



Cite this: DOI: 10.1039/c8cs00106e

## Single-molecule protein sensing in a nanopore: a tutorial

Nitinun Varongchayakul,<sup>id</sup><sup>a</sup> Jiaxi Song,<sup>id</sup><sup>a</sup> Amit Meller<sup>id</sup><sup>\*ab</sup> and Mark W. Grinstaff<sup>id</sup><sup>\*ac</sup>

Proteins are the structural elements and machinery of cells responsible for a functioning biological architecture and homeostasis. Advances in nanotechnology are catalyzing key breakthroughs in many areas, including the analysis and study of proteins at the single-molecule level. Nanopore sensing is at the forefront of this revolution. This tutorial review provides readers a guidebook and reference for detecting and characterizing proteins at the single-molecule level using nanopores. Specifically, the review describes the key materials, nanoscale features, and design requirements of nanopores. It also discusses general design requirements as well as details on the analysis of protein translocation. Finally, the article provides the background necessary to understand current research trends and to encourage the identification of new biomedical applications for protein sensing using nanopores.

Received 12th July 2018

DOI: 10.1039/c8cs00106e

rsc.li/chem-soc-rev

### Key learning points

- (1) Nanopore sensing allows detection and characterization of protein(s) at the single-molecule level.
- (2) Nanopore sensing provides key information on protein concentration, shape, size, secondary/tertiary structure, conformations and conformational changes, post-translational modifications, unfolding kinetics, diffusion coefficients, and reaction kinetics.
- (3) Nanopore design and assay conditions control the type of interactions a protein will have with a nanopore, which includes translocation, collision, deformation and partial unfolding, linearization, adsorption to the wall, or tumbling inside the pore.
- (4) Nanopore sensing provides unprecedented opportunities to study fundamental protein biochemistry at the molecular level, to characterize the biophysics of protein translocation, and to design and evaluate new diagnostic devices for rapid, quantitative detection of proteins.

## 1. P is for protein

Proteins play critical roles in all aspects of life. Proteins consist of amino acids as their fundamental building units, and the sequence of which is referred to as the primary structure. These chemically precise macromolecules fold locally and globally into secondary and tertiary three-dimensional structures in aqueous solution to afford biochemically active entities. Proteins engage with other biomolecules through specific physicochemical interactions at their surfaces, and these interactions are dependent on the protein's structure and composition. These interactions dictate the selectivity, the time, and the strength of interaction(s), which enable the diversity of protein functions from structural and

organizational support, manufacturing, to signaling. Study of the protein structure in solution is key to understanding its biological role as well as to identifying new diagnostics and treatments for diseases.

Proteins are abundant in our daily life, for example, we observe albumin proteins in egg white transform from a clear liquid to a white solid during cooking or heat-induced denaturing process. In a research setting, the structures and compositions of proteins are analyzed using a variety of biochemical and biophysical techniques. Edman sequencing is the gold standard to determine a protein's primary structure, and it involves digesting the protein one amino acid at a time followed by mass spectrometry analysis to determine the sequence and any post-translational modifications, such as glycosylation. X-Ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, circular dichroism, and infrared spectroscopy are all common methods to assess a protein's secondary structure including alpha helices, beta sheets, beta turns, and random coils. The spatial relationships between the secondary structures of a protein define its tertiary structure, which X-ray crystallography and NMR spectroscopy are often used to

<sup>a</sup> Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA.  
E-mail: mgrin@bu.edu

<sup>b</sup> Department of Biomedical Engineering, Technion – Israel Institute of Technology, Haifa 3200003, Israel. E-mail: ameller@technion.ac.il

<sup>c</sup> Departments of Chemistry and Medicine, Boston University, Boston, MA 02215, USA

characterize. Finally, the quaternary structure reflects multiple folded subunits into a larger oligomeric complex. Similarly, techniques like X-ray crystallography and NMR spectroscopy are extensively used along with newer methods such as cryo-electron microscopy to elucidate the quaternary structures. The above techniques directly determine the protein composition and structures. There are also indirect methods, whereby a spectroscopy probe(s) attached at a specific site(s) on a protein provides a signal in response to folding, binding with another protein, post-translational modification, or oligomeric states.

Although the above techniques are widely used and valued in protein characterization, these bulk measurement or ensemble-averaged techniques do not enable us to address a number of critical questions, such as obtaining dynamic structural information on a protein (or its reaction or binding with another protein) in real-time. Moreover, bulk analysis methods require a large copy number of identical proteins which is not readily available in many clinical and biomedical situations. Investigating proteins at the single-molecule level is of particular interest as it provides: (1) insights into molecular mechanisms; (2) a dynamic view of the

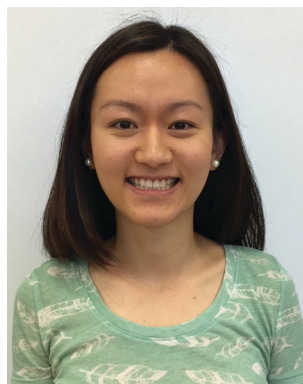
stochastic nature of chemical processes; (3) an overview of the heterogeneity across a population of molecules; and (4) opportunities to develop new protein based molecular diagnostics for research and clinical applications (Fig. 1).

To address these challenges as well as future questions, single-molecule protein sensing techniques have been developed, which include advanced optical microscopy techniques, optical/magnetic tweezers, atomic force microscopy, microcantilevers, nanochannels, and nanopores. For example, single-molecule force spectroscopy techniques such as optical tweezers, magnetic tweezers, and atomic force microscopy can provide information on the folding and unfolding of a protein structure as a function of applied external force and the binding strength between a protein and its partner ligand. Current emerging single-molecule techniques used to study proteins are reviewed elsewhere.<sup>1</sup> Among these techniques, nanopore sensing is at the forefront. The successful commercialization of nanopores for DNA sequencing (MinION by Oxford Nanopore Technologies) represents a significant milestone in the field, demonstrating the translatability of this technology.<sup>2</sup> The potential



**Nitinun Varongchayakul**

*Nitinun Varongchayakul received her Bachelor's degree from Chulalongkorn University in 2010 and her Master's degree in Material Science and Engineering from University of Maryland at College Park in 2012. She is pursuing a PhD degree under the tutelage of Professors Mark Grinstaff and Amit Meller. Her research interests lie in the use of nanopores for protein sensing.*



**Jiaxi Song**

*Jiaxi Song received her undergraduate degree in Chemistry at Mount Holyoke College in 2012. She is conducting her PhD studies under the mentorship of Professors Mark W. Grinstaff and Amit Meller. Her current research interests include using solid-state nanopores to identify biomolecules and developing novel photoinitiators for two-photon polymerization.*



**Amit Meller**

*Amit Meller is the Roy Matas Winnipeg Chair in Biomedical Engineering and full professor in the Department of Biomedical Engineering at the Technion-Israel Institute of Technology, Haifa, Israel. He is also director of the Israeli Center of Excellence in Biological Physics (i-Core). During his postdoctoral studies at Harvard University (1998–2000), he was fortunate to be among the first to develop and employ nanopore-based DNA and*

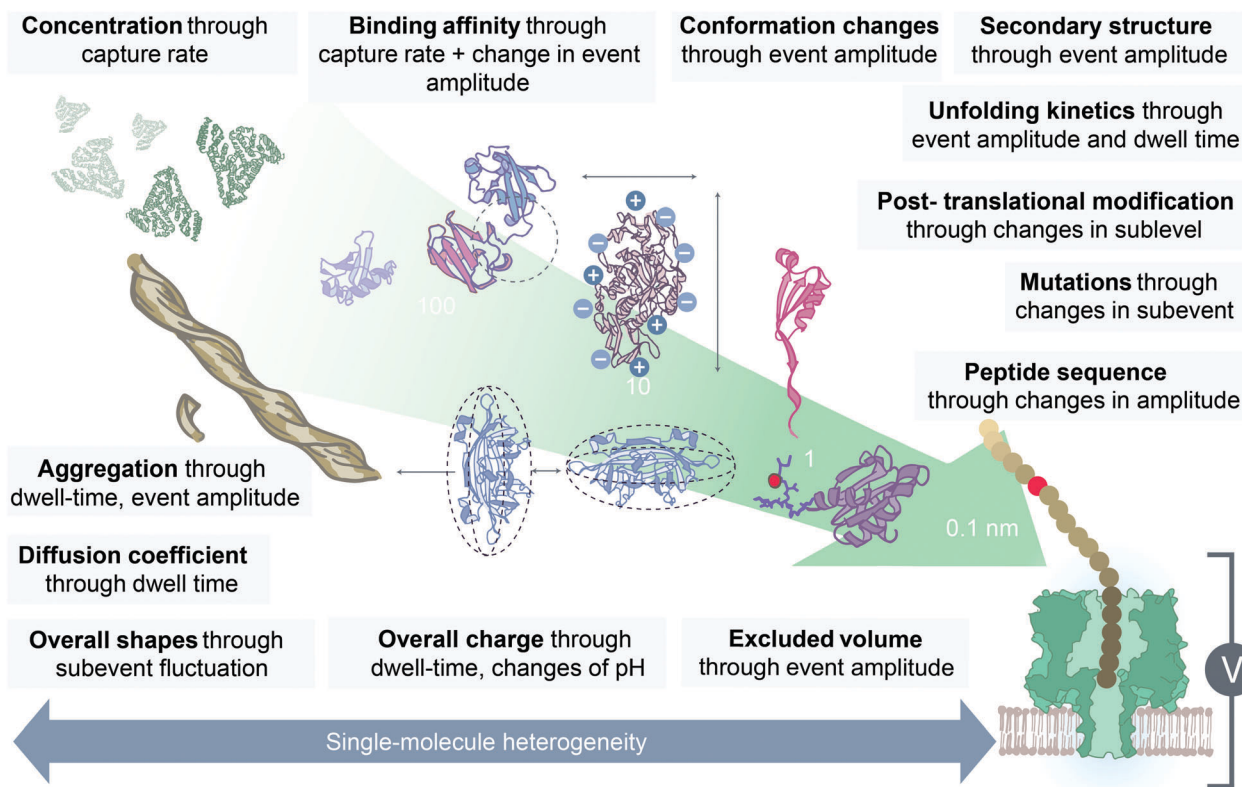
*RNA single molecule sensors, and publish some of the early and highly cited papers in this field. Dr Meller's current research interests include single-molecule sensing and single molecule biophysics, bio-optics, nano-sciences and nano-biotechnology.*



**Mark W. Grinstaff**

*Mark W. Grinstaff is the Distinguished Professor of Translational Research, a Professor of Biomedical Engineering, Chemistry, Materials Science and Engineering, and Medicine at Boston University, and the Director of BU's Nanotechnology Innovation Center. Grinstaff's groundbreaking research has yielded more than 290 peer-reviewed publications, and more than 200 patents and patent applications. He is a co-founder of five companies and his innovative*

*ideas and efforts have led to new commercialized products. His research interests cover a broad range of topics but always with an emphasis on new technologies, compositions, synthetic methods, and evaluation of properties.*



**Fig. 1** Biochemical/biophysical data on proteins are obtained from nanopore sensors. The information is ranked according to its length scale. For example, the concentration of the bulk solution protein is obtained by measuring the translocation rate. Nanopore sensing enables measurement of an enzyme and protein's binding kinetics,<sup>3–5</sup> oligomeric formation,<sup>6,7</sup> and aggregation kinetics.<sup>8</sup> Once the pool of events are collected, the information about a protein's diffusion coefficient, overall charge, volume and shapes is estimated.<sup>9,10</sup> Small nanopores are used to sense protein conformational changes,<sup>4,10</sup> unfolding pathway,<sup>11,12</sup> domain structures,<sup>3,5</sup> post-translational modifications,<sup>3,5,13</sup> and mutations,<sup>12</sup> as well as to identify peptide sequences.<sup>14</sup> A major advantage of using this single-molecule technique is the collection of data on individual biomacromolecules.

applications of native, unmodified, protein nanopore sensing include the opportunities to: (1) detect and quantify concentration of a protein(s) in solution; (2) analyze protein size and charge; (3) monitor binding interactions or reactions between proteins or a protein and a ligand to afford kinetic and equilibrium data; and (4) resolve conformational changes in structure (Fig. 1).

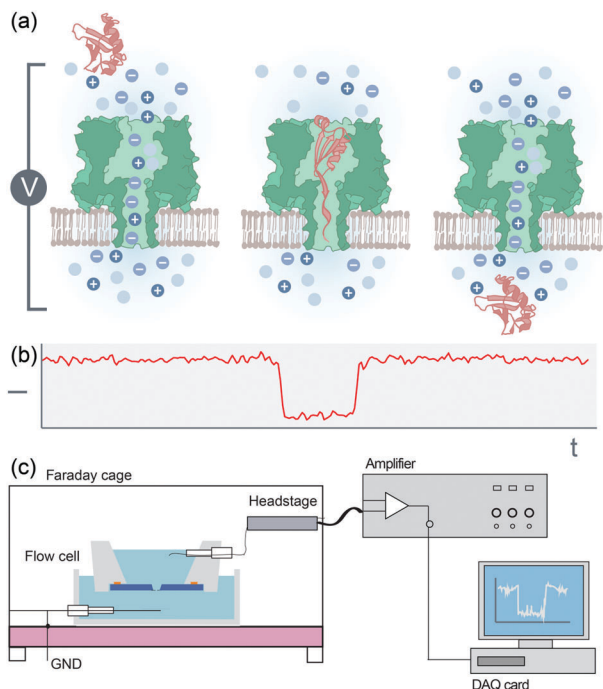
## 2. N is for nanopore sensor

A nanopore sensor consists of a nanometer-sized pore (*i.e.*, 1–100 nm in diameter) embedded or formed in an insulating membrane that separates two chambers containing an electrolyte solution. When an electrical bias is applied across the membrane, ions flow freely through the pore producing a constant open pore current. The flow of ions is partially impeded when a biomolecule diffuses through the pore or translocates from one side to the other under the influence of a driving force, thereby changing the ionic current. Fig. 2 depicts schematically the fundamentals of the nanopore sensing techniques. The principles of operation are reminiscent of a classic Coulter counter. However, the nanopore sensor provides enhanced sensing capability at the single-molecule level due to its nanoscale dimension. Proteins are

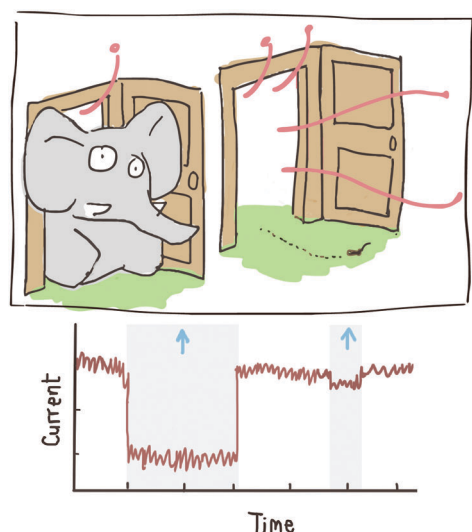
typically 2–10 nm in size. When the size of the pore is comparable to the size of the analyte, the change in ionic current is more prominent as compared to when using a larger pore. By analogy, pushing an elephant through a door significantly reduces the flow of air through the door while an ant quickly and easily crawls through the door barely, if at all, disturbing the airflow (Fig. 3).

### 2.1 Nanopore types – biological and synthetic

Nanopores are broadly categorized into two main groups: biological and synthetic pores (Table 1). The first nanopore for single-biomolecule detection used a biological protein pore,  $\alpha$ -hemolysin, for single-stranded RNA and DNA detection.<sup>15</sup> The same type of  $\alpha$ -hemolysin nanopores have subsequently been used to study proteins,<sup>16</sup> and protein/DNA complexes.<sup>17</sup> The device is composed of a single recombinant protein pore embedded in a lipid bilayer. More recently, additional protein pores such as ClyA,<sup>4,5</sup> aerolysin,<sup>14,18</sup> Nfp,<sup>19</sup> and FraC<sup>20</sup> have been evaluated. The advantage of protein nanopores include their well-characterized 3D structure and high reproducibility. Synthetic nanopores generally exhibit greater mechanical robustness, finer control over pore geometry, and opportunities for surface chemistry.<sup>21–25</sup> Synthetic solid-state nanopores are fabricated by ion beam sculpting,<sup>26</sup> electron-beam drilling,<sup>27</sup> controlled dielectric breakdown,<sup>28</sup> or direct laser drilling.<sup>29</sup>



**Fig. 2** A schematic view of the single-molecule nanopore sensing techniques. (a) Ions flow freely across the nanopore due to an applied electric field. Physical blockage of a protein of interest obstructs the ion flow, resulting in a drop in the pore's current. The current is restored once the protein leaves the pore. (b) A representative event as a result of a single protein translocation. The ionic current,  $I$ , is monitored over time,  $t$ . (c) Nanopore sensor experimental setup. The nanopore sensor resides in a Faraday cage to prevent external noise. The sensor is connected to a headstage, which transduces the current signal to an amplifier and to a data acquisition (DAQ) card. The data is then processed using a program to extract the nanopore events from the current trace as well as obtain the event characteristics.



**Fig. 3** Cartoon illustrating nanopore operation. Translocating an elephant vs. an ant through a door.

Advances in micro-nanofabrication are also enabling the fabrication of single-atom layer nanopores from 2D materials such

as graphene<sup>30</sup> and molybdenum disulfide (MoS<sub>2</sub>).<sup>25</sup> Last but not least, glass nanopores are fabricated by pulling a glass capillary to produce a taper end followed by electron beam sculpting to afford a narrow pore opening at the end.<sup>31</sup> The advantage of using a glass nanopore is the ease of fabrication compared to other synthetic pores, however, fine-tuning the diameter of the pore down to sub nanometer level is challenging. For additional information, we refer the reader to the following review which focus specifically on nanopore fabrication.<sup>32</sup>

## 2.2 General considerations for sensing single molecules: sensing region and resolution

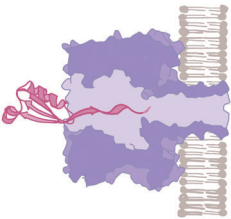
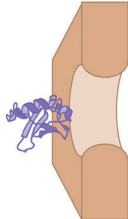
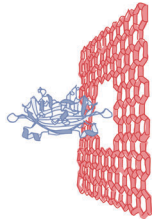
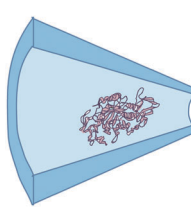
In single-molecule nanopore sensing, the pore's geometry is chosen to be comparable with the cross section of the bio-molecule of interest to maximize the change in ionic current during a translocation event and ultimately, maximize the signal-to-noise ratio. Typically, a translocation event is measured by the fractional blockade current ( $I_B$ ), defined as a ratio between the mean blocked current ( $i_b$ ) and the mean open pore current ( $i_o$ ):  $I_B = i_b/i_o$ . The fractional event amplitude ( $\Delta i/i_o$ ) is used interchangeably as it is  $1 - (i_b/i_o)$ . To a first approximation, the fractional event amplitude is correlated to the physical blocking volume of the analyte over the sensing volume of the pore. In general, the mean open pore current is expressed as:

$$i_o = V/R = V\sigma[(4l/\pi d^2) + (1/d)]^{-1} \quad (1)$$

where  $V$  is the voltage bias and  $R$  is the pore's resistance. The first term of the resistance ( $4l/\sigma\pi d^2$ ) arises from the geometrical constraint of the pore itself (assuming a cylindrical pore) while the second component so-called access resistance,  $R_{\text{access}} = 1/2\sigma d$ , results from the ionic current converging from the bulk solution into the pore's vicinity, hence  $R_{\text{total}} = R_{\text{pore}} + 2R_{\text{access}}$ . For instance, a synthetic pore of 4 nm in diameter and 10 nm in thickness, immersed in 1 M KCl solution (bulk conductivity  $\sigma = 10.5 \text{ S m}^{-1}$ ) results in  $R_{\text{total}}$  of 99.6 M $\Omega$ , which the access resistance contributes to over 24% of the total resistance. Under a bias of 300 mV, it will generate an open pore current of 3.0 nanoamperes.

The narrowest region of the pore is where the sensing occurs as it is the area with the largest electric field drop. In an  $\alpha$ -hemolysin pore, this sensing region is estimated to be  $\sim 5$  nm of the  $\beta$ -barrel's length, while the entire pore length is 10 nm.<sup>33</sup> An electron-beam sculpting solid-state nanopore usually possesses a double-conical shape, and thus, the effective pore's length is smaller than the nominal film's thickness. This dimension can be estimated, experimentally, by characterizing the translocation of a well-known analyte such as double-stranded DNA.<sup>34</sup> A narrow sensing region is desired in order to resolve the molecular features of the analyte such as a protein's primary or secondary structure.<sup>35</sup> However, a drawback is that thin pores exhibit lower mechanical stability and shorter translocation time. As a result, many of the events may not be resolved within the experimental limited temporal bandwidth.

**Table 1** The four main categories of nanopore sensors (drawings are not to scale)

Type of nanopores	Features	Advantages	Challenges
<b>Protein pore</b> 	<ul style="list-style-type: none"> <li>• A biological protein pore suspended on a lipid membrane</li> <li>• Recombinant protein synthesis technique</li> <li>• <i>e.g.</i>, <math>\alpha</math>-hemolysin,<sup>18</sup> ClyA,<sup>4,5</sup> aerolysin,<sup>14,18</sup> Nfp,<sup>19</sup> FraC20</li> </ul>	<ul style="list-style-type: none"> <li>• Well-defined sensing region (typically 1–2 nm in diameter, ~1 nm thickness)</li> <li>• Modification with atomic precision</li> <li>• Highly repeatable</li> <li>• Low electrical noise</li> </ul>	<ul style="list-style-type: none"> <li>• Constraint pore geometry and surface's chemical properties</li> <li>• Not compatible with the harsh chemical environment such as bleach cleaning and plasma treatment.</li> <li>• Poor mechanical stability (due to lipid membrane)</li> </ul>
<b>Solid-state pore</b> 	<ul style="list-style-type: none"> <li>• Small hole embedded in a free-standing membrane on a silicon chip</li> <li>• Milling by electron beam, ion beam, dielectric breakdown or laser-drilling</li> <li>• <i>e.g.</i>, SiN<sub>x</sub>,<sup>21,23</sup> HfO<sub>2</sub><sup>10,22</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Flexible geometrical design (1–100 nm in diameter, down to a few atomic layer thickness)</li> <li>• Surface chemical control</li> </ul>	<ul style="list-style-type: none"> <li>• E-beam drilling is expensive and slow</li> <li>• Difficulty in controlling and imaging the pore's shape</li> </ul>
<b>2D material pore</b> 	<ul style="list-style-type: none"> <li>• Small hole embedded in a free-standing membrane on a silicon chip</li> <li>• Milling by electron beam or ion beam</li> <li>• <i>e.g.</i>, graphene,<sup>30</sup> MoS<sub>2</sub><sup>25</sup></li> </ul>	<ul style="list-style-type: none"> <li>• High sensitivity (ideally at single-atom level)</li> </ul>	<ul style="list-style-type: none"> <li>• Limited surface chemical modification</li> <li>• Poor mechanical stability</li> <li>• The exact chemical composition of the edge of the nanopore is not known</li> </ul>
<b>Glass pore</b> 	<ul style="list-style-type: none"> <li>• A glass capillary with a tapered end</li> <li>• Pulling a glass capillary followed by electron beam sculpting</li> <li>• <i>e.g.</i>, borosilicate<sup>31</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Ease of fabrication</li> </ul>	<ul style="list-style-type: none"> <li>• Restricted pore size (usually &gt; 50 nm size-unless followed by e-beam sculpting)</li> </ul>

Capturing the nanopore signal requires a sufficient temporal bandwidth. For example, for a translocation dwell time of 1  $\mu$ s the system bandwidth should be  $\sim 1$  MHz. However, using a high bandwidth may complicate the measurement as the overall system noise increases rapidly at high frequencies and hence can obscure the protein translocation signals. Consequently, 10 to 100 kHz acquisition bandwidths (defined by the application of a low-pass filter) have often been used in nanopore sensing despite capturing only a fraction of the events.

In practice, the temporal resolution in nanopore sensing is dictated by a ratio of two factors: the event amplitude ( $\Delta i/i_0$ ) and the overall electrical noise ( $i_n$ ). The noise is approximated by the root-mean-square (RMS) of  $i_0$ . For example, a ratio of  $\sim 5$  is obtained, with a signal amplitude of  $\sim 0.5$  nA and an associated root-mean-square (RMS) noise of  $\sim 0.1$  nA, measured at a bandwidth of 100 kHz. When this ratio is smaller than  $\sim 2$  sensing might become impractical since the translocation events cannot be differentiated from the background noise.<sup>36</sup> The overall noise in the system originates from both the electrical sensing circuit (amplifier's internal circuit, choice of filters, charge transfer at the electrodes, *etc.*) and the physical characteristics (membrane material and composition, charge, capacitance, *etc.*) of the nanopore sensor.

### 3. A is for analysis

#### 3.1 Understanding the nanopore signal generated from proteins

Characterizing and understanding the physical principles governing protein translocation through a nanopore by the resulting ionic current signal is an ongoing challenge. The most straightforward model is that a protein translocates through the pore as a fully-intact protein, similar to a cell translocating through the several hundred micron-sized pore of a traditional Coulter counter (Table 2 – first entry). However, protein molecules are subjected to significant forces of physical, electrical, and chemical origins while traversing through the nanopore. As a result, the protein may: (1) translocate smoothly through the pore, (2) momentarily collide with the pore opening and never enter the pore, (3) temporarily or permanently adsorb onto the pore walls,<sup>37,38</sup> or (4) interacting with chemical moieties or ligands on the pore walls.<sup>4,39</sup> While the protein resides inside the pore, it may (5) tumble,<sup>4,9</sup> (6) bind/unbind with its ligands or (7) undergo transient or permanent structural changes such as unfolding.<sup>11,12,40</sup> Some, or all of these events may occur resulting in various characteristic signal signatures. Table 2 summarizes possible events, signal characteristics, and recommendations as to how to match a particular signal event with a translocation or other event(s). Note that first, these are generalized for all biological and synthetic pores. Second, multiple modes of interaction could happen in a single translocation. Lastly, this is just a set of guidelines and ultimately the analysis will require a thorough series of experiments.

#### 3.2 Regulating the type of interaction(s): voltage, pH, ionic strength, temperature, surfactant, protein concentration

Proteins are drawn and translocated through the nanopore as a result of electrophoretic, electroosmotic, and in some cases

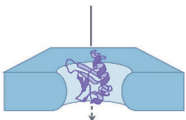

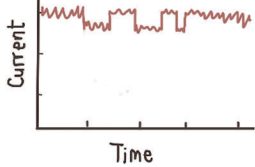
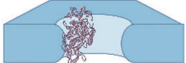
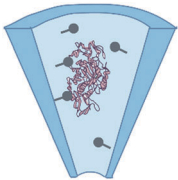
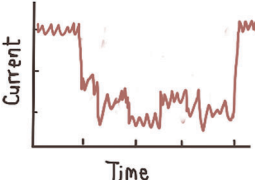
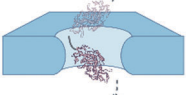
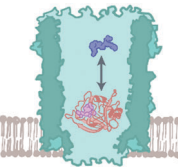
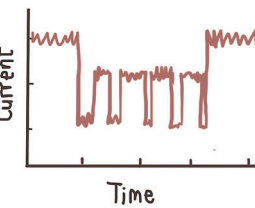
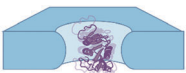
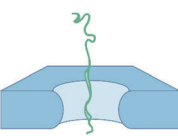
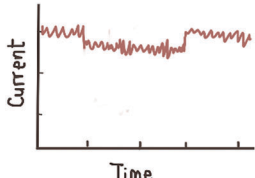
thermophoretic forces. The nanopore can also serve as a physical barrier to facilitate the unfolding of protein, especially when a pore's diameter is comparable or smaller than the size of a solvated protein.<sup>12</sup> The voltage bias directly determines the electrophoretic force: a larger voltage bias facilitates the translocation of a protein with the opposite charge with a shorter dwell time. The voltage bias also increases the flow of counter ions due to electroosmotic force, which may facilitate or hamper the translocation of a protein depending on the direction of water flow.<sup>20</sup> A stronger pulling force favors protein shearing or unfolding, resulting in a shallower event amplitude and a longer translocation time.<sup>12</sup> Nanopore-induced protein denaturation is naturally stochastic, resulting in broad signal distributions.

To promote protein linearization prior to translocating through the pore, an unfoldase protein has been anchored to the pore's opening to afford a signal with unique patterns corresponding to different proteins.<sup>11</sup> Strong denaturants are also used to denature the protein pre-emptively in solution before reaching the pore such as urea<sup>12,21</sup> or sodium dodecyl sulfate (SDS).<sup>42</sup> Surfactants will form a monolayer coating the pore's wall, and, thus, reduce the pore's effective diameter while minimizing non-specific adsorption of a protein onto the pore's wall similar to a lipid-coated pore.<sup>8</sup> In the later strategy, Yusko *et al.*, pre-treated 6–65 nm silicon nitride pores with an aqueous suspension of small unilamellar liposomes, creating a bilayer coating which increases the overall resistance of the nanopore. The protein analyte was anchored with the lipid molecule and diffused through the pore governed by the high viscosity of lipid membrane rather than the low viscosity of the aqueous electrolyte in the pore.

Temperature affects protein translocation through the nanopore in various aspects. First, a higher temperature promotes ion diffusion, hence the open pore current increases while the noise remains similar.<sup>18</sup> When a protein such as a maltose binding protein, MalE219, enters the nanopore, the observed event arrival rate increases as a function of temperature, which is likely due to the enhancement of thermal motion.<sup>18</sup> The increase in the protein's diffusion coefficient results in faster translocations through the nanopore. Further heating the nanopore system up to 70 °C affords a temperature-induced unfolding of the protein.<sup>18</sup>

Because a protein is an amphiphilic molecule with ionizable amino acids, its overall charge depends on the pH of the solution, which in turn affects the electrophoretic force acting on the protein. The isoelectric point (pI), or the pH at which the overall charge of the protein equals to zero, depends on its amino acid sequence and post-translational modifications. At a pH smaller than the pI, the protein is positively charged and therefore is drawn across the nanopore with an applied negative bias, and *vice versa*. Furthermore, as the pH approaches the pI, the protein's overall charge decreases, resulting in an increase in the translocation dwell time. Nir *et al.* reported that this small change can be significant – more than 2–3 orders of magnitude change in the dwell time when the pH is roughly 0.2–0.3 units away from the pI.<sup>3</sup> This has allowed a robust sensing

**Table 2** The possible events inside a nanopore as a protein molecule enters

Events	Signal characteristics	Events	Signal characteristics
Fully-intact, smooth translocation	Deep event amplitude  <ul style="list-style-type: none"><li>• Increase voltage/shorter dwell time</li></ul>	Collision  <ul style="list-style-type: none"><li>• Increase voltage/longer dwell time<sup>38</sup></li></ul>	Shallow events 
Adsorption to a pore's wall (non-specific interaction)	Deep events, millisecond- to second long  <ul style="list-style-type: none"><li>• Adsorption</li><li>• Electrostatic interaction</li></ul>	Specific interaction with the pore's wall  <ul style="list-style-type: none"><li>• ligand on the pore's wall</li></ul>	Deep events, millisecond- to second long <sup>8</sup> 
Tumbling inside a pore	Strong intra-event signal fluctuation, two preferred levels <sup>9,10</sup>  <ul style="list-style-type: none"><li>• Introduce strongly-charged tag like a short oligo<sup>4,11</sup></li></ul>	Binding/unbinding  <ul style="list-style-type: none"><li>• Due to coupling with other molecules (e.g. enzyme substrates)</li></ul>	Well defined sub-levels 
Conformational change (transient)	Contribute to the width of the event amplitude histogram  <ul style="list-style-type: none"><li>• May be induced by the pore's electric field or physical confinement</li></ul>	Unfolding  <ul style="list-style-type: none"><li>• Induced by the pore's electric field or physical confinement</li><li>• Introducing denaturants</li></ul>	Shallow events  <ul style="list-style-type: none"><li>• Introduce denaturing agents such as SDS or urea</li><li>• Increase voltage/shorter dwell time<sup>11,12,41</sup></li></ul>

of a 8.5 kDa protein, ubiquitin, as well as ubiquitin chains of various types. By performing nanopore sensing experiments across a series of pH values, the precise isoelectric point of each molecule is determined similar to the isoelectric focusing technique. However, as the pH of the solution approaches the pI, the protein has almost zero net charge and aggregations are likely to occur with few translocations.

In the case where the analyte has a strong dipole moment, a protein will enter the nanopore and translocate in a preferred

orientation. For example, a multilevel translocation signal is observed with the thioredoxin and DNA chimera as it translocates the pore under a positive bias. The event profile corresponds to a signal of the densely negatively charged DNA followed by a positively charged protein.<sup>12</sup> Another interesting translocation profile is observed with a linearized amphiphilic peptide which gives a ratcheting motion or molecular stalling effect as a likely result of the presence of localized charge on the adjacent segments affording a molecular tug-of-war effect.<sup>21</sup>

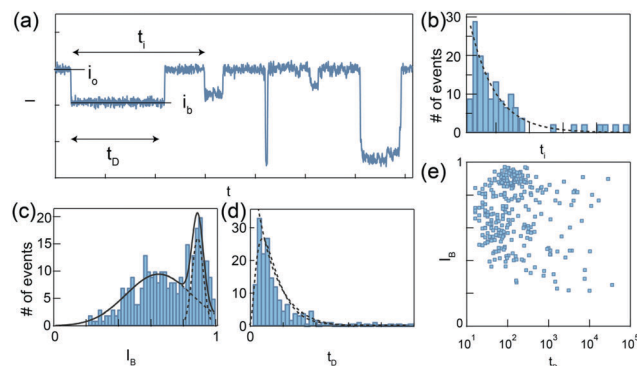
### 3.3 Controlling the protein capture rate

The signals obtained from one single molecule measurement are subject to statistical fluctuations since natural phenomena are usually governed by thermal noise. Thus, it is important to collect a significant number of events, usually from a few hundred to thousands of events, and perform statistical analysis. Large data set require longer data collection time, which also depends on the protein capture rate (the rate at which the nanopore events are detected).

There are multiple strategies to increase the detection rate of proteins. The first strategy is to enhance the electro-chemical gradient pulling the molecules towards the nanopore by the use of a salt-gradient.<sup>43</sup> A second strategy utilizes various types of ligands as carriers for the proteins of interest. Some common biomacromolecular carriers are DNA and RNA. In contrast to proteins, DNA and RNA molecules possess a uniform distribution of negative charges at neutral pH and therefore their translocation events are relatively easy to study and predict. Several studies have reported the successful use of these carriers to identify proteins using unfunctionalized solid-state nanopores. For example, Niedzwiecki *et al.* demonstrated the use of an RNA aptamer, stem-loop 3, to detect a protein biomarker of the human immunodeficiency virus 1, nucleocapsid protein 7 (NCp 7), using silicon nitride nanopores.<sup>44</sup> Upon applying a positive voltage, no translocation of NCp 7 was observed due to its net positive charge. When a negative voltage was applied, the proteins absorbed onto the silicon nitride surface, and translocation was thus hindered. However, when the NCp 7 protein was bound to an aptamer, well-defined translocation events with distinguishable current blockades from the free NCp 7 were observed, allowing measurement of the dissociation constant. In addition to RNA aptamers, Bell and co-worker reported the use of a DNA carrier to detect various proteins with high specificity.<sup>31</sup> The DNA carrier used in the study was a seven kbp double-stranded DNA with evenly spaced nicks on one strand, which allowed for chemical attachment of functional motifs at the 3' or 5' end. When biotin was conjugated to the DNA carrier, streptavidin was able to selectively bind to the DNA carrier, producing predominantly multi-step translocation events. Whereas unmodified DNA carriers and biotin-conjugated DNA carriers, without exposure to streptavidin, afforded mainly single-step translocation events. Beside controlling capture rate, protein/aptamer interaction can also be used to enhance signal discrimination, enabling a quantification of multiple protein biomarkers from relatively complex sample like blood.<sup>45</sup>

### 3.4 Basic characteristics in signal processing: capture rate, event amplitude, and dwell time

The three basic nanopore event characteristics are capture rate, event amplitude and dwell time. The capture rate is the rate at which the nanopore event occurs, obtained by measuring the inter-event interval ( $t_i$ , see Fig. 4a and b) and fitting its histogram with the exponential decay function. The nature of the protein capture rate is stochastic and is classified based on



**Fig. 4** Basic nanopore signal processing. Example translocations of a model protein. (a) Representative current trace ( $I$ ) as a function of time ( $t$ ). Each drop in current is defined as a nanopore event. (b) The time between the start of each event is defined as the inter-event interval ( $t_i$ ), which is fitted by an exponential decay function. (c) The ratio between event amplitude ( $i_b$ ) and open pore current ( $i_o$ ) is fitted by multimodal distribution. (d) The event dwell time ( $t_d$ ) histogram is fitted by either drift-diffusion model or exponential decay function. (e) The scatter plot of  $i_b$  vs.  $t_d$  is used to perform single-molecule protein classification.

two mechanisms: (1) a protein diffuses from bulk solution and is trapped in the electric field generated near the vicinity of the pore following the Smoluchowski's rate equation; and, (2) a protein overcomes an entropic barrier and enters the pore funnel under electrophoretic and electroosmotic forces.<sup>43,46</sup> The two mechanisms depend on the protein's bulk concentration as well as its biophysical properties such as diffusion coefficient, charge, and size relative to the pore. These dependencies imply that the capture rate can be used to measure the analyte bulk concentration (if the biophysical properties are known or the concentration/capture rate curve is established) or, *vice versa*.

The nanopore event amplitude ( $\Delta i$ ) is proportional to the physical blockage of ionic current flow through the pore. The event amplitude is often normalized to  $i_o$  to mitigate slow drifts in  $i_o$  during an experiment. It is widely accepted that the event amplitude is positively correlated to the occupied volume of the protein inside the sensing region. However, the exact relationship between protein's geometry and the event amplitude is still ambiguous. This is partly a consequence of the natural non-homogenous electrical charge distribution in the vicinity of each folded protein, which could give rise to significant deviations from the simple geometrical volume exclusion by the molecule.

Last but not least, the dwell time, defined by the time elapsed between the initial drop in the ion current to its return above the same level, is a reflection of the time the protein spends in the pores sensing volume. For small nanopores the translocation process can be generally viewed as an energy barrier crossing, hence the dwell time decays exponentially with the applied voltage (the applied voltage reduces the barrier by  $qV$ , where  $q$  is the total protein charge and  $V$  is the potential drop). In this case, performing measurements at several applied voltages is an effective way to discern the true translocation events from short collisional events, which are much less sensitive to the applied voltage. The dwell time distribution

contains information about the protein's charge, mobility, and diffusion coefficient. Typically, the dwell time distribution is tail-fitted by an exponential decay function to provide a characteristic time scale. In cases which there is an evidence that the protein translocate through the pore with a uniform velocity, a diffusion-drift model may be used to fit the data, although caution should be exercised when interpreting the meaning of the diffusion constant,  $D$ , and the proteins mobility,  $\nu$ , inside a narrow pore. Notably, these two fit parameters are largely influenced by interactions with the pore itself, in addition to the normal solution friction, hence complicating their physical meaning.<sup>22</sup>

$$P(t) = (h_{\text{eff}}/(4\pi Dt^3))^{\frac{1}{2}} e^{-\frac{(h_{\text{eff}} - \nu t)^2}{4Dt}} \quad (2)$$

The physical confinement and the interaction between the protein and pore's wall also can be theoretically related under some approximation of  $D$  and  $\nu$ .<sup>47</sup>

### 3.5 Advanced signal analysis: subevent information

Complex structural motifs affecting the local size (cross-section) of proteins may result in long ion-current blockade events consisting of multiple subevent current levels. To analyze such events, an all data point histogram is plotted to identify if there are multiple sublevels. Each difference in sublevel must be larger than the current noise to be resolved. The sublevel identification and pattern recognition is done manually<sup>11,12</sup> or by an algorithm such as edge-finding (similar to the one used in analyzing single-molecule FRET data).<sup>23</sup> More advanced signal processing techniques involving machine learning are emerging and will play a crucial role in analyzing nanopore signals generated from proteins. The current fluctuations during the event, which is measured by the noise of the residual current, is used to detect mutations or post-translational modifications on the protein.<sup>13</sup> The distribution of an event amplitude histogram describes the natural fluctuation of the protein inside the pore, which is related to its overall gross shapes (e.g., prolate or oblate ellipsoid), its dipole moment, and its rotational diffusion coefficient. Up to five parameters can be obtained to classify a protein.<sup>9,10</sup>

## 4. E is for examples

The following examples showcase the capabilities of nanopores for sensing and characterizing proteins. For a comprehensive review of single-molecule protein sensing in a nanopore, we refer the readers to two excellent reviews.<sup>35,40</sup> In the following examples, we describe the use of nanopores to: (1) probe binding affinity; (2) characterize protein domains and oligomeric states; (3) study protein unfolding; (4) detect specific sites of phosphorylation on a protein; and (5) identify biomarkers and peptide variants.

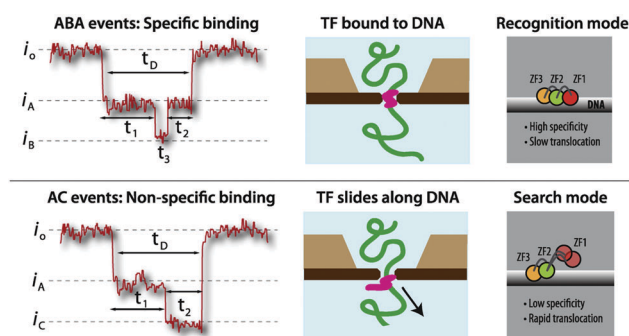
### 4.1 Probing binding affinity

Nanopore sensing is a useful tool for the study of biomolecular complexes consisting of two or more biomolecules. Squires *et al.* investigated the binding between DNA and a classical

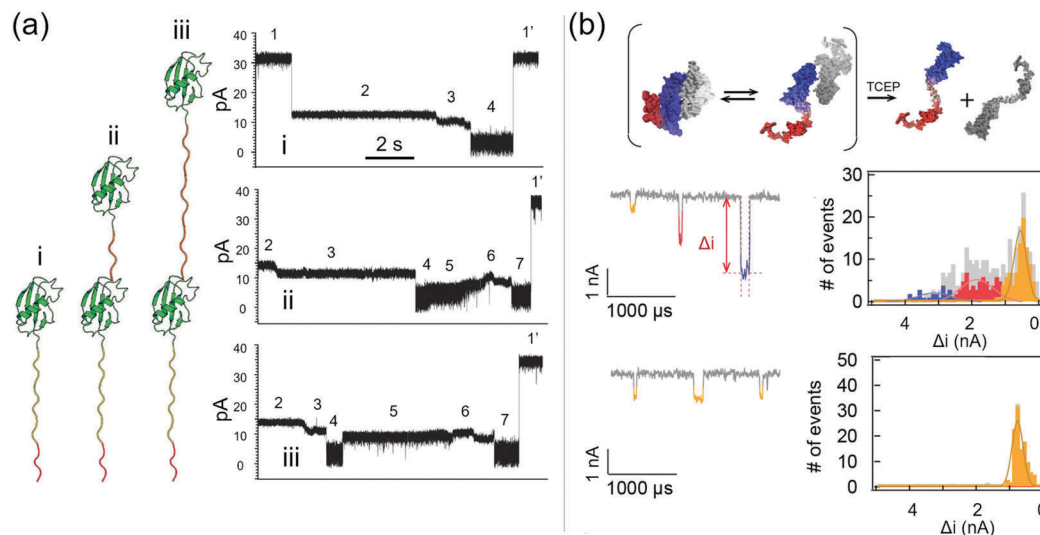
transcription factor (TF) using a solid-state nanopore.<sup>23</sup> Binding events between DNA and TFs play a critical role in the regulation of gene expression, and thus understanding these interactions provides valuable insights into genetic regulation. In this study, the well-characterized zinc finger protein, zif268, also known as early growth response protein 1 (EGR-1), binds a 1 kbp DNA fragment containing only one binding site for zif268. Upon translocation, five distinct event patterns, based on different steps found in subevent, were categorized. The researchers concluded that Zif268 binds to the DNA *via* two different states, a specific recognition state and a non-specific search state. When zif268 specifically binds to the DNA in the recognition state, all three zinc finger domains were bound to the major groove of the DNA. Whereas in the non-specific search state, one of the zinc finger domains was positioned away from the DNA, and the TF was able to slide easily and quickly along the DNA. Using the translocation data of the zif268 + DNA complex, different subgroups of translocation events are associated with the recognition and search mode of zif268 respectively. In the recognition state, translocation of the zif268 + DNA complex produced an ABA pattern event as shown in Fig. 5, top panel. Blockage at level A represented the strand of DNA and a deeper level at B resulted from the tightly bound zif268. In the search state, translocation of the DNA-TF complex led to the event with AC pattern (Fig. 5, bottom panel). Similar to the recognition state, the A level was associated with the DNA backbone and the deeper blockage of level C compared to level B was attributed to the less compact structure of zif268 in the search state.

### 4.2 Characterizing protein domains and oligomeric states

In biological systems, proteins can exist in different oligomeric states, and some proteins possess multiple domains or secondary/tertiary structures. Protein oligomers are often connected by disulfide bonds, and different domains within a protein exhibit distinct structures and functions. Several studies have elucidated



**Fig. 5** Probing binding states between DNA and a transcription factor (TF) using a solid-state nanopore. The 1 kbp DNA contains a single binding site for TF zif268. When the TF binds to the DNA specifically (top panel), the DNA backbone leads to a current drop from  $i_o$  to  $i_A$  and the deeper blockage level at  $i_B$  represents the tightly bound TF. In the non-specific binding state (bottom panel), the TF slides along the DNA backbone, creating a distinct blockage level at  $i_C$ . Adapted with permission from A. Squires, E. Atas, and A. Meller, *Sci. Rep.*, 2015, **5**, 11643, under Creative Commons Attribution 4.0 International license, Nature Publishing.



**Fig. 6** Characterization of protein (a) domains and (b) oligomeric states using nanopores. In the example shown in (a), three variants of Smt3 protein were labelled as i, ii and iii as shown on the left. Their corresponding translocation events revealed multiple ionic current blockage levels on the right. The sub-events were categorized into 7 stages, with stage 4 and stage 7 representing the translocation of the unfolded Smt7 protein. The events shown in (a) are also examples of nanopore-induced protein unfolding discussed in Section 4.3. Adapted with permission from J. Nivala, D. B. Marks, and M. Akeson, *Nat. Biotechnol.*, 2013, **31**(3), 247–250. Copyright 2013, Nature Publishing. In (b), the monomeric, dimeric and trimeric states of VEGF exhibited moderate (yellow), intermediate (red) and deep (blue) current blockage levels that correspond to their sizes. After the addition of TCEP, a disulfide bonds reducing agent, the dimeric and trimeric VEGF proteins were converted into the monomeric state. Adapted with permission from N. Varongchayakul, D. Huttner, M. W. Grinstaff, and A. Meller, *Sci. Rep.*, 2018, **8**(1), 1017, under Creative Commons Attribution 4.0 International license, Nature Publishing.

the different domains and oligomeric states of proteins using nanopores.<sup>3,7,11</sup> Nivala *et al.*,<sup>11</sup> examined three variants of an ubiquitin-like protein (Smt3), all of which were conjugated to a negatively-charged polypeptide tail to aid in nanopore translocation with an applied positive bias and an unfoldase protein coupled to an  $\alpha$ -hemolysin pore (Fig. 6a). Variant i had one unit of Smt3, and variant ii and iii each had two units of Smt3 separated by a short and long linker, respectively. Translocation of these three variants produced distinct sub-event levels and residence times which correlated with their three-dimensional structures. Another study by Varongchayakul *et al.*,<sup>7</sup> studied the dynamics of monomer, dimer, and trimer states of vascular endothelial growth factor (VEGF), which are held together *via* disulfide linkages. Each VEGF structure produced a unique event amplitude, with monomers affording the smallest blockage level and trimers giving the largest (Fig. 6b). After the addition of tris(2-carboxyethyl)phosphine (TCEP), a commonly used disulfide bond reducing agent, to a solution containing a mixture of VEGF structures, the majority of translocation events exhibited a small blockage level consistent with a population of VEGF monomer present. When the enzyme plasmin cleaved the receptor recognition and heparin-binding domain of VEGF, the positively charged heparin-binding domain gave single-level events with lower amplitude compared to events before cleavage when a negative potential was applied. When the potential was switched across the membrane, the negatively charge receptor recognition domain translocated with single-level events detected.

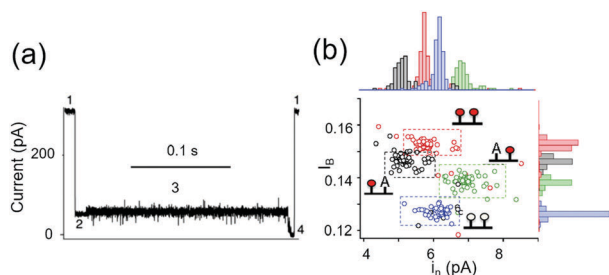
### 4.3 Exploring nanopore-induced protein unfolding

When the diameter of the nanopore is smaller than the diameter of the protein molecules, the protein may undergo

structural deformation such as partial or complete unfolding in order to traverse the pore. In the reported translocation events of the Smt3 protein, the translocation corresponding to construct i (Fig. 6a) shows several distinct sub-events with different amplitudes and dwell times.<sup>12</sup> These sub-events are correlated to different stages of translocation in the  $\alpha$ -hemolysin nanopore. The protein substrate, in this case, is tagged with a negatively-charged polypeptide segment to induce unidirectional movement of the molecule when a potential is applied. Once the polypeptide tail traverses through the pore, the unfoldase-targeting sequence on the tail is recognized by the ClpX unfoldase protein in the *trans* chamber and the protein is then unfolded and threaded through the pore. Stage 1 shows the open-pore current before the capture of the protein. When the linear polypeptide tail enters the pore, the ion flow decreases immediately, creating a sharp drop in current as shown in Stage 2. Upon recognition between ClpX unfoldase and the polypeptide, the current drops further in Stage 3. The potential and unfoldase-driven force continue to pull on the protein, causing it to gradually unfold, enter and exit the pore, leading to a period of deep current blockage in Stage 4. Once the protein leaves the pore, the open-pore current returns to the same level in Stage 1. The magnitude and duration of each level in these subevents reveals further structural and functional information on the proteins.<sup>48</sup>

### 4.4 Detecting post-translational modification

After translation, a protein may undergo post-translational modification(s) such as phosphorylation, acetylation, ubiquitination and glycosylation. These modifications are critical to the function of the mature, native protein. To determine if a

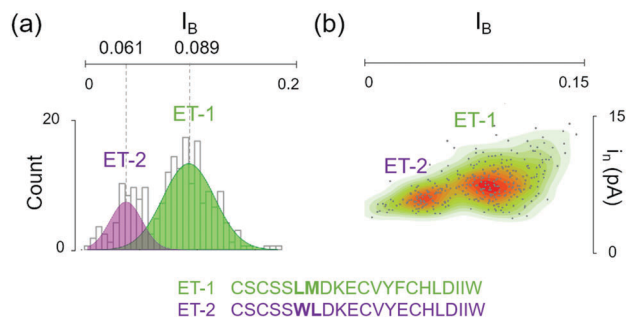


**Fig. 7** Detecting thioredoxin protein in native and three different phosphorylation states using an  $\alpha$ -hemolysin nanopore. (a) Representation event trace shows multistep translocation. (b) The scatter plot of the fractional blockade current ( $I_B$ ) vs. subevent noise ( $i_n$ ) at level 3 of the four thioredoxin constructs revealed four unique populations highlighted in blue, green, black and red, which correspond to the native and three different phosphorylation states of thioredoxin. The red circles signify the number and position of phosphorylated residues on the protein. Each protein is also tagged with a 30-mer oligonucleotide to the C-terminus to facilitate translocation. Adapted with permission from C. B. Rosen, D. Rodriguez-Larrea, and H. Bayley, *Nat. Biotechnol.*, 2014, **32**(2), 179–181. Copyright 2014, Nature publishing.

nanopore sensor detects these fine features of protein primary structure at the single-molecule level, Rosen and co-workers investigated the phosphorylation of a model protein thioredoxin (Fig. 7).<sup>13</sup> Thioredoxin contains two phosphorylatable sites at positions 107 and 112. To facilitate nanopore sensing, they conjugated a 30-mer oligonucleotide to the C-terminus of the protein. Upon applying a potential across the nanopore with a positive bias at the *trans* chamber, the negatively charged oligonucleotide guided the oligonucleotide–protein conjugate to the pore for subsequent translocation. During the translocation, thioredoxin unfolds, and a multi-level current pattern is observed. The four constructs of thioredoxin, containing the combinations of native and three different phosphorylation states, gave unique translocation signatures elucidated by the scatter plots of the fractional event amplitude *versus* subevent noise as shown in Fig. 7.

#### 4.5 Identifying biomarkers and peptide variants

Nanopore technologies have been widely explored for applications in nucleic acid sequencing in the past decade. Recent research efforts are directed at nanopore sensing to detect proteins or peptides. In 2017, Huang and co-workers described the use of FraC nanopores to identify a variety of proteins with size ranging from 1.3 kDa to 25 kDa.<sup>20</sup> When these proteins were analyzed independently, the translocation events were compared based on the dwell times and the fractional event amplitude. The events corresponding to each protein clustered around a certain region on the dwell time *versus* fractional event amplitude plot. When a mixture of three different proteins ( $\beta$ 2-microglobulin, human epidermal growth factor and endothelin 1) was investigated and analyzed, based on a dwell time *vs.* fractional event amplitude diagram, the event cluster corresponding to each protein produced a distinct population on the plot. Thus, these proteins are discriminated from each other based on the translocation data. Furthermore, when



**Fig. 8** Distinguishing endothelin 1 (ET-1) from endothelin 2 (ET-2) using a FraC nanopore. (a) The translocation data of ET-1 and ET-2 revealed distinguishable fractional blockade current ( $I_B$ ) at 0.089 for ET-1 and 0.061 for ET-2. (b) By adding ET-1 and ET-2 consecutively to the same pore, two distinct populations were observed by plotting the event amplitude standard deviation ( $i_n$ ) over the corresponding  $I_B$ . (inset) ET-1 and ET-2 are two nearly isomeric polypeptides differing from each other by 1 amino acid out of 21 as well as the position of a leucine residue (leucine 6 in ET-1 and leucine 7 in ET-2). Adapted with permission from G. Huang, K. Willems, M. Soskine, C. Wloka, and G. Maglia, *Nat. Commun.*, 2017, **8**, 935, under Creative Commons Attribution 4.0 International license, Nature Publishing.

comparing the translocation events of endothelin 1 (ET-1) and endothelin 2 (ET-2), it was possible to distinguish these two peptides from each other even though these two peptides only differ by 1 amino acid out of 21 in their sequences (Fig. 8). This remarkable sensing specificity highlights the potential of nanopore sensing for identifying proteins in a mixture. In 2018, Piguet *et al.*, used a wild-type aerolysin nanopore to identify a single amino acid variant in uniformly charged homopolymeric peptides.<sup>14</sup> First, arginine peptides of 5–10 amino acids, which differed by a single amino acid in length, were introduced to the nanopore. Unique fractional event amplitude and dwell-times were observed. The 10-amino acid long lysine homopeptides show different blockage current compared to the arginine homopeptides of the same length. Additionally, the event population shifts toward shorter species upon addition of trypsin enzyme. This result demonstrated that the mixture of the two homopeptides were discriminated at single-molecule resolution. Finally, a heteropolymer of 5 lysine amino acids and 5 arginine amino acids displayed unique event amplitude upon comparison with the homopolymers of arginine or lysine. These two reports reinforce the promising ability of a nanopore to sense differences at the single amino acid level.

## 5. O is for outlook

### 5.1 Robust signal processing and understanding the protein translocation signal

There are multiple approaches to analyse or extract the features of the nanopore events, including the average dwell time, event amplitude, steps, noise, *etc.* Future nanopore analysis will include more sophisticated signal processing algorithm such as Wavelet Transform (similar to EKG signal), and competent composite method (similar to data compression technique).

Classifying the signals using machine learning algorithms *via* supervised learning (such as support vector machine or neural network) or unsupervised learning (such as kmeans<sup>9</sup>) is becoming more common and likely to be the mainstream in the future. Computational simulation, both at atomic and higher level, will shed additional light on the mechanism of protein translocation or its interaction with the pore, however, such technique do require high computational power to account for the translation, rotation, as well as structural unfolding of the protein.

## 5.2 Device development

Although nanopore sensors are used in a number of laboratories and commercial available for DNA sequencing, significant opportunities exist for improvements including:

- (1) higher temporal bandwidth, while minimizing noise, in order to resolve nanopore events with shorter dwell time and to capture more proteins of low copy number present in solution;
- (2) innovative fabrication techniques to precisely create an ultra-thin and ultra-small sensing region, down to Angstroms, will enable identification of individual amino acids in a protein;
- (3) expansion of emerging detection modalities such as optical readout<sup>49</sup> or transverse current readout<sup>50</sup> to obtain additional information from the protein translocation event;
- (4) improved methods for high-throughput nanopore fabrication. Alternative nanopore fabrication techniques such as dielectric breakdown<sup>28</sup> and laser-based drilling<sup>29</sup> might be an answer to the issue of nanopore reproducibility and mass-production; and,
- (5) auxiliary devices to connect with the nanopore or multi-use integrated nanopore devices (*e.g.*, microfluidic nanopore device) to enable the sensing of biological and clinical samples wherein the sample is isolated, purified, and delivered to the nanopore for sensing in a single device.

## 5.3 Future applications

Nanopore sensing is at the forefront of the nanotechnology and nanomedicine revolution. The availability of low-cost, single-molecule protein nanopore sensing technology will undoubtedly advance basic research in molecular biology and medicine. Moreover, such technologies will likely be quickly adopted for commercial use in a broad range of industries, including biomedicine (molecular diagnostics and drug development), biotechnology (food industries and water quality control), cosmetics, and forensics. The application space is extensive. For example, the detection and quantification of a specific protein, including those present at a low concentration, will be accomplished using a tag, where the tag ensure a specific and unique signal from the translocation event for a given protein. Multiplexing this approach using different tags for different proteins along with signal processing advancements will enable many proteins to be detected from a single assay volume. Another specific example involves quantification of enzymatic activities. Given the ability to study protein binding and enzyme reaction kinetics, the nanopore provides the means to study protein reactivity at the single-molecule level or to screen for reactivity during drug discovery and development. As such, the nanopore will reduce the amount of material required for analysis and replace the use of laborious bulk measurements such as gel

electrophoresis or immunoassays. With regards to clinical applications, nanopore sensors require minimal sample, and, thus, screening or identifying a specific protein in a patient's blood, urine, or sweat sample could be accomplished in a minimally invasive manner. Alternatively, a nanopore sensor could be coupled with a controlled microfluidic device to create a single-molecule sorting device – analogous to a flow cytometer used to sort cells.

The purpose of this tutorial review is to provide a background for those interested in nanopore technology for sensing proteins and to instil a sense of the excitement, that we see, for this area. We encourage newcomers, from all science and engineering disciplines, to work in this field in order to advance nanopore materials and sensing methodologies, and to identify new applications for protein sensing. Today, the opportunities seem endless.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was supported in part by the NIH (Center for Future Technologies in Cancer Care (CFTCC U54 EB015403)), Boston University, BeyondSeq consortium (EC program 63489), and the i-Core program of the Israel Science Foundation (1902/12).

## References

- 1 J. J. Gooding and K. Gaus, *Angew. Chem., Int. Ed.*, 2016, **55**, 11354–11366.
- 2 M. Jain, S. Koren, K. H. Miga, J. Quick, A. C. Rand, T. A. Sasani, J. R. Tyson, A. D. Beggs, A. T. Dilthey, I. T. Fiddes, S. Malla, H. Marriott, T. Nieto, J. O'Grady, H. E. Olsen, B. S. Pedersen, A. Rhie, H. Richardson, A. R. Quinlan, T. P. Snutch, L. Tee, B. Paten, A. M. Phillippy, J. T. Simpson, N. J. Loman and M. Loose, *Nat. Biotechnol.*, 2018, **36**, 338–345.
- 3 I. Nir, D. Huttner and A. Meller, *Biophys. J.*, 2015, **108**, 2340–2349.
- 4 V. Van Meervelt, M. Soskine, S. Singh, G. K. Schuurman-Wolters, H. J. Wijma, B. Poolman and G. Maglia, *J. Am. Chem. Soc.*, 2017, **139**, 18640–18646.
- 5 C. Wloka, V. Van Meervelt, D. van Gelder, N. Danda, N. Jager, C. P. Williams and G. Maglia, *ACS Nano*, 2017, **11**, 4387–4394.
- 6 D. J. Niedzwiecki, C. J. Lanci, G. Shemer, P. S. Cheng, J. G. Saven and M. Drndić, *ACS Nano*, 2015, **9**, 8907–8915.
- 7 N. Varongchayakul, D. Huttner, M. W. Grinstaff and A. Meller, *Sci. Rep.*, 2018, **8**, 1017.
- 8 E. C. Yusko, J. M. Johnson, S. Majd, P. Prangkio, R. C. Rollings, J. Li, J. Yang and M. Mayer, *Nat. Nanotechnol.*, 2011, **6**, 253–260.
- 9 E. C. Yusko, B. R. Bruhn, O. M. Eggenberger, J. Houghtaling, R. C. Rollings, N. C. Walsh, S. Nandivada, M. Pindrus, A. R. Hall and D. Sept, *Nat. Nanotechnol.*, 2017, **12**, 360–367.

- 10 P. Waduge, R. Hu, P. Bandarkar, H. Yamazaki, B. Cressiot, Q. Zhao, P. C. Whitford and M. Wanunu, *ACS Nano*, 2017, **11**, 5706–5716.
- 11 J. Nivala, D. B. Marks and M. Akeson, *Nat. Biotechnol.*, 2013, **31**, 247–250.
- 12 D. Rodriguez-Larrea and H. Bayley, *Nat. Nanotechnol.*, 2013, **8**, 288–295.
- 13 C. B. Rosen, D. Rodriguez-Larrea and H. Bayley, *Nat. Biotechnol.*, 2014, **32**, 179–181.
- 14 F. Piguet, H. Ouldali, M. Pastoriza-Gallego, P. Manivet, J. Pelta and A. Oukhaled, *Nat. Commun.*, 2018, **9**, 966.
- 15 J. J. Kasianowicz, E. Brandin, D. Branton and D. W. Deamer, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 13770–13773.
- 16 T. C. Sutherland, Y. T. Long, R. I. Stefureac, I. Bediako-Amoa, H. B. Kraatz and J. S. Lee, *Nano Lett.*, 2004, **4**, 1273–1277.
- 17 B. Hornblower, A. Coombs, R. D. Whitaker, A. Kolomeisky, S. J. Picone, A. Meller and M. Akeson, *Nat. Methods*, 2007, **4**, 315–317.
- 18 L. Payet, M. Martinho, M. Pastoriza-Gallego, J.-M. Betton, L. Auvray, J. Pelta and J. Mathé, *Anal. Chem.*, 2012, **84**, 4071–4076.
- 19 P. R. Singh, I. Bárcena-Urbarri, N. Modi, U. Kleinekathöfer, R. Benz, M. Winterhalter and K. R. Mahendran, *ACS Nano*, 2012, **6**, 10699–10707.
- 20 G. Huang, K. Willems, M. Soskine, C. Wloka and G. Maglia, *Nat. Commun.*, 2017, **8**, 935.
- 21 D. S. Talaga and J. Li, *J. Am. Chem. Soc.*, 2009, **131**, 9287–9297.
- 22 J. Larkin, R. Y. Henley, M. Muthukumar, J. K. Rosenstein and M. Wanunu, *Biophys. J.*, 2014, **106**, 696–704.
- 23 A. Squires, E. Atas and A. Meller, *Sci. Rep.*, 2015, **5**, 11643.
- 24 G. Goyal, Y. B. Lee, A. Darvish, C. W. Ahn and M. J. Kim, *Nanotechnology*, 2016, **27**, 495301.
- 25 J. Feng, K. Liu, R. D. Bulushev, S. Khlybov, D. Dumcenco, A. Kis and A. Radenovic, *Nat. Nanotechnol.*, 2015, **10**, 1070–1076.
- 26 J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz and J. A. Golovchenko, *Nature*, 2001, **412**, 166–169.
- 27 A. J. Storm, J. H. Chen, X. S. Ling, H. W. Zandbergen and C. Dekker, *Nat. Mater.*, 2003, **2**, 537–540.
- 28 H. Kwok, K. Briggs and V. Tabard-Cossa, *PLoS One*, 2014, **9**, e92880.
- 29 T. Gilboa, A. Zrehen, A. Girsault and A. Meller, *Sci. Rep.*, 2018, **8**, 9765.
- 30 S. Garaj, W. Hubbard, A. Reina, J. Kong, D. Branton and J. A. Golovchenko, *Nature*, 2010, **467**, 190–193.
- 31 N. A. W. Bell and U. F. Keyser, *J. Am. Chem. Soc.*, 2015, **137**, 2035–2041.
- 32 B. N. Miles, A. P. Ivanov, K. A. Wilson, F. Doğan, D. Japrun and J. B. Edel, *Chem. Soc. Rev.*, 2013, **42**, 15–28.
- 33 L. Song, M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley and J. E. Gouaux, *Science*, 1996, **274**, 1859–1866.
- 34 M. Wanunu, J. Sutin, B. McNally, A. Chow and A. Meller, *Biophys. J.*, 2008, **95**, 4716–4725.
- 35 M. Chinappi and F. Cecconi, *J. Phys.: Condens. Matter*, 2018, **30**, 204002.
- 36 C. Plesa, S. W. Kowalczyk, R. Zinsmeister, A. Y. Grosberg, Y. Rabin and C. Dekker, *Nano Lett.*, 2013, **13**, 658–663.
- 37 D. J. Niedzwiecki, J. Grazul and L. Movileanu, *J. Am. Chem. Soc.*, 2010, **132**, 10816–10822.
- 38 K. J. Freedman, S. R. Haq, M. R. Fletcher, J. P. Foley, P. Jemth, J. B. Edel and M. J. Kim, *ACS Nano*, 2014, **8**, 12238–12249.
- 39 A. Asandei, M. Chinappi, J. Lee, C. Ho Seo, L. Mereuta, Y. Park and T. Luchian, *Sci. Rep.*, 2015, **5**, 10419.
- 40 B. Cressiot, A. Oukhaled, L. Bacri and J. Pelta, *Bionano-science*, 2014, **4**, 111–118.
- 41 W. Si and A. Aksimentiev, *ACS Nano*, 2017, **11**, 7091–7100.
- 42 L. Restrepo-Pérez, S. John, A. Aksimentiev, C. Joo and C. Dekker, *Nanoscale*, 2017, **9**, 11685–11693.
- 43 M. Wanunu, W. Morrison, Y. Rabin, A. Y. Grosberg and A. Meller, *Nat. Nanotechnol.*, 2010, **5**, 160–165.
- 44 D. J. Niedzwiecki, R. Iyer, P. N. Borer and L. Movileanu, *ACS Nano*, 2013, **7**, 3341–3350.
- 45 L. Liu, T. Li, S. Zhang, P. Song, B. Guo, Y. Zhao and H.-C. Wu, *Angew. Chem., Int. Ed.*, 2018, **57**, 11882–11887.
- 46 A. Asandei, I. Schiopu, M. Chinappi, C. H. Seo, Y. Park and T. Luchian, *ACS Appl. Mater. Interfaces*, 2016, **8**, 13166–13179.
- 47 M. Muthukumar, *J. Chem. Phys.*, 2014, **141**, 081104.
- 48 L. Mereuta, M. Roy, A. Asandei, J. K. Lee, Y. Park, I. Andricioaei and T. Luchian, *Sci. Rep.*, 2014, **4**, 3885.
- 49 T. Gilboa and A. Meller, *Analyst*, 2015, **140**, 4733–4747.
- 50 M. Di Ventra and M. Taniguchi, *Nat. Nanotechnol.*, 2016, **11**, 117–126.