot below shows a region that contains four rewinding events, where the trace jumps back to the level of the consensus displayed in shadow. The data in the boxes are extracted with permission from the references: (a): 177, (b): 179, (c): 180, and all corresponding schematic diagrams are created with Biorender.com.

3. Manipulating the motion of peptides

Prolonging peptide residence within a nanopore via electroosmotic flow (EOF) enhancement.

An important area of research within the domain of nanopore-based protein sequencing is the development of methods to slow down the high velocities of peptides as they traverse across pores (with the majority of transit times estimated to be <1 ms·AA^{-1}).[81] Sufficiently long residence within the nanopore can allow high signal-to-noise recording of time-series signals for the well-ordering of each AA.[81,161] As mentioned earlier, using mutant nanopores is an effective way to slow down the translocation speed of peptides. Another potentially powerful approach is manipulating the EOF. EOF in nanopore systems is the movement of electrolyte fluid through a nanopore in response to an applied electric field, which can affect the accuracy and resolution of nanopore-based sensing and sequencing. In general, a dipolar peptide can be trapped within a nanopore for a time by controlling the balance of the opposite electroosmotic and electrophoretic driving forces. To date, much of the nanopore research exploiting EOF has been devoted to controlling the translocation and trapping of peptides by: (1) modulating the magnitude and polarity of the applied voltages or the charge distribution of the analytes,[115,162,163] (2) introducing oppositely charged segments at the *N*- and *C*-termini of polypeptides to create a dipolar feature,[114] (3) tuning the solution pH,[164] (4) introducing electro-osmotic vortices into the inner surface of nanopores,[165] (5) immobilizing dihydrofolate reductase (DHFR) inside a ClyA biological nanopore,[166] and (6) binding guanidinium cations to the inner surface of the nanopore.[167]

Sequencing by controlled movement of peptides in a nanopore.

Nanopores have gained substantial recognition in genomics by demonstrating ultralong read lengths for DNA and RNA sequencing.[80,168,169] In this process, translocating motor enzyme is used to ratchet DNA/RNA through the pore in single-nucleotide steps, yielding a base-by-base sequence readout. In proteomics, many recent studies that focus on protein fingerprinting or sequencing apply a similar principle, using an enzyme motor.[170,171] For example, by employing a ClpX protein, a proteasome-like complex of *Escherichia coli*, as an unfoldase,[33] different proteins with fully-

characterized folded domains were shown to be driven through an α-HL nanopore and effectively yield distinguishable signals.[172] However, these reads are difficult to interpret, in part due to the irregular stepping behavior of ClpX and the low spatial sensitivity of the α-HL pore.[173,174,175]

The homo-octameric "globlet-like" configuration of the *Mycobacterium smegmatis* porin A (MspA) nanopore features a ~1.2 nm (in diameter) constriction located at the very bottom of the channel, which could procure better read lengths of some analytes and particularly high current levels, relative to other protein pores.[176] Bai *et al.* selected the MspA nanopore for its unique geometry to test the idea of conjugating a target peptide to a single-stranded DNA handle, regulated by a DNA helicase, to better control nanopore-based peptide transport through the pore (**Figure 4a**).[177] Their method was able to thread peptides of up to 17 AAs through the constriction of the MspA pore and differentiate AA residues based on their charge and position. It also demonstrated the capability of distinguishing between different phosphorylation sites, with the potential to expand its application for the detection of other PTMs. However, the blockade current given by the translocating peptide was somewhat irregular, failing to produce consistent step-like signals. Consequently, the researchers were unable to sequentially discriminate between individual AA residues. The authors attributed this observation to the smoothing effect of thermal motion on the blockade current generated by different AA residues within an unstrained peptide and suggested that engineering the MspA pore to reduce the size of its constriction zone and stretching the peptide chain to minimize thermal motion may be solutions to the issue.

Clear and step-like ionic current blockades have been achieved in nanopore-based nucleic acid sequencing.[168] For example, a nanopore-induced phase-shift sequencing (NIPSS) procedure was recently developed using an engineered MspA pore, in combination with a motor enzyme, to directly sequence 2'-deoxy-2'-fluoroarabinonucleic acid, a type of xeno-nucleic acid.[178] Inspired by this technique, Huang *et al.* took advantage of the high resolution of NIPSS to directly observe the ratcheting motion of a peptide within a nanopore, which had not been accomplished prior. Conjugating an oligonucleotide to the *N*- or the *C*-terminus of a peptide to form a peptide-oligonucleotide conjugate (POC) and using a wild-type phi29 DNA polymerase (phi 29 DNAP) as the ratcheting enzyme, the study reported the observation of the discrete steps of a peptide's ratcheting motion during a NIPSS measurement (**Figure 4b**).[179] Furthermore, the event patterns generated during NIPSS showed a clear sequence-dependence, enabling identification of event variations caused by a single AA substitution. The study did, however, face a problem in regard to the incompatibility of phi29 DNAP's enzymatic activity with the peptide segment of the POC, resulting in a limited read length, roughly equivalent to the length of ~14 nucleotides.

Dekker *et al*. have recently improved the accuracy of a single-molecule protein sequencer through multiple "rereads" of analyte.[180] The work used a DNA translocating motor to pull a peptide through a MspA nanopore, demonstrating the ability to obtain high-fidelity signals by rereading the same peptide molecule multiple times (**Figure 4c**). Specifically, three negatively charged synthetic peptides that differed by a single AA (Asp, Gly, Trp) were conjugated with an 80-nucleotide DNA strand to form DNA-peptide conjugates. A Hel308 DNA helicase, capable of half-nucleotide (~0.33 nm) stepping, was present in the solution and allowed to bind repeatedly to the DNA molecules. An individual peptide could thereby be read, pulled back into the MspA nanopore by DNA, and then reread with the help of a new helicase, improving read accuracy with each additional read. This process generated step-like pattern signals that enabled detection of the three ''mutant'' peptides, with an initial 87% single-read accuracy, that continuously improved with an increasing number of

rereads, ultimately resulting in an undetectably low error rate (<1 in 10^6).[181]

Similar to the aforementioned strategies, the application of this method to protein sequencing poses two main challenges. First, its efficacy in sequencing natural peptides comprising AAs of varying charge and polarity remains uncertain. Second, the maximum number of distinct AAs that can be mapped and differentiated on a peptide for sufficient sequencing or identification is unclear. Additionally, the peptide sequence tested in Dekker's strategy was atypical, with near-uniformly negatively charged target peptide chains. In a mixture of natural, heterogeneously charged peptides, conformational changes of the peptides within the nanopore may cause significant variations in the current blockages, a factor that must be properly accounted for to understand protein dynamics.[182]

4. Artificial intelligence aided signal processing and recognition.

Signal processing and artificial intelligence (AI) techniques, especially machine learning, have been increasingly utilized in the field of nanopore sensing.[183] The development of advanced machine learning algorithms has vastly improved the capacity to analyze complex patterns in pore signals, efficiently and accurately classify peptide and protein sequences with corresponding confidence scores, perform error correction, predict sequences based solely on the observed pore signals, and engage in continuous refinement and optimization through iterative model training. These algorithms can also rapidly process large amounts of data, enabling real-time analysis and high-throughput sequencing. **Figure 5a** shows a typical flow diagram of the machine learning training process to attain the best performing predictive model.[118,184-186] Relatively early in the field of machine learning-aided protein sequencing, Akeson *et al*. used a Gaussian Naïve Bayes classifier to recognize different variants of proteins using an α-HL nanopore, with the assistance of a ClpXP protease, achieving an 86-99% accuracy in the recognition of five of the S2-GT variants under a five-fold cross-validation setting.[172] A few years later, Aluru *et al*. investigated the use of molybdenum disulfide (MoS_2) as a nanopore material, to explore its ability to characterize the ionic current and residence time of the 20 proteinogenic AAs using MD simulations and machine learning techniques. They showed success in identifying individual AAs with high accuracy and found that, on average, the smaller the diameter of the pore, the (much) greater its sensitivity (a 20-fold increase in sensitivity for a 1.85 nm, relative to a 2.52 nm diameter MoS_2 pore), noting the size of the analyte of interest naturally defines the lower limit on pore diameter. They also demonstrated the competency of three different predictive machine learning models in predicting AA type, given the ionic current and residence time from nanopore analysis. These included logistic regression, k-nearest neighbor, and random forest, which displayed a 72.45, 94.55, and 99.6% recognition accuracy, respectively.[187]

These promising early results have motivated the continued application and advancement of machine learning-based tools in conjunction with simulation within the domain of nanopore research. De Angelis and Rocchia *et al*. have recently detailed a design for a protein construct, designated by the authors as an "adaptive biological nanopore", with a pore structure comparable to a traditional nanopore (*e.g.*, α-HL, MspA, *etc*.) but with a potentially superior sensing mechanism, optimized using MD simulation. In this proof-of-concept work, the atomistic interactions between a translocating polypeptide and the pore-construct, delineated by the MD simulations, cause a unique "reshaping" of the pore, specific to the "shape" of each AA. The definition of "shape" here includes the different types of interactions the molecule engages in (*i.e.,* steric, electrostatic). Machine learning analysis (support-vector machine and random forest) provides information about each morphological transformation of the adaptive nanopore, provoked by the translocation of each constitutive AA of the peptide sequence, and helps identify these AAs. The results of this

computational approach offer optimism for the ability to detect all 20 AAs when applied in an experimental setting, with further refinement of the method.[188]

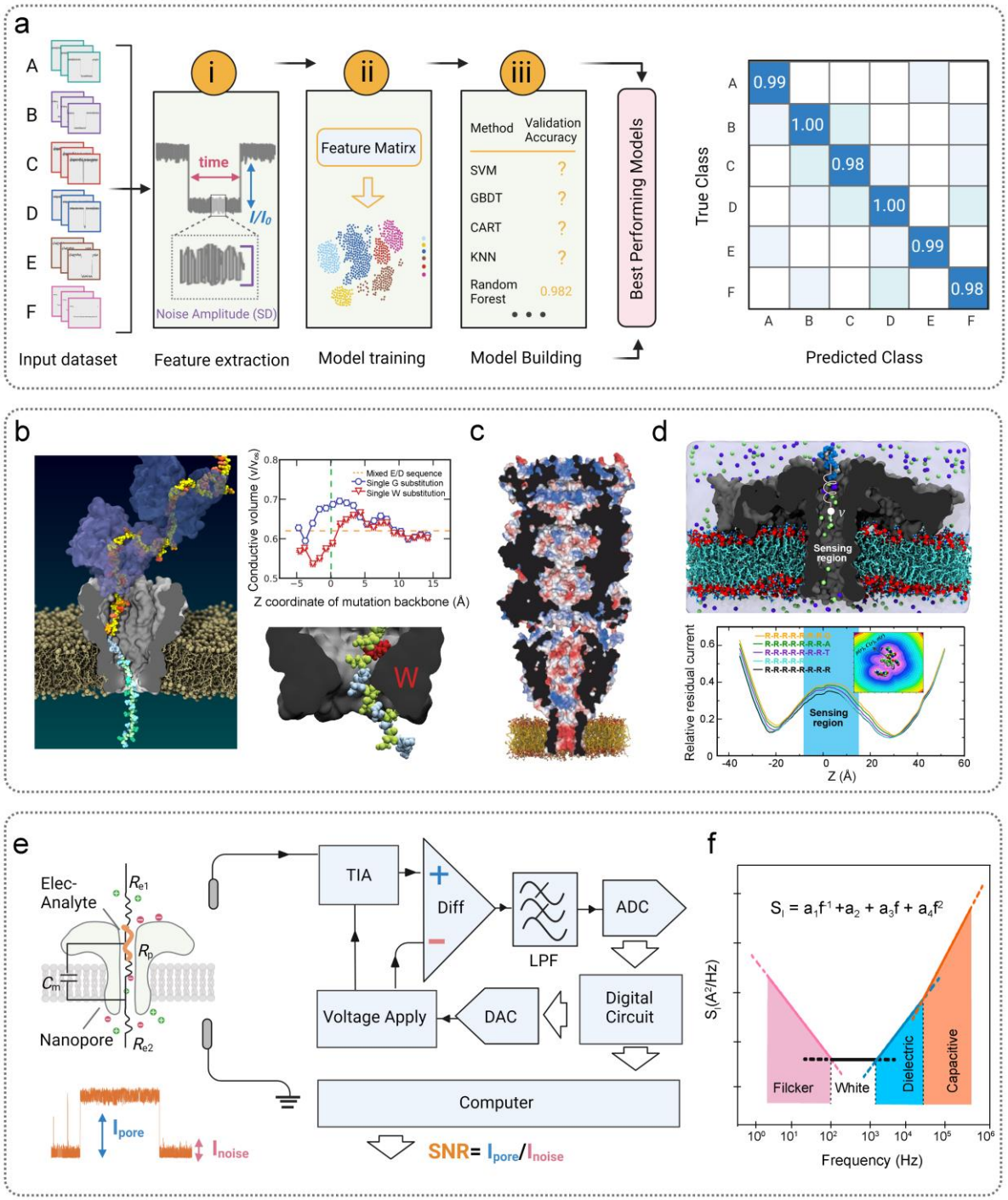


Figure 5. (a) AI aided signal processing and recognition. Left: Typical flow diagram of the training process. Different classes of events including A, B, C, D, E, F, were applied as the input dataset. The time, I/I_0, or standard deviation of each event were extracted to form a feature matrix. Results in the matrix were further randomly split into a training subset for model training and a validation subset for model validation. Right: The confusion matrix of classification generated using the best performing model. **(b-d) Understanding molecular dynamics by simulation.** b: All-atom model of a helicase-assisted protein sequencing platform recapitulates dependence of the ionic current blockade on peptide sequence and reveals its molecular origin. Images courtesy Jingqian Liu (UIUC). c: Multi-scale model of a cut-and-drop experimental system. d: Combining steered MD simulation of peptide transport with a steric exclusion model enables precise characterization of atom-scale modification on nanopore current. **(e-f) Instrumentation of nanopores.** e: Schematic representation of nanopore sensing instruments. TIA – trans-impedance amplifier, Diff – differential amplifier, LPF – low-pass filter, ADC – analogue-to-digital converter, DAC – digital-to-analogue converter. f: Noise power spectral density (PSD) as a function of frequency for a typical nanopore. Dominant noise sources at different frequency ranges. The data in the boxes are extracted and reproduced with permission from the references: (b): 180, (c):197, (d) 110, (f): 229. The schematic diagrams of (a) and (e) are created with Biorender.com.

In recent years, deep learning models, which are usually implemented by deep neural networks, have been employed to process and classify nanopore signals.[189] For example, a deep learning model implemented by long short-term memory (LSTM) networks and multilayer perceptron (MLP) has been trained to deal with nanopore current signals and classify polymer sequences.[190] Another work developed a deep neural network with four fully-connected layers to process ionic current signals of nine different single-residue mutants of the thioredoxin model protein collected using α-HL nanopores. After performing a series of optimization techniques to enhance the neural network's capability to generalize its training experience (*i.e.*, successfully interpret data not introduced in the training and validation sets), the method achieved 100% correct identification of the nine mutants in 54% of the experiments.[191] One study has reported the use of a bi-path network, specifically designed to process nanopore signals. Based on the fundamental architecture of residual neural networks, the bi-path network follows a two-way architecture and jointly predicts the number of pulses, as well as the average translocation amplitude duration.[189]

Current studies have leveraged artificial intelligence to facilitate the analysis of nanopore signals and have demonstrated promising potential. However, there are still several remaining research challenges in the computational domain of nanopore protein sequencing. First, the temporal dynamics of nanopore signals have not been sufficiently modeled in current studies. Many studies directly apply standard machine learning models, such as support vector machine and random forest, to classify the nanopore signals, yet these models cannot effectively exploit the dynamic patterns in sequential data. Second, the robustness of machine learning models should be carefully considered when dealing with nanopore signals in practice. The features extracted by current machine learning models might be incapable of distinguishing AAs in real-world scenarios, due to noise and variability in raw signals. Recent methods to ensure trustworthy machine learning, such as adversarial training, could be adopted for nanopore-based AA recognition.[192] By leveraging advanced machine learning paradigms, such as transfer learning and multi-task learning, multiple relevant tasks could be jointly optimized and work synergistically. The success of deep learning in other fields, such as natural language processing, protein structure prediction, and image generation, has demonstrated the effectiveness of large pre-trained models. Training large deep learning models for nanopore signal

analysis would require collaborative efforts from AI, mathematic modeling, physics, chemistry, and engineering communities, with the potential to fundamentally advance the field of nanopore sensing in proteomics.

5. Understanding nanopore transport by molecular dynamics simulation.

MD simulation of nanopore related translocation[193] has emerged as a powerful tool to complement experimental development of nanopore sensors for identifying protein structure and sequence.[84] Atom-by-atom, complete microscopic models of various experimental systems have been constructed,[180,194-197] typically featuring a single nanopore embedded in either a synthetic or biological membrane, a peptide or a protein molecule placed inside or near the nanopore, and electrolyte solution. An applied external electric field, directed normal to the membrane, produces an ionic current through the nanopore and its magnitude, together with the size of the system, governs the transmembrane voltage bias.[198] Repeating the simulations for various placements of the target peptide in the nanopore, and/or changing the AA sequence of the peptide, produces the dependence of the ionic current blockades on peptide conformation and/or sequence (**Figure 5b**).[180] This information can be directly compared to experiment to verify results. While the brute-force, all-atom MD method has the spatial resolution required to relate the protein sequence to its ionic current signature, the duration of processes that can be investigated by this method is presently limited to single-digit microseconds, which is often too short to directly observe the transport of a polypeptide chain through a nanopore. Although conventional coarse-grained simulations, such as those employing the MARTINI force field, can explore the equilibrium properties of multi-protein-nanopore assemblies (**Figure 5c**),[197] they lack the resolution to relate the protein sequence to the nanopore blockade current. To address the time-scale problem, hybrid methods that combine coarse-grained[199] or continuum[200] models with all-atom MD or enhanced sampling simulations[201] can be used to extend simulation time, while retaining a high level of atomic detail for molecules of most interest. Some studies (**Figure 5d**)[110] have combined steered MD simulations of polypeptide transport through an AeL pore with a steric exclusion model of the ionic current blockades[200] to resolve subtle differences in the chemical structure of the analyte, such as the type of individual AAs attached to a polyarginine carrier[110] or the placement of PTMs along a tail of a histone protein.[160] All-atom MD and hybrid simulation methods are imperative for studying biomolecular systems and uncovering the molecular mechanisms underlying puzzling experimental observations. Advances in computing power and simulation algorithms continue to push the boundaries of what can be achieved for these complex systems. For example, atomistic simulations were able to attribute the steady, unidirectional transport of proteins in guanidinium chloride solution[167] to a strong electro-osmotic effect produced by the binding of guanidinium ions to the surface of the nanopore.

Notwithstanding the ongoing progress in the development of software and hardware, enabling ever faster MD simulations, predictive modeling of nanopore blockade signatures in the context of protein sequencing remains difficult. The key challenge lies in statistical sampling of protein conformations within the nanopore, which typically occur at time-scales inaccessible to the brute-force approach. Additionally, enhanced sampling simulations have not been developed to operate under applied electric field conditions.[201] The advent of special purpose computing hardware, such as the D.E. Shaw Anton 3,[202] offers a promising means of simulating the behavior of complex biomolecular systems. With a simulation speed of ~200 microseconds per day, this technology can provide adequate statistical sampling of polypeptide conformations directly from a brute-force simulation. Another obstacle facing simulation-based protein sequencing is the propensity for errors

in results arising from imperfections in the applied molecular force field.[184] Fortunately, improvements to the force fields are ever ongoing,[203] producing more sophisticated and realistic models that make the prospect of accurately predicting blockade currents from all-atom simulations continuously more feasible.

6. Device and instrumentation

Detection and characterization of single biomolecules in bulk solution using biological or solid-state nanopores are crucially dependent on reliable instrumentation. With careful systems-level engineering, from the electrodes to the nanopore setup, bandwidths approaching 10 MHz can be achieved.[204] Multiple commercially available systems are currently available in the field from various companies.[135,205-208] Continuous improvements in equipment, refinements in the underlying theory, and optimization of experimental setups together further the achievements in single-entity recognition and biomedical diagnostics,[209-212] establishing nanopore technology as an ever-more powerful tool for DNA sequencing and a promising, emerging tool for protein sequencing.[109,132,213-215] However, in regard to complex protein analysis, this early-stage technology still encounters a number of hurdles that must be overcome. In addition to obstacles with fabrication, functionalization, or folded protein structures,[216-219] bottleneck instrumentation issues, such as low signal-to-noise ratio (SNR) or lack of suitability for multiplexed screening, can limit the real-world biotechnological applications.

Typical nanopore-based detection and characterization of analyte is relatively straightforward: when constant voltage across two chambers is applied, both ions in solution and the molecule of interest are driven through the nanoscale pore. The analyte's occupation of the pore and the resulting analyte-pore interactions block the ionic current, causing a single (sub-microsecond) event detectable by electronics. A general instrumental setup is universal for different single-entity measurement devices, including commercial or custom-made equipment, based on biological or solid-state nanopores (**Figure 5e**).[220-226] Typically, the picoampere current is amplified by a trans-impedance amplifier (TIA) and converted into an analog voltage signal, which is then digitalized with an analog-to-digital converter (ADC). The signal is then modulated *via* a conditioning circuit that includes a low-pass filter to suppress high-frequency noise and other application-specific integrated circuitry, thereby increasing the SNR. With the small ionic currents typical of nanopore measurements, it is essential to suppress the noise to avoid significant error and increase the sensitivity.[227,228] The nanopore system exhibits different noise sources that can be depicted as a noise power spectral density (PSD) function of frequency (**Figure 5f**).[224,229] The noise PSD rise at low frequency is dominated by the ionic friction inside the nanopore.[230] The frequency-independent white noise[231,232] is dominant in the moderate-frequency range. Finally, the noise PSD at high frequencies is attributed to dielectric loss or capacitive noise from the chip and the amplifier.[233,234] Considerable research has been carried out on the origins of the noise in biological and solid-state nanopores and various methods have been suggested for improvement.[235-240] While the origin of the low-frequency (flicker) noise is not yet fully clear, other noise components at higher frequency (**Figure 5f**) can be impacted by instrumentation, especially amplifiers.[234] Hence, improvement of the signal-to-noise ratio can be achieved by optimizing the membrane capacitance[241-243] and reducing the amplifier input capacitance with electronic filters. However, such filters limit the bandwidth of the ionic current, causing the loss of information when analyzing RNA, DNA, and proteins with very short translocation times. It was demonstrated, that the SNR can be improved by suppression of applied voltage noise.[244] Significant instrumentation upgrades can be achieved by using complementary

metal-oxide-semiconductor (CMOS) integration.[224] CMOS-amplifier chips allow incorporation of fluidics directly on the amplifier, reduced wiring, and extension of the measurement bandwidth into the MHz range for nanosecond translocation events with high SNR.[245-247]

The emergence of nanopore sensing from a niche curiosity to a potential workhorse for sequencing and protein identification has resulted in an explosion of purpose-built systems. No longer reliant on repurposed, low-current amplifiers designed for patch clamp measurements, these recent systems offer potential for portability and multiplex measurements. The incorporation of CMOS-integrated electronics together with high integration microfluidic systems[248] enables the fabrication of multiplexed, multi-nanopore chips for high-throughput, parallel analysis.[246,249] The concerted application of such instruments already allows nanopore array sensors with more than 65,000 electrodes[250] and a lower per-sample cost for DNA sequencing.[251]

In addition to the variety of instrumental advances that have been discussed here, there has been considerable work devoted to the coupling of optical fields with nanopore sensors for improved detection and characterization of proteins and peptides. These optical and photonic applications have been reviewed extensively.[252-255] Optical approaches offer several advantages to purely resistive pulse methods. Large scale nanopore arrays can be integrated with optical techniques, which enables multiplex detection with little to no crosstalk. By coupling wide-field microscopy with high-speed cameras capable of single photon detection, the photonic sensors can operate in a molecular counting mode, similar to a resistive pulse detector. This is a major advantage over single pore analysis because pore size is directly related to the on-rate of target analyte. Like with resistive pulse sensors, increasing sampling rates is necessary for these cameras to produce clear detection and analysis of individual molecular transits through the pores. Nevertheless, efforts are proceeding in this area with the long-term goal of developing these arrays for early-onset diagnosis, primarily in specialized facilities such as large hospitals and dedicated research institutions.

It is also possible to combine electrical (resistive pulse) and optical methods for single molecule detection.[256] By coupling nanopore interfaces to optical recording setups *via* a total internal reflection fluorescence microscope (TIRF), a confocal microscope, or a simple nanopipette positioned near a wide-field objective, optical signals can be recorded with time-synched ionic current measurements.[253] The large electromagnetic fields from focused laser illumination affect molecular transport through the pore, which can be controlled on the nanoscale through plasmonic nanostructures serving to enhance nanopore detection. This gives rise to applications related to the detection of a variety of different molecules (*e.g.*, DNA, proteins, and biomarkers). Improvements in fabrication techniques are still required to enable <20 nm precision for these arrays. Finally, a polymer unfolding step, prior to insertion into the pore, is required to enable protein sequencing. This might be achieved with plasmonic-based heating. Plasmonic modifications to nanopore sensing are discussed in detail by Garoli and Wang *et al*.[254] and Wanunu and Garoli *et al*.[255]

Plasmonic nanostructures have become a powerful tool for nanopore sensors since their introduction in 2015.[257] Aligning with other trends in the field, much of the focus has moved towards sequencing applications.[254,255] Due to the high optical fields produced between closely linked plasmonic nanostructures, plasmonic-modified nanopores can enhance single molecule detection *via* FRET and even surface enhanced Raman spectroscopy (SERS). The review by Garoli and collegues[254] suggests that future work might explore the use of optical, thermal, magnetic, and electro-osmotic forces to better control molecular transport through the pores. This can include denaturing nucleic acids by melting out base-pair stabilized duplexes or unfolding proteins through

carefully controlled temperature impulses.[258] Also, the ability to perform SERS on progressively transported polymers offers a strong candidate for single molecule analysis, given that SERS spectra exhibit limited spectral overlap between molecules and thus avoids the limitations of size-exclusion analysis, inherent in nanopore resistive pulses.

Plasmonic-based systems enable heating protocols, which are most likely present in solid-state nanopore systems. There has been some interest in implementing temperature control operations in biological nanopore systems to better understand transmembrane transport.[259-264] Interestingly, plasmonic heating has been demonstrated in conjunction with α-HL *via* DNA-tethered linkage of gold nanoparticles.[258] This work demonstrated the possibility of applying heating steps at time scales commensurate with intrapore molecular fluctuations. More recent efforts have used infrared laser heating to separate the entropy and enthalpy components of the free energy barrier to the escape of peptides and neutral polymers (PEG) from the nanopore environment.[265] This work could be used to help optimize nanopore designs for peptide and proteomic sensing applications.

7. Challenges and opportunities

Given the great success in genetic sequencing, nanopore-based single-molecule approaches have been one of the most robust potential methods for single-molecule proteomics.[266] Protein sequencing is an essential aspect of proteomics, serving as a fundamental tool for identifying and discovering proteins and characterizing PTMs that are key to understanding their functional dynamics. In recent years, biological nanopore-based single-molecule stochastic sensing for protein sequencing has been under intense investigation. This review summarizes the recent major advances in this field. Starting from the main elements that make up proteins, we present an overview of the application of nanopores for individual AA identification and peptide sensing, as well as strategies for controlling the movement of peptides through the nanopore. From the engineering perspective, we highlight the critical roles of machine learning and molecular simulation in data processing and interpretation. Finally, we discuss the design and development prospects of nanopore devices for protein sequencing in the real world.

Significant scientific progress is often met with significant hurdles. First, identifying single AAs with a universal method remains challenging. Although recently developed strategies such as chemical modification-assisted and peptide-assisted sensing have greatly promoted AA identification using nanopores, a feasible scheme capable of accomplishing the ultimate goal of sequencing a protein through accurate, individual identification of all component AA molecules is still lacking. The task of developing a uniform modification method for different AAs is further complicated by the need to incorporate target AAs into pre-designed, charged peptide chains while preserving the self-identity of the AAs. Furthermore, due to the diversity of AAs, current blockage measurements commonly used for analyte identification are insufficient for protein sequencing;[267] more parameters from current traces, such as standard variation of the current,[268,269] peak shape,[270] and the frequency in the current recording need to be considered.[22,227,228] Evidently, achieving nanopore protein sequencing through AA discrimination remains an ongoing endeavor that requires a multidisciplinary approach and systematic engineering solutions.

Second, PTMs add another layer of complexity to nanopore-based protein sequencing. Although recent advances suggest that PTM identification could be attainable with nanopore sensing directly, or in combination with PTM-specific labeling, current successful results are still limited to phosphorylation and glycosylation. Considering the hundreds of various PTMs, significant overlap of the signals they produce inside the nanopore is expected to reduce identification accuracy.[271]

Additionally, the molecular mass of some PTMs can substantially surpass that of AAs, preventing peptide translocation through the nanopore. Certain enzymes necessary for protein sequencing may also affect or change PTMs.

Third, unfolding proteins into polypeptides and controlling the translocation of peptides through nanopores are complicated tasks. To date, the peptides studied for nanopore sequencing have been deliberately designed to contain solely negatively charged AA residues, in order for them to be pulled into the pore and stretched under an applied potential for readout. In an ideal system, proteins should be translocated, in their entirety, as extended polypeptides through a nanopore for sequencing and variation mapping.[271] As natural peptides have random charge distributions, other mechanisms are urgently needed to provide the stable threading and pulling of peptides through the nanopore. Recently, a nanopore sensor based on peptide volume recognition that consists of three co-assembled proteins was proposed to control the unfolding and threading of individual proteins. The mechanism responsible for such control is an archaeal 20S proteasome, built directly into the nanopore. Fragmented peptides or intact polypeptides can be read by the nanopore sequentially.[197] This engineering development suggests that significant improvements in polypeptide translocation are on the horizon. Another important aspect regarding the controlled translocation of peptides is the effect(s) of transient changes in peptide conformation inside the nanopore, the understanding of which is essential to accurately describe folding/unfolding pathways of dynamic proteins.[272-275] For example, the conformations of small peptides that enable their entry into the nanopore may also lead to significant current fluctuations, which are only now beginning to be understood. Due to the short-lived transitions and sub-nanometer conformation differences involved, the experimental acquisition of the transient conformations and dynamics of peptides is largely beyond the ability of current instrumentation. Possible strategies to address this issue include improving the instrumentational resolution on the microsecond-to-millisecond scale and achieving high-throughput readouts of individual peptides in a way that retains their natural dynamic character.[182]

Finally, throughput remains a bottleneck, obstructing the full potential for proteome sequencing.[181] The number of proteins, even in a single cell, is massive (*e.g.*, a simple eukaryotic cell has ~40 million proteins).[276] The complete proteome contains 16 billion AAs, with an average length of 400 AAs in each protein. In addition, translocation event signals from whole proteins or protein fragments with the full set of 20 proteinogenic AAs is complex and very difficult to interpret.

The challenges described here are simply matters of engineering and can be surmounted with thoughtful engineering solutions. Clearly, the potential opportunities that can be realized with continued efforts to solve these problems are well worth the required resources. The development of higher-precision molecular tools will permit the tailoring of a nanopore's size and shape, beyond current engineering capabilities. Such control of the nanopore's morphology can be used to optimize its sensing ability, ensuring measurable changes in the ionic current amplitude above the noise level, for individual peptides, proteins, and PTMs. This fine-tuning can be applied to artificial nanopores: solid-state nanopores,[62,277-279] nanopipets,[280-282] chemosynthetic membrane channels,[283] and hybrid nanopores,[284,285] as well as biological nanopores: DNA-based channels,[206,286-290] peptide-based transmembrane pores,[291] helicase nanopores,[292] ligand-gated pores,[293-295] transmembrane β barrels,[296] voltage-dependent anion channels (VDAC) of the mitochondrion,[297,298] *etc*., as needed, depending on the experimental conditions and particular analytes of interest. Additionally, simultaneous sensing with a range of different nanopores could lead to a comprehensive understanding of proteome and protein isoform diversity.[299] Other than a nanopore's macroscopic morphology, advanced method-

based functionalization of the nanopore's inner surfaces can also be implemented. For example, chemical modification,[300-303] mutations,[176,304,305] and combination with aptamer molecules[306,307] *etc*., have the potential to provide the required sensitivity, selectivity, and capture efficiency by manipulating interactions between proteins/peptides/AAs and nanopores.[308-310] All of these advances will be complemented with improvements in protein-folding predictions. From the current recording and signal analysis perspective, some noteworthy investigations have been performed on current noise fluctuations and signal bandwidth,[229,230,233] calibration of instrument differences,[223] improving signal-to-noise ratio (SNR),[311] increasing signal specificity,[71,312] and introducing more dimensional parameters for data analysis,[313] *etc*., which have made significant strides in the nanobiotechnology arena.

On a last note, combining the nanopore-based approach with other methods of protein analysis is likely to prove beneficial. Several recent protein fingerprinting methods, based on the readout of a subset of residue types, have begun to adopt such an integrative approach towards biomolecule sequencing.[33,57,59,314,315] For example, the combination of MS and a nanopore ion source was developed for sequencing single proteins.[316] In this concept, the nanopore electrospray was used to guide a protein into a linear configuration by delivering individual AA ions directly into a mass spectrometer sequentially; thus, the ions could be efficiently detected using their mass-to-charge ratios.[317] The field of nanopore-based protein sequencing has been active, dynamic, and filled with innovation and opportunity, yet it is still in its early stages with many hurdles left to surmount. Through continuous efforts and interdisciplinary collaborations, we can expect to gradually overcome the current technical barriers and achieve further optimization of nanopore sensing systems, leading to broader applications of the methods outlined in this review. Such progress will significantly advance the field of protein research, providing valuable tools and insights for a deeper understanding of biological processes. In engineering, any great progress and success stems from the intersection and fusion of multiple disciplines. The field of nanopore-based protein sequencing is no exception.

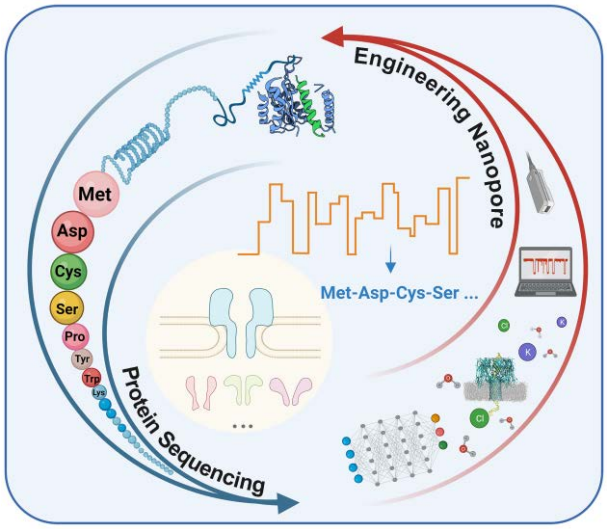
Competing Interests

Certain commercial materials, equipment, and instruments may be identified in this work to describe the experiments as completely as possible. In no case does such an identification imply a recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials, equipment, or instrument identified are necessarily the best available for the purpose. The authors declare no other competing interest.

Acknowledgments

C.L. acknowledges supports from the National Institute of Allergy and Infectious Diseases (NIAID) grants K22AI136686 and R61AI174295, as well as the National Science Foundation (NSF) CAREER Award 2047503. A.A. acknowledges supports from the National Human Genome Research Institute (NHGRI) through grant R21HG011741 and the NSF through grant PHY-1430124. J.E.R. was supported by the NSF under grant CBET-2011173. X.W. acknowledges Dr. Lihong Jing from the Institute of Chemistry, Chinese Academy of Sciences, for her kind discussion in producing figures.

Table of Contents



Vocabulary Section

- Nanopore: a nanoscale-sized pore, or channel, typically on the order of a few nanometers in diameter, that allows the passage of molecules or ions through it, often on an individual basis.
- Translocation (in nanopore sensing): the process in which an analyte, such as a biomolecule or nanoparticle, passes through a nanopore, which causes changes in the ionic current flowing through the pore.
- Current blockade: the temporary interruption or reduction of the ionic current flowing through a nanopore when analyte passes through it, which provides a characteristic signal used to gain information about the analyte.
- Dwell time: the period during which an analyte remains within the nanopore. It represents the duration of the analyte's interaction with the nanopore before it translocates through or detaches from the pore.
- Protein sequencing: the process of determining the precise order of amino acids within a protein molecule.
- Post-translational modification (PTM): the chemical or structural modifications that occur on a protein molecule after it has been synthesized through the process of translation. PTMs can involve a wide range of modifications, diversifying the structure and function of proteins.
- Machine learning: a subset of artificial intelligence that focuses on developing algorithms and models that enable computers or machines to learn from and make predictions or decisions based on data, without being explicitly programmed.

References

(1) Alfaro, J. A.; Bohländer, P.; Dai, M.; Filius, M.; Howard, C. J.; van Kooten, X. F.; Ohayon, S.; Pomorski, A.; Schmid, S.; Aksimentiev, A. The emerging landscape of single-molecule protein sequencing technologies. *Nature methods* **2021**, *18*, 604-617.

(2) Restrepo-Pérez, L.; Joo, C.; Dekker, C. Paving the way to single-molecule protein sequencing. *Nat. Nanotechnol.* **2018**, *13*, 786-796.

(3) Squires, A. H.; Gilboa, T.; Torfstein, C.; Varongchayakul, N.; Meller, A. Single-molecule characterization of DNA–protein interactions using nanopore biosensors. *Methods in enzymology* **2017**, *582*, 353-385.

(4) Timp, W.; Timp, G. Beyond mass spectrometry, the next step in proteomics. *Science Advances* **2020**, *6*, eaax8978.

(5) Nicholson, J. A nanopore distance away from next-generation protein sequencing. *Chem* **2022**, *8*, 17-19.

(6) Steen, H.; Mann, M. The abc's (and xyz's) of peptide sequencing. *Nat. Rev. Mol. Cell Bio* **2004**, *5*, 699-711.

(7) Motone, K.; Cardozo, N.; Nivala, J. Herding cats: Label-based approaches in protein translocation through nanopore sensors for single-molecule protein sequence analysis. *Iscience* **2021**, *24*, 103032.

(8) Brocchieri, L.; Karlin, S. Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Research* **2005**, *33*, 3390-3400.

(9) Steinke, L.; Cook, R. G. Identification of phosphorylation sites by Edman degradation. *Protein Sequencing Protocols* **2003**, 301-307.

(10) MacCoss, M. J.; Alfaro, J. A.; Faivre, D. A.; Wu, C. C.; Wanunu, M.; Slavov, N. Sampling the proteome by emerging single-molecule and mass spectrometry methods. *Nature Methods* **2023**, *20*, 339-346.

(11) Smith, L. M.; Agar, J. N.; Chamot-Rooke, J.; Danis, P. O.; Ge, Y.; Loo, J. A.; Paša-Tolić, L.; Tsybin, Y. O.; Kelleher, N. L.; Proteomics, C. f. T.-D. The human proteoform project: defining the human proteome. *Science advances* **2021**, *7*, eabk0734.

(12) O'Bryon, I.; Jenson, S. C.; Merkley, E. D. Flying blind, or just flying under the radar? The underappreciated power of de novo methods of mass spectrometric peptide identification. *Protein Science* **2020**, *29*, 1864-1878.

(13) Tran, N. H.; Qiao, R.; Xin, L.; Chen, X.; Liu, C.; Zhang, X.; Shan, B.; Ghodsi, A.; Li, M. Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry. *Nature methods* **2019**, *16*, 63-66.

(14) Kelly, R. T. Single-cell proteomics: progress and prospects. *Molecular & Cellular Proteomics* **2020**, *19*, 1739-1748.

(15) Zubarev, R. A. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics* **2013**, *13*, 723-726.

(16) Wang, H.-Y.; Ying, Y.-L.; Li, Y.; Kraatz, H.-B.; Long, Y.-T. Nanopore analysis of β-amyloid peptide aggregation transition induced by small molecules. *Analytical chemistry* **2011**, *83*, 1746-1752.

(17) Li, P.; Chen, Y.; Wang, B.; Li, M.; Xiang, D.; Jia, C.; Guo, X. Single-molecule optoelectronic devices: physical mechanism and beyond. *Opto-Electronic Advances* **2022**, *5*, 210094-210094-210021.

(18) Bowen, W.; Vollmer, F.; Gordon, R. Single Molecule Sensing Beyond Fluorescence (Eds.: W.

Bowen, F. Vollmer, R. Gordon). *Springer International Publishing, Cham* **2022**, pp. 5–36.
(19) Boynton, P.; Di Ventra, M. Sequencing proteins with transverse ionic transport in nanochannels. *Sci. Rep.* **2016**, *6*, 1-10.
(20) Ryu, J.; Komoto, Y.; Ohshiro, T.; Taniguchi, M. Single-Molecule Classification of Aspartic Acid and Leucine by Molecular Recognition through Hydrogen Bonding and Time-Series Analysis. *Chemistry–An Asian Journal* **2022**, e202200179.
(21) Filius, M. Next-Generation Protein Identification: Advancing Single-Molecule Fluorescence Approaches. *Doctoral Thesis* **2022**, ISBN 978-990-8593-8510-8590.
(22) Ying, Y.-L.; Long, Y.-T. Nanopore-based single-biomolecule interfaces: from information to knowledge. *Journal of the American Chemical Society* **2019**, *141*, 15720-15729.
(23) Cui, M.; Ge, Y.; Zhuge, X.; Zhou, X.; Xi, D.; Zhang, S. Recent Advances in Nanopore Sensing†. *Chinese Journal of Chemistry* **2021**, *39*, 2035-2043.
(24) Macchia, E.; Torricelli, F.; Bollella, P.; Sarcina, L.; Tricase, A.; Di Franco, C.; Österbacka, R.; Kovács-Vajna, Z. M.; Scamarcio, G.; Torsi, L. Large-Area Interfaces for Single-Molecule Label-free Bioelectronic Detection. *Chemical Reviews* **2022**, *122*, 4636-4699.
(25) Reed, B. D.; Meyer, M. J.; Abramzon, V.; Ad, O.; Ad, O.; Adcock, P.; Ahmad, F. R.; Alppay, G.; Ball, J. A.; Beach, J. Real-time dynamic single-molecule protein sequencing on an integrated semiconductor device. *Science* **2022**, *378*, 186-192.
(26) Zwolak, M.; Di Ventra, M. Electronic signature of DNA nucleotides via transverse transport. *Nano letters* **2005**, *5*, 421-424.
(27) Lagerqvist, J.; Zwolak, M.; Di Ventra, M. Fast DNA sequencing via transverse electronic transport. *Nano letters* **2006**, *6*, 779-782.
(28) Zwolak, M.; Di Ventra, M. Colloquium: Physical approaches to DNA sequencing and detection. *Rev. Mod. Phys.* **2008**, *80*, 141.
(29) 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.
(30) Zhao, Y.; Ashcroft, B.; Zhang, P.; Liu, H.; Sen, S.; Song, W.; Im, J.; Gyarfas, B.; Manna, S.; Biswas, S. Single-molecule spectroscopy of amino acids and peptides by recognition tunnelling. *Nat. Nanotechnol.* **2014**, *9*, 466-473.
(31) Püntener, S.; Rivera-Fuentes, P. Single-Molecule Peptide Identification Using Fluorescence Blinking Fingerprints. *Journal of the American Chemical Society* **2023**, *145*, 1441-1447.
(32) Swaminathan, J.; Boulgakov, A. A.; Hernandez, E. T.; Bardo, A. M.; Bachman, J. L.; Marotta, J.; Johnson, A. M.; Anslyn, E. V.; Marcotte, E. M. Highly parallel single-molecule identification of proteins in zeptomole-scale mixtures. *Nature biotechnology* **2018**, *36*, 1076-1082.
(33) Van Ginkel, J.; Filius, M.; Szczepaniak, M.; Tulinski, P.; Meyer, A. S.; Joo, C. Single-molecule peptide fingerprinting. *Proceedings of the National Academy of Sciences* **2018**, *115*, 3338-3343.
(34) Hernandez, E. T.; Swaminathan, J.; Marcotte, E. M.; Anslyn, E. V. Solution-phase and solid-phase sequential, selective modification of side chains in KDYWEC and KDYWE as models for usage in single-molecule protein sequencing. *New Journal of Chemistry* **2017**, *41*, 462-469.
(35) Bowen, W.; Vollmer, F.; Gordon, R. Single Molecule Sensing Beyond Fluorescence(Eds.: W. Bowen, F. Vollmer, R. Gordon). *Springer International Publishing* **2022**, pp. 5–36.
(36) Deamer, D.; Akeson, M.; Branton, D. Three decades of nanopore sequencing. *Nature biotechnology* **2016**, *34*, 518-524.

(37) Ying, Y.-L.; Hu, Z.-L.; Zhang, S.; Qing, Y.; Fragasso, A.; Maglia, G.; Meller, A.; Bayley, H.; Dekker, C.; Long, Y.-T. Nanopore-based technologies beyond DNA sequencing. *Nat. Nanotechnol.* **2022**, 1-11.
(38) Wu, Y.; Gooding, J. J. The application of single molecule nanopore sensing for quantitative analysis. *Chemical Society Reviews* **2022**, *51*, 3862-3885.
(39) Nouri, R.; Jiang, Y.; Tang, Z.; Lian, X. L.; Guan, W. Detection of SARS-CoV-2 with Solid-State CRISPR-Cas12a-Assisted Nanopores. *Nano Letters* **2021**, *21*, 8393-8400.
(40) Liu, L.; Li, T.; Zhang, S.; Song, P.; Guo, B.; Zhao, Y.; Wu, H. C. Simultaneous Quantification of Multiple Cancer Biomarkers in Blood Samples through DNA-Assisted Nanopore Sensing. *Angewandte Chemie International Edition* **2018**, *57*, 11882-11887.
(41) Ying, Y.-L.; Cao, C.; Hu, Y.-X.; Long, Y.-T. A single biomolecule interface for advancing the sensitivity, selectivity and accuracy of sensors. *National Science Review* **2018**, *5*, 450-452.
(42) Mayer, S. F.; Cao, C.; Dal Peraro, M. Biological nanopores for single-molecule sensing. *Iscience* **2022**, 104145.
(43) Tanimoto, I. M. F.; Cressiot, B.; Greive, S. J.; Le Pioufle, B.; Bacri, L.; Pelta, J. Focus on using nanopore technology for societal health, environmental, and energy challenges. *Nano Res.* **2022**, 1-15.
(44) Nouri, R.; Jiang, Y.; Lian, X. L.; Guan, W. Sequence-Specific Recognition of HIV-1 DNA with Solid-State CRISPR-Cas12a-Assisted Nanopores (SCAN). *ACS Sensors* **2020**, *5*, 1273-1280.
(45) Hu, W. H.; Zhou, K.; Liu, L.; Wu, H. C. Construction of a pH-Mediated Single-Molecule Switch with a Nanopore-DNA Complex. *Small* **2022**, *18*, 2201650.
(46) Lv, P.; Zhang, W.; Yang, Y.; Gao, H.; Li, S.; Tan, C. S.; Ming, D. Recent Advances in Aptamer-Based Nanopore Sensing at Single-Molecule Resolution. *Chemistry–An Asian Journal*, *17*, e202200364.
(47) Gao, C.; Ding, S.; Tan, Q.; Gu, L.-Q. Method of creating a nanopore-terminated probe for single-molecule enantiomer discrimination. *Analytical Chemistry* **2009**, *81*, 80-86.
(48) Dong, Z.; Kennedy, E.; Hokmabadi, M.; Timp, G. Discriminating Residue Substitutions in a Single Protein Molecule Using a Sub-nanopore. *ACS Nano* **2017**, *11*, 5440-5452.
(49) Lisauskas, T.; Rayabharam, A.; Murkute, P.; Almonte, L.; Rigo, E.; Robles, J. O.; Dong, Z.; Joseph, J.; Aluru, N.; Timp, G. Decoding Proteoforms with Single Acid Resolution Using a Sub-nanometer Diameter Pore. *bioRxiv* **2022**, *12*, DOI：10.1101/2022.1112.1122.521660.
(50) Liu, X.; Dong, Z.; Timp, G. Calling the amino acid sequence of a protein/peptide from the nanospectrum produced by a sub-nanometer diameter pore. *Sci. Rep.* **2022**, *12*, 17853.
(51) Talaga, D. S.; Li, J. Single-molecule protein unfolding in solid state nanopores. *Journal of the American Chemical Society* **2009**, *131*, 9287-9297.
(52) Yusko, E. C.; Johnson, J. M.; Majd, S.; Prangkio, P.; Rollings, R. C.; Li, J.; Yang, J.; Mayer, M. Controlling protein translocation through nanopores with bio-inspired fluid walls. *Nat. Nanotechnol.* **2011**, *6*, 253-260.
(53) Yusko, E. C.; Bruhn, B. R.; Eggenberger, O. M.; Houghtaling, J.; Rollings, R. C.; Walsh, N. C.; Nandivada, S.; Pindrus, M.; Hall, A. R.; Sept, D. Real-time shape approximation and fingerprinting of single proteins using a nanopore. *Nat. Nanotechnol.* **2017**, *12*, 360-367.
(54) Sha, J.; Si, W.; Xu, B.; Zhang, S.; Li, K.; Lin, K.; Shi, H.; Chen, Y. Identification of spherical and nonspherical proteins by a solid-state nanopore. *Analytical chemistry* **2018**, *90*, 13826-13831.

(55) Fologea, D.; Ledden, B.; McNabb, D. S.; Li, J. Electrical characterization of protein molecules by a solid-state nanopore. *Applied physics letters* **2007**, *91*, 053901.
(56) Hu, R.; Rodrigues, J. o. V.; Waduge, P.; Yamazaki, H.; Cressiot, B.; Chishti, Y.; Makowski, L.; Yu, D.; Shakhnovich, E.; Zhao, Q. Differential enzyme flexibility probed using solid-state nanopores. *ACS Nano* **2018**, *12*, 4494-4502.
(57) Zhao, Y.; Iarossi, M.; De Fazio, A. F.; Huang, J.-A.; De Angelis, F. Label-Free Optical Analysis of Biomolecules in Solid-State Nanopores: Toward Single-Molecule Protein Sequencing. *ACS photonics* **2022**, *9*, 730-742.
(58) Ohayon, S.; Girsault, A.; Nasser, M.; Shen-Orr, S.; Meller, A. Simulation of single-protein nanopore sensing shows feasibility for whole-proteome identification. *PLoS Comp Biol* **2019**, *15*, e1007067.
(59) Yao, Y.; Docter, M.; Van Ginkel, J.; de Ridder, D.; Joo, C. Single-molecule protein sequencing through fingerprinting: computational assessment. *Physical biology* **2015**, *12*, 055003.
(60) Wang, F.; Zhao, C.; Zhao, P.; Chen, F.; Qiao, D.; Feng, J. MoS2 nanopore identifies single amino acids with sub-1 Dalton resolution. *Nat. Commun.* **2023**, *14*, 2895.
(61) Peng, W.; Yan, S.; Zhou, K.; Wu, H.-C.; Liu, L.; Zhao, Y. High-resolution discrimination of homologous and isomeric proteinogenic amino acids in nanopore sensors with ultrashort single-walled carbon nanotubes. *Nat. Commun.* **2023**, *14*, 1-8.
(62) Xue, L.; Yamazaki, H.; Ren, R.; Wanunu, M.; Ivanov, A. P.; Edel, J. B. Solid-state nanopore sensors. *Nat. Rev. Mater.* **2020**, *5*, 931-951.
(63) Dekker, C. Solid-state nanopores. *Nat. Nanotechnol.* **2007**, *2*, 209-215.
(64) Waugh, M.; Briggs, K.; Gunn, D.; Gibeault, M.; King, S.; Ingram, Q.; Jimenez, A. M.; Berryman, S.; Lomovtsev, D.; Andrzejewski, L. Solid-state nanopore fabrication by automated controlled breakdown. *Nat. Protoc.* **2020**, *15*, 122-143.
(65) Zhang, M.; Chen, C.; Zhang, Y.; Geng, J. Biological nanopores for sensing applications. *Proteins: Structure, Function, and Bioinformatics* **2022**, *90*, 1786-1799.
(66) Wei, X.; Wang, Q.; Liu, C. Nanopore sensing of γ-cyclodextrin induced host-guest interaction to reverse the binding of perfluorooctanoic acid to human serum albumin. *Proteomics* **2022**, *22*, 2100058.
(67) Wei, X.; Zhang, Z.; Wang, X.; Lenhart, B.; Gambarini, R.; Gray, J.; Liu, C. Insight into the effects of electrochemical factors on host-guest interaction induced signature events in a biological nanopore. *Nanotechnology and Precision Engineering* **2020**, *3*, 2-8.
(68) Houghtaling, J.; List, J.; Mayer, M. Nanopore-Based, Rapid Characterization of Individual Amyloid Particles in Solution: Concepts, Challenges, and Prospects. *Small* **2018**, *14*, 1802412.
(69) Robertson, J. W.; Kasianowicz, J. J.; Banerjee, S. Analytical approaches for studying transporters, channels and porins. *Chemical reviews* **2012**, *112*, 6227-6249.
(70) Yu, R.-J.; Chen, K.-L.; Ying, Y.-L.; Long, Y.-T. Nanopore electrochemical measurement for single-molecular interactions and beyond. *Current Opinion in Electrochemistry* **2022**, 101063.
(71) Zhang, Z.; Wang, X.; Wei, X.; Zheng, S. W.; Lenhart, B. J.; Xu, P.; Li, J.; Pan, J.; Albrecht, H.; Liu, C. Multiplex quantitative detection of SARS-CoV-2 specific IgG and IgM antibodies based on DNA-assisted nanopore sensing. *Biosensors and Bioelectronics* **2021**, *181*, 113134.
(72) Lenhart, B.; Wei, X.; Watson, B.; Wang, X.; Zhang, Z.; Li, C.-z.; Moss, M.; Liu, C. In vitro biosensing of β-Amyloid peptide aggregation dynamics using a biological nanopore. *Sensors Actuators B: Chem* **2021**, *338*, 129863.

(73) Lee, I.; Razaghi, R.; Gilpatrick, T.; Molnar, M.; Gershman, A.; Sadowski, N.; Sedlazeck, F. J.; Hansen, K. D.; Simpson, J. T.; Timp, W. Simultaneous profiling of chromatin accessibility and methylation on human cell lines with nanopore sequencing. *Nature Methods* **2020**, *17*, 1191-1199.
(74) Wei, X.; Wang, X.; Zhang, Z.; Luo, Y.; Wang, Z.; Xiong, W.; Jain, P. K.; Monnier, J. R.; Wang, H.; Hu, T. Y.; Tang, C.; Albrecht, H.; Liu, C. A click chemistry amplified nanopore assay for ultrasensitive quantification of HIV-1 p24 antigen in clinical samples. *Nat. Commun.* **2022**, *13*, 6852.
(75) Lu, H.; Giordano, F.; Ning, Z. Oxford Nanopore MinION sequencing and genome assembly. *Genomics, proteomics & bioinformatics* **2016**, *14*, 265-279.
(76) Jain, M.; Fiddes, I. T.; Miga, K. H.; Olsen, H. E.; Paten, B.; Akeson, M. Improved data analysis for the MinION nanopore sequencer. *Nature methods* **2015**, *12*, 351-356.
(77) Castro-Wallace, S. L.; Chiu, C. Y.; John, K. K.; Stahl, S. E.; Rubins, K. H.; McIntyre, A. B. R.; Dworkin, J. P.; Lupisella, M. L.; Smith, D. J.; Botkin, D. J.; Stephenson, T. A.; Juul, S.; Turner, D. J.; Izquierdo, F.; Federman, S.; Stryke, D.; Somasekar, S.; Alexander, N.; Yu, G.; Mason, C. E.; Burton, A. S. Nanopore DNA Sequencing and Genome Assembly on the International Space Station. *Sci. Rep.* **2017**, *7*, 18022.
(78) Jain, M.; Olsen, H. E.; Paten, B.; Akeson, M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome biology* **2016**, *17*, 1-11.
(79) Gamaarachchi, H.; Samarakoon, H.; Jenner, S. P.; Ferguson, J. M.; Amos, T. G.; Hammond, J. M.; Saadat, H.; Smith, M. A.; Parameswaran, S.; Deveson, I. W. Fast nanopore sequencing data analysis with SLOW5. *Nature biotechnology* **2022**, 1-4.
(80) Garalde, D. R.; Snell, E. A.; Jachimowicz, D.; Sipos, B.; Lloyd, J. H.; Bruce, M.; Pantic, N.; Admassu, T.; James, P.; Warland, A. Highly parallel direct RNA sequencing on an array of nanopores. *Nature methods* **2018**, *15*, 201-206.
(81) Hu, Z. L.; Huo, M. Z.; Ying, Y. L.; Long, Y. T. Biological nanopore approach for single-molecule protein sequencing. *Angewandte Chemie* **2021**, *133*, 14862-14873.
(82) Seet, B. T.; Dikic, I.; Zhou, M.-M.; Pawson, T. Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Bio* **2006**, *7*, 473-483.
(83) Mann, M.; Jensen, O. N. Proteomic analysis of post-translational modifications. *Nature biotechnology* **2003**, *21*, 255-261.
(84) Asandei, A.; Di Muccio, G.; Schiopu, I.; Mereuta, L.; Dragomir, I. S.; Chinappi, M.; Luchian, T. Nanopore-based protein sequencing using biopores: Current achievements and open challenges. *Small Methods* **2020**, *4*, 1900595.
(85) Belmonte, G.; Cescatti, L.; Ferrari, B.; Nicolussi, T.; Ropele, M.; Menestrina, G. Pore formation by Staphylococcus aureus alpha-toxin in lipid bilayers: dependence upon temperature and toxin concentration. *Eur Biophys J* **1987**, *14*, 349-358.
(86) Bezrukov, S. M.; Kasianowicz, J. J. Current noise reveals protonation kinetics and number of ionizable sites in an open protein ion channel. *Physical review letters* **1993**, *70*, 2352.
(87) Kasianowicz, J. J.; Bezrukov, S. M. Protonation dynamics of the alpha-toxin ion channel from spectral analysis of pH-dependent current fluctuations. *Biophysical Journal* **1995**, *69*, 94-105.
(88) Guo, B.-Y.; Zeng, T.; Wu, H.-C. Recent advances of DNA sequencing via nanopore-based technologies. *Science bulletin* **2015**, *60*, 287-295.
(89) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences* **1996**, *93*, 13770-13773.

(90) Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J. E. Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore. *Science* **1996**, *274*, 1859-1865.
(91) Edman, P. A method for the determination of the amino acid sequence in peptides. *Arch. Biochem.* **1949**, *22*, 475-476.
(92) Checa-Moreno, R.; Manzano, E.; Mirón, G.; Capitán-Vallvey, L. F. Revisitation of the phenylisothiocyanate-derivatives procedure for amino acid determination by HPLC-UV. *Journal of separation science* **2008**, *31*, 3817-3828.
(93) Wei, X.; Ma, D.; Zhang, Z.; Wang, L. Y.; Gray, J. L.; Zhang, L.; Zhu, T.; Wang, X.; Lenhart, B. J.; Yin, Y.; Wang, Q.; Liu, C. N-terminal derivatization-assisted identification of individual amino acids using a biological nanopore sensor. *ACS sensors* **2020**, *5*, 1707-1716.
(94) Wei, X.; Ma, D.; Jing, L.; Wang, L. Y.; Wang, X.; Zhang, Z.; Lenhart, B. J.; Yin, Y.; Wang, Q.; Liu, C. Enabling nanopore technology for sensing individual amino acids by a derivatization strategy. *J. Mater. Chem. B* **2020**, *8*, 6792-6797.
(95) Kennedy, E.; Dong, Z.; Tennant, C.; Timp, G. Reading the primary structure of a protein with 0.07 nm3 resolution using a subnanometre-diameter pore. *Nat. Nanotechnol.* **2016**, *11*, 968-976.
(96) Darby, N.; Creighton, T. Protein structure: In focus Oxford University Press. **1993**.
(97) Zamyatnin, A. Protein volume in solution. *Prog Biophys Mol Biol* **1972**, *24*, 107-123.
(98) Huo, M.-Z.; Li, M.-Y.; Ying, Y.-L.; Long, Y.-T. Is the Volume Exclusion Model Practicable for Nanopore Protein Sequencing? *Analytical Chemistry* **2021**, *93*, 11364-11369.
(99) Li, M.-Y.; Ying, Y.-L.; Yu, J.; Liu, S.-C.; Wang, Y.-Q.; Li, S.; Long, Y.-T. Revisiting the Origin of Nanopore Current Blockage for Volume Difference Sensing at the Atomic Level. *JACS Au* **2021**, *1*, 967-976.
(100) Lu, S. M.; Wu, X. Y.; Li, M. Y.; Ying, Y. L.; Long, Y. T. Diversified exploitation of aerolysin nanopore in single-molecule sensing and protein sequencing. *View* **2020**, *1*, 20200006.
(101) 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.
(102) Qiu, K.; Wu, X.-Y.; Yang, J.; Ying, Y.-L.; Long, Y.-T. Pore-forming confined space for the innovative electrochemical methods. *Current Opinion in Electrochemistry* **2018**, *10*, 46-53.
(103) Tang, L. Sensing proteinogenic amino acids. *Nature Methods* **2020**, *17*, 126-126.
(104) Baaken, G.; Halimeh, I.; Bacri, L.; Pelta, J.; Oukhaled, A.; Behrends, J. C. High-Resolution Size-Discrimination of Single Nonionic Synthetic Polymers with a Highly Charged Biological Nanopore. *ACS Nano* **2015**, *9*, 6443-6449.
(105) 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. *Nature nanotechnology* **2016**, *11*, 713-718.
(106) Yu, J.; Cao, C.; Long, Y.-T. Selective and Sensitive Detection of Methylcytosine by Aerolysin Nanopore under Serum Condition. *Analytical Chemistry* **2017**, *89*, 11685-11689.
(107) Yuan, B.; Li, S.; Ying, Y.-L.; Long, Y.-T. The analysis of single cysteine molecules with an aerolysin nanopore. *Analyst.* **2020**, *145*, 1179-1183.
(108) Howorka, S.; Siwy, Z. S. Reading amino acids in a nanopore. *Nature biotechnology* **2020**, *38*, 159-160.
(109) Piguet, F.; Ouldali, H.; Pastoriza-Gallego, M.; Manivet, P.; Pelta, J.; Oukhaled, A. Identification of single amino acid differences in uniformly charged homopolymeric peptides with

aerolysin nanopore. *Nature communications* **2018**, *9*, 1-13.
(110) 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. *Nature biotechnology* **2020**, *38*, 176-181.
(111) Cuervo, A.; Pulido-Cid, M.; Chagoyen, M.; Arranz, R.; González-García, V. A.; Garcia-Doval, C.; Castón, J. R.; Valpuesta, J. M.; van Raaij, M. J.; Martín-Benito, J. Structural Characterization of the Bacteriophage T7 Tail Machinery*♦. *Journal of Biological Chemistry* **2013**, *288*, 26290-26299.
(112) Ji, Z.; Kang, X.; Wang, S.; Guo, P. Nano-channel of viral DNA packaging motor as single pore to differentiate peptides with single amino acid difference. *Biomaterials* **2018**, *182*, 227-233.
(113) Ji, Z.; Guo, P. Channel from bacterial virus T7 DNA packaging motor for the differentiation of peptides composed of a mixture of acidic and basic amino acids. *Biomaterials* **2019**, *214*, 119222.
(114) Asandei, A.; Chinappi, M.; Lee, J.-k.; Ho Seo, C.; Mereuta, L.; Park, Y.; Luchian, T. Placement of oppositely charged aminoacids at a polypeptide termini determines the voltage-controlled braking of polymer transport through nanometer-scale pores. *Scientific reports* **2015**, *5*, 1-13.
(115) Asandei, A.; Chinappi, M.; Kang, H.-K.; Seo, C. H.; Mereuta, L.; Park, Y.; Luchian, T. Acidity-mediated, electrostatic tuning of asymmetrically charged peptides interactions with protein nanopores. *ACS applied materials & interfaces* **2015**, *7*, 16706-16714.
(116) Restrepo-Pérez, L.; Huang, G.; Bohländer, P. R.; Worp, N.; Eelkema, R.; Maglia, G.; Joo, C.; Dekker, C. Resolving chemical modifications to a single amino acid within a peptide using a biological nanopore. *ACS Nano* **2019**, *13*, 13668-13676.
(117) Ge, Y.; Cui, M.; Zhang, Q.; Wang, Y.; Xi, D. Aerolysin Nanopore-Based Identification of Proteinogenic Amino Acids Using a Bipolar Peptide Probe. *Nanoscale Advances* **2022**, *4* 3883-3891.
(118) Jia, W.; Hu, C.; Wang, Y.; Liu, Y.; Wang, L.; Zhang, S.; Zhu, Q.; Gu, Y.; Zhang, P.; Ma, J. Identification of Single-Molecule Catecholamine Enantiomers Using a Programmable Nanopore. *ACS Nano* **2022**, *16*, 6615-6624.
(119) Boersma, A. J.; Bayley, H. Continuous stochastic detection of amino acid enantiomers with a protein nanopore. *Angewandte Chemie* **2012**, *124*, 9744-9747.
(120) Cooper, J. A.; Borsley, S.; Lusby, P. J.; Cockroft, S. L. Discrimination of supramolecular chirality using a protein nanopore. *Chem. Sci.* **2017**, *8*, 5005-5009.
(121) Du, X.; Zhang, S.; Wang, L.; Wang, Y.; Fan, P.; Jia, W.; Zhang, P.; Huang, S. Single-Molecule Interconversion between Chiral Configurations of Boronate Esters Observed in a Nanoreactor. *ACS Nano* **2023**, *17*, 2881-2892.
(122) Guo, Y.; Niu, A.; Jian, F.; Wang, Y.; Yao, F.; Wei, Y.; Tian, L.; Kang, X. Metal–organic complex-functionalized protein nanopore sensor for aromatic amino acids chiral recognition. *Analyst.* **2017**, *142*, 1048-1053.
(123) Wang, J.; Prajapati, J. D.; Gao, F.; Ying, Y.-L.; Kleinekathöfer, U.; Winterhalter, M.; Long, Y.-T. Identification of single amino acid chiral and positional isomers using an electrostatically asymmetric nanopore. *Journal of the American Chemical Society* **2022**, *144*, 15072-15078.
(124) Sutherland, T. C.; Long, Y.-T.; Stefureac, R.-I.; Bediako-Amoa, I.; Kraatz, H.-B.; Lee, J. S. Structure of Peptides Investigated by Nanopore Analysis. *Nano Letters* **2004**, *4*, 1273-1277.
(125) Long, Y.; Zhang, M. Self-assembling bacterial pores as components of nanobiosensors for the detection of single peptide molecules. *Science in China Series B: Chemistry* **2009**, *52*, 731-733.
(126) Movileanu, L.; Schmittschmitt, J. P.; Scholtz, J. M.; Bayley, H. Interactions of peptides with a protein pore. *Biophysical journal* **2005**, *89*, 1030-1045.

(127) Stefureac, R.; Long, Y.-t.; Kraatz, H.-B.; Howard, P.; Lee, J. S. Transport of α-helical peptides through α-hemolysin and aerolysin pores. *Biochemistry-us.* **2006**, *45*, 9172-9179.
(128) Wang, H.-Y.; Gu, Z.; Cao, C.; Wang, J.; Long, Y.-T. Analysis of a Single α-Synuclein Fibrillation by the Interaction with a Protein Nanopore. *Analytical Chemistry* **2013**, *85*, 8254-8261.
(129) Huang, G.; Willems, K.; Soskine, M.; Wloka, C.; Maglia, G. Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores. *Nature communications* **2017**, *8*, 1-11.
(130) Wloka, C.; Mutter, N. L.; Soskine, M.; Maglia, G. Alpha-Helical Fragaceatoxin C Nanopore Engineered for Double-Stranded and Single-Stranded Nucleic Acid Analysis. *Angewandte Chemie International Edition* **2016**, *55*, 12494-12498.
(131) Tanaka, K.; Caaveiro, J. M.; Morante, K.; González-Mañas, J. M.; Tsumoto, K. Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. *Nat. Commun.* **2015**, *6*, 1-11.
(132) Huang, G.; Voet, A.; Maglia, G. FraC nanopores with adjustable diameter identify the mass of opposite-charge peptides with 44 dalton resolution. *Nature communications* **2019**, *10*, 1-10.
(133) Robertson, J. W. F.; Rodrigues, C. G.; Stanford, V. M.; Rubinson, K. A.; Krasilnikov, O. V.; Kasianowicz, J. J. Single-molecule mass spectrometry in solution using a solitary nanopore. *Proceedings of the National Academy of Sciences* **2007**, *104*, 8207-8211.
(134) Baaken, G.; Ankri, N.; Schuler, A.-K.; Rühe, J. r.; Behrends, J. C. Nanopore-based single-molecule mass spectrometry on a lipid membrane microarray. *ACS Nano* **2011**, *5*, 8080-8088.
(135) Reiner, J. E.; Kasianowicz, J. J.; Nablo, B. J.; Robertson, J. W. Theory for polymer analysis using nanopore-based single-molecule mass spectrometry. *Proceedings of the National Academy of Sciences* **2010**, *107*, 12080-12085.
(136) Heaven, M. W.; Dass, A.; White, P. S.; Holt, K. M.; Murray, R. W. Crystal structure of the gold nanoparticle $[N(C_8H_{17})_4][Au_{25}(SCH_2CH_2Ph)_{18}]$. *Journal of the American Chemical Society* **2008**, *130*, 3754-3755.
(137) Chavis, A. E.; Brady, K. T.; Hatmaker, G. A.; Angevine, C. E.; Kothalawala, N.; Dass, A.; Robertson, J. W.; Reiner, J. E. Single molecule nanopore spectrometry for peptide detection. *ACS sensors* **2017**, *2*, 1319-1328.
(138) Cox, B. D.; Woodworth, P. H.; Wilkerson, P. D.; Bertino, M. F.; Reiner, J. E. Ligand-Induced Structural Changes of Thiolate-Capped Gold Nanoclusters Observed with Resistive-Pulse Nanopore Sensing. *Journal of the American Chemical Society* **2019**, *141*, 3792-3796.
(139) Cox, B. D.; Ghimire, M. L.; Bertino, M. F.; Reiner, J. E. Resistive-Pulse Nanopore Sensing of Ligand Exchange at the Single Nanocluster Limit for Peptide Detection. *ACS Applied Nano Materials* **2020**, *3*, 7973-7981.
(140) Ghimire, M. L.; Cox, B. D.; Winn, C. A.; Rockett, T. W.; Schifano, N. P.; Slagle, H. M.; Gonzalez, F.; Bertino, M. F.; Caputo, G. A.; Reiner, J. E. Selective Detection and Characterization of Small Cysteine-Containing Peptides with Cluster-Modified Nanopore Sensing. *ACS Nano* **2022**, *16*, 17229-17241.
(141) Aebersold, R.; Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **2016**, *537*, 347-355.
(142) Li, Z.; He, B.; Feng, W. Evaluation of bottom-up and top-down mass spectrum identifications with different customized protein sequences databases. *Bioinformatics* **2020**, *36*, 1030-1036.
(143) Lucas, F. L. R.; Versloot, R. C. A.; Yakovlieva, L.; Walvoort, M. T.; Maglia, G. Protein identification by nanopore peptide profiling. *Nat. Commun.* **2021**, *12*, 1-9.

(144) Afshar Bakshloo, M.; Kasianowicz, J. J.; Pastoriza-Gallego, M.; Mathé, J.; Daniel, R.; Piguet, F.; Oukhaled, A. Nanopore-Based Protein Identification. *Journal of the American Chemical Society* **2022**, *144*, 2716-2725.
(145) Afshar Bakshloo, M.; Yahiaoui, S.; Ouldali, H.; Pastoriza-Gallego, M.; Piguet, F.; Oukhaled, A. On possible trypsin-induced biases in peptides analysis with aerolysin nanopore. *Proteomics* **2022**, 2100056.
(146) Bakshloo, M. A.; Yahiaoui, S.; Piguet, F.; Pastoriza-Gallego, M.; Daniel, R.; Mathé, J.; Kasianowicz, J. J.; Oukhaled, A. Polypeptide analysis for nanopore-based protein identification. *Nano Res.* **2022**, *15*, 9831-9842.
(147) Niu, H.; Li, M.-Y.; Ying, Y.-L.; Long, Y.-T. An engineered third electrostatic constriction of aerolysin to manipulate heterogeneously charged peptide transport. *Chemical science* **2022**, *13*, 2456-2461.
(148) Versloot, R. C. A.; Straathof, S. A. P.; Stouwie, G.; Tadema, M. J.; Maglia, G. β-Barrel Nanopores with an Acidic–Aromatic Sensing Region Identify Proteinogenic Peptides at Low pH. *ACS Nano* **2022**, *16*, 7258-7268.
(149) 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*, 1-11.
(150) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A. How many human proteoforms are there? *Nat Chem Biol* **2018**, *14*, 206-214.
(151) Hoogerheide, D. P.; Gurnev, P. A.; Rostovtseva, T. K.; Bezrukov, S. M. Effect of a post-translational modification mimic on protein translocation through a nanopore. *Nanoscale* **2020**, *12*, 11070-11078.
(152) Deribe, Y. L.; Pawson, T.; Dikic, I. Post-translational modifications in signal integration. *Nat Struct Mol Biol* **2010**, *17*, 666-672.
(153) Harashima, T.; Egami, Y.; Homma, K.; Jono, Y.; Kaneko, S.; Fujii, S.; Ono, T.; Nishino, T. Unique Electrical Signature of Phosphate for Specific Single-Molecule Detection of Peptide Phosphorylation. *Journal of the American Chemical Society* **2022**.
(154) Liu, F.; Iqbal, K.; Grundke-Iqbal, I.; Hart, G. W.; Gong, C.-X. O-GlcNAcylation regulates phosphorylation of tau: a mechanism involved in Alzheimer's disease. *Proceedings of the National Academy of Sciences* **2004**, *101*, 10804-10809.
(155) Mulagapati, S.; Koppolu, V.; Raju, T. S. Decoding of O-linked glycosylation by mass spectrometry. *Biochemistry-us.* **2017**, *56*, 1218-1226.
(156) Restrepo-Pérez, L.; Wong, C. H.; Maglia, G.; Dekker, C.; Joo, C. Label-free detection of post-translational modifications with a nanopore. *Nano letters* **2019**, *19*, 7957-7964.
(157) 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*.
(158) Versloot, R. C. A.; Lucas, F. L. R.; Yakovlieva, L.; Tadema, M. J.; Zhang, Y.; Wood, T. M.; Martin, N. I.; Marrink, S. J.; Walvoort, M. T. C.; Maglia, G. Quantification of Protein Glycosylation Using Nanopores. *Nano Letters* **2022**, *22*, 5357-5364.
(159) Li, S.; Wu, X. Y.; Li, M. Y.; Liu, S. C.; Ying, Y. L.; Long, Y. T. T232K/K238Q aerolysin nanopore for mapping adjacent phosphorylation sites of a single tau peptide. *Small Methods* **2020**, *4*, 2000014.

(160) Ensslen, T.; Sarthak, K.; Aksimentiev, A.; Behrends, J. C. Resolving Isomeric Posttranslational Modifications Using a Biological Nanopore as a Sensor of Molecular Shape. *Journal of the American Chemical Society* **2022**, *144*, 16060-16068.
(161) Kasianowicz, J. J.; Balijepalli, A. K.; Ettedgui, J.; Forstater, J. H.; Wang, H.; Zhang, H.; Robertson, J. W. Analytical applications for pore-forming proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2016**, *1858*, 593-606.
(162) Li, S.; Cao, C.; Yang, J.; Long, Y. T. Detection of peptides with different charges and lengths by using the aerolysin nanopore. *ChemElectroChem* **2019**, *6*, 126-129.
(163) Zhao, S.; Restrepo-Pérez, L.; Soskine, M.; Maglia, G.; Joo, C.; Dekker, C.; Aksimentiev, A. Electro-Mechanical Conductance Modulation of a Nanopore Using a Removable Gate. *ACS Nano* **2019**, *13*, 2398-2409.
(164) Asandei, A.; Schiopu, I.; Chinappi, M.; Seo, C. H.; Park, Y.; Luchian, T. Electroosmotic trap against the electrophoretic force near a protein nanopore reveals peptide dynamics during capture and translocation. *ACS Appl. Mater. Interfaces* **2016**, *8*, 13166-13179.
(165) Huang, G.; Willems, K.; Bartelds, M.; van Dorpe, P.; Soskine, M.; Maglia, G. Electro-Osmotic Vortices Promote the Capture of Folded Proteins by PlyAB Nanopores. *Nano Letters* **2020**, *20*, 3819-3827.
(166) Willems, K.; Ruić, D.; Biesemans, A.; Galenkamp, N. S.; Van Dorpe, P.; Maglia, G. Engineering and Modeling the Electrophoretic Trapping of a Single Protein Inside a Nanopore. *ACS Nano* **2019**, *13*, 9980-9992.
(167) 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. *Nature Biotechnology* **2023**, 1-10.
(168) Zhang, J.; Yan, S.; Chang, L.; Guo, W.; Wang, Y.; Wang, Y.; Zhang, P.; Chen, H.-Y.; Huang, S. Direct microRNA sequencing using nanopore-induced phase-shift sequencing. *Iscience* **2020**, *23*, 100916.
(169) 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. Decoding long nanopore sequencing reads of natural DNA. *Nature biotechnology* **2014**, *32*, 829-833.
(170) Nivala, J.; Marks, D. B.; Akeson, M. Unfoldase-mediated protein translocation through an α-hemolysin nanopore. *Nature biotechnology* **2013**, *31*, 247-250.
(171) Rodriguez-Larrea, D.; Bayley, H. Multistep protein unfolding during nanopore translocation. *Nat. Nanotechnol.* **2013**, *8*, 288-295.
(172) Nivala, J.; Mulroney, L.; Li, G.; Schreiber, J.; Akeson, M. Discrimination among Protein Variants Using an Unfoldase-Coupled Nanopore. *ACS Nano* **2014**, *8*, 12365-12375.
(173) Cordova, J. C.; Olivares, A. O.; Shin, Y.; Stinson, B. M.; Calmat, S.; Schmitz, K. R.; Aubin-Tam, M.-E.; Baker, T. A.; Lang, M. J.; Sauer, R. T. Stochastic but highly coordinated protein unfolding and translocation by the ClpXP proteolytic machine. *Cell* **2014**, *158*, 647-658.
(174) Hoogerheide, D. P.; Gurnev, P. A.; Rostovtseva, T. K.; Bezrukov, S. M. Real-Time Nanopore-Based Recognition of Protein Translocation Success. *Biophysical Journal* **2018**, *114*, 772-776.
(175) Hoogerheide, D. P. PPDiffuse: A quantitative prediction tool for diffusion of charged polymers in a nanopore. *J. Res. Natl. Inst. Stand* **2020**, *125*, 125018.
(176) Pavlenok, M.; Yu, L.; Herrmann, D.; Wanunu, M.; Niederweis, M. Control of subunit stoichiometry in single-chain MspA nanopores. *Biophysical Journal* **2022**, *121*, 742-754.

(177) Chen, Z.; Wang, Z.; Xu, Y.; Zhang, X.; Tian, B.; Bai, J. Controlled movement of ssDNA conjugated peptide through Mycobacterium smegmatis porin A (MspA) nanopore by a helicase motor for peptide sequencing application. *Chem. Sci.* **2021**, *12*, 15750-15756.
(178) Yan, S.; Li, X.; Zhang, P.; Wang, Y.; Chen, H.-Y.; Huang, S.; Yu, H. Direct sequencing of 2′-deoxy-2′-fluoroarabinonucleic acid (FANA) using nanopore-induced phase-shift sequencing (NIPSS). *Chem. Sci.* **2019**, *10*, 3110-3117.
(179) Yan, S.; Zhang, J.; Wang, Y.; Guo, W.; Zhang, S.; Liu, Y.; Cao, J.; Wang, Y.; Wang, L.; Ma, F. Single molecule ratcheting motion of peptides in a Mycobacterium smegmatis Porin A (MspA) nanopore. *Nano letters* **2021**, *21*, 6703-6710.
(180) Brinkerhoff, H.; Kang, A. S.; Liu, J.; Aksimentiev, A.; Dekker, C. Multiple rereads of single proteins at single–amino acid resolution using nanopores. *Science* **2021**, *374*, 1509-1513.
(181) Wanunu, M. Back and forth with nanopore peptide sequencing. *Nature Biotechnology* **2022**, *40*, 172-173.
(182) Liu, S. C.; Ying, Y. L.; Li, W. H.; Wan, Y. J.; Long, Y. T. Snapshotting the transient conformations and tracing the multiple pathways of single peptide folding using a solid-state nanopore. *Chem Sci* **2021**, *12*, 3282-3289.
(183) Wen, C.; Dematties, D.; Zhang, S.-L. A Guide to Signal Processing Algorithms for Nanopore Sensors. *ACS Sensors* **2021**, *6*, 3536-3555.
(184) Wang, Y.; Zhang, S.; Jia, W.; Fan, P.; Wang, L.; Li, X.; Chen, J.; Cao, Z.; Du, X.; Liu, Y.; Wang, K.; Hu, C.; Zhang, J.; Hu, J.; Zhang, P.; Chen, H.-Y.; Huang, S. Identification of nucleoside monophosphates and their epigenetic modifications using an engineered nanopore. *Nat. Nanotechnol.* **2022**, *17*, 976-983.
(185) Liu, Y.; Zhang, S.; Wang, Y.; Wang, L.; Cao, Z.; Sun, W.; Fan, P.; Zhang, P.; Chen, H.-Y.; Huang, S. Nanopore Identification of Alditol Epimers and Their Application in Rapid Analysis of Alditol-Containing Drinks and Healthcare Products. *Journal of the American Chemical Society* **2022**, *144*, 13717-13728.
(186) Liu, Y.; Wang, K.; Wang, Y.; Wang, L.; Yan, S.; Du, X.; Zhang, P.; Chen, H.-Y.; Huang, S. Machine Learning Assisted Simultaneous Structural Profiling of Differently Charged Proteins in a Mycobacterium smegmatis Porin A (MspA) Electroosmotic Trap. *Journal of the American Chemical Society* **2022**, *144*, 757-768.
(187) Barati Farimani, A.; Heiranian, M.; Aluru, N. R. Identification of amino acids with sensitive nanoporous MoS2: towards machine learning-based prediction. *npj 2D Materials and Applications* **2018**, *2*, 14.
(188) Spitaleri, A.; Garoli, D.; Schütte, M.; Lehrach, H.; Rocchia, W.; De Angelis, F. Adaptive nanopores: A bioinspired label-free approach for protein sequencing and identification. *Nano Res.* **2021**, *14*, 328-333.
(189) Dematties, D.; Wen, C.; Pérez, M. D.; Zhou, D.; Zhang, S.-L. Deep Learning of Nanopore Sensing Signals Using a Bi-Path Network. *ACS Nano* **2021**, *15*, 14419.
(190) 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. *Science Advances* **2020**, *6*, eabc2661.
(191) Rodriguez-Larrea, D. Single-aminoacid discrimination in proteins with homogeneous nanopore sensors and neural networks. *Biosensors and Bioelectronics* **2021**, *180*, 113108.
(192) Kolmogorov, M.; Kennedy, E.; Dong, Z.; Timp, G.; Pevzner, P. A. Single-molecule protein

identification by sub-nanopore sensors. *PLoS Computational Biology* **2017**, *13*, e1005356-e1005356.
(193) Aksimentiev, A.; Heng, J. B.; Timp, G.; Schulten, K. Microscopic kinetics of DNA translocation through synthetic nanopores. *Biophysical journal* **2004**, *87*, 2086-2097.
(194) MJ, C. C. C. N. M. Buglakova E. Duperrex A. Radenovic A. Dal Peraro M. *Nat. Commun* **2019**, *10*, 4918.
(195) Di Muccio, G.; Rossini, A. E.; Di Marino, D.; Zollo, G.; Chinappi, M. Insights into protein sequencing with an α-Hemolysin nanopore by atomistic simulations. *Sci. Rep.* **2019**, *9*, 1-8.
(196) Lucas, F. L. R.; Sarthak, K.; Lenting, E. M.; Coltan, D.; Van Der Heide, N. J.; Versloot, R. C. A.; Aksimentiev, A.; Maglia, G. The manipulation of the internal hydrophobicity of FraC nanopores augments peptide capture and recognition. *ACS Nano* **2021**, *15*, 9600-9613.
(197) Zhang, S.; Huang, G.; Versloot, R. C. A.; Bruininks, B. M. H.; de Souza, P. C. T.; Marrink, S.-J.; Maglia, G. Bottom-up fabrication of a proteasome–nanopore that unravels and processes single proteins. *Nature chemistry* **2021**, *13*, 1192-1199.
(198) Aksimentiev, A.; Schulten, K. Imaging α-hemolysin with molecular dynamics: ionic conductance, osmotic permeability, and the electrostatic potential map. *Biophysical journal* **2005**, *88*, 3745-3761.
(199) Choudhary, A.; Maffeo, C.; Aksimentiev, A. Multi-resolution simulation of DNA transport through large synthetic nanostructures. *Physical Chemistry Chemical Physics* **2022**, *24*, 2706-2716.
(200) Wilson, J.; Sarthak, K.; Si, W.; Gao, L.; Aksimentiev, A. Rapid and accurate determination of nanopore ionic current using a steric exclusion model. *Acs Sensors* **2019**, *4*, 634-644.
(201) Golla, V. K.; Prajapati, J. D.; Joshi, M.; Kleinekathöfer, U. Exploration of free energy surfaces across a membrane channel using metadynamics and umbrella sampling. *Journal of Chemical Theory and Computation* **2020**, *16*, 2751-2765.
(202) Shaw, D. E.; Adams, P. J.; Azaria, A.; Bank, J. A.; Batson, B.; Bell, A.; Bergdorf, M.; Bhatt, J.; Butts, J. A.; Correia, T. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* 2021, p 1-11.
(203) Sarthak, K.; Winogradoff, D.; Ge, Y.; Myong, S.; Aksimentiev, A. Benchmarking Molecular Dynamics Force Fields for All-Atom Simulations of Biological Condensates. *Journal of Chemical Theory and Computation* **2023**.
(204) Lin, C.-Y.; Fotis, R.; Xia, Z.; Kavetsky, K.; Chou, Y.-C.; Niedzwiecki, D. J.; Biondi, M.; Thei, F.; Drndić, M. Ultrafast Polymer Dynamics through a Nanopore. *Nano Letters* **2022**, *22*, 8719-8727.
(205) Lathrop, D. K.; Ervin, E. N.; Barrall, G. A.; Keehan, M. G.; Kawano, R.; Krupka, M. A.; White, H. S.; Hibbs, A. H. Monitoring the escape of DNA from a nanopore using an alternating current signal. *Journal of the American Chemical Society* **2010**, *132*, 1878-1885.
(206) Xing, Y.; Dorey, A.; Jayasinghe, L.; Howorka, S. Highly shape-and size-tunable membrane nanopores made with DNA. *Nature Nanotechnology* **2022**, 1-6.
(207) 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. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature biotechnology* **2018**, *36*, 338-345.
(208) Niedzwiecki, D. J.; Chou, Y.-C.; Xia, Z.; Thei, F.; Drndić, M. Detection of single analyte and environmental samples with silicon nitride nanopores: antarctic dirt particulates and dna in artificial seawater. *Review of Scientific Instruments* **2020**, *91*, 031301.
(209) Miles, B. N.; Ivanov, A. P.; Wilson, K. A.; Doğan, F.; Japrung, D.; Edel, J. B. Single molecule sensing with solid-state nanopores: novel materials, methods, and applications. *Chemical Society*

Reviews **2013**, *42*, 15-28.
(210) Sheng, Y.; You, Y.; Cao, Z.; Liu, L.; Wu, H.-C. Rapid and selective DNA-based detection of melamine using α-hemolysin nanopores. *Analyst.* **2018**, *143*, 2411-2415.
(211) Weckman, N. E.; Ermann, N.; Gutierrez, R.; Chen, K.; Graham, J.; Tivony, R.; Heron, A.; Keyser, U. F. Multiplexed DNA identification using site specific dCas9 barcodes and nanopore sensing. *ACS sensors* **2019**, *4*, 2065-2072.
(212) Morin, T. J.; McKenna, W. L.; Shropshire, T. D.; Wride, D. A.; Deschamps, J. D.; Liu, X.; Stamm, R.; Wang, H.; Dunbar, W. B. A handheld platform for target protein detection and quantification using disposable nanopore strips. *Sci. Rep.* **2018**, *8*, 1-12.
(213) Carter, J.-M.; Hussain, S. Robust long-read native DNA sequencing using the ONT CsgG Nanopore system. *Wellcome open research* **2017**, *2*.
(214) Baker, L. A. Perspective and prospectus on single-entity electrochemistry. *Journal of the American Chemical Society* **2018**, *140*, 15549-15559.
(215) Ying, Y.-L.; Wang, J.; Leach, A. R.; Jiang, Y.; Gao, R.; Xu, C.; Edwards, M. A.; Pendergast, A. D.; Ren, H.; Weatherly, C. K. T. Single-entity electrochemistry at confined sensing interfaces. *Science China Chemistry* **2020**, *63*, 589-618.
(216) Hu, R.; Tong, X.; Zhao, Q. Four Aspects about Solid-State Nanopores for Protein Sensing: Fabrication, Sensitivity, Selectivity, and Durability. *Adv. Healthcare Mater.* **2020**, *9*, 2000933.
(217) Loudwig, S.; Bayley, H. Photoisomerization of an individual azobenzene molecule in water: an on-off switch triggered by light at a fixed wavelength. *Journal of the American Chemical Society* **2006**, *128*, 12404-12405.
(218) Tan, S.; Wang, L.; Liu, H.; Wu, H.; Liu, Q. Single nanoparticle translocation through chemically modified solid nanopore. *Nanoscale Res. Lett.* **2016**, *11*, 1-10.
(219) Merstorf, C. l.; Cressiot, B.; Pastoriza-Gallego, M.; Oukhaled, A.; Betton, J.-M.; Auvray, L.; Pelta, J. Wild type, mutant protein unfolding and phase transition detected by single-nanopore recording. *ACS chemical biology* **2012**, *7*, 652-658.
(220) Crescentini, M.; Bennati, M.; Saha, S. C.; Ivica, J.; De Planque, M.; Morgan, H.; Tartagni, M. A low-noise transimpedance amplifier for BLM-based ion channel recording. *Sensors* **2016**, *16*, 709.
(221) Balan, A.; Machielse, B.; Niedzwiecki, D.; Lin, J.; Ong, P.; Engelke, R.; Shepard, K. L.; Drndić, M. Improving signal-to-noise performance for DNA translocation in solid-state nanopores at MHz bandwidths. *Nano letters* **2014**, *14*, 7215-7220.
(222) Ma, H.; Zhong, C. B.; Ying, Y.-L.; Long, Y.-T. Seeing Is Not Believing: Filtering Effects on Random Nature in Electrochemical Measurements of Single-Entity Collision. *ACS Measurement Science Au* **2022**, *2*, 325-331.
(223) Zhang, L.-L.; Zhong, C.-B.; Li, J.-G.; Niu, H.-Y.; Ying, Y.-L.; Long, Y.-T. A two-step calibration method for evaluation high bandwidth electrochemical instrument. *Journal of Electroanalytical Chemistry* **2022**, *915*, 116266.
(224) Rosenstein, J. K.; Wanunu, M.; Merchant, C. A.; Drndic, M.; Shepard, K. L. Integrated nanopore sensing platform with sub-microsecond temporal resolution. *Nature methods* **2012**, *9*, 487-492.
(225) Fahie, M. A.; Yang, B.; Chisholm, C. M.; Chen, M. In *Nanopore Technology*; Springer: 2021, p 77-94.
(226) Gu, Z.; Ying, Y.-L.; Long, Y.-T. Nanopore sensing system for high-throughput single molecular analysis. *Science China Chemistry* **2018**, *61*, 1483-1485.

(227) Liu, S.-C.; Li, M.-X.; Li, M.-Y.; Wang, Y.-Q.; Ying, Y.-L.; Wan, Y.-J.; Long, Y.-T. Measuring a frequency spectrum for single-molecule interactions with a confined nanopore. *Faraday Discussions* **2018**, *210*, 87-99.
(228) Li, X.; Ying, Y. L.; Fu, X. X.; Wan, Y. J.; Long, Y. T. Single-Molecule Frequency Fingerprint for Ion Interaction Networks in a Confined Nanopore. *Angewandte Chemie* **2021**, *133*, 24787-24792.
(229) Fragasso, A.; Schmid, S.; Dekker, C. Comparing current noise in biological and solid-state nanopores. *ACS Nano* **2020**, *14*, 1338-1349.
(230) Sun, J.; Thakur, A. K.; Movileanu, L. Current noise of a protein-selective biological nanopore. *Proteomics* **2022**, *22*, 2100077.
(231) Fragasso, A.; Pud, S.; Dekker, C. 1/f noise in solid-state nanopores is governed by access and surface regions. *Nanotechnology* **2019**, *30*, 395202.
(232) Zhu, Z.; Duan, X.; Li, Q.; Wu, R.; Wang, Y.; Li, B. Low-noise nanopore enables in-situ and label-free tracking of a trigger-induced DNA molecular machine at the single-molecular level. *Journal of the American Chemical Society* **2020**, *142*, 4481-4492.
(233) Uram, J. D.; Ke, K.; Mayer, M. Noise and bandwidth of current recordings from submicrometer pores and nanopores. *ACS Nano* **2008**, *2*, 857-872.
(234) Roelen, Z.; Bustamante, J. A.; Carlsen, A.; Baker-Murray, A.; Tabard-Cossa, V. Instrumentation for low noise nanopore-based ionic current recording under laser illumination. *Review of Scientific Instruments* **2018**, *89*, 015007.
(235) Tabard-Cossa, V.; Trivedi, D.; Wiggin, M.; Jetha, N. N.; Marziali, A. Noise analysis and reduction in solid-state nanopores. *Nanotechnology* **2007**, *18*, 305505.
(236) Edel, J. B.; Albrecht, T. Engineered nanopores for bioanalytical applications. *William Andrew* **2013**, 121-140.
(237) Smeets, R. M.; Keyser, U. F.; Dekker, N. H.; Dekker, C. Noise in solid-state nanopores. *Proceedings of the National Academy of Sciences* **2008**, *105*, 417-421.
(238) Lee, K.; Park, K. B.; Kim, H. J.; Yu, J. S.; Chae, H.; Kim, H. M.; Kim, K. B. Recent progress in solid-state nanopores. *Advanced materials* **2018**, *30*, 1704680.
(239) Liang, S.; Xiang, F.; Tang, Z.; Nouri, R.; He, X.; Dong, M.; Guan, W. Noise in nanopore sensors: sources, models, reduction, and benchmarking. *Nanotechnology and Precision Engineering* **2020**, *3*, 9-17.
(240) Hoogerheide, D. P.; Garaj, S.; Golovchenko, J. A. Probing surface charge fluctuations with solid-state nanopores. *Physical review letters* **2009**, *102*, 256804.
(241) Gornall, J. L.; Mahendran, K. R.; Pambos, O. J.; Steinbock, L. J.; Otto, O.; Chimerel, C.; Winterhalter, M.; Keyser, U. F. Simple reconstitution of protein pores in nano lipid bilayers. *Nano letters* **2011**, *11*, 3334-3340.
(242) Schibel, A. E.; Edwards, T.; Kawano, R.; Lan, W.; White, H. S. Quartz nanopore membranes for suspended bilayer ion channel recordings. *Analytical chemistry* **2010**, *82*, 7259-7266.
(243) Balan, A.; Chien, C.-C.; Engelke, R.; Drndić, M. Suspended solid-state membranes on glass chips with sub 1-pF capacitance for biomolecule sensing applications. *Sci. Rep.* **2015**, *5*, 1-8.
(244) Gu, Z.; Wang, H.; Ying, Y.-L.; Long, Y.-T. Ultra-low noise measurements of nanopore-based single molecular detection. *Science Bulletin* **2017**, *62*, 1245-1250.
(245) Hartel, A.; Ong, P.; Schroeder, I.; Giese, M.; Shekar, S.; Clarke, O.; Marks, A.; Hendrickson, W.; Shepard, K. Ion channel recordings of the ryanodine receptor RyR1 at microsecond temporal resolution. *Proc Natl Acad Sci USA* **2018**, *115*, 1789-1798.

(246) Shekar, S.; Niedzwiecki, D. J.; Chien, C.-C.; Ong, P.; Fleischer, D. A.; Lin, J.; Rosenstein, J. K.; Drndic, M.; Shepard, K. L. Measurement of DNA translocation dynamics in a solid-state nanopore at 100 ns temporal resolution. *Nano letters* **2016**, *16*, 4483-4489.
(247) Kim, J.; Dunbar, W. B. High-precision low-power DNA readout interface chip for multichannel nanopore applications. *Sensors Actuators B: Chem* **2016**, *234*, 273-277.
(248) Fu, J.; Wu, L.; Qiao, Y.; Tu, J.; Lu, Z. Microfluidic systems applied in solid-state nanopore sensors. *Micromachines* **2020**, *11*, 332.
(249) Kleefen, A.; Pedone, D.; Grunwald, C.; Wei, R.; Firnkes, M.; Abstreiter, G.; Rant, U.; Tampé, R. Multiplexed parallel single transport recordings on nanopore arrays. *Nano letters* **2010**, *10*, 5080-5087.
(250) Tsai, D.; Sawyer, D.; Bradd, A.; Yuste, R.; Shepard, K. L. A very large-scale microelectrode array for cellular-resolution electrophysiology. *Nat. Commun.* **2017**, *8*, 1-11.
(251) Emiliani, F. E.; Hsu, I.; McKenna, A. Multiplexed assembly and annotation of synthetic biology constructs using long-read Nanopore sequencing. *ACS synthetic biology* **2022**, *11*, 2238-2246.
(252) Fried, J. P.; Wu, Y.; Tilley, R. D.; Gooding, J. J. Optical Nanopore Sensors for Quantitative Analysis. *Nano Letters* **2022**, *22*, 869-880.
(253) Lu, W.; Hu, R.; Tong, X.; Yu, D.; Zhao, Q. Electro-Optical Detection of Single Molecules Based on Solid-State Nanopores. *Small Structures* **2020**, *1*, 2000003.
(254) Li, W.; Zhou, J.; Maccaferri, N.; Krahne, R.; Wang, K.; Garoli, D. Enhanced optical spectroscopy for multiplexed DNA and protein-sequencing with plasmonic nanopores: challenges and prospects. *Analytical Chemistry* **2022**, *94*, 503-514.
(255) Garoli, D.; Yamazaki, H.; Maccaferri, N.; Wanunu, M. Plasmonic nanopores for single-molecule detection and manipulation: toward sequencing applications. *Nano letters* **2019**, *19*, 7553-7562.
(256) Heron, A. J.; Thompson, J. R.; Cronin, B.; Bayley, H.; Wallace, M. I. Simultaneous measurement of ionic current and fluorescence from single protein pores. *Journal of the American Chemical Society* **2009**, *131*, 1652-1653.
(257) Belkin, M.; Chao, S.-H.; Jonsson, M. P.; Dekker, C.; Aksimentiev, A. Plasmonic nanopores for trapping, controlling displacement, and sequencing of DNA. *ACS Nano* **2015**, *9*, 10598-10611.
(258) Reiner, J. E.; Robertson, J. W. F.; Burden, D. L.; Burden, L. K.; Balijepalli, A.; Kasianowicz, J. J. Temperature Sculpting in Yoctoliter Volumes. *Journal of the American Chemical Society* **2013**, *135*, 3087-3094.
(259) Yao, J.; Liu, B.; Qin, F. Rapid temperature jump by infrared diode laser irradiation for patch-clamp studies. *Biophysical journal* **2009**, *96*, 3611-3619.
(260) Payet, L.; Martinho, M.; Pastoriza-Gallego, M.; Betton, J.-M.; Auvray, L.; Pelta, J.; Mathé, J. Thermal unfolding of proteins probed at the single molecule level using nanopores. *Analytical chemistry* **2012**, *84*, 4071-4076.
(261) Payet, L.; Martinho, M.; Merstorf, C.; Pastoriza-Gallego, M.; Pelta, J.; Viasnoff, V.; Auvray, L.; Muthukumar, M.; Mathé, J. Temperature effect on ionic current and ssDNA transport through nanopores. *Biophysical journal* **2015**, *109*, 1600-1607.
(262) Crick, C. R.; Albella, P.; Ng, B.; Ivanov, A. P.; Roschuk, T.; Cecchini, M. P.; Bresme, F.; Maier, S. A.; Edel, J. B. Precise attoliter temperature control of nanopore sensors using a nanoplasmonic bullseye. *Nano letters* **2015**, *15*, 553-559.

(263) Piguet, F.; Ouldali, H.; Discala, F.; Breton, M.-F.; Behrends, J. C.; Pelta, J.; Oukhaled, A. High temperature extends the range of size discrimination of nonionic polymers by a biological nanopore. *Sci. Rep.* **2016**, *6*, 1-10.
(264) Angevine, C. E.; Seashols-Williams, S. J.; Reiner, J. E. Infrared laser heating applied to nanopore sensing for DNA duplex analysis. *Analytical chemistry* **2016**, *88*, 2645-2651.
(265) Angevine, C. E.; Robertson, J. W.; Dass, A.; Reiner, J. E. Laser-based temperature control to study the roles of entropy and enthalpy in polymer-nanopore interactions. *Science Advances* **2021**, *7*, eabf5462.
(266) Zheng, W.; Saliba, J. G.; Wei, X.; Shu, Q.; Pierson, L. M.; Mao, L.; Liu, C.; Lyon, C. J.; Li, C.-Z.; Wimley, W. C.; Hu, T. Y. Nanopore-based disease diagnosis using pathogen-derived tryptic peptides from serum. *Nano Today* **2022**, *45*, 101515.
(267) Subramanian, N.; Watson, B.; Li, C.-Z.; Moss, M.; Liu, C. Patterning amyloid-β aggregation under the effect of acetylcholinesterase using a biological nanopore - an in vitro study. *Sensors and Actuators Reports* **2023**, *6*, 100170.
(268) Wang, Y.; Zhang, S.; Jia, W.; Fan, P.; Wang, L.; Li, X.; Chen, J.; Cao, Z.; Du, X.; Liu, Y. Identification of nucleoside monophosphates and their epigenetic modifications using an engineered nanopore. *Nat. Nanotechnol.* **2022**, *17*, 976-983.
(269) Xin, K. L.; Hu, Z. L.; Liu, S. C.; Li, X. Y.; Li, J. G.; Niu, H.; Ying, Y. L.; Long, Y. T. 3D Blockage Mapping for Identifying Familial Point Mutations in Single Amyloid-β Peptides with a Nanopore. *Angewandte Chemie* **2022**, *134*, e202209970.
(270) Wei, Z.-X.; Ying, Y.-L.; Li, M.-Y.; Yang, J.; Zhou, J.-L.; Wang, H.-F.; Yan, B.-Y.; Long, Y.-T. Learning shapelets for improving single-molecule nanopore sensing. *Analytical chemistry* **2019**, *91*, 10033-10039.
(271) Khoury, G. A.; Baliban, R. C.; Floudas, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* **2011**, *1*, 90.
(272) Liu, S.-C.; Ying, Y.-L.; Li, W.-H.; Wan, Y.-J.; Long, Y.-T. Snapshotting the transient conformations and tracing the multiple pathways of single peptide folding using a solid-state nanopore. *Chem. Sci.* **2021**, *12*, 3282-3289.
(273) Van Meervelt, V.; Soskine, M.; Maglia, G. Detection of two isomeric binding configurations in a protein–aptamer complex with a biological nanopore. *ACS Nano* **2014**, *8*, 12826-12835.
(274) Van Meervelt, V.; Soskine, M.; Singh, S.; Schuurman-Wolters, G. K.; Wijma, H. J.; Poolman, B.; Maglia, G. Real-time conformational changes and controlled orientation of native proteins inside a protein nanoreactor. *Journal of the American Chemical Society* **2017**, *139*, 18640-18646.
(275) Galenkamp, N. S.; Biesemans, A.; Maglia, G. Directional conformer exchange in dihydrofolate reductase revealed by single-molecule nanopore recordings. *Nature Chemistry* **2020**, *12*, 481-488.
(276) Ho, B.; Baryshnikova, A.; Brown, G. W. Unification of protein abundance datasets yields a quantitative Saccharomyces cerevisiae proteome. *Cell systems* **2018**, *6*, 192-205. e193.
(277) Graf, M.; Lihter, M.; Thakur, M.; Georgiou, V.; Topolancik, J.; Ilic, B. R.; Liu, K.; Feng, J.; Astier, Y.; Radenovic, A. Fabrication and practical applications of molybdenum disulfide nanopores. *Nature protocols* **2019**, *14*, 1130-1168.
(278) Fried, J. P.; Swett, J. L.; Nadappuram, B. P.; Fedosyuk, A.; Gee, A.; Dyck, O. E.; Yates, J. R.; Ivanov, A. P.; Edel, J. B.; Mol, J. A. Localised solid-state nanopore fabrication via controlled breakdown using on-chip electrodes. *Nano Res.* **2022**, 1-9.
(279) Qiu, H.; Zhou, W.; Guo, W. Nanopores in Graphene and Other 2D Materials: A Decade's

Journey toward Sequencing. *ACS Nano* **2021**, *15*, 18848-18864.
(280) Hu, K.; Jia, R.; Hatamie, A.; Le Vo, K. L.; Mirkin, M. V.; Ewing, A. G. Correlating molecule count and release kinetics with vesicular size using open carbon nanopipettes. *Journal of the American Chemical Society* **2020**, *142*, 16910-16914.
(281) Yu, R.-J.; Hu, Y.-X.; Chen, K.-L.; Gu, Z.; Ying, Y.-L.; Long, Y.-T. Confined Nanopipet as a Versatile Tool for Precise Single Cell Manipulation. *Analytical Chemistry* **2022**, *94*, 12948-12953.
(282) Zhou, Y.; Sun, L.; Watanabe, S.; Ando, T. Recent Advances in the Glass Pipet: from Fundament to Applications. *Analytical Chemistry* **2021**, *94*, 324-335.
(283) Qiao, D.; Joshi, H.; Zhu, H.; Wang, F.; Xu, Y.; Gao, J.; Huang, F.; Aksimentiev, A.; Feng, J. Synthetic macrocycle nanopore for potassium-selective transmembrane transport. *Journal of the American Chemical Society* **2021**, *143*, 15975-15983.
(284) Hall, A. R.; Scott, A.; Rotem, D.; Mehta, K. K.; Bayley, H.; Dekker, C. Hybrid pore formation by directed insertion of α-haemolysin into solid-state nanopores. *Nat. Nanotechnol.* **2010**, *5*, 874-877.
(285) Cressiot, B.; Greive, S. J.; Mojtabavi, M.; Antson, A. A.; Wanunu, M. Thermostable virus portal proteins as reprogrammable adapters for solid-state nanopore sensors. *Nat. Commun.* **2018**, *9*, 1-7.
(286) Bell, N. A.; Engst, C. R.; Ablay, M.; Divitini, G.; Ducati, C.; Liedl, T.; Keyser, U. F. DNA origami nanopores. *Nano letters* **2012**, *12*, 512-517.
(287) Lanphere, C.; Ciccone, J.; Dorey, A.; Hagleitner-Ertuğrul, N.; Knyazev, D.; Haider, S.; Howorka, S. Triggered Assembly of a DNA-Based Membrane Channel. *Journal of the American Chemical Society* **2022**, *144*, 4333-4344.
(288) Dey, S.; Dorey, A.; Abraham, L.; Xing, Y.; Zhang, I.; Zhang, F.; Howorka, S.; Yan, H. A reversibly gated protein-transporting membrane channel made of DNA. *Nat. Commun.* **2022**, *13*, 1-12.
(289) Fennouri, A.; List, J.; Ducrey, J.; Dupasquier, J.; Sukyte, V.; Mayer, S. F.; Vargas, R. D.; Pascual Fernandez, L.; Bertani, F.; Rodriguez Gonzalo, S. Tuning the diameter, stability, and membrane affinity of peptide pores by DNA-programmed self-assembly. *ACS Nano* **2021**, *15*, 11263-11275.
(290) Langecker, M.; Arnaut, V.; Martin, T. G.; List, J.; Renner, S.; Mayer, M.; Dietz, H.; Simmel, F. C. Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science* **2012**, *338*, 932-936.
(291) Krishnan R, S.; Jana, K.; Shaji, A. H.; Nair, K. S.; Das, A. D.; Vikraman, D.; Bajaj, H.; Kleinekathöfer, U.; Mahendran, K. R. Assembly of transmembrane pores from mirror-image peptides. *Nature communications* **2022**, *13*, 1-13.
(292) Sun, K.; Zhao, C.; Zeng, X.; Chen, Y.; Jiang, X.; Ding, X.; Gou, L.; Xie, H.; Li, X.; Zhang, X. Active DNA unwinding and transport by a membrane-adapted helicase nanopore. *Nature communications* **2019**, *10*, 1-11.
(293) Niedzwiecki, D. J.; Mohammad, M. M.; Movileanu, L. Inspection of the engineered fhua δc/δ4l protein nanopore by polymer exclusion. *Biophysical journal* **2012**, *103*, 2115-2124.
(294) Fahie, M. A.; Pham, B.; Li, F.; Chen, M. A selective activity-based approach for analysis of enzymes with an OmpG nanopore. *Nanopore Technology: Methods and Protocols* **2021**, 115-133.
(295) Wolfe, A. J.; Gugel, J. F.; Chen, M.; Movileanu, L. Kinetics of Membrane Protein–Detergent Interactions Depend on Protein Electrostatics. *J. Phys. Chem. B* **2018**, *122*, 9471-9481.
(296) Vorobieva, A. A.; White, P.; Liang, B.; Horne, J. E.; Bera, A. K.; Chow, C. M.; Gerben, S.;

Marx, S.; Kang, A.; Stiving, A. Q. De novo design of transmembrane β barrels. *Science* **2021**, *371*, eabc8182.
(297) Hoogerheide, D. P.; Rostovtseva, T. K.; Jacobs, D.; Gurnev, P. A.; Bezrukov, S. M. Tunable Electromechanical Nanopore Trap Reveals Populations of Peripheral Membrane Protein Binding Conformations. *ACS Nano* **2021**, *15*, 989-1001.
(298) Hoogerheide, D. P.; Rostovtseva, T. K.; Bezrukov, S. M. Exploring lipid-dependent conformations of membrane-bound α-synuclein with the VDAC nanopore. *Biochim Biophys Acta* **2021**, *1863*, 183643.
(299) Bošković, F.; Keyser, U. F. Toward single-molecule proteomics. *Science* **2021**, *374*, 1443-1444.
(300) Wang, Y.; Zhang, S.; Jia, W.; Fan, P.; Wang, L.; Li, X.; Chen, J.; Cao, Z.; Du, X.; Liu, Y. Identification of nucleoside monophosphates and their epigenetic modifications using an engineered nanopore. *Nat. Nanotechnol.* **2022**, 1-8.
(301) Jia, W.; Hu, C.; Wang, Y.; Liu, Y.; Wang, L.; Zhang, S.; Zhu, Q.; Gu, Y.; Zhang, P.; Ma, J.; Chen, H.-Y.; Huang, S. Identification of Single-Molecule Catecholamine Enantiomers Using a Programmable Nanopore. *ACS Nano* **2022**, *16*, 6615-6624.
(302) Lu, W.; Cao, Y.; Qing, G. Recent advance in solid state nanopores modification and characterization. *Chemistry–An Asian Journal* **2022**, *17*, e202200675.
(303) Pardehkhorram, R.; Andrieu-Brunsen, A. Pushing the limits of nanopore transport performance by polymer functionalization. *Chemical Communications* **2022**, *58*, 5188-5204.
(304) Dal Peraro, M.; Cao, C. Mutant aerolysin and uses thereof. *Patent number: US2022127310 (A1).* **2022**.
(305) Wang, Y.-Q.; Li, M.-Y.; Qiu, H.; Cao, C.; Wang, M.-B.; Wu, X.-Y.; Huang, J.; Ying, Y.-L.; Long, Y.-T. Identification of essential sensitive regions of the aerolysin nanopore for single oligonucleotide analysis. *Analytical chemistry* **2018**, *90*, 7790-7794.
(306) Gu, L.-Q.; Braha, O.; Conlan, S.; Cheley, S.; Bayley, H. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature* **1999**, *398*, 686-690.
(307) Ayub, M.; Stoddart, D.; Bayley, H. Nucleobase recognition by truncated α-hemolysin pores. *ACS Nano* **2015**, *9*, 7895-7903.
(308) Qing, Y.; Tamagaki-Asahina, H.; Ionescu, S. A.; Liu, M. D.; Bayley, H. Catalytic site-selective substrate processing within a tubular nanoreactor. *Nat. Nanotechnol.* **2019**, *14*, 1135-1142.
(309) Astumian, R. D. Microscopic reversibility as the organizing principle of molecular machines. *Nat. Nanotechnol.* **2012**, *7*, 684-688.
(310) Wang, Y.; Zhang, Y.; Chen, X.; Guan, X.; Wang, L. Analysis with biological nanopore: On-pore, off-pore strategies and application in biological fluids. *Talanta* **2021**, *223*, 121684.
(311) Bandara, Y. M. N. D. Y.; Freedman, K. J. Enhanced Signal to Noise Ratio Enables High Bandwidth Nanopore Recordings and Molecular Weight Profiling of Proteins. *ACS Nano* **2022**, *16*, 14111-14120.
(312) Li, T.; Liu, L.; Li, Y. R.; Xie, J. N.; Wu, H. C. A Universal Strategy for Aptamer-Based Nanopore Sensing through Host-Guest Interactions inside -Hemolysin. *Angew. Chem., Int. Ed.* **2015**, *54*, 7568-7571.
(313) Xin, K.-L.; Hu, Z.-L.; Liu, S.-C.; Li, X.-Y.; Li, J.-G.; Niu, H.; Ying, Y.-L.; Long, Y.-T. 3D Blockage Mapping for Identifying Familial Point Mutations in Single Amyloid-β Peptides with a Nanopore. *Angewandte Chemie International Edition* **2022**, *61*, e202209970.
(314) De Lannoy, C. V.; Filius, M.; Van Wee, R.; Joo, C.; De Ridder, D. Evaluation of FRET X for

single-molecule protein fingerprinting. *Iscience* **2021**, *24*, 103239.
(315) de Lannoy, C.; Lucas, F. L. R.; Maglia, G.; de Ridder, D. In silico assessment of a novel single-molecule protein fingerprinting method employing fragmentation and nanopore detection. *iScience* **2021**, *24*, 103202.
(316) Bush, J.; Maulbetsch, W.; Lepoitevin, M.; Wiener, B.; Mihovilovic Skanata, M.; Moon, W.; Pruitt, C.; Stein, D. The nanopore mass spectrometer. *Review of Scientific Instruments* **2017**, *88*, 113307.
(317) Maulbetsch, W.; Wiener, B.; Poole, W.; Bush, J.; Stein, D. Preserving the sequence of a biopolymer's monomers as they enter an electrospray mass spectrometer. *Physical Review Applied* **2016**, *6*, 054006.