**Fingerprinting of Peptides with a Large Channel of Bacteriophage Phi29 DNA Packaging Motor**

Zhouxiang Ji, Shaoying Wang, Zhengyi Zhao, Zhi Zhou, Farzin Haque, and Peixuan Guo*

**Introduction**

Transport of peptides and proteins across inter- or intracellular compartments through biological pores is an ubiquitous process in cells, such as protein transport into endoplasmic reticulum, mitochondria, chloroplast, and nucleus.[1,2] A wide assortment of secreted peptides are involved in the onset of several diseases, such as neurodegenerative disorders,[3,4] cancer[5] and viral infections.[6] Peptides can serve as valuable biomarkers due to the different expression levels in disease-associated conditions.[7] Several approaches have been developed over the years for characterizing peptides and proteins, including ELISA,[8] mass spectrometry,[9] fluorescence spectroscopy,[10] real-time PCR using chimeric protein covalently linked to ssDNA[11] and immuno-staining.[12] However, limitations include ensemble averaging, low sensitivity to detect rare subpopulations, labor intensive sample preparations, requirement of high concentrations, and relatively low throughput nature. Nanopore has emerged as a powerful alternative method due to its intrinsic capability of single molecule detection at ultralow concentrations.[13] The approach requires no labeling, no amplification, low sample amounts, and can be translated into a high-throughput device for detection of multiple analytes. The detection principle is based on the resistive pulse technique, whereby charged analytes are driven through nanoscale pores under an applied potential and the analytes of interest are detected based on their characteristic current signature profiles. Several works also show that molecules do not have to be charged (e.g., cyclodextran or PEG (polyethylene glycol)) and can be driven by electroosmotic forces into the pore.[14]
While the focus of both solid-state and biological nanopores has been on characterizing nucleic acids,[13,15–20] several studies have shown the potentials for peptide sensing (reviewed in refs. [21–24]). For instance, synthetic nanopores for studying folding and unfolding of proteins,[6,25–32] detecting protein analytes[33–35] protein binding to nucleic acids,[36–38] and sizing of proteins and peptides.[39–43] Similarly, several biological pores, such as α-hemolysin, ClyA, and aerolysin have been employed for studying peptide fibrillation,[44–46] unraveling protein and peptide folding,[47–50] studying kinetics of protein binding,[51] as well as using aptamers for peptide sensing.[52–54] Solid-state pores generally can be fabricated with various pore sizes and geometry, but generating pores with reproducible thickness and dimensions with sub-nanometer precision is challenging.[55] Biological pores on the other hand have reproducible pore sizes at the atomic level and are amenable to site-directed mutagenesis for introducing functional modules for detection. However, most commonly used protein nanopores, such as α-hemolysin, and MspA have an internal channel diameter of around 1.3 nm,[60,61] and thus not suitable for structural analysis of typical peptides in functional forms except for protein channels equipped with aptamers.[52,58] Larger protein channels such as ClyA with an internal diameter of 3.3 nm have recently been utilized for sensing peptides and for studying protein folding.[59,62] However, ClyA nanopores are often a mixture of three different oligomeric states and specific channels need to be isolated for downstream applications. Bacteriophage phi29 motor channel is an attractive alternative since it has a larger pore diameter that can potentially be used for peptide fingerprinting and more importantly, display homogenous pore sizes. The portal channel, termed connector self-assembles into a truncated cone-shaped structure from 12 subunits (Figure 1a). The exterior of the channel is 13.8 nm at the wider end and 7.8 nm at the narrower end. The interior of the channel is 6 nm in diameter at the wider end and 3.6 nm in diameter at the narrower constriction. The channel has previously been inserted into a planar lipid membrane[63] and the resulting nanopore exhibits well-behaved conductance[64] under a wide range of experimental conditions for sensing DNA,[63–68] RNA,[69] chemicals,[70] and antibodies.[71,72] The crystal structure of the channel is available[73,74] and therefore mutagenesis can be made with relative ease to tailor the pore size and selectively introduce probes at desired location for added functionalities. In addition, large-scale production and purification of the channel has been well-established.[75–77]

Herein we investigated the feasibility of single molecule fingerprinting of different peptides (TAT(transactivator of transcription), MAR, and EpCAM (epithelial cell adhesion molecule)) using the membrane-embedded phi29 channel. Parameters derived from single channel conduction assays such as, current blockage and characteristic signatures were used to characterize peptides with high sensitivity. The translocation of peptides was confirmed by single molecule fluorescence imaging. The results demonstrate the potentials of phi29 nanopore for fingerprinting peptide biomarkers for early disease diagnosis, structural characterizations of peptides including charge, hydrophobicity, and chirality, kinetic and thermodynamic analysis of peptide folding, and eventually peptide sequencing.

2. Results
2.1. Characterization of Membrane-Embedded phi29 Channels

The C-terminal His-tagged channel was expressed and purified to homogeneity[63,75] and verified by 10% SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) (Figure 1b). The purified channel was reconstituted into liposomes during dehydration/rehydration step and the resulting proteoliposomes were fused with a planar lipid bilayer to generate membrane-embedded nanopores.[63,75] The insertion of each channel resulted in discrete step size increase in current (Figure 1c). The conductance (current step of one channel insertion/applied voltage) based on multiple insertion events was determined to be 4.22 ± 0.56 nS in presence of 1 m KCl, 5 × 10⁻³ m HEPES, pH 7.9 (Figure 1d). The current trace was quiescent without displaying any gating phenomena in the voltage range of ±50 mV.

![Image](https://www.small-journal.com/DOI:10.1002/smll.201601157)
2.2. Characterization of Peptides Based on Current Blockage

We investigated translocation properties of three different peptides as model systems: TAT (CYG RKK RRO RRR; $M_w$ 1.66 kDa; pI 12.0), a cell penetrating peptide; MAR (MAR KRS NTY RSI NEC; $M_w$ 1.83 kDa; pI 10.0), a fragment of gp10 connector at the N-terminal end; and EpCAM (ELK HKA REK PYD SKS LRT HHH HHH; $M_w$ 3.01 kDa; pI 9.8), a small polypeptide fragment that binds to EpCAM antibody.

In presence of peptides in the conducting buffer, a burst of transient current blockage events were observed (Figure 2) under an applied potential of –50 mV. One parameter used to characterize these transient events is the current blockage percentage, defined as \[ \left( I_p - I_0 \right) / I_0 \times 100\% \], where $I_0$ is the current of open channel and $I_p$ is the current level after peptide blockage. The current blockage percentage was deduced by fitting the event distribution with a Gaussian function. Since the conductance buffer has a pH of 7.9, all three peptides are positively charged. Translocation of TAT peptide showed two current blockages [31.02% ± 0.96% and 51.63% ± 4.39% (Figure 2a)]. MAR peptide also showed two current blockages [30.58% ± 4.27% and 44.34% ± 7.43%] (Figure 2b). EpCAM was characterized by the channel with a single current blockage peak at 31.33% ± 4.30% (Figure 2c). Finally, unfolded TAT peptide displayed a single peak at 31.72% ± 0.70% (Figure 2d). In our analysis, we excluded events with very short residence times or low amplitude spikes, which are most likely collision events or nonspecific noise.

By comparing the current blockage profiles of these four analytes, it is evident that all these four profiles display a common peak with a smaller blockage of ≈31% (Figure 2; gray box). This peak likely represents general unfolded peptides that are about 2 nm in diameter. This is supported by the profile of unfolded TAT peptide (Figure 2d) compared to folded peptides (Figure 2a). The blockage of 32% has been reported previously to represent the translocation of dsDNA,\(^{[63-68]}\) which also displays a diameter of 2 nm. From Figure 2, it is also evident that each of the four blockage profiles show a unique blockage peak with large but different blockage percentages that is representative of the characteristic fingerprint or signature of each peptide.

![Figure 2](link-to-image)
The electrophoretic transport time, \( \tau_{\text{ele}} \), can be estimated using the relation\(^{83}\),

\[
\tau_{\text{ele}} = \frac{\pi \eta r}{6 \pi \eta r L} \frac{Q}{V} = \frac{6 \pi \eta r}{(QL^2/V)},
\]

where \( L \) is the pore length, \( r \) is the electrophoretic velocity, \( \eta \) is the viscosity of the solution, \( r \) is the peptide radius, \( Q \) is the protein net charge, \( E \) is the electric field strength, and \( V \) is the applied electric voltage. The \( \tau_{\text{ele}} \) is calculated to be in the order of a few microseconds, while our experimentally determined \( \tau \) is on the order of at least a hundred microseconds. Our results are consistent with several studies in the literature showing that protein net charge, \( Q \), the peptide radius, \( r \), and current blockage distribution showing TAT peptide characterization profile before b) and after c) addition of reducing agent TCEP. Applied voltage: –50 mV. Buffer: 1 M KCl, 5 \times 10^{-3} M HEPES, pH 7.9.

\[ A \cdot \frac{dQ}{dt} = \frac{4}{3} \pi r^3 \frac{dL}{dQ} \]

\[ A \cdot \frac{dL}{dt} = \frac{4}{3} \pi r^3 \frac{dQ}{dL} \]

2.3. Characterization of Peptides Based on Dwell Time Distributions

Another parameter used to characterize the translocation events is the dwell time (\( \tau \)), defined as the time taken for the peptide to pass through the channel. The dwell time distribution typically follows an exponential decay and the lifetime data was obtained by fitting the duration events with an exponential function. Folded TAT peptide had a dwell time of 0.89 ± 0.04 ms (Figure 2a, right panel). MAR peptide had a faster average dwell time of 0.15 ± 0.01 ms (Figure 2b, right panel). EpCAM passed through the channel with a dwell time of 0.11 ± 0.01 ms (Figure 2c, right panel). Finally, unfolded TAT had a dwell time of 0.35 ± 0.03 ms (Figure 2d, right panel). Our data showed that folded peptides, such as TAT indeed stayed in the channel longer than unfolded TAT under the same electric potentials. This highlights the fact that interactions between the peptide and the nanopore wall influence the dwell time due to several factors such as the surface charge distribution and shape of the folded and unfolded states of the peptide.

The electrophoretic transport time, \( \tau_{\text{ele}} \), can be estimated using the relation\(^{83}\),

\[
\tau_{\text{ele}} = \frac{\pi \eta r}{6 \pi \eta r L} \frac{Q}{V} = \frac{6 \pi \eta r}{(QL^2/V)},
\]

where \( L \) is the pore length, \( r \) is the electrophoretic velocity, \( \eta \) is the viscosity of the solution, \( r \) is the peptide radius, \( Q \) is the protein net charge, \( E \) is the electric field strength, and \( V \) is the applied electric voltage. The \( \tau_{\text{ele}} \) is calculated to be in the order of a few microseconds, while our experimentally determined \( \tau \) is on the order of at least a hundred microseconds. Our results are consistent with several studies in the literature showing that event dwell times for peptides and proteins are many orders of magnitude faster than the theoretical electrophoretic transport times\(^{83}\). The anomalous dynamics of protein transport can be explained by several factors including adsorption of peptides to nanopore inner walls due to attractive interactions; electro-osmotic effects, which can significantly slow down the passage of peptides; and finally a more subtle case of heterogeneous charge distribution along the peptide, which can generate a disordered energy landscape as the peptide traverses through the pore.

2.4. Characterization of Oligomerization States of Peptide in Real Time

The current blockage percentage depends on the excluded volume and charge of the peptides.\(^{14}\) Different current amplitudes indicate different conformations or different oligomerization states adopted by the peptide. To delineate the oligomerization states, we investigated TAT peptide. Structural analysis of TAT peptide sequence (CYG RKK RRQ RRR) using PEP-FOLD\(^{84}\) revealed that folded TAT can exist as a dimer mediated by disulfide bond formation between cysteines of two peptides. To confirm whether the two current blockages observed for TAT [31.02% ± 0.96% and 51.63% ± 4.39%] correspond to monomer and dimer forms respectively, we added reducing agent TCEP [[tris(2-carboxyethyl)phosphine]] into the chamber, mixed it and analyzed the current blockage events (Figure 3a). Over the course of the reaction, the predominantly 51.63% ± 4.39% dimer peak (Figure 3b) transformed to 31.02% ± 0.96% monomer peak (Figure 3c), thus confirming the reduction of the disulfide bond.

2.5. Verification of Peptide Translocation Using Single Molecule Fluorescence Imaging

To prove that the observed transient current blockage events are indeed peptide translocation rather than peptide binding to the channel or collision events, we used fluorescent peptides as analytes. We labeled TAT peptide with a Cy3 fluorophore at the terminal cysteine residue using maleimide-sulfhydryl coupling and purified the labeled peptide by HPLC. After insertion of multiple channels in the membrane, Cy3-TAT peptide (1.0 μg mL\(^{-1}\)) was added to the cis-chamber. We note that the conjugation of Cy3 prevented the dimer formation and yielded similar blockage percentage as observed with TCEP reduction assay. At specific time points (0, 20, 40, and 60 min), 50 μL aliquots were extracted from the trans-chamber and deposited on glass substrate for single molecule fluorescence imaging. The number of fluorescence spots increased over time indicating accumulation of Cy3-TAT in the trans-chamber and therefore translocation through the channels (Figure 4a, b). The control experiment without the presence of...
The translocation of peptides through nanopores is a complicated process at the microscopic level. Initially, the peptide diffuses and electrophoretically drifts toward the vicinity of the pore. The electric potential needs to be strong enough to overcome the energetic and entropic costs of capturing the peptide inside the nanopore. Peptide-pore interactions and electrosomotic flow play critical roles in modulating these free-energy landscapes. Herein, we observed that peptides characterized by the channel with different capture rates. These differences can possibly be attributed to the charge density and the conformational entropy of the peptides. Computational studies are necessary to unravel the mechanistic of peptide translocation and experimental methods need to be developed to slow down the peptide translocation process to enable in-depth characterization with high spatial and temporal resolution.

Many studies have indicated peptide as valuable biomarkers for early disease diagnosis and for monitoring responses to therapy. Detecting peptide conformational changes at single molecule level can be valorous insight into the structural transitions that occur during pathogenesis of certain diseases, such as protein misfolding diseases. The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases. The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases.

The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases.

The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases.

The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases.

The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases.

The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases.

The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases.

The challenge is always to detect the peptides at ultralow concentrations. Conventional spectroscopy and immuno-detection methods have been the benchmark in the proteomics arena, but lack single molecule sensitivity and specificity and often require high amount of samples. Nanopores offer an attractive alternative for proteomic profiling. To date, the majority of biological nanopore studies have focused on small channels like α-hemolysin, aerolysin, OmpF, and MspA, which only allow ssDNA, small peptides, and some unfolded proteins to pass through. Their limited pore aperture restricts analysis of certain diseases, such as protein misfolding diseases.
purified by His-tag column and eluted using His-tag elution buffer. Fractions were collected and analyzed by 10% SDS-PAGE. One modification from the previous procedure[63,75–77] was the removal of the unassembled connector protein by sucrose gradient sedimentation. Around 0.2 mL fractions of protein were loaded onto 5%–20% sucrose gradient and then centrifuged at 35 000 rpm for 3 h at 4 °C. Fractions were collected and analyzed by 10% SDS-PAGE. Fractions with faster migration rate corresponding to the assembled connector were collected for further experiment.

**Insertion of Channel into Liposomes:** As described previously,[63,75] 1 mL DPhPC lipids (1 mg mL⁻¹) in chloroform was placed in a round bottomed flask and the chloroform was evaporated under vacuum using a rotary evaporator (Buchi). Rehydration buffer containing 250 × 10⁻³ mL sucrose was added along with phi29 channel (0.5 mg mL⁻¹) and thoroughly vortexed. The mixture was then extruded through 400 nm polycarbonate membrane filters (Avanti polar lipids) at 45 °C (above the phase transition temperature of DPhPC) for 21 times to generate homogenous-sized proteolioposomes. The extruded proteolioposomes were then fused with planar lipid bilayer to generate membrane-embedded channel.

**Electrophysiological Measurements:** Bilayer lipid membranes (BLM) were formed by painting technique (using DPhPC lipids in n-hexane and n-decan) on the Teflon partition (housing 200 μm aperture) which was used to separate the BLM chamber into cis- and trans-compartments, as described previously.[63,75] A pair of Ag/AgCl electrodes were placed in the cis- and trans-chambers. The Axopatch 200B patch clamp amplifier was connected to the Axon DigiData 1440A analog-digital converter. All voltages reported were those of the trans-compartment. Data was low band-pass filtered at a frequency of 1–10 KHz and acquired at a sampling frequency of 20–200 KHz. The PClamp 9.1 software (Axon Instruments) was used to collect the data, and Clampfit (Axon Instruments) and Origin Pro 8.0 (Origin Lab corporation) software was used for data analysis.

**Peptide Translocation Experiments:** TAT, EpCAM, or MAR peptide with suitable concentration was premixed in the conducting buffer. The signals were recorded using Andor iXon 887 V electron multiplied charge coupled devive (CCD) camera. Images were taken with an exposure time of 500 ms. The number of spots in the images were counted using ISMS software.[91]

**Acknowledgements**

*The research was supported by NIH grant RO1EB012135. Peixuan Guo’s Sylvan G. Frank Endowed Chair position in Pharmaceutics and Drug Delivery is funded by the CM Chen Foundation. P.G. is a co-founder of Biomotor and RNA Nanotechnology Development Corp. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH.*
