2016

# Development of New Biological Nanopores and Their Application for Biosensing and Disease Detection

Shaoying Wang
*University of Kentukcy*, umkcwang@gmail.com
Digital Object Identifier: https://doi.org/10.13023/ETD.2016.428

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Shaoying Wang, Student

Dr. Peixuan Guo, Major Professor

Dr. David Feola, Director of Graduate Studies

# DEVELOPMENT OF NEW BIOLOGICAL NANOPORES AND THEIR APPLICATION FOR BIOSENSING AND DISEASE DETECTION

---

DISSERTATION

---

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Pharmacy at the University of Kentucky

By

Shaoying Wang

Lexington, Kentucky

Co-Directors: Dr. Robert Lodder, Professor of Pharmaceutical Sciences, Lexington, Kentucky

and Dr. Peixuan Guo, Professor of Pharmaceutics & Pharmaceutical Chemistry,

Columbus, Ohio

2016

ABSTRACT OF DISSERTATION

# DEVELOPMENT OF NEW BIOLOGICAL NANOPORES AND THEIR APPLICATION FOR BIOSENSING AND DISEASE DETECTION

Nanopore technology has recently emerged as a new real-time single molecule sensing method. The current dominant technologies, such as mass spectrometry and immunoassay, for protein analysis is still slow and complex, which can't meet the urgent need and fields of use. Development of a highly simple, portable and sensitive detection system for pathogen detection, disease diagnosis, and environmental monitoring is in great need. Membrane embedded Phi29 connector nanopore, the first protein nanopore coming from bacteriophage, was mainly focusing on DNA and RNA translocation in previous studies. Here, Phi29 connector nanopore was first time established for antibody detection by engineering Epithelial Cell Adhesion Molecule peptide as a probe. The results demonstrate that the specific antibody can be detected in presence of diluted serum or non-specific antibody. To enable detecting more different types of analytes with high sensitivity, developing new nanopore with various properties, such as size, charge, hydrophilic/hydrophobic and physical dimension, is needed. In this work, besides Phi29 nanopore, several new protein nanopores that derived from T3, T4, and SPP1 bacteriophages were developed. A shared property of three step conformational change among these portal channel has been discovered. Elucidating the sequence and oligomeric states of proteins and peptides is critical for understanding their biological functions. Here, SPP1 nanopore was used to characterize the translocation of TAT peptide with dimer and monomer forms. Translocation of the peptide was confirmed by optical single molecule imaging for the first time, and analyzed quantitatively. The dynamics of peptide oligomeric states were clearly differentiated based on their characteristic electronic signatures. Main challenge for probing protein structure, folding, detection and sequencing using nanopore is the ultra-fast translocation speed which normally beyond electronic detection limit. In this work, the peptides translocation was slowed in SPP1 nanopore by changing the charge shielding of the channel. A 500-fold reduction was observed for TAT peptide translocation. By using this method, arginine chain peptide as short as two arginine can be detected first time. Further improving the bandwidth may lead to single amino acid detection and has the

potential for protein sequencing. Compared with protein nanopore, *de novo* designed nanopore can provide numerous advantages, such as tunable size and functionality, ease of construction, scale up and modification. In the final study, an RNA-based biomimetic nanopore was first time constructed. The insertion of RNA nanopore into lipid bilayer and cell membrane were characterized and translocation of short amino acids through RNA nanopore was detected. This new artificial nanopore has the potential to be used for sensing, disease diagnosis, and even protein sequencing.

KEYWORDS: Biological nanopore, peptide translocation, bacteriophage, RNA Nanotechnology, protein sequencing

Shaoying Wang
Student's Signature
October 4th, 2016
Date

# DEVELOPMENT OF NEW BIOLOGICAL NANOPORES AND THEIR APPLICATION FOR BIOSENSING AND DISEASE DETECTION

By

Shaoying Wang

Dr. Peixuan Guo
Co-Director of Dissertation

Dr. Robert Lodder
Co-Director of Dissertation

Dr. David.Feola
Director of Graduate Studies

October 4th, 2016

To my wife, parents and family for their continued support

ACKNOWLEDGMENTS

Firstly, I would like to express my deepest gratitude and best regards to my thesis advisor, Dr. Peixuan Guo, for his continual guidance and support during my PhD studies. I sincerely appreciate Dr. Guo's invaluable advice, experience, and guidance.

Also, I would greatly appreciate my committee members, Dr. Robert Lodder, Dr. David Rodgers, Dr. Paul Bummer, Dr.Chang-Guo Zhan, and Dr. Jianhang Jia. Their instruction through my graduate studies proved indispensable in all aspects of my graduate work. Their advices and instructions through my research, qualifying exam, and dissertation defense are indispensable to my study.

I would like to give special thanks to Dr.Dan Shu, Dr. Farzin Haque, Dr.Hui Zhang, Dr. Jia Geng for their training and valuable insight in my experiments and data throughout my research. In addition, I would like to specially thanks to Hui Li and Fengmei Pi for their encouragement, and support through my PhD studies. I would have not made it through this endeavor without their assistance and kindness. I also would like to thank all past and present members of the Guo laboratory; Dr.Yi Shu, Dr.Zhi Zhou, Dr. Mario Vieweger, Dr. Ashwani Sharma, Dr. Taek Lee, Dr. Zhanxi Hao, Dr. Chad Schwartz, Dr. Huaming Fang, Fengmei Pi, Hui Li, Zhengyi Zhao, Danny Jasinski, Erfu Yan, Yanxi Xie, Zhouxiang Ji, Concong Xu, Hongran Yin, Sijin Guo, Zhefeng Li, Le Zhang, Hongzhi Wang, and Megan Heitkemper. Without their support I truly believe my success would not have happened.

I would like to acknowledge all other faculty and staff members from College of Pharmacy at University of Kentucky throughout my graduate career, especially graduate coordinator Catina Rossol and current director of graduate studies Dr. David Feola and previous director of graduate studies Dr. Jim Pauly.

Finally, I would like to show my appreciation to my wife Dr.Rongrong Zhang, my parents for their continued encouragement, love and support through my everyday struggles.

The work in the thesis dissertation was supported by the National Institute of Health under grants R01 EB012135 to Peixuan Guo.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| EpCAM | Epithelial Cell Adhesion Molecule |
| antibody | Ab |
| MspA | Mycobacterium smegmatis porin |
| OmpG | Outer membrane protein G |
| FhuA | Ferric hydroxamate uptake component A |
| ClyA | Cytolysin A |
| PKIP | protein kinase inhibitor peptide |
| pRNA | packaging RNA |
| DPhPC | phospholipid 1,2-diphytanoyl-sn-glycerol-3-phosphocholine |
| DNA | Deoxy Nucleic Acid |
| dsDNA | Double Stranded Deoxy Ribonucleic Acid |
| RNA | Ribonucleic Acid |
| ΔG° | Change in Gibbs Free Energy |

# Chapter 1: Introduction and Literature Review

**BRIEF SUMMARY:**

Chapter 1 begins this thesis with an overview of the principle of nanopore and current state of this field. First, the structure and properties of the most commonly used biological nanopore are summarized. Then, recent progresses in biological nanopores for DNA sequencing and single molecule sensing has been reviewed from a protein engineering standpoint. Several milestone studies illustrating the logic behind the engineering approaches have also been discussed.

Chapter 2 is a proof of concept study to demonstrate the feasibility of phi29 connector nanopore for disease diagnosis. In the study, an Epithelial Cell Adhesion Molecule (EpCAM) peptide was engineered into the C-terminal of nanopore as a probe to specifically detect EpCAM antibody (Ab) in nano-molar concentration at the single molecule level. The binding of Abs sequentially to each peptide probe induced step-wise blocks in current. The signal of EpCAM antibody can be discriminated from the background events in the presence of non-specific antibody or serum. The results demonstrate the feasibility of generating a highly sensitive platform for detecting antibodies at low concentration in the presence of contaminants.

Chapter 3 explores more biological nanopore from other bacteriophages and discovers a three step gating property that is common to DNA packaging motors of bacterial viruses. The DNA packaging motor of dsDNA bacterial viruses contains a head-tail connector with a channel for genome to enter during assembly and to exit during host infection. In this study, a three step conformational change of the portal channel that is common among DNA translocation motors of bacterial viruses T3, T4, SPP1, and phi29.

The channels of these motors exercise three discrete steps of gating, as revealed by electrophysiological assays. It is proposed that the three step channel conformational changes occur during DNA entry process, resulting in a structural transition in preparation of DNA movement in the reverse direction during ejection.

Chapter 4 expands the applications of portal protein to study the translocation of peptides and the dynamics of peptide oligomeric states in real-time at single molecule level. In this study, to acquire protein channels with larger pore sizes, we engineered a portal channel from bacteriophage SPP1 that allowed the translocation of peptides with higher ordered structures. Utilizing single channel conduction assay and optical single molecule imaging, we observed translocation of peptides quantitatively and analyzed the dynamics of peptide oligomeric states in real-time at single molecule level. The oxidative and the reduced states of peptides were clearly differentiated based on their characteristic electronic signatures. A similar Gibbs free energy ($\Delta G^0$) was obtained when different concentrations of substrates were applied, suggesting that the use of SPP1 nanopore for real-time quantification of peptide oligomeric states is feasible. With the intrinsic nature of size and conjugation amenability, the SPP1 nanopore has the potential for development into a tool for the quantification of peptide and protein structures in real time.

Chapter 5 optimizes the condition for short peptide detection based on the work from the forth chapters. Using nanpore for probing protein structure, folding, detection and sequencing is in early stage because it is still lack of an effective and simple method to slow down the ultra-fast translocation speed of peptide. In this work, the peptides translocation was slowed down in SPP1 DNA packaging motor with low NaCl concentration buffer. The underlying mechanism is that reduced charge screening effect

resulted from low salt concentration enhanced electrostatic attractive force between the peptide and pore, which therefore slow down peptide translocation speed. A factor of 500 fold reduction was observed for TAT translocation. Appling such condition in detecting arginine chain peptide allows us to measure as short peptide as two arginine amino acids. Further improve the bandwidth may lead to single amino acid detection and protein sequencing.

Chapter 6 constructs a new biomimetic RNA-based nanopore and applies for peptide detection. Creating a new pore *de novo* with the ability to accurately tune the size and its functionality is still of great interest in fundamental and applied science. One of the key challenges in the *de novo* design of membrane channel is, to achieve a defined architecture in atomic scale with defined size, shape and properties. Over the last decade, RNA nanotechnology has advanced rapidly and attracted widespread attention due to the potential for applications in material and medical sciences. Compared to DNA, RNA has higher diversity in structure and function. Due to their diversity in structures and functions, RNA molecules were used as building blocks for engineering various *de novo* architectures with unique physical/biochemical properties *via* bottom-up self-assembly. To expand biological nanopores and explore new features of artificial nanopores, in this study we constructed and assembled RNA nanopores with two different designs. The insertion of RNA nanopore into lipid bilayer and cell membrane were further characterized. Translocation of short amino acids through RNA nanopore was detected and characterized. Translocation of short amino acids is first time observed in artificial nanopore. This new artificial nanopore has the potential to be used for sensing, disease diagnosis and even protein sequencing.

Chapter 7 briefly summarizes the major findings and advancements discussed in this thesis dissertation. Furthermore, the future direction of this work and perspectives of nanopore for sensing and sequencing is also discussed.

**HYPOTHESIS:**

Biological nanopore particularly portal protein channels can be used for biosensing, disease diagnosis and even protein sequencing.

**INTRODUCTION:**

The intriguing design and elegant architecture of biological machines have inspired the development of nanopores for versatile applications in nanotechnology. The proof-of-concept was established in 1996 using α-hemolysin pores, which can spontaneously insert into a planar lipid bilayer to generate a membrane-embedded nanopore system(1). In the electrochemical chamber, the lipid membrane separated the chamber into *cis*- and *trans* compartments filled with conducting buffer. Under an applied potential, single-stranded DNA, which is negatively charged can be electrically driven through the α-hemolysin channels to the positive terminal and the resulting translocation process was characterized using several parameters and verified by qRT-PCR(1). Since the landmark paper, in the last two decade, nanopore-based technology has demonstrated significant potentials for sensing a wide range of analytes and DNA sequencing, and is well positioned to bring a revolution to medical diagnostics and DNA sequencing arena (2). Compared with other technologies, nanopore offers numerous advantages (3-6): (1) It is a label-free and amplification-free approach; (2) It offers real time recognition at single molecule level; (3) It requires very low sample volumes with minimal sample processing; (4) Electrical

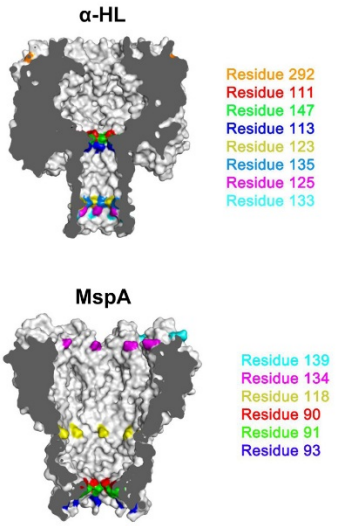
detection method is fast and can be carried out in parallel with multiple samples; and (5) Can detect target analyte with high specificity and sensitivity in the presence of a large amount of contaminants.

To date, several protein channels such as α-hemolysin (7), MspA (8), Aerolysin (9), FluA (10), Omp F or G (11,12), and viral connectors Phi29 (13), SPP1 (14), T3(14) and T4(14) have been successfully used for incredibly versatile applications including detection of small molecules, macromolecule, polymers, polypeptides, as well as DNA and RNA (15-17). Protein engineering techniques, such as site directed mutagenesis, insertion, and deletion of amino acids, and introduction of functional modules have been extensively employed to tune the nanopore properties for target applications(18,19). Herein, we will summarize the recent progresses in engineering biological nanopores for sensing and sequencing applications and provide perspectives for future studies. We note that due to page limitations, we omitted the discussion on DNA-based biological nanopores (20,21) and solid state nanopore (22) counterparts, which have been reviewed in detail elsewhere (23-25).

**1. General strategies used in protein nanopore engineering**

Although biological systems contain a wide variety of channel forming proteins (18,26), only a handful have been deemed to be suitable for nanopore studies. A major limitation is the availability of the crystal structure. Without knowing the structure, it is nearly impossible to engineer the channel precisely and rationally by genetic or chemical modifications. Another challenge is our limited knowledge of how the structure of the protein channel relates to its function. Generally, to be a good protein nanopore candidate, wild type or modified protein channel needs to meet at least three criteria. First, it should

be very robust and retain channel forming capability after mutagenesis. Second, it should have appropriate geometry for either biosensing or sequencing applications. The physical dimensions and the geometry of the nanopore including the entrance, the vestibule, and the narrowest constriction, as well as biophysical properties of the pore such as the charge distribution and hydrophilic/hydrophobic residues can significantly affect the functioning of the nanopore. Third, it should be electrically quiescent, which means that in presence of conducting buffer and applied electrical potential, the current signal should be stable and not display any voltage gating phenomena. So far, several different engineering approaches have been employed either singly or in combination to change the intrinsic properties of biological nanopores such as for increasing the stability, changing the oligomeric state, introducing recognition sites and enhancing sensitivity. The first approach is based on site-directed mutagenesis that substitutes specific amino acids at certain locations in the primary sequence and this method is widely used due to its inherent simplicity and widespread adaptability. The availability of a wide range of unnatural amino acids with unique side chain and functional properties have vastly expanded nanopore reengineering options. Common mutation sites of a-hemolysin and MspA and their effects are summarized (**Fig1.1**). The second approach is to attach specific ligands either through protein expression or chemical conjugation to the different sites of the pore. The third approach is to introduce adapter ligands either covalently or non-covalently into the pore. The fourth approach is a bit more aggressive and involves truncation or insertion of several amino acids up to a chain segment.




| Mutation | Effect | Ref |
|---|---|---|
| E111N<br>K147N | Break the electrostatic interactions between Glu-11 and Lys-147, better discrimination polyA and polyC | 28 |
| M113Y M113W | The recognition capability was enhanced | 62 |
| E111N<br>K147N<br>M113Y | Epigenetic DNA modification, 5-methylcytosine and 5-hydroxymethylcytosine, could be identified | 63 |
| M113C | Covalently attached cyclodextrin | 109 |
| M113 F/W/Y | TNT can only be detected | 91 |
| 123/125<br>133/135 → His | Detect the concentrations of two or more divalent metal ions simultaneously | 90 |

| Mutation | Effect | Ref |
|---|---|---|
| D90N/D91N<br>D93N/D11R<br>D134R/E139K | Displayed minimal spontaneous blockades for voltages above 180mV and optimized ssDNA translocation signature | 8 |
| L88R/T83R<br>S116R/A96R | slow DNA translocation by a fact 10-30 without eliminating nucleotide specific current | 80 |
| D90N/D91N<br>D93N. | spontaneous blockades was considerably reduced. Able to detect ssDNA translocation | 8 |

**Figure 1.1 Common mutation sites of a-hemolysin and MspA.** Left: Structure of a-hemolysin and MspA labeled with common mutation sites. Right: Mutations sites and the effect after mutation.

## 2. Structural basis of commonly used protein nanopores

### *2.1 α-hemolysin (PDB: 3ANZ)*

α-hemolysin is a self-assembled heptamer with a molecular weight of 232 kDa (**Fig.1.1**). It was discovered from *Staphylococcus aureus* and has a hollow mushroom-like shape. The external dimensions of the pore is 10×10 nm (27). It is composed of "stem", "rim" and "cap" domains. The 14-stranded trans-membrane β-barrel is in the stem part, assembled from seven hairpins. As revealed in the crystal structure (7), the inner diameter of the central channel is 2.6 nm at the *cis* entrance and 2.2 nm at the *trans* side. The vestibule is 3.6 nm at its widest part and the narrowest constriction is 1.4 nm in diameter. Several sites within the β-barrel has exhibited higher nucleotide-specific sensitivity with immobilized ssDNA (28). α-hemolysin can withstand temperatures close to 100°C and a wide range of pH from 2 to 12 while maintaining structural integrity (29). The conductance of α-hemolysin under 1M KCl is 1 nS per pore (30). Membrane insertion occurred spontaneously when monomeric α-hemolysin protein was added to the chamber.

### *2.2 Mycobacterium smegmatis porin (MspA) (PDB: 1UUN)*

MspA is an octameric porin with a molecular weight of 160 kDa (31)(**Fig 1.2**). It was discovered from *Mycobacterium smegmatis*, and has a funnel shaped structure (8). The external dimension of the pore are 8.8×9.6 nm. As revealed in the crystal structure (32), the inner diameter of the central channel is 4.8 nm at the *cis* entrance and it has a ~1.2 nm long and ~0.6 nm wide constriction site that is flanked by regions with significantly large space (32). It has been proved that such physical geometry provides improved spatial resolution for DNA sequencing. As each nucleotide is threaded through the narrow part of the pore, it produces a unique fingerprint of ionic current blockades (33). This pore is very

robust and maintains channel forming activity at different harsh conditions, including incubation at 80°C in presence of 2% SDS, pH 0–14; and even 100°C for 30 min after extraction (8,34). The wild-type MspA suffered from spontaneous gating without ssDNA under an applied voltage of >60 mV. The conductance of wildtype MspA under 1M KCl is 4.9 nS per pore (8) and the peak conductance of M1 MsPA mutant (D90N/D91N/D93N) is only 1.6 nS (35). Membrane insertion occurred spontaneously when monomeric MspA protein was added to the chamber.

### *2.3 Aerolysin*

Aerolysin is a heptamer pore with a molecular weight of 332 kDa (**Fig.1.2**). It was discovered from the bacterium *Aeromonas hydrophila* (36,37). It is an exotoxin secreted as a single polypeptide chain (38). While a crystal structure is not yet available, a high resolution electron microscopy study revealed that aerolysin is 1–1.7 nm diameter at the narrowest constriction (9,39). Overall, aerolysin is negatively charged (Z= –52e), while α-hemolysin has slightly positive net charge (Z=+7e) (40). Unlike α-hemolysin, the vestibule domain is not present in aerolysin. Thus aerolysin has a greater effective length, which enabled the wild-type pore to resolve individual short oligonucleotides that are 2 to 10 bases long without any modifications (39). Conductance studies revealed that the pore is slightly larger under positive applied voltage than under negative voltage, highlighting a geometric asymmetry because of a large extra-membranous vestibule domain (40). The channel is resistant to strong denaturing conditions, such as 1M Guanidinium-HCL (40). The conductance of Aerolysin under 1M KCl is 0.39 nS per pore (40). Membrane insertion occurred spontaneously when monomeric aerolysin was added to the chamber.

### ***2.4 CsgG*** *(PDB: 4UV3)*

CsgG exists as octameric or nonameric pore with a molecular weight of 244 kDa (**Fig.1.2**). It was discovered from outer membrane of *Escherichia coli* (41). As revealed in the crystal structure (41), the external dimension of the pore are 8.5×12 nm. The inner diameter of the pore is 4 nm and 3.5 at the *cis* and *trans* entrance. A 36-stranded β-barrel traverses the bilayer with a height of 3.5 nm and is linked to a cage-like vestibule in the periplasm with a height of 4 nm. The transmembrane and periplasmic domain are linked by a 0.9 nm channel constriction that formed by the side chains of residues Tyr51, Asn 55 and Phe56. Functional and electrophysiological study showed that CsgG is a non-selective and ungated protein secretion channel. The conductance of CsgG under 1M KCl is 0.88 nS per pore (41).Because CsgG is an outer membrane channel, membrane insertion occurred spontaneously by adding CsgG to the *cis* chamber. In March 2016, Oxford Nanopore announced the identity of the R9 pore which is based on a CsgG pore and came from screening of 700 mutants. The company indicated the new R9 pore will be available for MinION and PromethION systems soon.

***2.5 Outer membrane protein G*** *(OmpG) (PDB: open 2IWV and closed 2IWW)*

OmpG is a monomeric pore with a molecular weight of 35 kDa (**Fig.1.2**). It was discovered from *Escherichia coli*. The diameter of extracellular opening is 0.8 nm and the periplasmic side is 1.4 nm (42). The pore consists of 14 β-strands linked by 7 long loops on the extracellular side and 7 short turns on the periplasmic side. Structural studies showed that OmpG adopts an open and closed conformation at neutral pH 7.5 and acidic pH 5.6, respectively (43). Single channel assays revealed that wild-type OmpG gates spontaneously during applied potentials (42,44). However, the monomeric structure makes it straightforward for chemical or genetic modifications. Fluctuations in current signal

occurred due the presence of a flexible peptide loop inside the pore. Two approaches were employed to reduce the fluctuation by 95%: reduce the mobility of loop 6 by introducing a disulfide bond between the extracellular ends of strands β12 and β13; and optimize hydrogen bonding between strands β12 and β13 by deleting residue D215 (42). The conductance of OmpG monomeric wild-type and mutant under 1M KCl is 1.2 nS per pore in the fully open states (42). Membrane insertion occurred when OmpG protein with a potential of +150 mV applied to induce protein insertion. OmpG orients in the lipid bilayer with either the extracellular domain facing the *cis* or the *trans* chamber (42).

***2.6 Ferric hydroxamate uptake component A (FhuA)*** *(PDB: 1BY5)*

FhuA is a monomeric protein with a molecular weight of 82 kD (**Fig.1.2**). It was derived from the outer membrane of *E.coli.* As revealed in crystal structure, the pore has an elliptical shape with a cross-section of 3.1×4.4 nm composed of 22-stranded β-barrel domain. The β strands lie anti-parallel to one another and exhibit an exceptional robustness (45). The FhuA pores retain their pore forming ability and robust attributes even after removal of one third of the amino acids (10). On the other hand, upon addition of two β-strands to the protein primary sequence, it still maintains the pore forming ability with increased size (46). The channel diameter of the pore can be increased simply by doubling the amino acid sequence of the first two N-terminal β-strands resulting in 24 β-strands (46). The conductance of FhuA mutant ΔC/Δ 4L in which 160-residue cork domain and four long extracellular loops were deleted is ~4 nS under 1M KCl (47). Membrane insertion was aided by preparing FluA/Liposome (46).

***2.7 Cytolysin A (ClyA)*** *(PDB: 2WCD)*

ClyA is a dodecameric pore with a molecular weight of 408 kDa. It was discovered from *Salmonella typhi*. The geometry of this pore is still unknown, although the orthologue from *E.coli* has a crystal structure available (48). The dimension of ClyA nanopore constructed by homology modeling from the *E.coli* ClyA is 7.0 nm at its wider end and 3.8 nm at its narrower end with a length of 13.0 nm (49). ClyA nanopore did not show gating significantly between +60 mV and -90 mV (49). Beyond this range of potentials, ClyA nanopore showed long lasting current blockades, which could be relieved by reversal of the applied voltage (49). The interior surface of ClyA is negative charged. The conductance of ClyA is 1.8 nS under 150 mM NaCl. Membrane insertion occurred by adding oligomeric ClyA to the cis chamber (49).

***2.8 $PA_{63}$*** *(PDB: 1V36)*

$PA_{63}$ is a heptameric pore with a molecular weight of 442 kDa (**Fig.1.2**). It is a fragment cleaved from protective antigen, which is a three component part of anthrax toxin secreted by *Bacillus anthracis*. After binding to the receptor, the full length $PA_{83}$ (83 kDa) is cleaved by extracellular protease to generate $PA_{63}$ (50). $PA_{63}$ then self-assembles to form heptameric and octameric channels, which is a large cation-selective pore and inserts into the bilayer irreversibly (51,52). In the acidic endosomal environment, the oligomer of $PA_{63}$ undergoes dramatic structural changes, forming a mushroom-shaped ion channel. The crystal structure revealed that the dimension of $PA_{63}$ is 12 nm at the ring-like cap domain, and 2.8 nm at the narrow end with 7 nm long cap and 10 nm long stem (53). When inserted into a lipid bilayer, fast flickering between open and closed states was found in this pore (52,54). This fluctuation disappeared after mutagenesis of F427A (55). PA63 conductance is sublinear in salt concentration from 0.1 to 2.5 M KCl, changing from ~80 pS to ~320 pS

correspondingly under 20 mV of applied voltage (52). Membrane insertion occurred by adding $PA_{63}$ directly to the *cis* chamber.

### *2.9 Viral protein channels: Phi29, SPP1, T3 and T4 connectors*

***Phi29 connector (PDB:1H5W)***: Phi29 connector is a funnel shaped dodecameric channel with a molecular weight of 420 kDa (**Fig.1.2**). It is a key component of bacteriophage phi29 DNA packaging motor and acts as a pathway for the translocation of the genomic DNA. The diameter of the exterior part of the channel is 13.8 nm and 6.6 nm at the wide and narrow ends, respectively. The interior of the channel is 6 nm and 3.6 nm in diameter at the wide end and at the narrowest constriction, respectively. The channel is the first example of a nanopore that is neither an ion channel nor a membrane protein that has been inserted into a lipid bilyer(13). Phi29 connector displays robust properties and can withstand harsh mutagenesis, including site mutagenesis, insertion, and truncation. The membrane embedded channel is stable in a wide range of solution condition, including pH 2-12, and ionic strengths of 0.1-3M (13,56). Unlike other membrane channels, the insertion of phi29 connector channel is not spontaneous, and it requires a two-step approach. The connector channel is first incorporated into lipid vesicles during reconstitution process and then the resulting proteoliposomes are fused with a planar lipid membrane(13). The conductance is uniform and the Current-Voltage trace is perfectly linear and is about 4.5 nS under 1M KCl(13).

***SPP1 (PDB:2JES):*** SPP1 connector is also a funnel shaped dodecameric channel which is from bacteriophage SPP1 that infects *Bacillus subtilis.* The connector has an overall diameter of 16 nm and a height of 10.5 nm. The narrowest constriction of the internal channel is ~3 nm (57,58). The conductance is about 4.1 nS under 1M KCl. Similar

to phi29 connector channel, the insertion of SPP1 connector channel is not spontaneous, and it requires a two-step approach (14). Under a scanning voltage (–50 mV to +50 mV; 2.2 mV/s), SPP1 displays a linear Current-Voltage (I-V) relationship without voltage gating.

***T3:***The T3 portal is a mixture of 12 and 13 subunits, depending on the protein expression conditions and other factors. The 12-mer version of the T3 portal is 14.9 nm in width, 8.5 nm in height and 3.7 nm in diameter for the internal channel (59). Under a scanning voltage (–50 mV to +50 mV; 2.2 mV/s), T3 display a linear Current-Voltage (I-V) relationship without voltage gating. T3 conductance distribution appeared as two peaks: 2.6 nS and 3.9 nS. Like phi29 connector channel, the insertion of T3 connector channel is not spontaneous, and it requires a two-step approach (14).

***T4:*** The T4 portal exists as a dodecameric ring that is 14 nm long and 7 nm wide, and an interior channel of ~3 nm in diameter (60). Unlike other membrane channels, the insertion of phi29 connector channel is not spontaneous, and it requires a two-step approach. Under a scanning voltage (–50 mV to +50 mV; 2.2 mV/s), T4 display a linear Current-Voltage (I-V) relationship without voltage gating. The conductance is about 3 ns under 1M KCl. Like phi29 connector channel, the insertion of T4 connector channel is not spontaneous, and it requires a two-step approach (14).

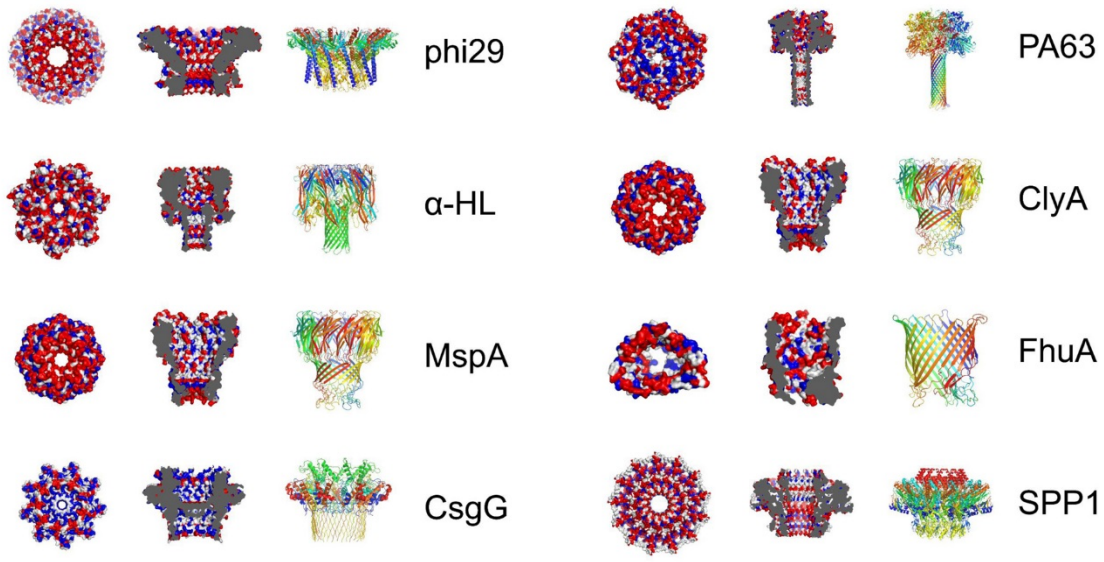



**Figure 1.2 Structure of commonly biological nanopore.** Bottom view, intersection view and side view of phi29 connector channel, PA63, α-hemolysin, ClyA, MspA, FhuA, CsgG and SPP1.

| **Name** | **Source** | **Oligomer State** | **PDB ID** | **Monomer MW(kDa)** |
|---|---|---|---|---|
| α-HL | *Staphylococcus aureus* | heptamer | 3ANZ | 33.2 |
| MspA | *Mycobacterium smegmatis* | octamer | 1UUN | 20 |
| Aerolysin | *Aeromonas hydrophila* | heptamer | | 47 |
| CsgG | *Escherichia coli* | octamer | 4UV3 | 31 |
| OmpG | *Escherichia coli* | Monomer | 2IWV(open) 2IWW(closed) | 35 |
| FhuA | *Escherichia coli* | Monomer | 1BY5 | 82 |
| GrA | *Bacillus Brevis* | dimer | 1GRM,1MAG | ~2 |
| ClyA | *Salmonellatyphi* | dodecamer | 2WCD | 34 |
| $PA_{63}$ | *Bacillus anthracis* | heptamer | 1V36 | 63 |
| Phi29 connector | *Bacteriophage Phi29* | dodecamer | 1H5W | 35 |
| SPP1 connector | *Bacteriophage SPP1* | dodecamer | 2JES | 56 |
| T3 connector | *Bacteriophage T3* | dodecamer | | 59 |
| T4 connector | *Bacteriophage T4* | dodecamer | | 60 |

**Table 1.1. Structure basis of different protein nanopore (mw, structure feature, source, geometry)**

## 3.Engineering protein nanopores for DNA and RNA sequencing

The capability of differentiating individual nucleotides of an intact DNA or RNA strand using protein nanopore is largely determined by the characteristics of the protein nanopore employed, and typically relies on nanopores with chemical modification or molecular motor incorporation. Without engineering, effective and high-resolution discrimination of nucleotides using wild-type biological nanopores remains elusive. The idea of using nanopores for DNA sequencing was proposed several decades ago since the landmark paper showing the translocation of single stranded DNA in α-hemolysin in 1996(1). Several follow-up studies have shown that DNA/RNA homo-or co-polymers could be readily distinguished by α-hemolysin (22,61). Upon site directed mutagenesis at E111N and K147N, the electrostatic interactions between Glu-11 and Lys-147 was interrupted and resulted in enhanced blockade and better discrimination of poly(dA) and poly(dC) compared with wild-type (28). In addition, through immobilizing ssDNA *via* biotin-streptavidin conjugation on the channel entrance, three recognition sites were identified in this mutant that are located near the constriction, in the middle of the barrel, and the trans-entrance, respectively (28). In follow-up studies, the recognition capability was further enhanced by extensive mutagenesis of residues, including M113Y and M113W. The data showed that the mutant with substitution Gly residue actually reduced recognition ability (62). The authors suggested some additional substitution with unnatural amino acids, particularly those with aromatic side chain can further improve nucleobase recognition (62). With modification of the three recognition sites E111N/K147N/M113Y in α-hemolysin, epigenetic DNA modification, 5-methylcytosine and 5-hydroxymethylcytosine, could be identified (63).

The next challenge was to slow down the translocation of DNA through α-hemolysin. Several passive approaches have been tested including increasing viscosity (64), decreasing temperature (65), and adding organic salts (66), as well as more active approaches by attaching complementary DNA sequences (67), and enzyme (68). So far, incorporation of enzyme has turned out to be the most efficient way to slow down the DNA translocation speed in order to achieve DNA sequencing. Initial success was achieved by introducing T7 DNA polymerase into α-hemolysin. The blockage current signal caused by sequential nucleotide incorporation were recorded when DNA-polymerase complex was held at the entrance of the nanopore (68). However, the observed three sequential ionic current steps before the T7 DNA Polymerase dissociation from the template made this method impractical. Also the need for precise control of the replication rate and strand motion is critical for polymerase based DNA sequencing(69). Later, it was shown that phi29 DNA polymerase could overcome these weaknesses because it can maintain normal elongation and remain bound to the DNA template while captured in the nanopore (69) (**Fig 1.2C**). With phi29 DNA polymerase, α-hemolysin is able to ratchet DNA through the pore at 5-Å precision (69). By using an exonuclease approach, a DNA strand was digested and the released nucleotides were then captured and sensed with non-covalent adapter (4)(**Fig 1.3B**). Studies showed that individual RNA bases and even the modified bases can be distinguished by immobilizing ssRNA with α-hemolysin mutant E11N/K147N/M113Y (70). Later, the authors realized that direct RNA sequencing is feasible by introducing adaptor and polynucleotide phosphorylase (PNPase) which cut ssRNA in the 3'to 5' direction processively and liberate rNDPs(71)(**Fig 1.3E**). In the study, kinetic parameters for rNDP identification was thoroughly compared and optimized in different mutant system

with different types of adaptor, including M113R-RL2 mutant of α-hemolysin with non-covalently linked heptakis-(6-deoxy-6-amino)-β-cyclodextrin and heptakis-(6-deoxy-6-guanidino)-β-cyclodextrin, and α-hemolysin-$(N139Q)_6$(N139Q/L135C)1 with covalently linked heptakis-(6-deoxy-6-amino)-6- N-mono(2-pyridyl)dithiopropanoyl-β-cyclodextrin, thereby demonstrating that the mutant α-hemolysin with covalently attached cyclodextrin adapters is able to continuously identify RNA nucleotides (71).

Recently, a different approach for DNA sequencing has been demonstrated that combines sequencing by synthesis (SBS) with α-hemolysin nanopore (72)( **Fig 1.3D**). The principle is that as DNA polymerase elongates the template, the incoming complementary-tagged nucleotides is cleaved and the nucleotide can be differentiated by different-sized polymer tags (72). Unlike previous non-covalent polymerase, in this platform a single polymerase molecule is covalently attached to a single subunit of α-hemolysin using two strategies at an optimized distance that permits fast capture of the tag. First, the conjugation of Phi29 polymerase to α-hemolysin was accomplished using either SpyTag and SpyCatcher peptide domains, or chemical conjugation. By controlling the ratio of native and SpyTag monomer (1:6), α-hemolysin was assembled with only one SpyTag which was then covalently linked to Phi29 polymerase. Second, the conjugation of polymerase to α-hemolysin was realized using Diels-Alder fast click chemical conjugation. Briefly, a *Bst* DNA polymerase molecule containing a single cysteine was reacted with excess of 6-methyl-tetrazine-$PEG_4$-maleimide in one tube, and an α-hemolysin heptamer containing a single cysteine at position 46 of only one monomer react with excess of *trans*-cyclooctene-$PEG_3$-maleimide in a separate tube. The polymerase-nanopore conjugate was obtained after mixing the two reactions after purification (72).

In MspA nanopore platform, the wild-type exhibits spontaneous blockades without ssDNA under an applied potential above 60 mV. To circumvent this problem, a mutant channel was generated whereby the negatively charged aspartate residues were replaced with asparagines in the constriction zone resulting in D90N/D91N/D93N mutant. Inspired by the unique structure of MspA (73,74), radical protein engineering rather than point mutagenesis was done in order to generate a sharper reading head and improve the resolution for base identification (75). In this study, truncated barrel mutant TBMΔ6 was generated based on NN mutant HL by PCR mutagenesis and ligation-free *in vivo* recombination that pair-wise removed amino acids from both β–strands (NN/Δ120-125/Δ133-138). The resulting mutant TBMΔ6 has ~1.6 nm shortened barrel and showed a weakened R1 recognition site. Further mutagenesis Met113→Gly demonstrated sharpened base recognition. In another mutant of TBMΔ6 (Met113→Phe), it was found that positively charged β-cyclodextrin and am7βCD can reside in the interior of the mutant, thereby allowing continuous recognition of mononucleotides, which is a prerequisite for exonuclease sequencing. The capability of the mutant pore to discriminate between the four DNA and RNA nucleotides has been significantly improved compared to previous studies using the full-length RHL pore (71,76). The capability was measured as the products of the sequential differences (δ) between each residual current levels in the histograms. δ is defined as 0 when a pore that is unable to discriminate between all NMPs. In this study, $\delta^{dNMP} = 5.4$ and $\delta^{rNMP} = 4.4$ were obtained demonstrating a high differential capability (75).

Several early studies have demonstrated that recognition of single nucleotide is possible with short ssDNA (77) without using DNA polymerase or exonuclease. To further

improve the MspA analysis capabilities in DNA sequencing, the authors further replaced negatively charged residues around the entrance and in the vestibule with positively charged residues, resulting in a mutant D90N/D91N/D93N/D118R/D134R/E139K. This mutant also had similar conductance as the first mutant, but is smaller than wild type. The studies demonstrate that proper substitution of charged residues provides a simple means to dramatically tailor the interaction between channel and analyte.

By combining the reengineered MspA nanopore with the phi29 DNA polymerase, MspA nanopore is able to read DNA at single-nucleotide resolution (73)(**Fig 1.3A**). The same system is capable of detecting and mapping 5-methylcytosine and 5-hydroxymethylcytosine, and proved to be extremely sensitive to abasic residues, which is one of the most common DNA lesions (78). Another study demonstrated that the engineered MspA nanopore with Phi29 DNA polymerase is capable of discriminating cytosine, hydroxymethylcytosine, and methylcytosine along individual DNA strands (79). To further push the platform into practical use for long reads of DNA sequence, Laszlo modified the platform by introducing two asymmetric adaptors: a cholesterol-tailed adaptor and a nicked, hairpin adaptor to the ends of linearized phi X174 dsDNA. Using this platform, they first constructed a current blockade map of all 256 four-nucleotide combinations because each four nucleotide combination has its own unique current fingerprint determined by this MspA platform. Using this map as a reference, this MspA platform is able to read 4500-bp sequence corresponding to natural DNA sequences (74). The current levels obtained from nanopore sequencing of the phi X 174 DNA strongly match the predicted current levels from quadromer map (r=0.9905, 95% confidence

bounds.). The authors speculated that the error can result from the shifts in the positioning of the DNA within the pore's constriction and nucleotides outside of the quadromer.

The speed of DNA translocation MspA is faster than α-hemolysin (8). The difference in translocation speed is possibly because of the difference between the widths of the channel regions flanking the narrowest constrictions. About 10-20 bases are highly confined in the constrict site and forced to interact with the channel wall in α-hemolysin. Whereas for MspA, there is only 2-4 bases in the constrict site. Another reason is the charge distribution within the constriction. The constrict site of α-hemolysin is formed by E111, K147, and M113. However, there are no charged residues in the constriction of MspA mutant, which may be responsible for the faster translocation speed in contrast to α-hemolysin (8). A further arginine mutation in M1MspA slows DNA translocation by a factor of 10-30 without eliminating nucleotide specific current blockades (80). Recent studies have shown that by introducing a net positive charge on the interior surfaces, the rate of capture of DNA is improved in α-hemolysin and MspA (8,81). These results demonstrate that a better understanding of nanopore geometry and charge location is critical for improving their utility for sequencing and sensing.

R9 pore, based on a CsgG pore as Oxford nanopore revealed the identity in March 2016, is widely integrated into MinIon and PromethION systems for DNA and RNA sequencing and has replaced earlier used R6 pore. It is claimed that the accuracy in 1D reads (linear reads) and 2D (read both the template and complement strands) reads can be 90% and 95% respectively (82).

***Oxford Nanopore Technologies:*** Although significant progress has been made in DNA sequencing in academic arena, *Oxford Nanopore Technologies*, is the first

commercial company to release a prototype in the market. MinION, the first product announced in 2012, measures electrical current changes as single strand of DNA pass through the pore. This device can be used to perform field-based work because it is portable (87g and flash drive like size) and can be operated through the USB port of a laptop computer. It contains up to 512 channels inside the chip and only needs 10pg-1ug DNA sample without PCR amplification. The flow cell input volume is 50-160ul with a lifetime up to 72 hours. Read length can be up to 200kb and run time can be 1 minute to 48 hours. The sequencing data can be analyzed in real time online or offline. The company launched MinION early access program in May 2014, which allow selected researchers to test this device for sequencing applications. Now it is available for purchase on its website starting for just $1000. The accuracy of sequencing was continually improved from June (66%), July (70%), October (78%) and November (85%) due to sequencing chemistry updates. In a recent study from Mark Akeson's lab reports the sequencing of the 7.2 kb genome of M13mp18 with 99% of 2D MinION reads (83). The results demonstrate a single-nucleotide polymorphisms precision with accuracy up to 99% (84). Meanwhile, the company is developing various platforms, including PromethION and SmidgION. PromethION is a high throughput, tablet-sized instrument designed to allow the analysis of DNA, RNA, proteins or small molecules. It allows users to increase or decrease the flow cell numbers depending on experiment needs, and enable users to run an experiment until sufficient data, or sufficient quality has been met for high accuracy. SmidgION the same core technology as MinION and PromethION but will be designed for use with smartphones or other mobile with low power supply. It is designed to meet the broad needs of ield-based analyses such

as remote monitoring of pathogens in a breakout, the on-site analysis of environmental samples, and analysis of agricultural environments.

***Other Companies:*** **Genia Technologies**, a nanopore based sequencing company acquired by Roche in 2014, employed Nanotag sequencing approach. Briefly, four different length PEG-coumarin tags were attached to the terminal phosphate of 2'-deoxyguanosine-5'-tetraphosphate first. Then, during the polymerase reaction as each modified nucleotide is incorporated into the growing strand, each tag is cleaved and enters a nanopore in order. Unique ionic current blockade of four different tags is determined and therefore decoding the sequencing of DNA (85). Recently, the company combined the approach, sequencing by synthesis (SBS), with α-hemolysin nanopore as discussed in section of Engineering protein nanopores for DNA and RNA sequencing. It has demonstrated high base resolution (72). **Noblegen Biosciences**, uses an approach that combines fluorescent labels and nanopores. In the approach, DNA is first converted into a longer 'expanded synthetic representation' (ESR) molecule through a process named 'circular DNA conversion'. Each expanded nucleotides has a corresponding color-coded 'beacon'. As the hybrid DNA passes through nanopores, beacon complex is pulled off and the fluorescent signal is detected (86). **Illumina**, as the dominant player in Next Generation Sequencing market, has also made substantial investments to develop their own nanopore technology since the commercialization

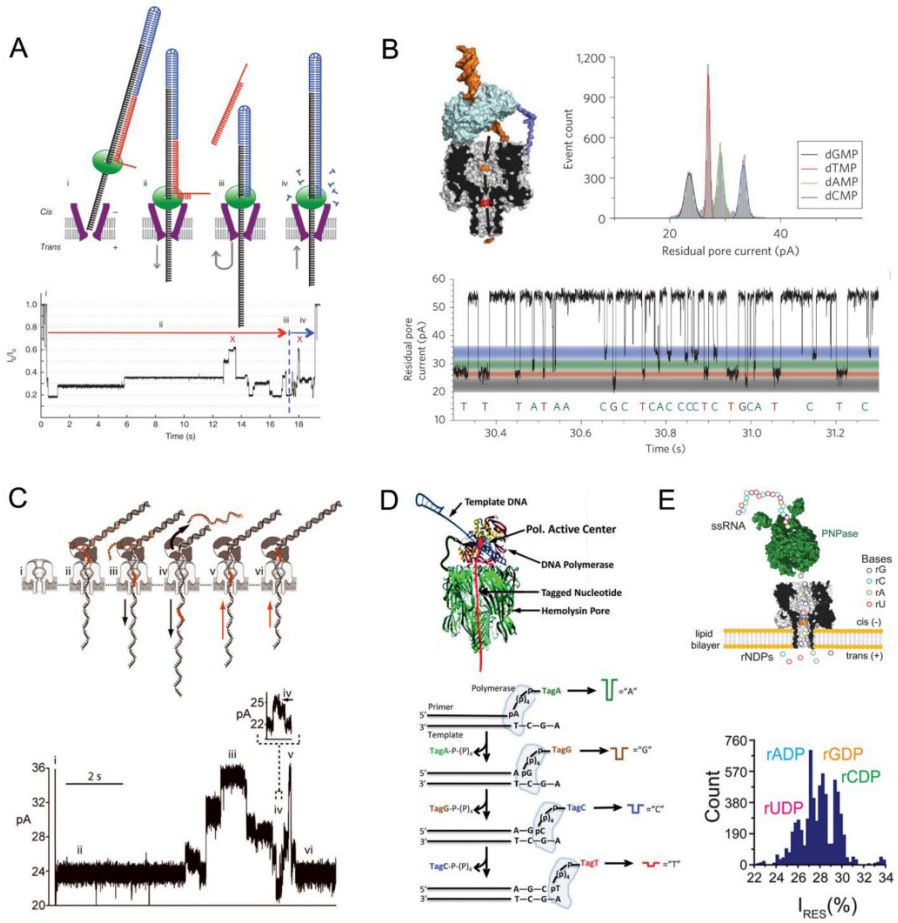



**Figure 1.3 Strategies for sequencing DNA and RNA at single-nucleotide resolution using DNA polymerase or exonuclease.** (A) MspA nanopore (73) (B) Proposed exonuclease based sequencing model (4). An exonuclease (pale blue) linked to the top of a nanopore by chemical linker or genetic fusion. After cleaved by exonuclease, the identity of individual nucleotide is determined sequentially by the level of the current blockaded. Single-channel recording from a-hemolysin with adapter and residual current histogram of nucleotide binding events showed the discrimination of dGMP, dTMP, dAMP and dCMP. (76) (C) Forward and reverse ratcheting of DNA through a-hemolysin equipped with phi29 DNA polymerase (69). (D) Principle of nanopore DNA sequencing by syntheses (72). A DNA polymerase molecule is linked to an a-hemolysin. Template DNA and primer along

with the tagged nucleotide from a ternary complex with the polymerase. The tagged nucleotides is captured and detected sequentially with specific current blockades as they are incorporated into the growing DNA strands. (E) Nanopore RNA sequencing. A ssRNA oligonucleotide (circles) cleaved by polynucleotide phosphorylase (green) sequentially. The released nucleotides are detected by mutant a-hemolysin equipped with a cyclodextrin adapter(71). Figures reproduced with permissions from: (A) Ref. (73); (B) Ref. (4); (C) Ref. (69); (D) Ref. (72);(E) Ref. (71)

agreement with Oxford Nanopore has been terminated. It has licensed the MspA patent from the University of Alabama at Birmingham and University of Washington in 2013. However, there are no development progress released from the company so far.

**4. Engineering protein nanopore for single molecule sensing**

Another area that nanopore has been widely applied to is single molecule detection of ions, small molecules, macromolecules, and biopolymers. The sensitivity, specificity and stability of the sensing is largely determined by the characteristics of the nanopore employed, chemical modifications, and/or incorporation of probe or adaptor to the nanopore. Various approaches have been used to introduce probe, including site directed mutagenesis, fused expression, and covalent or non-covalent attachment of adaptors. Compared with solid state nanopore, one of the great advantages of protein pores is amenability for selective modification with various probes that can interact with individual target molecules with high sensitivity and selectivity. Normally, the probe has two states: occupied (by analyte) and unoccupied (vacant). The concentration of the analyte can be revealed by the occurrence frequency of the events, whereas the identity can be revealed by the characteristic binding and distinctive current signatures such as dwell time and amplitude of the events (87,88).

**Sensing directly using site directed mutagenesis**

Without specific amino acids modification, protein channels are generally able to detect quite a few analytes. However, by introducing artificial molecular recognition sites, it can greatly broaden the sensing spectrum of biological nanopores. A variety of engineered protein nanpore have been used to detect several ionic species. In one of the examples, four histidines were introduced at position 123, 125, 133, and 135 of α-

hemolysin in which the side chains were projecting into the lumen. This mutant subunit was incorporated into a heteromeric pore composed of six wild-type α-hemolysin subunits and one mutant α-hemolysin ($WT_6 4H_1$). The mutant heteromer form was able to detect the concentrations of two or more divalent metal ions simultaneously (89,90).

Guan generated a series of mutants on Met 113 in α-hemolysin, including Met113Phe7, Met113Tyr7, Met113Tyr7, Met113His7, Met113Asp7, Met113Glu7, Met113Arg7, Met113Lys7, Met113Pro7, and Met113Val7, which can be classified into three groups: aromatic (Phe,Trp,Tyr), hydrophobic (Met, Pro, Val), and charged(Asp, Glu, Arg, Lys). The data showed that TNT can only be detected by three aromatic Met113 mutants (91). In phi29 connector nanopore system, ethane, thymine and benzene with reactive thioester moieties were clearly distinguished upon binding with the cysteine group mutated from lysine 234 (15)(**Fig 1.6 D,E**). DNA translocation with one-way property has also been reported in this system (**Fig 1.6 A-C**) (56,92).

Recently, it is reported that the fingerprints of various peptides were detected by phi29 motor channel (**Fig.1.5**). The peptides were distinguished by current blockage percentage and current signatures. In addition, the oligomeric states of peptides can be studied in real time (93). One way traffic of peptide translocation has been observed in both SPP1 and phi29 motor channel (94) **(Fig.1.5)**. Also, the study found that the orientation of the insertion into lipid membrane was governed by the hydrophobicity and hydrophilicity of the termini of the protein channel **(Fig.1.5)**.

**Sensing *via* probes introduced through covalent linker**

Engineering of nanopore using probes at specific locations can greatly increase the possibilities for stochastic sensing of macromolecules. Normally, protein is too large to

pass or fit inside the pore, so it is necessary to couple external binding modules to generate a signal. One end of a single 3.4 kDa polyethylene glycol chain was attached to the mutant S106C monomer. A biotin molecule was tethered covalently to the free end of the polymer. Nanomolar concentration of streptavidin was detected through reversible capture observed as changes in ionic current. This approach might be modified for detecting a wide variety of protein ligands (95).

By hybridization to an oligonucleotide, which is attached to N17C monomer through a disulfide bond near the *cis*-entrance of α-hemolysin, a DNA aptamer can be attached to α-hemolysin and used for quantification of nanomolar concentrations of thrombin (96). Using a similar strategy, Howorka covalently attached a single DNA oligonucleotide to a mutant single cysteine residue at position 17 of α-hemolysin, located at the cis-opening of the pore. This modification enable the system to detect complementary strands (67). This platform was further exploited to detect a resistance conferring mutation in an HIV-derived DNA sequence (67).

A similar strategy was used in ClyA nanopore which is wide enough to sample and differentiate large analyte. In the study, DNA aptamers containing an activated thiol group, were conjugated to ClyA monomers comprising of a single cysteine at position 110. The results showed that the engineered ClyA nanopore is able to distinguish human and bovine thrombin despite 86% sequence similarities (49). Howorka reported an approach to study the kinetics of binding of tetravalent lectin to a ligand engineered pore (97). In the report, thiopyridyl derivative of the disaccharide Gal-β-1,3-GalNac, a ligand for protein lectin, was coupled to a cysteine residue engineered at position 9, located at the narrower part of the cis-opening of the α-hemolysin pore. The engineered subunit was then assembled into

three different heptamers containing one, two and seven modified subunits (97). In another study, a ligand, PKIP5-24, was tethered to *trans*-entrance of α-hemolysin at position Cysteine 129 through disulfide bond, to study the binding interaction of cAMP-dependent protein kinase to the ligand (98).

Beside the study of macromolecule interaction, a reversible covalent chemistry was also studied using protein nanopores. In one study, a single photolabile carbamate group 3,4-dimethoxy-6-nitrobenzylcarbamate group was tethered to the interior wall at Thr117Cys of α-hemolysin. Upon UV light illumination the fragments of protecting group, the intermediates of the decomposition can be detected because the current depend on the shape, size, and polarity of the reactants tethered within the pore (99). At the same mutant position, Loudwig attached azobenzene via disulfide bond and studied the light induced isomerisation at the single molecule level (100). Chemical transformation such as covalent bond formation was also studied using this strategy (101).

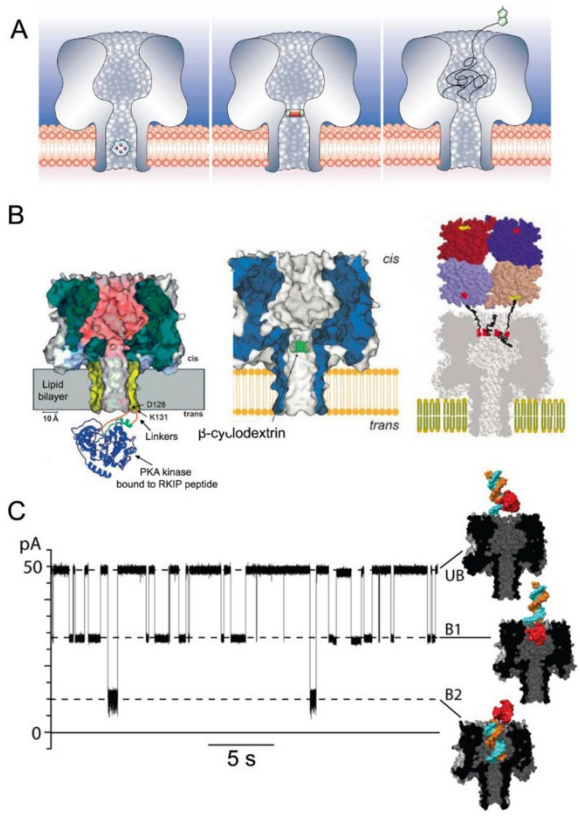



**Figure 1.4 Main approaches for biological nanopore sensing and detection. (A)** Illustration of the three main approaches to engineer a-hemolysin pores for sensing. Left: The sensor element is introduced by genetic modification within the lumen which is able to bind metal ions. Middle: Enhancing host-guest interaction by equipping with a ring-shaped molecular adaptors inside the lumen such as cyclodextrins. Right: A ligand is covalent attached poly (ethylene glycol) chain at one end, which is linked to the lumen of the pore at the other end (87). **(B)** Left: a genetically engineered a-hemolysin fused with a single protein kinase inhibitor peptide (PKIP), which is able to bind cAMP-dependent protein kinase and cause current blockade change. Transmembrane β -barrel (yellow) and the PKIP sequence (orange and green) (102). Middle: schematic of the wild type a-hemolysin displaying β -cyclodextrin resident in the lumen of the channel (103). Right: a

disaccharide natural ligand is covalently linked to modified a-hemolysin to stochastically detect the interaction with the tetrameric lectin (97). **(C)** A thrombin aptamer is linked to a-hemolysin by hybridization to an oligonucleotide that is attached to a single cysteine residue near the mouth of the pore through a disulfide bond. From the current blockade level, the interactions of thrombin and aptamer can be detected (96). Figures reproduced with permissions from: (A) Ref. (87); (B) Ref. (97,102,103); (C) Ref. (96).

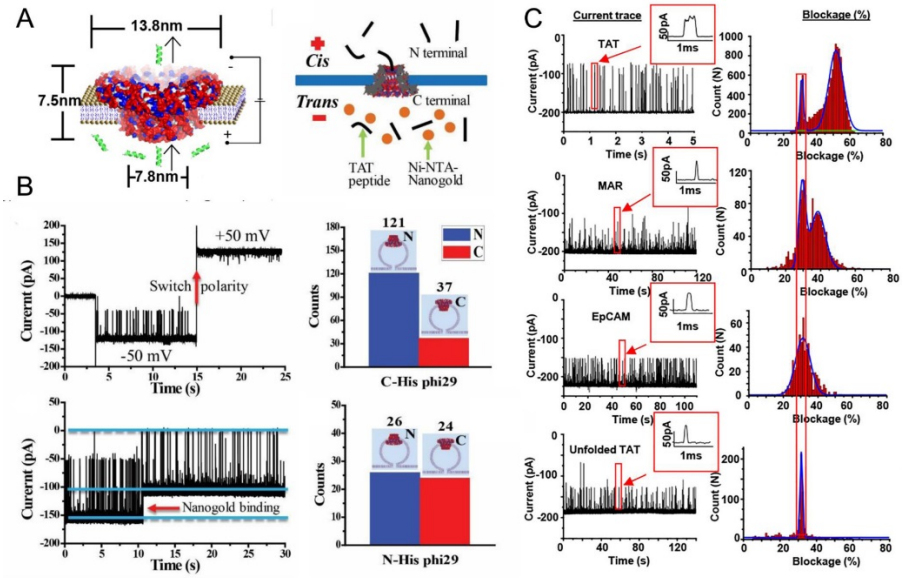



**Figure 1.5 Application of phi29 connector nanopore for peptide sensing. (A)** Schematic of peptide translocation through phi29 connector nanopore and the Ni-NTA-nanogold binding assay for determining the orientation of C-His phi29 connector nanopore. **(B)** The orientation of phi29 connector nanopore when inserted into lipid membrane is governed by hydrophobicity at the N- or C-terminus. **(C)** Current trace and blockade of different peptides when passing through phi29 connector nanopore (93,94). Figures reproduced with permissions from (93,94).

***Sensing via probes introduced through fusion protein expression***

Introducing new functions through fused expression of the encoded genetic sequence of the proteins is always challenging, because the folding of proteins cannot be reliably predicted. One method often adopted by nature to solve this challenge is combining evolutionarily preserved regions of proteins with generally independent protein domains. In one study, the functional elements of co-chaperonin GroES was introduced to α-hemolysin nanopore to study the interaction with GroEL at the single molecule level. The fused flexible loops of GroES with α-hemolysin showed correct folding with enzymatic activity mediating GroEL-assisted protein folding efficiently as native GroES (104). In a recent study, epithelial cell adhesion molecule (EpCAM) peptide was co-expressed with phi29 connector nanopore at the C-terminus (105). The binding of EpCAM antibodies sequentially to each peptide probe induced stepwise blocks in current. The docking and undocking kinetics of probe-antibody can be analyzed in real time at single molecule level by the distinctive current signatures. Importantly, the signal of EpCAM antibody can be distinguished from the background events such as serum or nonspecific antibody. The results demonstrate the feasibility of generating a highly sensitive platform for detecting antibodies at extremely low concentrations in the presence of contaminants (14) (**Fig 1.6F**).

***Sensing with non-covalent adaptors***

A variety of sensing functions can be achieved by introducing an adapter into protein nanopore. Lodging of an adapter in the lumen of protein nanopores can reduce the current flow and provide host-guest interactions. One class of adapters are ring-shaped molecules, which can lodge inside a pore and mediate the binding of small analytes. β-cyclodextrin, which has a hydrophobic interior cavity and a hydrophilic exterior, is the

most commonly used adapter for biosensing. Gu first introduced β-cyclodextrins to wild-type α-hemolysin non-covalently (106). Due to weak interactions, the residence time of β-cyclodextrins on the pore is only several hundred microseconds, although it can be extended to several milliseconds by adjusting the pH and trans-membrane potential. This largely limits widespread applications as stochastic biosensors. The possible lodging sites of β-cyclodextrins at the internal end of the β-barrel were further enhanced by site directed mutagenesis at Met 113, Lys 147 and Glu 111. The binding affinity was thereby increased $10^4$ fold compared to wild type, resulting in prolonged residence time of about tens of seconds (103,107). Organic ligands, such as 1-adamantanecarboxylic acid, 2-adamantanamine, or drugs imipramine and promethazine reversibly bound to the β-yclodextrin adapter and resulted in analyte characteristic blockade fingerprints (107).

The advantages of non-covalent adaptors have been further demonstrated by sensing individual nucleotide analytes. The adapter heptakis-(6-deoxy-6-amino)-β-cyclodextrin was lodged into the α-hemolysin pore, and reversible binding of adaptor to 2'-deoxyribonucleoside 5'-monophosphaes resulted in distinct current levels for each base (108). Another protein nanopore that employed β-cyclodextrin is outer membrane protein G. Heptakis-(6-deoxy-6-amino)-β-cyclodextrin, a positively charged β-cyclodextrin derivative, was introduced to double mutant OmpG (42).

***Sensing with covalent adaptors***

Covalently attachment of cyclodextrin was achieved by Wu and his colleagues by linking engineered Met 113 cys to cyclodextrin-propionate through disulfide bond (109). One of the advantages of covalent attachment is that the adapter cannot dissociate, which will not cause gaps during stochastic detection. This is especially useful for single molecule

exonuclease DNA sequencing which require the continuous presence of a molecular adapter. Another advantage is that the orientation of the adapter will be fixed which will allow analytes to bind through only one of the entrances to the cyclodextrin cavity (109). In a more recent study, the unlabeled nucleoside 5'-mono-phosphate molecules can be continuously identified by a covalently attached adapter with accuracies averaging 99.8% (76). Methylated cytosine could also be differentiated from the four standard bases (76).

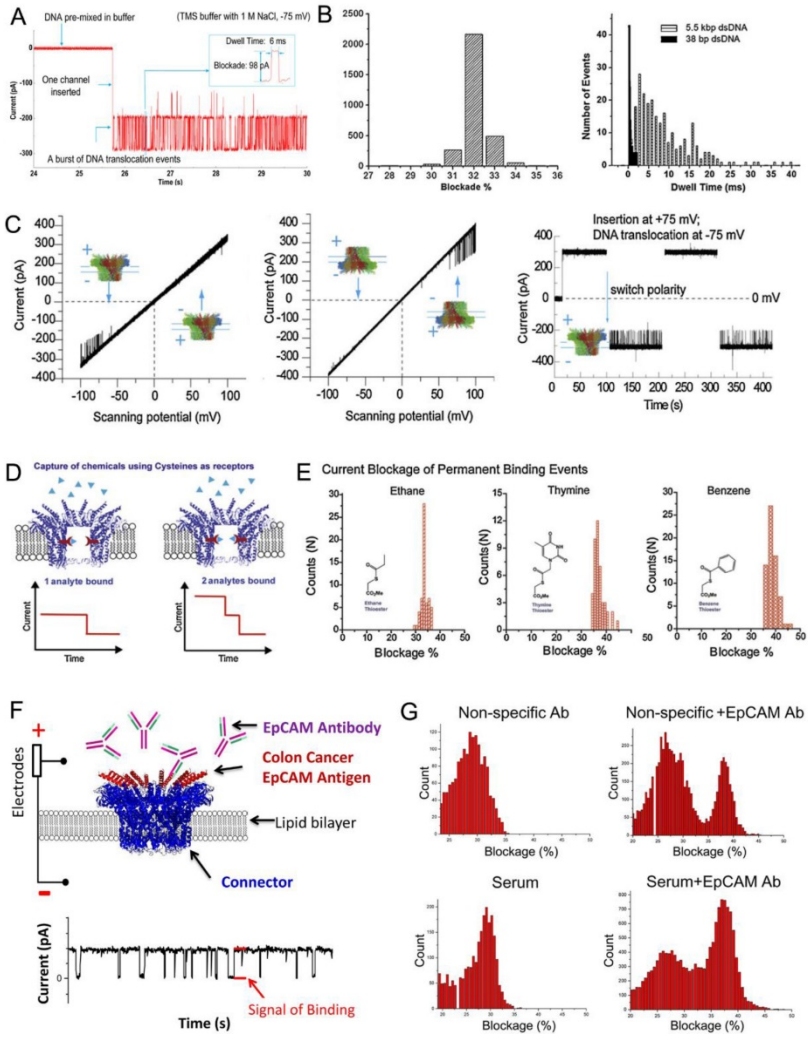



**Figure 1.6 Application of phi29 connector nanopore for dsDNA, single chemical and antibody detection.** (A) Current trace of dsDNA translocation through phi29 connector nanopore (92). (B) Histogram of current blockade and dwell time of dsDNA(56,92). (C) One-way traffic of dsDNA translocation through a single phi29 connector nanopore under a ramping potential and switching polarity. DNA is mixed in both cis- and trans-chambers (110). (D) Schematic of conjugation of chemical ligands to the channel wall caused the reduction of channel size as reflected by current blockade (15). (E) Analysis of current blockage events for the binding of thioester groups containing ethane, thymine, and benzene (15). (F) Schematic of phi29 connector nanopore for colon cancer specific

antibody detection (14). (G) Histogram of current blockage events caused by nonspecific antibody with or without EpCAM antibody, and diluted serum with or without EpCAM(14). Figures reproduced with permissions from: (A-B) Ref. (56,92);(C) Ref. (110);(D-E) Ref. (15);(F-G) Ref. (14).

***Changing oligomeric state and capability of forming channel***

In biological system, pores normally evolve with fixed stoichiometry. Control of subunit stoichiometry is critical to tailor protein nanopore for various applications since it can change the constriction zone size, and later the physical and chemical properties of the protein channel. In the studies of α-hemolysin, either homo- or hetero-heptameric form have been achieved. Because both N- and C-termini of single subunit of α-hemolysin are on the same side of the lipid membrane, the N- and C-termini of two subunits was linked to form dimer to generate an even number of subunits of α-hemolysin. Interestingly, the dimer still can assemble into heptamer form indicating that one of more subunits are expelled from the central pore (111). In MspA, the oligomer state was changed by connecting two mutated M1-MspA subunit through a liker ranging from 17-62 amino acids in length. A follow-up study demonstrated that the linked dimer formed functional channels with similar property as M1 MspA, which is suitable for DNA sequencing (35).

In phi29 connector channel, it has been reported that by deleting a loop in the central part of the channel, the conductance of the channel can be reduced about 40% compared to the wild-type indicating that the oligomeric state has been changed. The resulting mutant can be used to distinguished ssRNA and ssDNA (112). It is believed that the connector channel exists with fixed oligomeric states, which is mostly dodecamers, (113), the stoichiometry of the connector of different bacteriophages has been reported to vary from 11-mer to a 14-mer *in vitro* following ectopic expression and assembly, and heterogeneous population have also been found (59,114-117). It has been reported that the conformational changes occurring in specific segments, such as helix α6 of the tunnel loop and the crown region may be responsible for the different oligomeric states (57).

**CONCLUSIONS:**

In the last two decade, biological nanopores-based technology has made significant progress for DNA sequencing and single molecule sensing, and is well positioned to bring a revolution to medical diagnostics and DNA sequencing market. To date, several protein channels such as α-hemolysin, MspA, Aerolysin, FluA, Omp G, and viral connectors Phi29, SPP1, T3 and T4 have been successfully used for incredibly versatile applications including detection of small molecules, macromolecule, polymers, polypeptides, as well as DNA and RNA. Although significant progress has been made, several challenges still remains in nanopore area, such as limited knowledge of structure-function of protein channel, and the analytes that can be analyzed by available nanopores are still limited because different analytes requires unique protein nanopore with different shape, size, physical dimensions and hydrophilic/hydrophobic properties. Therefore, it still requires continuous improvement of current available nanopore and exploration new channels and pores by protein engineering.

For DNA /RNA sequencing, several key challenges that used be the bottlenecks in nanopore-based sequencing has been resolved such as the ultra-fast translocation of DNA through nanopore, spatial resolution and platform stability. However, the accuracy which is still low compared with current sequencing technologies. Many efforts are still needed to optimize the platform by using protein engineering, large scale screening, exploration new nanopore and improving the algorithmic in the future.

A wide range of analytes ranging from ions, small molecules, macromolecules, to biopolymers have been detected successfully using nanopore. However, most of these development were still in the laboratory stage. How to improve the sensitivity, stability and

specificity especially using clinical sample which is often in the presence of many impurities is still the major bottleneck to transform nanopore for clinical use. Also, more effort should be made to explore detectable capability such as protein and peptide sequencing, protein structure and folding studies.

# Chapter 2: Engineered Nanopore of Phi29 DNA-Packaging Motor for Real-Time Detection of Single Colon Cancer Specific Antibody in Serum

This chapter was reproduced (with some modification) with permission from Wang S, Haque F, Rychahou P, Evers BM, and Guo P. "Engineered Nanopore of Phi29 DNA-Packaging Motor for Real-Time Detection of Single Colon Cancer Specific Antibody in Serum." *ACS Nano*, 2013 Nov 26;7(11):9814-22. DOI: 10.1021/nn404435v. 

**INTRODUCTION:**

In biological systems, transmembrane and biomotor channels play critical roles in all aspects of life, such as regulating the traffic of macromolecules and ions into and out of nuclei, organelles, and cells, segregating chromosomes, and translocating and transporting single- or double-stranded DNA.(118-122) In bacterial virus phi29, with the aid of an ATP-driven motor, double-stranded DNA (dsDNA) viruses package their genome into pre-formed protein shells called procapsids.(123,124) The motor consists of the protein enzyme gp16, which functions as a part of ATPase, six copies of packaging RNA (pRNA),(125-128) and a central protein core called connector.(129,130) The connector is composed of 12 protein subunits that encircle to form a dodecameric channel, which enables dsDNA to translocate into the procapsid of the phage and exit during maturation and infection.(123-126) The crystal structure of the phi29 connector was determined at atomic resolution.(130) The ring is 13.8 nm at its wide end and 6.6 nm at its narrow end. The internal channel is 6 nm in diameter at the wide end and 3.6 nm in diameter at the narrow end (**Fig. 2.1C, D, E**). The wider end (C-terminal) of the connector is located within the capsid, while the narrow end (N-terminal) partially protrudes out of the capsid.

Recently, the connector channel has been incorporated into lipid bilayers to serve as a membrane-embedded nanopore with robust properties. This system has shown sensitive and unique conductance signatures when DNA or ions pass through the channel.(13,56) The channel conductance was uniform, as demonstrated by a perfectly correlated linear response to applied voltage. The connector channel is stable under a wide range of solution conditions, including a pH range of 2 to 12 and different salt species and concentrations.(56) A one-way traffic property for dsDNA translocation from the N-terminal to the C-terminal with a valve mechanism for DNA packaging has also been observed.(110) The connector channel has shown three discrete steps, gating at higher trans-membrane potentials associated with conformational change in the channel subunits.(131) The availability of the crystal structure of this connector channel has enabled explicit engineering with atomic precision for added functionality. Recently, a reengineered connector channel with a reduced channel size has shown the capability to discriminate single strand DNA or RNA from double strand(112). By selectively functionalizing a probe in the interior of the channel, single chemicals with reactive thioesters or maleimide groups can be identified based on their distinct fingerprints(15). Procedures for large-scale production and purification of the connector have already been developed.(13,132-136) Taken together, these features make the connector channel an ideal system for sensing and diagnostic applications.

Nanopore-based sensory techniques have been extensively studied for the detection of a variety of macromolecules and chemicals based on modulations of the individual current blockage events.(4,17,137-140) Engineered transmembrane channels with various probes are capable of stochastic detection by observing (in real time) the individual binding

events between single ligands and receptors with high selectivity and sensitivity.(87,141,142) The unique current signatures and characteristic binding can reveal the identity and concentration of the target analyte.(143,144) Moreover, the dynamic interaction between the analyte and the probes can be studied in real-time at high resolution by using single channel conduction assays. One advantage of the single molecule techniques is the low limit of detection, which is ideal for the early detection of biomarkers that exist in ultra-low concentrations, making it possible to diagnose specific diseases at an earlier (i.e., asymptomatic) stage.

In this study, we introduced an Epithelial Cell Adhesion Molecule peptide (EpCAM) that was 18 amino acids long as a probe into the C-terminal of each subunit (Fig. 2.1A, B). EpCAM is a cell surface molecule known to be overexpressed by the majority of human epithelial carcinomas, including prostate, breast, colorectal and head and neck cancers.(145-151) We incorporated this reengineered channel into a lipid bilayer and characterized its conductance properties. We then detected the binding of EpCAM antibody at a single molecule level in nano-molar concentrations based on the unique current signatures. The docking and undocking kinetics of antibody-probe interactions enabled us to determine the Kd. To further test the detection capability of this engineered phi29 channel and push nanopore techniques for clinical utilization, we performed the EpCAM antibody detection in the presence of either non-specific antibody or diluted serum. We showed that the EpCAM antibody can be distinguished from the background events that are present. Our results demonstrate the feasibility of reengineering the connector channels with a wide range of probes for medical diagnostic applications in the future.

## MATERIALS AND METHODS:

### *Materials*

The phospholipid 1,2-diphytanoyl-sn-glycerol-3-phosphocholine (DPhPC) was purchased from Avanti Polar Lipids, Inc. Organic solvents (n-decane and chloroform) were obtained from Fisher Scientific, Inc. and TEDIA, Inc., respectively. The EpCAM antibody was purchased from Abcam Company. All other reagents, if not specified, were purchased from Sigma-Aldrich, Inc.

### *Serum sample preparation*

Male athymic NCr nude mice between 6-8 weeks of age were acquired from Taconic (Hudson, NY) and kept in an animal house with 12 hour of light and dark cycled. Food and water was given ad libitum. Blood serum was collected from vena cava and allowed to clot for 30 min at room temperature. The clotted material was removed by centrifugation at 2000 rpm for 10 min in a refrigerated centrifuge. Hemolytic material was not observed. The sera obtained from the blood samples were frozen immediately without any further treatment in liquid nitrogen and stored at -80°C until further use.

### *Cloning and Purification of the Engineered EpCAM Phi29 Connector Protein*

The construction of the plasmid containing the gp-10 gene and the expression and purification of the phi29 connector have been reported previously.(132,134) The new plasmid was constructed first by introducing an EpCAM probe (ELKHKAREKPYDSKSLRT) to the C-terminal of the connector channel, just downstream of gp-10 gene; a His tag was inserted into the N-terminal for purification. The newly constructed clone was transformed into HMS174 (DE3) E. coli bacteria. The

successfully transformed bacteria were cultured in 10 mL Luria-Bertani medium overnight at 37°C. These cultured bacteria were transferred to 500 mL of fresh LB medium. When OD600 reached 0.5-0.6, 0.5 mM IPTG was added to the cultured medium to induce protein expression. The bacteria were collected after 3 hr, post-centrifugation induction. A French press was used to lyse the bacterial wall, and the protein and other components were differentiated by centrifugation.

An Ni-NTA His bind resin with a His tag was applied to purify the mutant protein. Briefly, 2 ml of regenerated His resin was packed into a column First. The supernatant differentiated by centrifugation was loaded into the column. The column was then washed with washing buffer to remove any contaminant proteins. The His-tagged mutated gp10 protein was eluted using elution buffer that contained 500 mM imidazole.

***Preparation of Lipid Vesicles Containing EpCAM Engineered Connector Channels.***

The incorporation of connectors into liposomes has been reported previously.(13) Briefly, 1 mL of 1 mg/mL DPhPC in chloroform was poured in a round-bottomed flask. The chloroform in DPhPC lipid was removed under vacuum. Then, the lipid film that formed was rehydrated with EpCAM engineered connector channel buffer, which contained 1M NaCl, 10 mM Tris/pH7.9, and 250 mM sucrose to bud off vesicles into the solution. The unilammellar lipid vesicles were generated by passing the lipid solution through a 400 nm polycarbonate membrane filter.

***Incorporation of the Connector Channel into a Planar Bilayer Lipid Membrane***

The method of inserting the connector with reconstituted liposomes into a lipid bilayer has been reported previously.(13) Briefly, a Teflon film partition (aperture 200 μm

in diameter) was used to separate a bilayer lipid membrane chamber into *cis* and *trans*-compartments. The aperture was painted two times with 0.5 uL of 3% (w/v) DPhPC n-decane solution, and the two compartments were filled with conducting buffer (0.2 M NaCl, 1mM HEPES, pH 7.4). After the formation of the lipid bilayer on the aperture, the lipid/connector complexes were added to the chamber and allowed to fuse with the planar lipid bilayer.

***Single Channel Conduction Assays for Each Membrane Inserted Connector Channels***

A pair of Ag/AgCl electrodes was connected directly into the cis and trans-compartments to measure the current traces across the lipid bilayer membrane. The current trace was recorded using an Axopatch 200B patch clamp amplifier coupled with the Axon DigiData 1322A analog-digital converter (Axon Instruments) or the BLM workstation (Warner Instruments). All voltages reported were those of the trans-compartment. Data was low band-pass filtered at a frequency of 1 kHz, and acquired at a sampling frequency of 10-100 kHz. The Patch clamp 9.1 software (Axon Instruments) was used to collect the data, and the software Origin Pro 8.0 was used to analyze all the data.

All antibody binding experiments were conducted under 0.2M NaCl, 1 mM HEPES pH 7.4 buffer. EpCAM antibody was added only after the definite insertion of EpCAM engineered connector channels in the lipid bilayer membrane. For selectivity of the EpCAM connector channel study, mouse serum was diluted to 100-fold using the electrolyte buffer; 10 ng/ul of non-specific antibody was pre-mixed with the buffer. Then EpCAM antibody was added only after the addition of EpCAM engineered connector channels. The current traces were recorded over a period of 1 to 2 hrs, and all the experiments were repeated at least three times.

**RESULTS:**

***Incorporation of the EpCAM Peptide to the C-terminal End of the Connector***

To facilitate connector purification, an N-terminal His tag was inserted just upstream of the gp-10 connector gene. An EpCAM peptide (18 amino acids in length) was inserted downstream of the gp-10 connector channel gene at the C-terminal end. To provide end flexibility, a linker with 6 glycines was included between the gp-10 connector and the EpCAM peptide probe (**Fig. 2.1A-B**). The reengineered connectors were purified to homogeneity and run on an SDS-PAGE gel (**Fig. 2.1F**) with wild-type and C-His tagged connectors as controls. The molecular weights of C-His tagged and N-His-C-EpCAM modified connector subunits were 36.7 kDa and 40.6 kDa, respectively. After purification, the reengineered channels with the EpCAM probe assembled into a ring composed of 12 evenly spaced probes in the same plane within the dodecameric connector channel. In addition, we compared the conductance of EpCAM connector channels and C-his connector channels under the same buffer conditions. The similar conductance and histogram distribution that was noted indicates they have similar cross-sectional areas. The accessibility and specificity of EpCAM probes on the engineered connector was further verified by Western blot assays (**Fig. 2.1G**). Compared with the controls (EpCAM peptide, C-His connector and bovine serum albumin [BSA]), the EpCAM antibody was observed to bind specifically to the EpCAM reengineered connector. The intensity of the band corresponding to EpCAM Ab/EpCAM connector complex increased as the concentration of EpCAM engineered connector channel was increased from 0.15 μg to 1.2 μg (**Fig. 2.1G**).

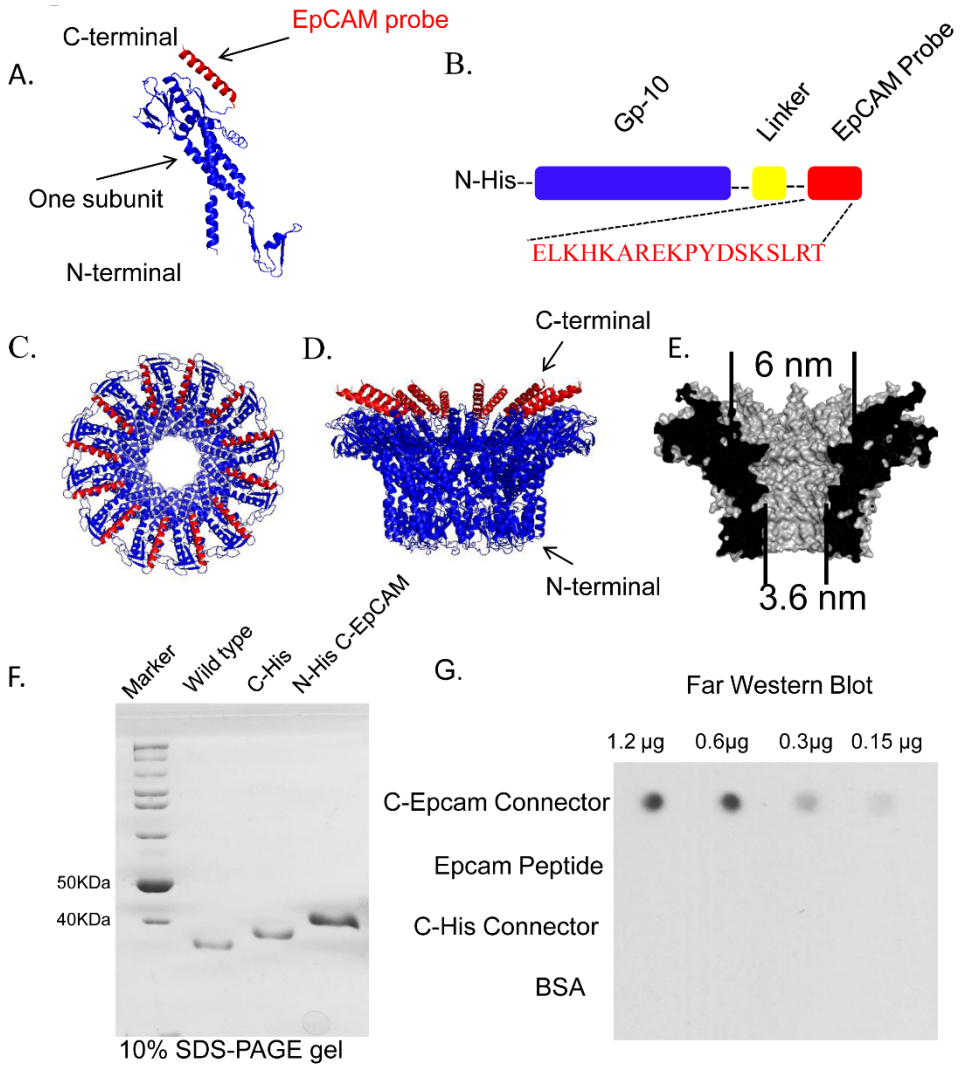



**Figure 2.1: Illustration of the phi29 connector channel structure. (A)** Structure of one subunit showing the location of the EpCAM probe (red). **(B)** Construction of the modified gp10 gene by insertion of His tag at the N-terminus, 6-Gly linker and EpCAM probe at the C-terminus. **(C)** Top view and **(D)** side view. **(E)** Section view of the phi29 connector channel showing the size and location of the EpCAM probe (red) and incorporation of the EpCAM receptor into the phi29 DNA packaging motor channel. **(F)** Molecular weight of wild type, C-His and N-His C-EpCAM connector on 10% SDS-PAGE. **(G)** Demonstration of the accessibility and specificity of EpCAM probe in the C-terminal of the phi29 connector channel by Far Western blot.

***Characterization of the Membrane-Embedded EpCAM Reengineered Connector Channels***

The EpCAM engineered connector channel was incorporated into planar lipid membranes in two steps, as previously described.(13) The first step was to incorporate the connector into liposomes, followed by insertion into planar lipid membranes via vesicle fusion of liposome/connector complexes, to form a membrane-embedded nanopore.(13,56) The insertion of the EpCAM engineered connector channels generated a stepwise increase of the current, as shown in the continuous current trace (**Fig. 2.2A**). The insertion of EpCAM probe at the C-terminus did not affect the membrane signal stability, voltage gating properties, membrane durability, or the membrane insertion efficiency of the connector channel. The current step size of the EpCAM engineered connector channels was homogeneous (**Fig. 2.2C**), and the channels showed equal conductance under both positive and negative transmembrane voltages. The average current jump was 55± 6 pA (0.74±0.09 nS) in 0.2 M NaCl, 1 mM HEPES, pH 7.4. Conductance was derived at specific, constant holding potentials (+75 mV or -75 mV) after the phi29 connector channel was incorporated into a lipid membrane and was calculated as the ratio of the current jump induced by a discrete step to the applied voltage. Occasionally, two connector channels were observed to insert simultaneously, as demonstrated by a conductance of 1.43±0.03 nS (**Fig. 2.2C**). Under the same buffer condition, the reengineered N-his C-EpCAM connector channel showed similar conductance with C-his one (0.76±0.08 nS), indicating the modification did not change the conductance and size of the channel (**Fig 2.2 C**). The uniformity of this engineered connector channel was further demonstrated by applying

ramping voltage which showed a nearly linear I-V relationship without exhibiting any voltage gating phenomenon under the reported conditions of ±100 mV (**Fig. 2.2B**).

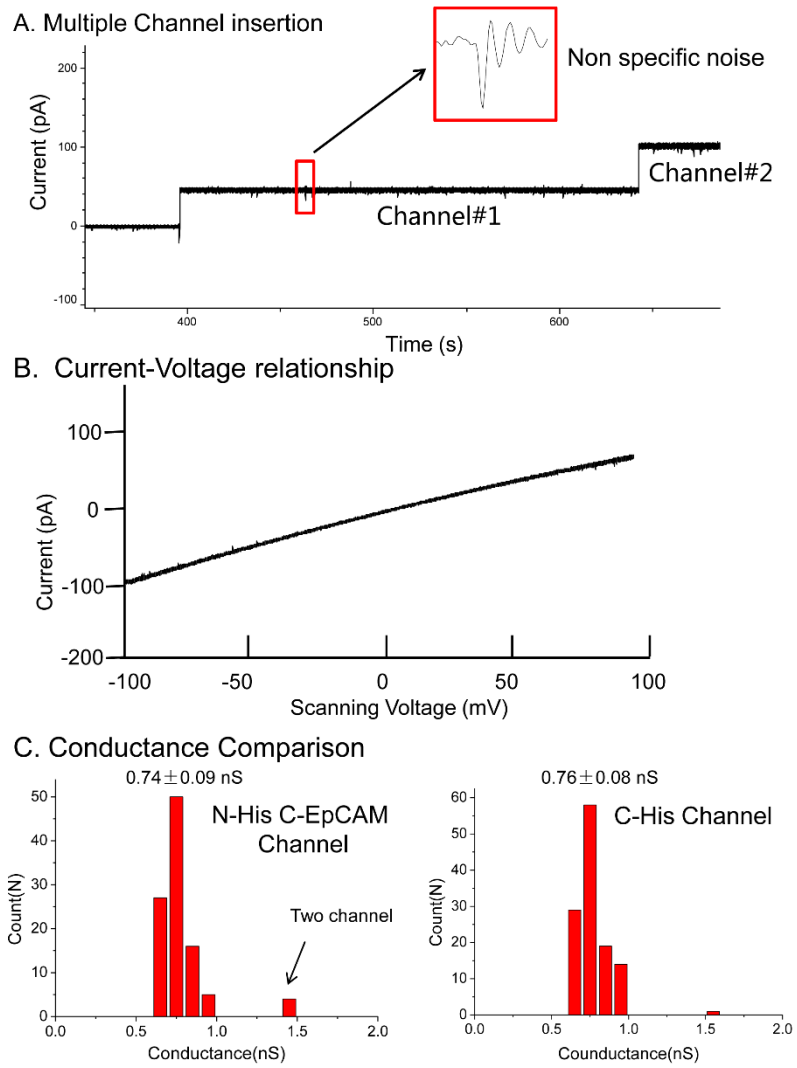



**Figure 2.2. Characterization of membrane-embedded EpCAM engineered phi29 connector channels. (A)** Continuous current trace showing multiple insertions of EpCAM reengineered connector channels into planar lipid bilayer. **(B)** Current voltage trace under a ramping voltage of -100 to 100 mV. **(C)** Conductance comparison of N-his C-EpCAM and C-his connector channel under the same buffer condition 75 mV.

***Real-Time Sensing of EpCAM Antibody Interactions with EpCAM Reengineered Connector Channels***

Under 0.2 M NaCl, 1 mM HEPES, pH 7.4 buffer solution, a series of blocking events was observed in the presence of EpCAM antibody in the cis-chamber. The binding of an EpCAM antibody to each probe induced stepwise blocks (every block represented a single molecule binding) in current (**Fig 2.3B-C**), with a corresponding decrease in conductance because of the physical blocking of the channel. One parameter used to characterize the binding events was the current blockage percentage, which represents the difference between the open connector channel and the current after EpCAM antibody binding. Current blockage percentage was calculated as follows: size of current blockage resulting from binding one EpCAM antibody to one connector channel divided by step size of the current for one connector insertion.

Two classes of current blockage signals were observed. The first class represented transient binding events, which may be induced by the temporary and reversible binding of an EpCAM antibody with the probe, shown as recoverable blockage signals (**Fig 2.3B**). The second class was the permanent binding events (**Fig 2.3C**), which represented tight binding between the EpCAM probe on the connector and the antibody. This second class was observed as discrete stepwise augmentation of current blockage. Both classes of blockage events resulted in ~20 pA reduction in current, which corresponds to 36.8±1.8% for transient events and 34.7±2.6% for permanent binding events (**Fig 2.4A-B**). The current blockage was not caused by the translocation of antibody, because the molecular weight of antibody, about 150 kDa, was too large to pass through the connector channel (487 kDa).

In addition, given the dimensions of an antibody (14.2 nm ×8.5 nm × 3.8 nm)(152), it is highly unlikely to translocate through the 3.6 nm connector channel.

Another parameter to characterize the antibody and probe interaction is the dwell time (τ), which is the duration of a blockage event. The transient events, on average, had a dwell time of ~30 ms, implying that the transient events are indeed caused by the temporary binding of the EpCAM antibody to the probe on the connector, rather than the unlikely event of translocation through the pore. The dwell time distribution can be used to determine the docking and undocking kinetics of antibody-probe interactions. The binding time of the antibody to the probe is defined as τoff, whereas the duration time between two consecutive blockade events is defined as τon. Figure 5 represents a statistical analysis of the binding kinetics based on more than one thousand transient binding events. The histograms of τon and τoff were fitted with a single-exponential distribution function, where is the time constant indicating the association and dissociation processes. The frequencies of association and dissociation are shown as a function of antibody concentration [P] (**Fig 2.5C-D**). The association rate constant was determined using the equation . The association frequency increases linearly with antibody concentration (Fig. 5C). The dissociation rate constant was determined using the equation . The dissociation frequency was independent of antibody concentration (**Fig 2.5D**). Based on the two rate constants, the equilibrium dissociation constant was (5.8±3.4) ×10-7 M.

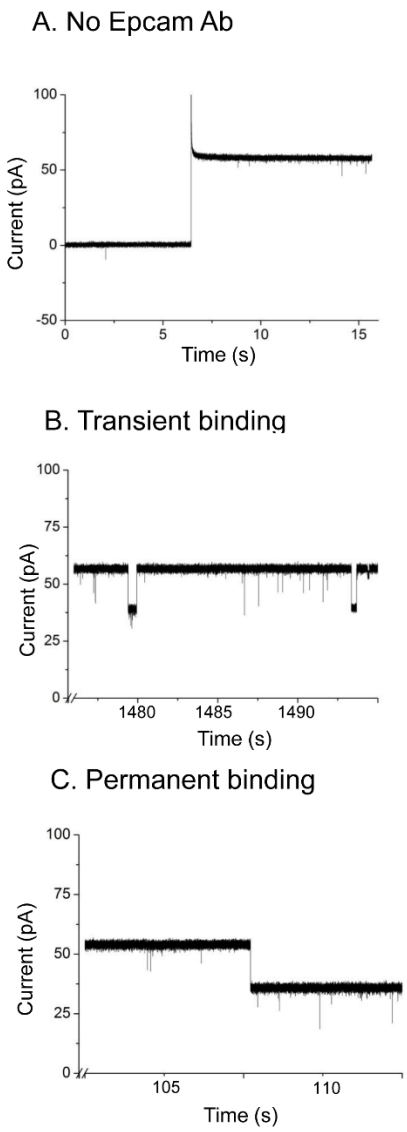



**Figure 2.3. Real time sensing of EpCAM antibody interactions with EpCAM engineered phi29 connector channels.** (A) Before addition of EpCAM antibody. (B) Transient binding events. (C) Permanent binding events.

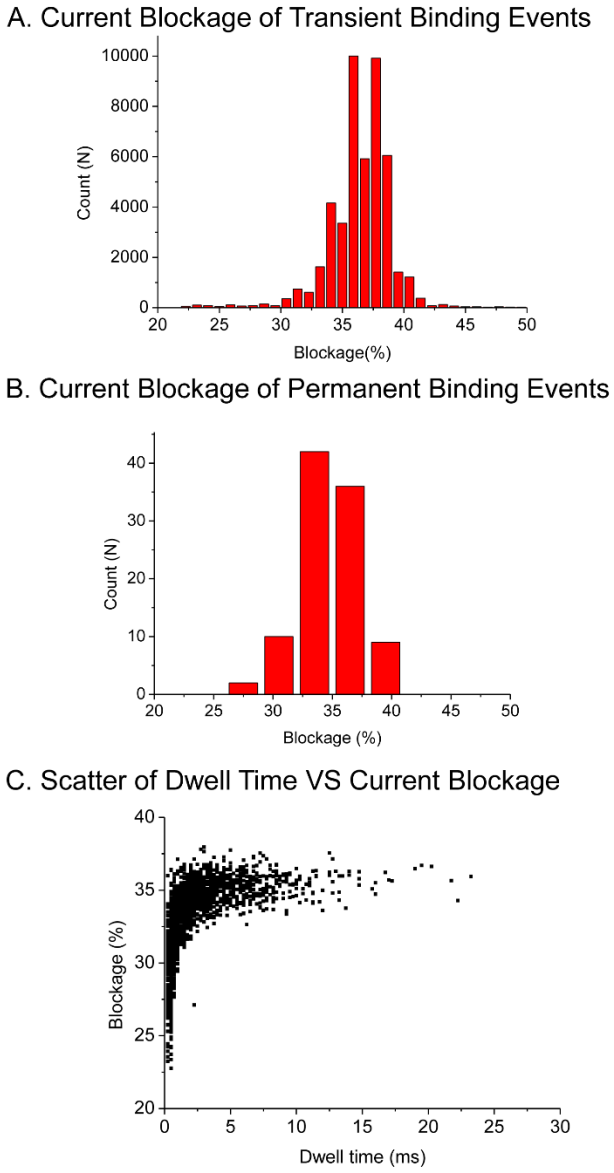



**Figure 2.4. Analysis of current blockage induced by EpCAM probe/antibody interaction. (A)** Histograms of transient current blockage. **(B)** Histograms of permanent current blockage events. **(C)** Scatter of dwell time versus current blockage.

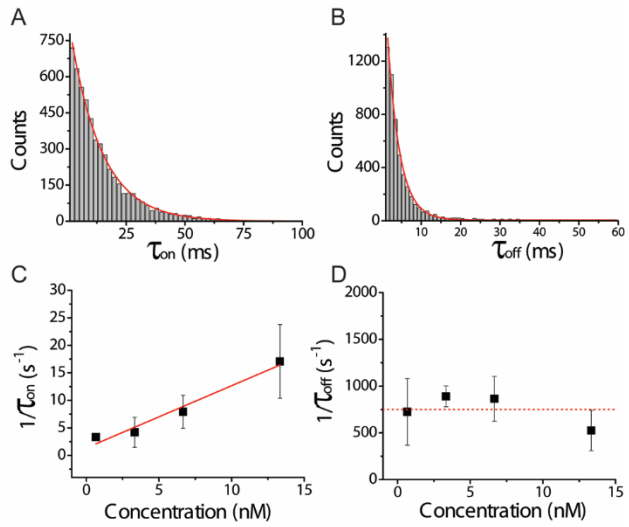


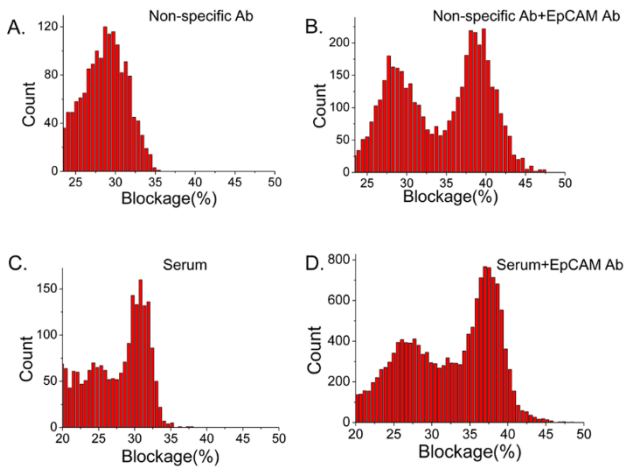

**Figure 2.5. Kinetic studies of EpCAM probe/antibody interaction based on transient dwell time events. (A)** Event histograms of current time traces τon, which is the time between two consecutive binding events. **(B)** Event histograms of current time traces $\tau_{off}$, which is the dwell time. **(C)** Frequency of association as a function of antibody concentration [P], which is linear fit to $(1/\bar{\tau}_{on}) = k_{on} \times [P]$, yielding the association rate $K_{on}$. **(D)** Frequency of association as a function of antibody concentration [P], which is a constant fit to $(1/\bar{\tau}_{off}) = k_{off}$, yielding the dissociation rate $K_{off}$.

***Selectivity of the EpCAM Probe for EpCAM Antibody in the Presence of Serum and Protein Impurities***

In almost all clinical scenarios, a patient's sample will not be highly purified and, most likely, will consist of serum protein and other components (e.g., a blood sample). To investigate how the impurities in the sample affect EpCAM antibody detection and to further push nanopore sensing techniques into the realm of clinical disease diagnosis, we performed the following experiments: 1) EpCAM antibody in the presence of a high concentration of non-specific antibody; and 2) EpCAM antibody in the presence of diluted serum sample (**Fig 2.6**). As the concentration of non-specific antibodies increased to 10 ng/ul, the current signal caused by non-specific interaction between the antibody and probe or channel was noted, and the current blockage distribution was determined to be 28.8±2.7%. Under the same buffer conditions containing a high concentration of non-specific antibody, we increased the EpCAM concentration to 4 ng/ul and noted another blockage peak, centered at 38.9±2.3%. The position of this new blockage peak was consistent with the previous blockage distribution of EpCAM antibody using the same buffer without non-specific antibody, indicating that this new blockage peak was caused by EpCAM interaction with the probe.

Serum sample is commonly used in clinical diagnosis. However, most diagnosis studies using nanopore techniques, particulally protein nanopore, have not been conducted in the presence of serum due to the fragile nature of the lipid bilayer, which greatly impedes the pace of applying the nanopore technique for clinical use. Here, we used 100-fold diluted serum with an electrolyte buffer that can maintain its stability for at least one hour. We first tested the noise level caused by impurity in the diluted sample in the presence of diluted

serum only. The current blockage distribution was 30.4±1.5 %. As the concentration of EpCAM antibody increased to 4 ng/ul, a new blockage peak was noted centered on 37.4±1.9 %, which corresponds to the previous EpCAM antibody blockage distribution. This data demonstrates that phi29 nanopore channels are capable of detecting individual molecular species in a complex mixture and can discriminate the signal from background events that are present due to the serum component.




| | Non specific peak (%) | EpCAM peak (%) |
|---|---|---|
| EpCAM Ab only (Transient) | N/A | 36.8±1.8 |
| Non-specific antibody | 28.8±2.7 | N/A |
| EpCAM Ab with Non-specific Ab | 28.8±2.6 | 38.9±2.3 |
| Serum only | 30.4±1.5 | N/A |
| Epcam Ab in serum | 27.6±4.9 | 37.4±1.9 |

**Figure 2.6: EpCAM antibody detection in the presence of high concentration of non-specific antibody and diluted serum. (A)** Histogram of current blockage events caused by high concentration of non-specific antibody. **(B)** Histogram of current blockage events caused by high concentration of non-specific antibody with EpCAM Ab. **(C)** Histogram of current blockage events caused by diluted serum. **(D)** Histogram of current blockage events caused by diluted serum with EpCAM Ab.

## DISCUSSION

The first step for detection of proteins or small molecules is to capture the analyte and gather evidence, such as fingerprints. Many sensing, detection, and diagostic techniques, such as biotin/streptavidin interactions, and microarray-based technologies, have been well developed, but also have a disadvantage: it is not possible to detect a single substrate molecule, regardless of the strength of the antibody affinity. In addition, the background noise will override the specific signal from the antigen/antibody complex at very low concentrations. Recent nanopore studies have detected single protein molecules and distinguished them from the other molecules with engineeredparticular receptor.(96,153,154) Compared to conventional protein detection methods, nanopore-based technology has the advantage of label-free, real-time sensing of individual molecules and in the presence of contaminants, all at a relatively low cost.

For the kinetic and thermodynamic studies of macromolecular interactions, nanopore offers several advantages over traditional methods, such as surface Plasmon resonance (SPR) and capillary electrophoresis (CE). For example, steric hindrance due to crowding of the binding sites occurs on the surface of an SPR chip, but does not occur in the nanopore setup.(155) Moreover, the molecular flow rate may interfere with the accuracy of the measurement in SPR and CE. For example, an analyte can rebind to the surface of a chip if the flow rate is too slow in SPR, and this can complicate the kinetic analysis.(155)

EpCAM is a transmembrane glycoprotein, which is highly overexpressed on many cancers. Its ectodomains can be released at levels ranging from 2-78 ng/ml into sera of cancer patients.(156) Studies have shown that the expression of EpCAM depends on many

factors, such as tumor type, disease stage, tumor microenvironment and host anti-tumor immunity.(157) In this study, we incorporated an EpCAM peptide into the phi29 connector channel. A linker with 6-glycines was included between the gp-10 connector and the EpCAM peptide probe to provide end flexibility. EpCAM antibody binding was then detected at nano-molar concentrations using single-channel conduction assays. The distinctive current signatures enabled us to analyze the docking and undocking kinetics of antibody-probe interactions and determine the Kd. However, one critical problem limits the application of nanopores for kinetic studies of single molecules: it is time-consuming to study high-affinity interactions. For example, if the ligand remains bound to the receptor for several minutes, several hours may be needed to acquire enough unbinding events. In our study, we observed two classes of current blockage signals: permanent binding and transient binding. For transient blockage events, dwell time lasts tens of milliseconds, suggesting weak interactions between EpCAM antibody and the probe. Therefore, in the kinetic assessment, we based our analysis on the transient events data, which may account for the relatively low Kd (10-7 M) compared with the value of typical EpCAM antibodies (10-8- 10-11 M). Here, the modification on the C-terminal end of phi29 connector channel has not changed the conductance and channel size compared with C-his connector channel (Fig. 2C). The channel has a diameter of 3.6 nm at the narrow end (N-terminal) and 6 nm at the wide end (C-terminal). Our small peptide antigen was fused to the wide end (C-terminal). Fusing the small peptide antigen to the wide end did not change the conductance of the channel, since conductance is determined by the narrowest end, which was not affected by the C-terminal conjugation with a peptide that sticks out.

Most recent research on applying nanopore techniques for sensing and detection has been performed under ideal conditions using only pure analytes in the presence of the detection system. However, in the majority of clinical samples, the trace analyte is always present in combination with a large proportion of impurities or in serum. To further test the detection capability of this engineered phi29 channel and push nanopore techniques for clinical utilization, we performed the EpCAM antibody detection in the presence of either high concentration non-specific antibody or diluted serum. We showed that the EpCAM antibody can be distinguished from the background events presented as either non-specific antibody or serum components.

**CONCLUSIONS:**

Our results demonstrate the feasibility of reengineering the phi29 connector channel for sensing probe-antibody interactions in real-time using single channel conduction assays. The Kd of EpCAM has been calculated from the docking and undocking kinetics of Ab-probe interactions. The signal of EpCAM antibody can be discriminated from the background events in the presence of non-specific antibody or serum. Our novel findings will inspire future studies to construct more sophisticated connector sensor systems capable of recognizing multiple analytes from ions, small molecules, peptides and small proteins.

**ACKNOWLEDGMENT:**

We thank YingYing Guo for the preparation of the phi29 connector channel pdb file. Research was supported by the NIH grant EB012135 (P.G.). The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH, NIBIB, or NIGMS. Funding to Peixuan Guo's Endowed Chair in Nanobiotechnology

position is by the William Fairish Endowment Fund. P. Guo is a co-founder of Kylin Therapeutics, Inc., and Biomotor and Nucleic Acid Nanotechnology Development Corp. Ltd.

## Chapter 3: Three-step Channel Conformational Changes Common to DNA Packaging Motors of Bacterial Viruses T3, T4, SPP1, and Phi29

This chapter was reproduced (with some modification) with permission from Wang S, Ji Z, Yan E, Haque Fand Guo P. "Three-step Channel Conformational Changes Common to DNA Packaging Motors of Bacterial Viruses T3, T4, SPP1, and Phi29" *Virology. 2016 May 12. pii: S0042-6822(16)30073-3. doi: 10.1016/j.virol.2016.04.015.* 

**INTRODUCTION:**

DNA translocation motors are ubiquitous in living systems (158). During replication, the genome of double stranded DNA (dsDNA) viruses is packaged into a preformed protein shell, referred to as the procapsid. This process requires a powerful, ATP-driven packaging motor. In many viruses, the motor involves a pair of DNA packaging proteins, a smaller auxiliary subunit is usually a protein oligomer that comes into contact with the dsDNA, and a larger one is an ATPase protein (159). In many dsDNA bacterial viruses and herpes viruses, the motor docks onto a structure called the portal or connector. Structural studies have shown that the portals in herpes virus and a variety of tailed bacterial viruses, such as phi29, SPP1, T4, and T3, share a similar cone-shaped dodecameric structure (**Fig 3.1**), even though their primary sequences do not display any significant homology.

In bacterial virus phi29, the portal is comprised of 12 protein subunits assembled into a truncated cone structure, with a diameter of 13.8 nm and 6.6 nm at the wide and narrow ends, respectively. The interior channel is 3.6 nm at the narrowest constriction (130). In SPP1, the assembled channel is a 13-mer structure, and the narrowest part is 2.77 nm in diameter (**Fig 3.1**) (57,58). The T4 portal exists as a dodecameric ring that is 14 nm

long and 7 nm wide, and an interior channel of ~3 nm in diameter (60). The T3 portal is a mixture of 12 and 13 subunits, depending on the protein expression conditions and other factors. The 12-mer version of the T3 portal is 14.9 nm in width, 8.5 nm in height and 3.7 nm in diameter for the internal channel (59).

The portal plays a critical role in genome packaging and the ejection process. During assembly, it acts as a docking point for the motor ATPase and a conduit for dsDNA entry. After DNA packaging, the portal serves as a binding site for the tail components in order to complete virion assembly. When bacterial viruses initiate infection, DNA is ejected through the coaxial channel of the portal and tail channel into the host cell. In the bacterial virus SPP1, the portal protein undergoes a concerted structural conformational change during its interaction with DNA (160). Recent results obtained using the membrane-embedded phi29 portal connector demonstrated that dsDNA moves in only one direction, i.e. from the external narrow end to the internal wide end, referred to as "one-way traffic"(110). Biological data from the ATPase studies combined with single molecule studies led to the conclusion that the DNA translocation of bacterial virus phi29 takes place via a "Push through a one-way valve" (161) or a "One-way revolution mechanism" in order to package DNA (162-164). The meaning of "Push" is in accordance with the findings in T4 that indicate a compression mechanism (165-167). Since the channels act like a one-way valve, an obvious question arises: how is dsDNA ejected during the course of infection if the channel is a one-way inward valve? Earlier studies demonstrated that the portal exercises conformational changes during the respective DNA packaging and ejection processes. For example, in the phi29 portal, conformational changes are inducible by DNA, pRNA or divalent metal ions (112,168,169). It was also reported that, the channel loop of

phi29 DNA packaging motor plays an important function near the end of packaging to retain the DNA (170). Cryo-EM imaging also revealed conformational changes of the connector in infectious virion in comparison with the free connector in vitro (171). However, none of these studies have yet shown conformational changes at the single molecule level.

Nanopore-based single molecule detection has attracted considerable attention across a number of disciplines due to its versatility of application. Examples include the detection of small molecules of chemicals, nucleotides, drugs and enantiomers, as well as larger polymers, such as PEG, polypeptides, RNA and DNA. One novel application was the insertion of the phi29 portal into an artificial membrane in order to serve as a robust nanopore (13) for single molecule detection (15) and disease diagnosis (105). The phi29 portal channel displays voltage-induced channel gating as well as a one-way traffic for dsDNA translocation during the course of DNA packaging (110,131,172). It has been reported that interaction of ligand with the C-terminal of the connector leads to the conformational changes in the phi29 connector channel, resulting in an altered current signal that have been utilized for detecting single antibodies as a very sensitive method for disease diagnosis (105). Here we report that the discrete conformational changes in the channel are common in bacterial viruses T3, T4, SPP1 and phi29. These observations support the idea that the one-way inbound channel is transformed into an outbound channel in preparation for DNA ejection (173).

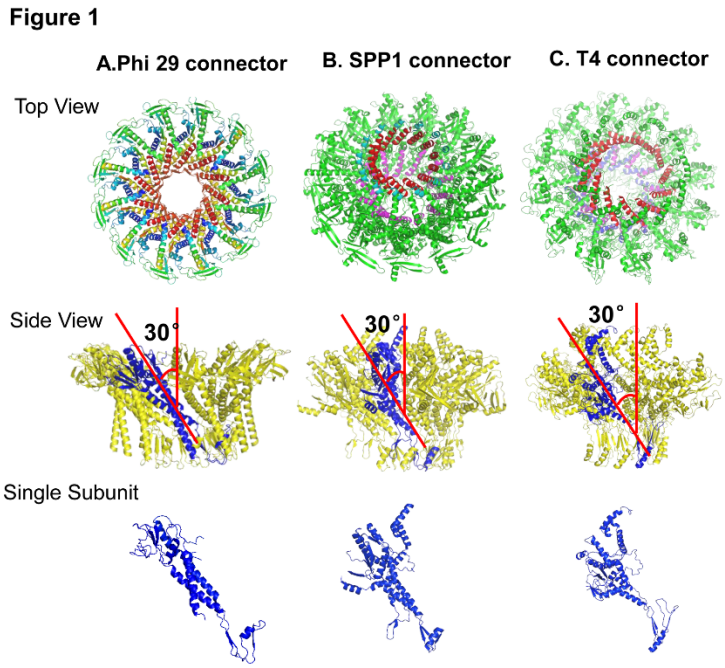



**Figure 3.1: Structures of phi29, SPP1, and T4 and T3 portal channels.** Top view, side view and single subunit of phi29 **(A)**; SPP1 **(B)**; and T4 **(C)** portal protein. Phi29 gp10 PDB: 1FOU; SPP1 gp6 PDB: 2JES;T4 gp20 PDB: 3JA7.

## MATERIALS AND METHODS:

### *Materials and reagents*

The phospholipid, 1,2-diphytanoyl-sn glycerol-3-phosphocholine (DPhPC) (Avanti Polar Lipids), n-Decane (Fisher), chloroform (TEDIA) were used as instructed by the vendors. If not specified, other reagents were purchased from Sigma. His binding buffer: 15% glycerol, 0.5 M NaCl, 5 mM Imidazole, 10 mM ATP, 50 mM Tris-Cl, pH 8.0. His washing buffer: 15% glycerol, 500 mM NaCl, 50 mM Imidazole, 10 mM ATP, 50 mM Tris-HCl, pH 8.0. His elution buffer: 15% glycerol, 500 mM NaCl, 500 mM Imidazole, 50 mM ATP, 50 mM Tris-Cl, pH 8.0.

### *Expression and purification of phi29, SPP1, T3, and T4 portals*

The expression and purification of phi29 portal followed the procedure reported previously (13,174). Gene 6 encoding for SPP1 portal protein gp6, and gene 8 encoding for T3 portal protein gp8 were synthesized (Genescript) and then cloned separately into pET3c vector between the NdeI and BamHI sites. A 6×His-tag was inserted at the C-terminus for purification. The resulting plasmids harboring gene 6 or gene 8 were transformed separately into E. Coli BL21(DE3) and a single colony was cultured in 10 mL Luria-Bertani (LB) medium overnight at 37 °C. The culture was transferred to 1 L of fresh LB medium and 0.5 mM IPTG was added to induce protein expression after the OD600 reached 0.5-0.6. After 3 hrs, cells were collected by centrifugation at 6000×rpm for 15 min and the pellet was resuspended in His Binding Buffer. Bacteria were lysed by passing through a French press and the clear supernatant was collected after centrifugation at 12000×rpm for 20 min and then loaded onto a His•Bind® Resin Column. SPP1 or T3 portal

protein was eluted from the His•Bind® Resin Column with His Elution buffer after several rounds of washing.

Gene 20 encoding for the T4 portal protein gp20 was amplified from the T4 genome and cloned into pET3c at the NdeI and BamHI sites (Keyclone). A 6×His-tag was introduced at the C-terminus for purification. Due to its hydrophobicity, T4 portal had a tendency to easily aggregate. Protein expression and purification methods were therefore modified (175). Plasmid pET3c harboring gene 20 was transformed into *E.Coli* HMS174(DE3) and a single colony was cultured in 10 ml LB medium overnight at 37 °C. The culture was transferred to 1 L of fresh LB medium and cultured until OD600 reached 0.5-0.6. IPTG (0.5 mM final concentration) was then added to induce T4 portal protein expression. The culture was transferred to 18 °C and incubation continued overnight. Cells were harvested by centrifugation at 6000×rpm for 15 min and resuspended in His binding buffer. Cells were lysed by passing through a French press. The cell pellet was collected after centrifugation at 12000×rpm for 20 min and resuspended in His binding buffer containing 1% N-Lauroylsarcosine for 20 min. The supernatant was collected after centrifugation at 12000×rpm for 1 hr and loaded to His•Bind® Resin Column and eluted after several rounds of washing. The purity of all final protein products was verified by 10% SDS-PAGE gel.

***Preparation of liposomes containing the phi29, SPP1, T4 and T3 portals***

All portal/liposome complexes were prepared following our reported procedures (13,174). Briefly, 0.5 mL of 1 mg/mL DPhPC in chloroform was added to a round bottom flask and the chloroform was evaporated under vacuum using a Rotary Evaporator (Buchi). The dried lipid film was rehydrated with 0.5 mL of portal protein solution containing 250

mM sucrose. Unilamellar lipid vesicles were obtained by extruding the lipid suspension through a 400 nm polycarbonate membrane filter (Avanti Polar Lipids).

***Portal insertion into planar lipid bilayer***

Procedures for inserting the portal connector into a lipid bilayer have been reported previously (13,136,174). Briefly, a thin Teflon partition with an aperture of 200 μm was used to separate the Bilayer Lipid Membrane (BLM) chamber into cis- and trans- compartments. The aperture was pre-painted with 5% (wt/vol) DPhPC in hexane solution. The cis and trans- chambers were filled with conducting buffer, 1 M KCl, 5 mM HEPES, pH 7.8. Then 20% (wt/vol) DPhPC in decane solution was used to form a planar lipid bilayer. After confirming the formation of the lipid bilayer, the portal/liposome complexes were added to the cis-chamber to fuse with the planar lipid bilayer to form the membrane embedded nanopore.

***Electrophysiological measurements***

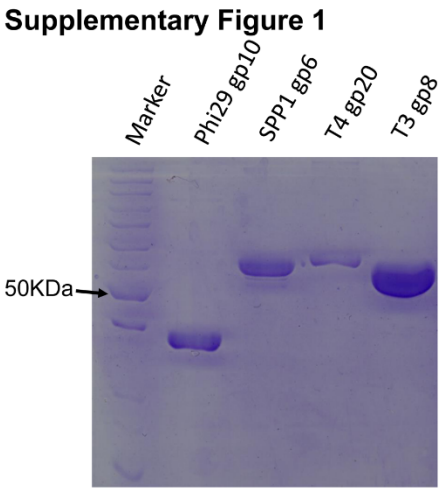
The stochastic nanopore sensing technique is based on the principle of the classical Coulter Counter or the 'resistive-pulse' technique (176). The portal is located in an electrochemical chamber, which is separated into two compartments filled with conducting buffers. Under an applied voltage, ions passing through the portal channel will generate current in pico-Ampere (pA) scale (137). When a charged molecule passes through the channel, it will generate transient current blockages due to volumetric exclusion of ions from the pore. Various parameters, such as the event dwell time, current amplitude, and unique electrical signature of the current blockages can be used either individually or in combination for detection.

A pair of Ag/AgCl electrodes inserted into both compartments was used to measure the current traces across the BLM. The current trace was recorded using an Axopatch 200B patch clamp amplifier coupled with the BLM workstation (Warner Instruments) or the Axon DigiData 1440A analog-digital converter (Axon Instruments). All voltages reported were those of the trans-compartment. Data was low band-pass filtered at a frequency of 5 kHz or 1 kHz and acquired at a sampling frequency of 2-20 KHz. PClamp 9.1 software (Axon Instruments) was used to collect the data, and Origin Pro 8.0 was used for data analysis.

**RESULTS:**

***Cloning, expression and purification of the portals of phi29, SPP1, T4 and T3***

Following the strategy previously used for the purification of phi29 portal (13,136,174), a 6×glycine linker was introduced between the portal coding sequence and 6×His-tag to provide end flexibility. Both SPP1 and T3 portals were soluble in the cytoplasm of E. Coli. The T4 portal showed a strong tendency to aggregate due to its hydrophobic nature. Therefore, 1% N-Lauroylsarcosine surfactant was added to the purification buffer to solubilize the protein (175). After purification to homogeneity, proteins were analyzed by 10% SDS-PAGE. The single protein subunit of the phi29, SPP1, T4 and T3 portals corresponded to their predicted sizes of 36 kDa, 56 kDa, 60 kDa and 59 kDa, respectively (**Supp. Fig 3. 1**).




**Supplemental Figure 3.1** Coomasie-blue stained 10% SDS-PAGE showing the molecular weight differences of single subunit of phi29 (36 kDa), SPP1 (56 kDa), T4 (60 kDa) and T3 (59 kDa) portal channels.

***Insertion of portal channels into lipid membrane for determining channel size using conductance measurements***

To incorporate phi29, SPP1, T4 and T3 portal proteins into a planar lipid bilayer, we adopted a two-step procedure described previously (13): reconstitution of the portal in liposomes, followed by fusion of protein/liposomes with the planar lipid bilayer to form the membrane-embedded portal channel. Experiments were carried out using 1 M KCl, 5 mM HEPES, pH 7.8 conduction buffer and 50 mV applied potential. Each current jump represented the insertion of one channel into the lipid bilayer. Since the fusion of the portal protein/liposome with the planar lipid bilayer is a random event, the time between independent insertion events varies. **Fig. 3.2A-D** provides representative results for the portals of the four phages. The channel conductance (derived from the ratio of measured current jump to the applied voltage) of phi29, SPP1 and T4 was determined to be 4.52±0.33 nS, 4.10±0.22 nS, and 3.03±0.37 nS, respectively (**Fig 3.2E-G**). T3 conductance distribution appeared as two peaks: 2.65±0.31 nS and 3.90±0.38 nS (**Fig 3.2H**). The conductance values correspond to the respective pore sizes of phi29, SPP1, T3 and T4 portal channels (**Fig. 3.2**).

Under a scanning voltage (–50 mV to +50 mV; 2.2 mV/s), the phi29, SPP1, T4 and T3 portal channels all display a linear Current-Voltage (I-V) relationship without voltage gating (**Fig 3.2I-L**). When 100 mV was applied, the phi29, T4 and T3 portal channels remained stable, but the SPP1 portal channel started to gate (data not shown). In addition, the SPP1 portal channel had a stronger tendency to close the gate under negative voltages compared to positive potentials (data not shown).

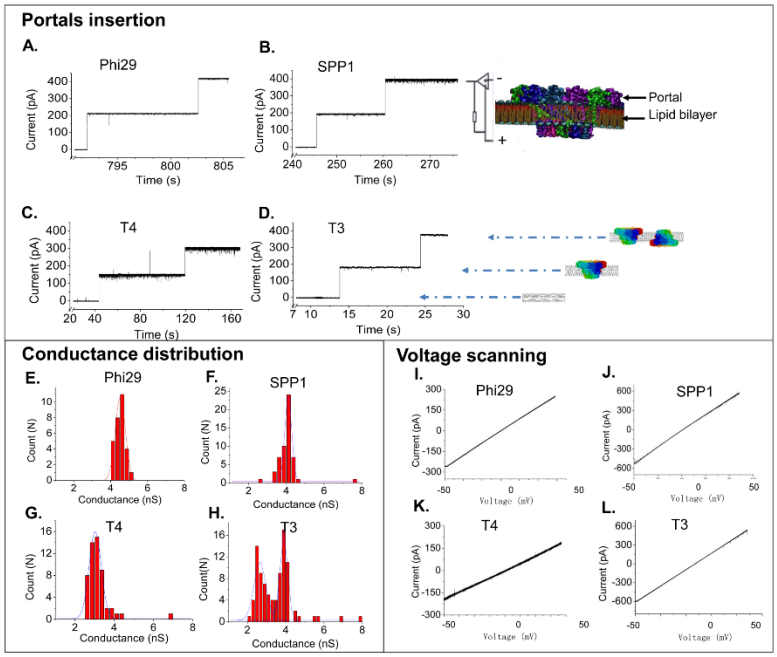



**Figure 3.2:** Representative current traces showing insertion of phi29 **(A)**, SPP1 **(B)**, T4 **(C)** and T3 **(D)** portal channels into planar lipid membrane. Applied voltage: +50 mV. Histogram showing the conductance distribution of phi29 **(E)**, SPP1 **(F)**, T4 **(G)** and T3 **(H)** portal channels. Applied voltage: +50 mV. Current-Voltage trace under a ramping potential (–50 mV to +50 mV; 2.2 mV/s) for phi29 (single channel) **(I)**, SPPI (two channels) **(J)**, T4 (one channel) **(K)** and T3 (three channels) **(L)** portals. Conducting buffer: 1 M KCl, 5 mM HEPES, pH 7.8

***Three step gating of phi29, SPP1, T4 and T3 portal channels***

When a higher voltage (>100 mV) was applied, three distinct steps of conformational changes of the channel were observed in all four portal channels. Conformational changes were reflected by a reduction in electrical current of 33%, 66% and 99% for the first, second, and third step, respectively (**Fig 3.3**). Three discrete step gating of the phi29, SPP1, T4 and T3 portal channels were observed under an applied positive voltage of +150 mV, +150 mV, +170 mV and +150 mV, respectively (**Fig 3.3A-D**). Similar phenomena were observed under negative voltages of –125 mV, –100 mV, –175 mV and –125 mV for the four portals, respectively (**Fig 3.3E-H**). These are the minimum voltages required for channel gating.

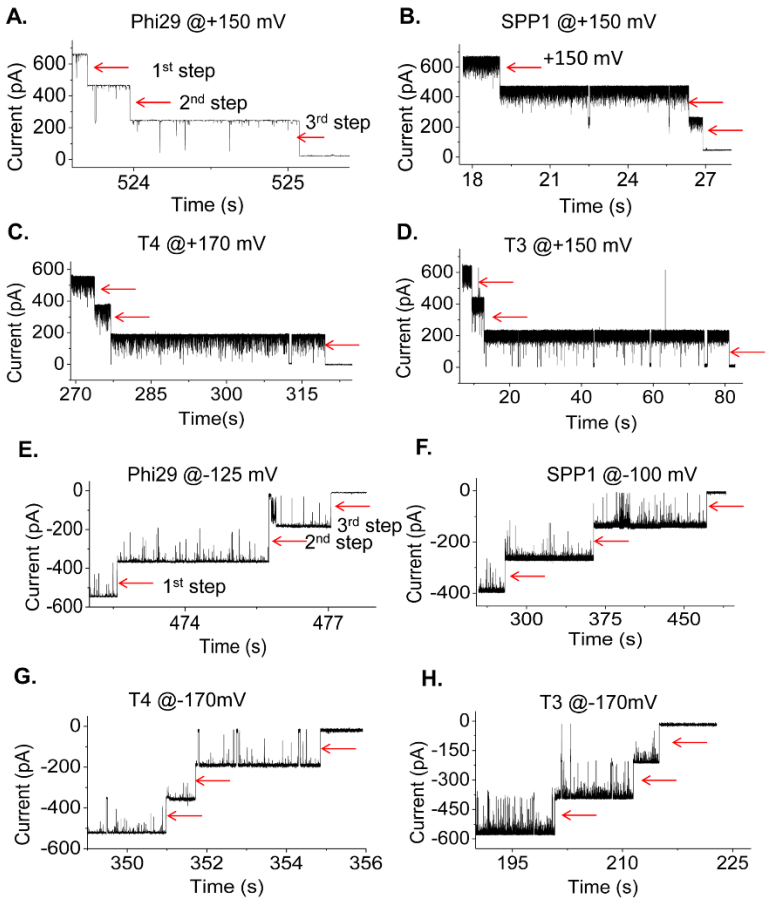



**Figure 3.3:** Three step gating associated with conformational changes of phi29 **(A)**, SPP1 **(B)**, T4 **(C)**, and T3 (**D**) portal channel under positive trans-membrane voltages. Three steps gating associated with conformational changes of phi29 **(E)**, SPP1 **(F)**, T4 **(G)** and T3 (**H**) portal channels under negative trans-membrane voltages.

## DISCUSSION

The polymorphism of portal complexes assembled from overexpressed genes of bacterial viruses has been reported for many years. Although it is believed that the T4 and SPP1 portals exist as dodecamers in their biologically active state , the stoichiometry of the overexpressed portal gene products in different bacterial viruses has been reported to vary from 11-mer to 14-mer (114,115,117,177,178). Several studies revealed that the T3 portal structure is a mixed population of 12 and 13 subunits (59). The diverse distribution of conductance for phi29, SPP1, T4, and T3 portals might represent various oligomeric states in these portal complexes. This is reflected by the two major peaks observed in the T3 conductance distribution (**Fig 3.2H**).

It has been shown that all portal channels of dsDNA bacterial viruses display a left-handed channel wall configuration to facilitate the one-way traffic of dsDNA into procapsid by a revolution mechanism without rotation (110,121,162,164). The one-way valve mechanism is consistent with the findings of genome gating in SPP1, albeit gating mechanism proposed by these authors is based on the analysis of the channel structure after the completion of DNA packaging instead of during translocation (160). The finding of the "push through a one-way valve" mechanism (158,161) raises the question of how dsDNA is ejected during infection if the channel only permits dsDNA to translocate in one direction. We believe that during dsDNA translocation, the interaction of the dsDNA with the channel wall and the procapsid component next to the portal will trigger conformational changes of the portal. Therefore, the left-handed portal channel, which facilitates dsDNA advancement in one direction, will transition to a neutral or right-handed configuration in three steps to facilitate DNA ejection after DNA packaging is complete (164).

Such conformational changes of portal proteins, as proposed above for ejection of the packaged dsDNA, have previously been proposed (131,134,171). Portal gate closing has been reported in SPP1 (179) and speculated in T4(60). Moreover, it was reported that SPP1 portal undergoes a concerted reorganization of the structural elements of its central channel during interaction with DNA. Structural rearrangements and gate closing were reported to associate with protein-protein and protein-DNA interactions, and a diaphragm-like mechanism for channel reduction and gate closing has been proposed (160). The changes with discrete steps might be considered as the analogy of a camera lens by suggesting discrete f-stops, like f4.5, f8, f16, f32. However, the diaphragm proposal is difficult to reconcile with the data implying a right-handed twisting of the connector structure while comparing the free connector with the structure of the connector in the DNA-filled virion (171). The finding of the common discrete 3-step conformational change in T3, T4, SPP1 and phi29 implies a universal property of all portals. It is possible that the three gating steps may also correspond to the quantized steps of partial genome ejection observed in T3 (180), and the partial packaging intermediates observed in phi29 (181).

**CONCLUSIONS**

The motor channel of T3, SPP1, T4, and phi29 all display three discrete steps of voltage gating resulting from channel conformational changes. The result suggests that the one way inbound channel during the DNA packaging process is transformed into an outbound channel prepared for DNA ejection during the host infection.

**ACKNOWLEDGMENT:**

We thank Zhengyi Zhao, Daniel Binzel, Daniel Jasinski and Mario Vieweger for insightful comments. The research was supported by NIH grant R01EB012135 and funding to Guo's

Endowed Chair in Nanobiotechnology position at University of Kentucky is from the William Fairish Endowment Fund. Peixuan Guo is cofounder Biomotor and RNA Nanotechnology Development Corp. Ltd.

# Chapter 4: Channel of SPP1 DNA Packaging Motor for Real Time Kinetic Analysis of Peptide States

This chapter (with some modification) is in submission at *Biomaterials*. Special thanks to Zhengyi Zhao for help and assistance in preparation of data for figures 4.4 and Dr. Zhi Zhou for help in preparation of data for figure 4.6 and Table 4.1.

## INTRODUCTION:

Living systems contain wide varieties of nanomachines with diverse structures and functions. The ingenious design of viral DNA packaging motors and their intriguing mechanism of action has provoked a wide range of interests among scientists in many different areas(158,182,183). Tailed bacteriophages are viruses that infect and reproduce within a bacterium, and widely exist in the biosphere. The portal protein is one of the essential components of the viral packaging motors with a turbine-like shape(184,185). SPP1 is a dsDNA phage that infects Bacillus subtilis. The DNA packaging motor of SPP1 consists of a terminase composed of small (gp1) and large (gp2) subunits, portal protein gp6, and a two head completion proteins gp15 and gp16(179,186) that powers the encapsulation of 45.9 kbp genomic DNA(177). The central core of the SPP1 motor is a portal channel, termed connector, which is composed of 12 or 13 copies of protein gp6 (57,186). Explicit engineering of the SPP1 portal protein is possible due to its available crystal structure (57,58). The connector has an overall diameter of 16 nm and a height of 10.5 nm. The narrowest constriction of the internal channel is ~3 nm (Fig. 1)(57,58).

Nanopore technology has recently emerged as a real-time and high-throughput single molecule detection method, holding great potential for sensing a wide range of

analytes, molecular diagnostics and DNA sequencing applications(4,16,17,138,139,187-190). Solid state nanopores generated by microfabrication generally have less reproducible pore sizes and lack chemical and location selectivity(137,191). Protein nanopores harvested in bacteria are homogenous in size and can be functionalized with probes, but commonly used nanopores such as α-hemolysin, MspA, and aerolysin have an internal channel diameter of only ~1.3 nm(40,73,192). Larger protein pores such as ClyA(193) and FhuA(194) are being developed for single molecule analysis. In search for alternative larger sized channels, we previously developed membrane-embedded phi29 motor channels(13,56,174) for single molecule sensing of nucleic acids(13,56,110,112,172,195), chemicals(15), or binding assays of antibody based on channel conformational change(105,110). Herein, we engineered the SPP1 motor channel and inserted it into a lipid bilayer to serve as a new nanopore platform for detecting the translocation and dynamics of peptide oligomeric states at single molecule level.

Elucidating the oligomeric states of proteins and peptides is critical for understanding their biological functions. A wide range of biophysical methods, such as X-ray crystallography, NMR, Circular Dichroism, Dual Polarization Interferometry, and Mass Spectrometry have been used to investigate folding and dynamic structural changes of peptides and proteins. However, these methods require expensive instrumentation and specialized labor. Nanopores offer an attractive alternative as they are intrinsically single molecule in nature requiring ultra-low sample volumes, are label free, amplification free, and function using a simple detection process requiring no specialized expertise. While the translocation of DNA and RNA have been studied extensively in biological nanopores(73,137,138,196), studies on translocation of protein or peptides are beginning

to emerge(197-201). The proteome can be an accurate and direct indicator of current health status of patients(202). For example, early diagnosis and monitoring the changes of amyloid-β peptide and α-synuclein are critical for the management of Alzheimer's and Parkinson's disease(203).

Several studies demonstrating peptide translocation using nanopores have been published(9,49,204-215), but quantitative analysis and translocation validation remain challenging, since there is no technique available for amplifying protein substrates as in DNA or RNA qualification using PCR. The availability of only trace amounts of peptide or protein for analysis after translocation is far beyond the sensitivity threshold of classical protein detection methods. In this report, we used single molecule fluorescence microscopy to validate data obtained from resistive pulse technique to quantitatively study peptide translocation through SPP1 connectors and elucidate the structural conformations of peptides at the single molecule level. Due to the available of crystal structure, this new nanopore can be explicitly engineered via site directed mutagenesis, and has the potential to be applied for biomarker analysis, early disease diagnosis and even protein sequencing.

## MATERIALS AND METHODS:

### *Materials*

The phospholipid 1, 2-diphytanoyl-sn-glycerol-3-phosphocholine (DPhPC) was obtained from Avanti Polar Lipids, Inc. Organic solvents (n-decane and chloroform) were purchased from Fisher Scientific, Inc. and TEDIA, Inc., respectively. TAT peptide was custom-ordered from GenScript, Inc. All other reagents were purchased from Sigma or Fisher, if not specified.

### *Cloning and Purification of the SPP1 connector*

Gene gp6 encoding SPP1 portal was synthesized and cloned into PET3 vector between NhdeI and BamHI by GenScript, Inc. His-tag was inserted into the C-terminal of the connector for purification. Then the plasmid was transformed into BL21 (DE3) for expression and purification. The purification procedure has been described previously(13,174). Briefly, the transformed bacteria were cultured in 10 mL LB medium overnight at 37°C. The bacteria were transferred to 1 L of fresh LB medium. When the OD600 reached ~0.5-0.6, 5 mM IPTG was added to the medium to induce protein expression. The bacteria were collected by centrifugation after 3 hours continuous culture. Bacteria was lysed by passing through French press. Lysed cell was separated by centrifugation and supernatant containing expressed protein was collected. The supernatant was purified with Nickel affinity chromatography (Novagen) (133). His Binding Buffer (15% glycerol, 0.5 M NaCl, 5 mM Imidazole, 10 mM ATP, 50 mM Tris-HCl, pH 8.0), and the cleared lysate was loaded onto a His•Bind® Resin Column and washed with His Washing Buffer (15% glycerol, 0.5 M NaCl, 50 mM Imidazole, 50 mM Tris-HCl, pH 8.0). The His-tagged connector was eluted by His Elution Buffer (15% glycerol, 0.5 M NaCl, 0.5 M Imidazole, 50 mM Tris-HCl, pH 8.0). The purified connector from chromatography was further purified by a 15-35% glycerol gradient ultracentrifugation at 35K for 2 hrs. Before incorporating into liposomes, the purified connector was dialyzed against buffer (0.5 M NaCl, 50 mM Tris-HCl, pH 8.0) to remove excess glycerol.

***Insertion of the connector into planar lipid bilayer***

The protocol for the incorporation of connectors into lipid bilayer has been reported(13,56,110,131,174). Briefly, 150 μL of 10 mg/mL DPhPC lipids in chloroform was placed in a round bottomed flask and the chloroform was evaporated under vacuum

using a rotary evaporator (Buchi). The dehydrated lipid film was then rehydrated with 300 μL buffer containing 250 mM sucrose and purified connectors (0.5-1 mg/mL). The suspension was thoroughly vortexed and maintained at 45 °C (above the phase transition temperature of DPhPC) followed by extrusion through 400 nm polycarbonate membranes (Avanti Polar Lipids) to generate uniform sized liposome/connector complexes.

Planar bilayer lipid membranes (BLMs) were generated in a BCH-1A horizontal BLM cell (Eastern Scientific). A Teflon partition with a 200 μm aperture was placed in the apparatus to separate the BLM cell into cis- (top) and trans- (bottom) compartments. The aperture was pre-painted with 0.5 μL of 3% (w/v) DPhPC in n-hexane. A conducting buffer (1 M KCl, 5 mM HEPES, pH 7.8) was added to both the top and bottom compartments of the BLM cell, and Ag/AgCl electrodes were placed in the buffer of each compartment. The electrode in the trans-compartment was connected to the headstage of an Axopatch 200B amplifier (Axon Instruments), and the electrode in the top compartment was grounded. A planar lipid bilayer was formed by painting the aperture with 0.5 μL of 3% (w/v) DPhPC in n-decane. 1 μL of the diluted Liposome/connector complex was added to the cis-compartment directly to fuse with the planar lipid bilayer to generate membrane embedded nanopores.

***Electrophysiological measurements***

The headstage and Axopatch 200B patch clamp amplifier were connected to a DigiData 1440 analog-digital converter (Axon Instruments, Inc.) to monitor and record electrochemical currents through BLMs(13,56,174). The current recordings were low-pass filtered at a frequency of 5 kHz. The sampling frequency was 200 kHz in all experiments, unless otherwise specified. The data were recorded a with pClamp 9.1 software (Axon

Instruments, Inc.), and analyzed with the Clampfit module of pClamp 9.1 and OriginPro 8.1 (OriginLab Corporation).

***Peptide translocation experiments***

TAT peptide (Cys-Tyr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg) with a final concentration of 23 μg/mL was premixed with the conducting buffer before the insertion of connector. For quantitative validation, Cy3 fluorophore was conjugated to TAT peptide by sulfhydryl-maleimide chemistry and purified by HPLC. Multiple channels were inserted into bilayer lipid membrane and then Cy3-TAT peptide was added into cis-chamber. Samples were collected from the trans-chamber after 0, 20, 40, and 60 min. The errors represent mean ± standard determined by Gaussian fit of the data.

Single molecule fluorescence imaging: Samples collected from the patch clamp were incubated with the ozone pre-treated glass substrate for 10 min before imaging to ensure sufficient adsorption. A 532 nm green laser was used for the excitation of Cy3. A 60× objective (N.A. = 1.4, oil immersion) was used for fluorescence imaging. The signals were recorded using Andor iXon 887 V electron multiplied CCD camera. Images were taken with an exposure time of 500 ms. The number of spots in the images were counted using iSMS software(216). The errors represent mean ± standard deviation from three independent imaging from one experiment. Three independent experiments were performed and similar trend was obtained as shown in **Fig. 4.5.**

**RESULTS AND DISCUSSION**

***Characterization of SPP1 reengineered nanopore embedded in a lipid bilayer***

Structural analysis revealed that the central region of the SPP1 connector shows slight hydrophobicity compared with the flanking regions at the N- and C-terminal ends which

are more hydrophilic (**Fig 4.1a-b**). We reengineered the SPP1 connector by inserting a 6×His-tag at the C-terminal end along with a 6×glycine linker for end-flexibility. The presence of His-tag enhanced the hydrophilicity of the C-terminal, thus making the hydrophilic–hydrophobic–hydrophilic layers of the connector more distinct, which is necessary to mimic the lipid bilayer architecture. After His-tag column purification, the purified protein was further purified by 15-35% glycerol gradient ultracentrifugation to verify whether SPP1 subunit could assemble into the channel complex (**Suppl. Fig 4.1**). Bovine serum albumin (BSA) with molecular weight of 60 KDa, which is almost equal to a single subunit (58 KDa) of the SPP1 connector, centered at fraction 27, whereas the SPP1 portal complex was at fraction 17, indicating the formation of higher molecular weight complex.

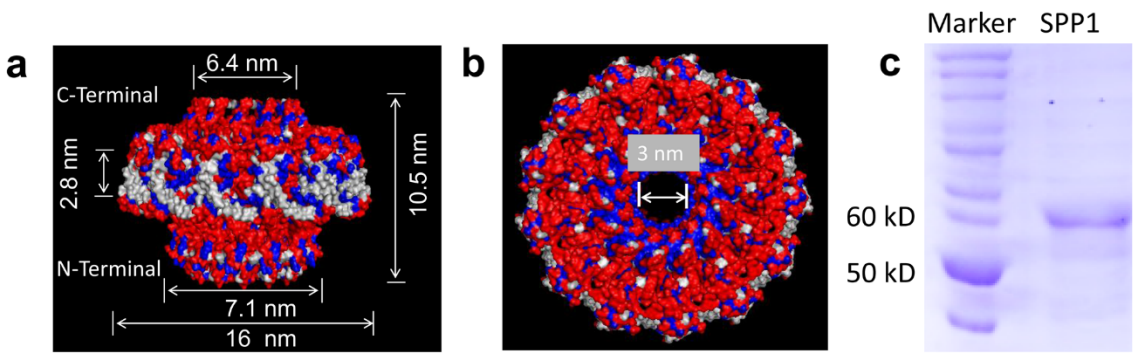
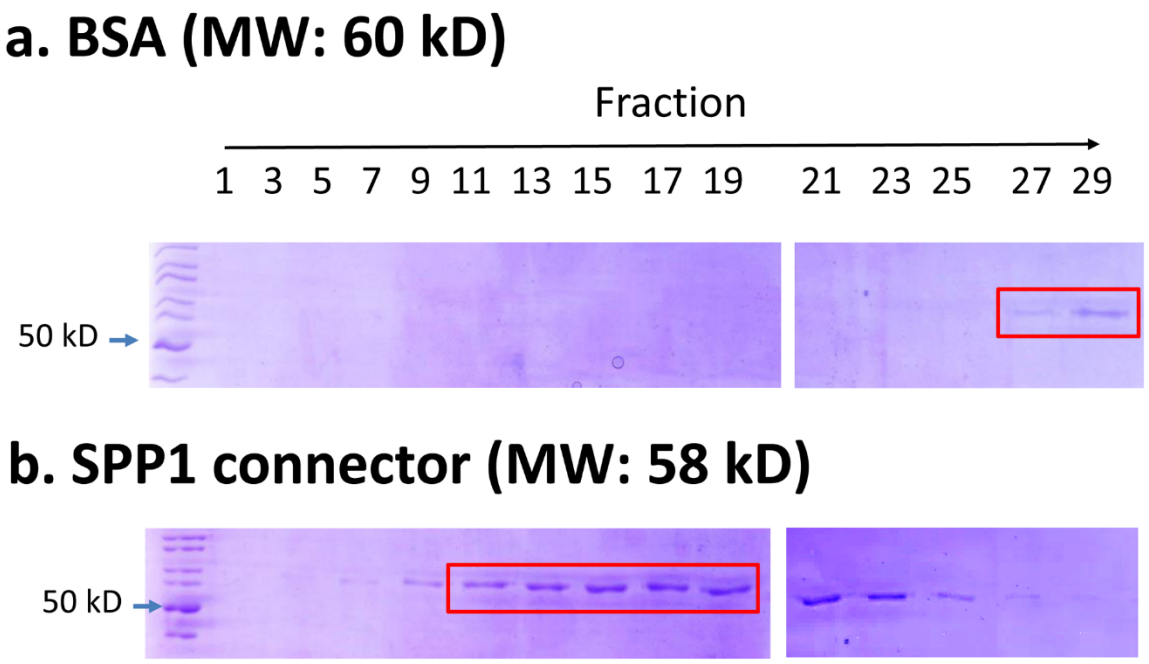



**Figure 4.1**: **Structure of the channel of SPP1 DNA packaging motor. (a)** Side view and **(b)** top view showing hydrophilic (red), hydrophobic (blue) and neutral (white) amino acids; and dimensions of the channel. PDB: 2JES. **(c)** Coomasie-blue stained SDS PAGE showing the purified SPP1 channel subunits gp6.




**Supplemental Figure 4.1**: Assembly of SPP1 gp6 subunits into 12 subunit connector complexes as shown by 15-35% glycerol gradient ultracentrifugation. BSA was used for size control.

To incorporate SPP1 portal channel into planar lipid membranes, a two-step procedure was employed(13). DPhPC lipids in chloroform were first dehydrated to remove organic solvents and then rehydrated with buffer containing 250 mM sucrose and purified connectors. The multilamellar lipid-connector suspension was then extruded through 400 nm polycarbonate membrane filters to generate uniform unilamellar liposomes with the connector embedded in the membrane. The resulting liposome-connector complex was fused with a planar lipid membrane to generate planar membrane-embedded SPP1 nanopore. Since the connector is neither a membrane protein nor an ion channel, direct incubation of the connector with a planar lipid bilayer did not result in connector insertion into the lipid bilayer. Single channel conductance assay was performed to measure the electrophysiological properties of membrane-embedded SPP1 connectors. The connector insertion steps were observed as distinct stepwise increase in conductance as revealed in a continuous current trace (**Fig 4.2a**). The insertion of single portal channel results in ~200 pA in current jump under an applied potential of 50 mV in conducting buffer (1 M KCl, 5 mM HEPES, pH 8). Occasionally ~400 pA current jumps were observed, attributed to simultaneous insertion of two connectors. The average conductance of reengineered SPP1 connectors is 4.27±0.27nS (**Fig 4.2b**). The conductance is uniform without displaying any voltage gating phenomena under the reported conditions of ±50 mV (Fig 4.2c). At voltages greater than ±100 mV, SPP1 connector displayed discrete stepwise gating of the channel (217).

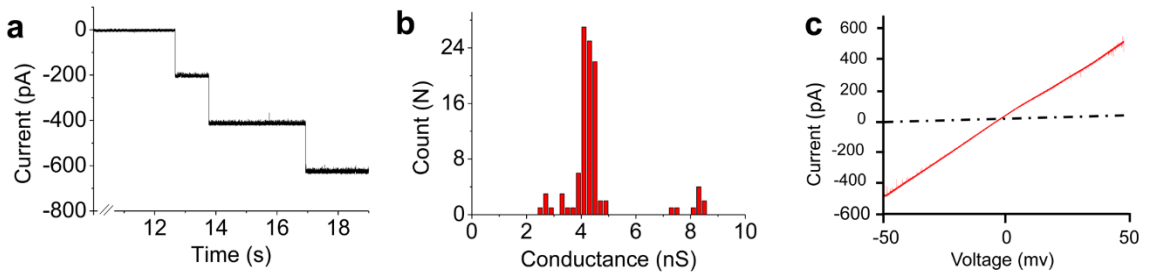



**Figure 4.2**: **Electrophysiological properties of membrane-embedded SPP1 connector**. **(a)** Current trace showing the insertion of SPP1 connector into the planar membrane with a characteristic step size of ~200 pA at –50 mV. **(b)** Conductance distribution based on 104 insertion events. **(c)** Current-Voltage trace acquired from –50 → +50 mV. Buffer: 1 M KCl, 5 mM HEPES, pH 8.

***Characterization of peptide translocation through reengineered SPP1 connector***

A positively charged 12 amino acid TAT peptide was used in the translocation studies with sequence Cys-Tyr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg. Due to the presence of cysteine at the N-terminus, the peptide forms a dimer by disulfide bond under physiological conditions, which is confirmed by mass spectroscopy. In the absence of peptide, the current trace was quiescent. In contrast, when the peptide was premixed with the conducting buffer in both *cis*- and *trans*-chambers, a burst of transient blockage events were observed immediately after the insertion of connector in the lipid membrane (**Fig 4.3a**). As the peptide concentration was increased from 1.25, 2.5, 3.75, to 5 μg/mL, the density of current blockage events increased with a corresponding linear increase in peptide translocation rate from 2.33±1.54, 7.5 ±1.66, 10.17±3.58, to 16.96±6.01 events per second (**Fig. 3b**). One parameter used to characterize the translocation was the current blockage percentage, calculated as the ratio of current blockade resulting from peptide translocation to the open current of one portal channel, expressed as [ 1 – ($Current_{peptide}$ / $Current_{open\_channel}$) ]. The distribution of current blockage was relatively broad with a major peak centered around 55.1±3.0%, determined by Gaussian fit of the data (**Fig 4.3c**). Another parameter was the dwell time (τ), the time taken for the peptide to traverse from one end of the connector to the other end. The dwell time distribution followed an exponential decay with a rate constant of 0.84±0.09 $ms^{-1}$ (**Fig 4.3d**).

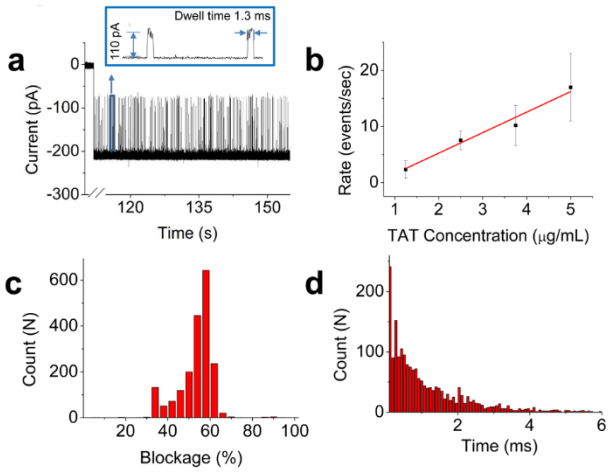



**Figure 4.3**: **Peptide translocation through SPP1 connector**. **(a)** Current trace showing a burst of current blockage events with characteristic current amplitude and dwell time indicating the translocation of TAT peptides. A representative event is shown in insert. TAT peptide concentration: 0.5 μg/mL **(b)** Rate of peptide translocation as a function of peptide concentration (n = 3). **(c)** Histogram of current blockage percentage from 1978 events. **(d)** Dwell time of peptide translocation events fitted with a single exponential function. Applied voltage: 50 mV; Buffer: 1 M KCl, 5 mM HEPES, pH 8.

***Quantitative validation of peptide translocation by single molecule fluorescence imaging***

To validate the translocation of peptide through SPP1 connector, we conducted single molecule fluorescence imaging of samples collected from patch clamp experiments. HPLC purified Cy3 conjugated TAT peptide at a final concentration of 2.5 ng/μL was added to the *cis*-chamber after stable insertion of SPP1 connectors in the lipid bilayer. Under an applied negative trans-membrane voltage, the positively charged peptide translocated through the portal channel to the *trans*-chamber. 50 μL samples were collected from the *trans*-chamber at 0, 20, 40, and 60 minutes after addition of Cy3-TAT peptide and loaded onto glass coverslips. The positively charged peptide can bind to the negatively charged glass surface through charge-charge interactions and appear as individual fluorescent spots (**Fig 4.4a**). Fluorescence imaging revealed that the number of Cy3 spots in the field of view increased over the time course of 60 mins (**Fig 4.4b**). In contrast, in a control experiment under the same conditions but in the absence of portal channel, very few Cy3 spots were observed, compared to the sample containing SPP1 connector (**Fig 4.4b**). Since TAT is a membrane penetrating peptide, it is conceivable that a small fraction could potentially pass through the lipid bilayer, contributing to the small increase in the fluorescence background signal.

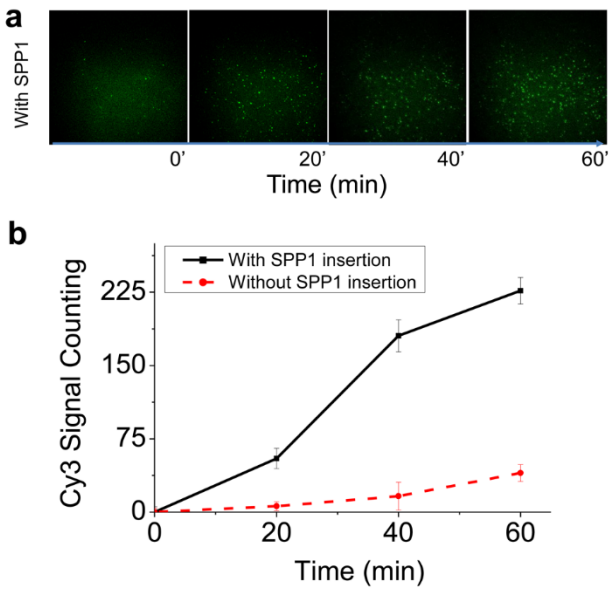



**Figure 4.4**: **Single molecule fluorescent images validating TAT peptide translocation**. **(a)** The upper row is the image showing the detection of Cy3-labeled TAT peptide from the fractions collected from patch clamp at 0, 20, 40 and 60 mins. Excitation λ: 532 nm; laser power: 5 mW; 60× objective (N.A. = 1.4, oil immersion); Exposure time: 500 ms. **(b)** Quantitative analysis showing the increase in Cy3-TAT peptide signal in presence of SPP1 connector compared to control without connector. The errors represent mean ± standard deviation from three independent imaging from one experiment. Three independent experiments were performed and similar trend was observed. Applied voltage: 50 mV; Buffer: 1 M KCl, 5 mM HEPES, pH 8.

**Kinetic analysis of TAT conformational status in real time**

The focus of this study was to investigate whether the SPP1 motor channel can be used for kinetic analysis of peptides in real time. TAT peptide with and without a cysteine was used as a model system to study oligomeric states. Structural analysis of TAT sequence using the computer program PEP-FOLD (218) revealed that without forming disulfide bond, the TAT only adopts one α-helical conformation (**Fig 4.5**). However, in the presence of a cysteine at the N-terminus, the peptide can be oxidized into a dimer by forming a disulfide bond between two N-terminal cysteines and exhibiting two parallel α-helices or extended linear conformation. The presence of two states, the oxidized and the reduced conformations, were confirmed by translocation studies with the SPP1 portal channel. Under an oxidized state, the current blockage distribution by TAT displayed a major peak, centered at 55.1±3.0% (**Fig 4.5a**), with a minor peak centered at 28.5±1.9%. However, after adding the reducing agent TCEP [(tris(2-carboxyethyl)phosphine)] to break the disulfide bond, the blockage distribution significantly changed. The peak of 55.1±3.0% decreased significantly over time, while the majority of the blockage distribution shifted to 28.5±1.9%. This indicated that most of the peptide passing through the channel is in its single α-helical conformation (**Fig 4.5b**). Current blockage distribution of the reduced TAT peptide was similar to Cy3-TAT signature with TCEP treatment, since the conjugation of Cy3 prevented the disulfide bond formation and resulted in similar blockage, representing the single TAT helix (**Fig 4.5c**).

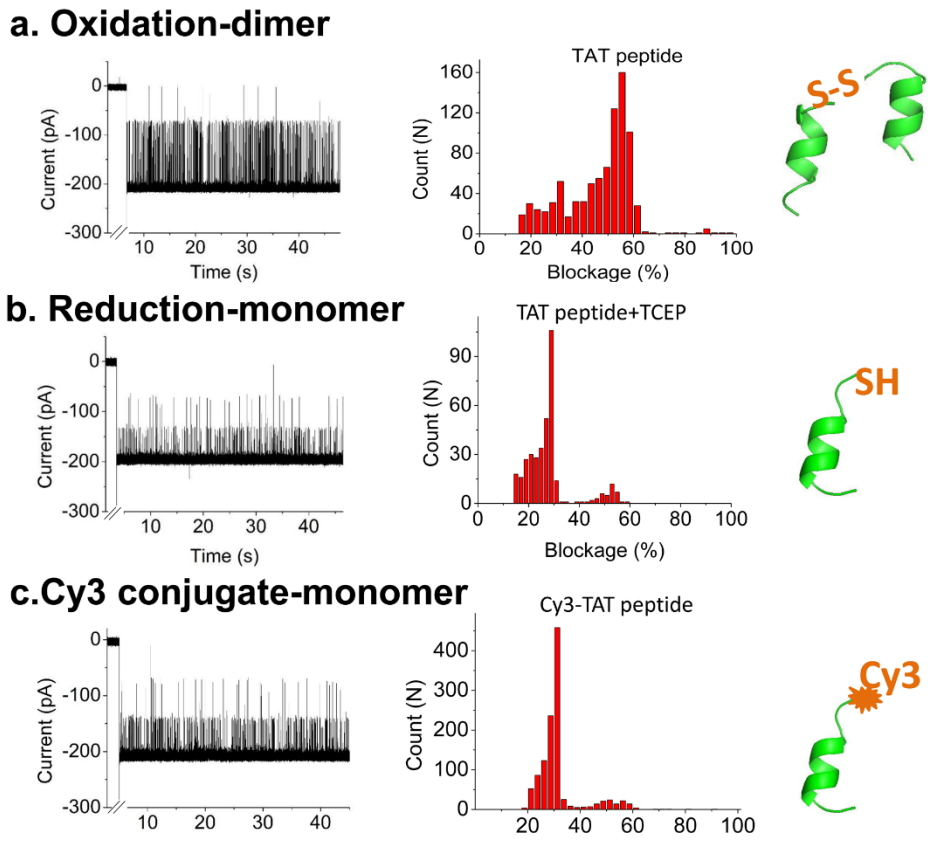



**Figure 4.5**: **Determining the conformational states of TAT peptide**. Current trace (left), current blockage distribution (middle) and conformation (right) for **(a)** oxidized dimer state of TAT peptide; **(b)** reduced monomer state of TAT peptide; and **(c)** Cy3-conjugated TAT monomer. Applied voltage: 50 mV; Buffer: 1 M KCl, 5 mM HEPES, pH 8. Total number of events: 858 in A; 367 in B and 1128 in C.

To quantify the kinetics parameter, we further examined the conformational changes of TAT peptide upon addition of TCEP and recorded the change of the current blockage profile in real time (**Fig 4.6**). It was found that current blockade signature progressively shifted from predominantly 55.1±3% (oxidized dimer) to 28.5±1.9% (reduced monomer), representing the real-time kinetic process of disulfide bond reduction (500 μM TCEP, **Fig 4.6A**). The TCEP reduction process can be described as a pseudo-first order reaction due to the nearly infinite TCEP amount and the reaction equation is described as follows:

$$\begin{array}{cccc} \mathrm{TAT-TAT+TCEP} & \xrightarrow{H_2O} & \mathrm{2\,TAT+O\text{-}TCEP} & \\ C_0 & C(t) & \varepsilon C_0 & 0 \\ C_t & C(t)\text{-}(C_0\text{-}C_t) & 2(C_0\text{-}C_t)+\varepsilon C_0 & C_0\text{-}C_t \end{array}$$

$$\frac{dC_t}{dt} = -kC_t[C(t)-(C_0-C_t)] \approx -kC_tC(t)$$

$$\text{thus, } C_t \approx C_0 * \exp(-kC(t)t) \qquad (1)$$

$C_0$ and $\varepsilon C_0$ are TAT dimer and monomer concentration prior to adding TCEP; $C_t$ is TAT dimer concentration at any giving reaction time $t$; $k$ is the rate constant. $C(t)$ represents the TCEP concentration at the pore entrance. Due to the limited capture radius of nanopores, only chemicals at the pore entrance can be effectively captured and detected (219,220). After adding TCEP into *cis*-chamber, the reductant gradually diffused to the pore entrance and reduced TAT dimer to monomer.

According to three dimensional diffusion equation, $C(t)$ has the following equation:

$C(t) = \frac{N}{(4\pi Dt)^{1.5}} \exp(-\frac{r^2}{4Dt})$, where $D$ is the diffusion constant of TCEP and $r$ is the distance from the pore entrance to TCEP injection location. To simplify the representation

of this reaction kinetics, two parameters $\gamma$ (defined as the ratio of the number of dimer events versus the sum of dimer and monomer events, $\gamma \equiv \frac{[TAT-TAT]}{[TAT-TAT]+[TAT]}$) and reaction quotient $Q_r$ ($Q_r \equiv \frac{[TAT]^2 * [O-TCEP]}{[TAT-TAT]*[TCEP]}$, unit is normalized) were investigated. Solving the above equations yields:

$$\gamma = \frac{1}{2\exp(\frac{Nk}{(4\pi D)^{1.5} t^{0.5}} \exp(-\frac{r^2}{4Dt})) - 1 + \varepsilon} \quad (2)$$

$$Q_r = \frac{2(1-\gamma)^3 C_0^{\ 2}}{\gamma C_T (1+\gamma)^2}. \quad (3)$$

Here, $C_T$ is the final TCEP concentration at diffusion equilibrium state. Experimentally, $\gamma$ and $Q_r$ can be derived from the concentration of TAT dimer and monomer, which is represented by the number of the dimer (with 40% ~ 70% blockade, 5σ) and monomer (with 19% ~ 38% blockade, 5σ) translocation events(220). By calculating those two parameters at each reaction time, the transition of $\gamma$ and $Q_r$ as shown in **Fig. 4.6B** and **Fig. 4.6D** represents the temporal evolution of the reduction reaction. The red line in **Fig. 4.6B** is the fitting curve with equation (2) (500 μM TCEP as that in **Fig 4.6A**). $\gamma$ decreased over the course of reaction time, indicating a decrease in oxidized dimer conformation and concurrent increase in reduced monomeric state. Fitting the data revealed a rate constant k = 0.989 ± 0.096 $min^{-1}$. The trend is also demonstrated in current blockage percentage *vs.* dwell time distribution of events over the course of the reaction time (**Fig 4.6C**). $Q_r$ (500 μM TCEP) in **Fig. 4.6D**, derived from $\gamma$ with equation (3), reached a plateau when the reaction approached equilibrium. The plateau $Q_r$ is equal to standard equilibrium

constant $K^e$, estimated to be $1.4\times10^4$ and corresponding to –23.3 kJ/mol Gibbs free energy ($\Delta G^o = -RT\ln K$). All the reduction data from different concentration of TCEP show the same trend to approach equilibrium status (**Fig 4.6D**). When the concentration of TCEP was increased from 100, 250, to 500 μM, the $\Delta G^e$ were –24.5, –23.1, and –23.3 kJ/mol (**Table 4.1**), respectively. The $\Delta G^e$ derived from our assay is very close to the $\Delta G^e$ of –27 kJ/mol for protein disulfide bond reduction by TCEP obtained by conventional Raman and computational methods (221).

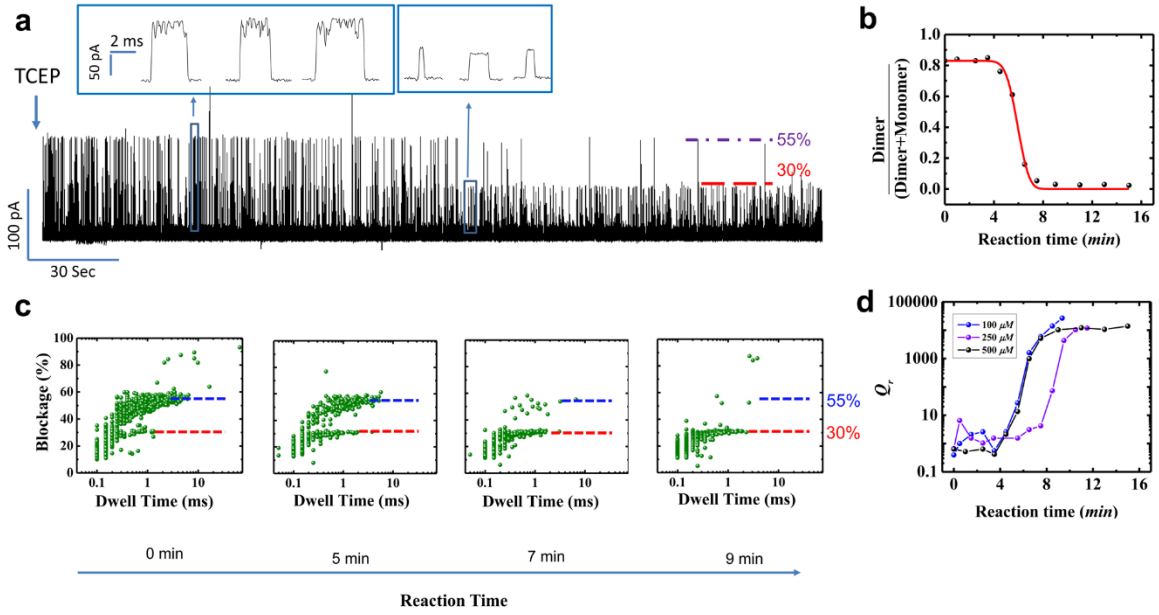



**Figure 4.6**: **Real-time assessment of the conformational states of TAT peptide**. **(a)** Continuous current trace showing transition of oxidized dimer states to reduce monomer states after addition of reducing agent TCEP. **(b)** Quantitative analysis showing the fraction ($\gamma$) of dimer and monomer states as a function of reaction time. **(c)** Current blockage *versus* dwell time distribution over the course of reaction time. Applied voltage: 50 mV; Buffer: 1 M KCl, 5 mM HEPES, pH 8. (**d)** Quantitative analysis showing the reaction quotient $Q_r$ as a function of reaction time.

**Table 4.1:** Parameters for the quantification of the oligomeric states of TAT peptide in real time

| [TAT] (nM) | [TCEP] (μM) | $Q_r (10^3)$ | $K^0 (10^3)$ | $\Delta G^0$ (kJ/mol) |
|---|---|---|---|---|
| 300 | 100 | 26.6 | 26.6 | -24.5 |
| 300 | 250 | 11.8 | 11.8 | -23.1 |
| 300 | 500 | 13.9 | 13.9 | -23.1 |

**Discussion**

In this study, we showed that the SPP1 portal channel can be successfully incorporated into planar lipid bilayer to serve as a nanopore with robust electrophysiological properties. With a conductance of 4.27 *nS*, it is one of the larger channels and only the second viral protein channel (apart from phi29 connector)(13) to be incorporated into a lipid membrane. In the presence of TAT peptide, a burst of current blockage events with characteristic current amplitude and dwell time were observed indicating the translocation of the peptides through the SPP1 connector. In this study, we further demonstrated that this new protein nanopore is capable of detecting kinetics of TCEP mediated reduction of disulfide bond and demonstrating peptide conformational changes in real time. From **Fig. 4.6**, the frequency of dimer forms (current blockage of ~55%) decreased gradually to monomer forms (current blockage of 30%), upon addition of TCEP, as the reaction time increased. The oligomerization states of peptides was determined in real-time at single molecule level, which is not possible by ensemble techniques like NMR, Circular Dichroism, and optical second harmonic generation.

Direct evidence of DNA translocation through α-hemolysin was demonstrated twenty years ago by quantitative PCR(222). Although several studies utilizing peptide and protein translocation through nanopores, have been published over the last several years (204-206,210-214), direct method for validating peptide translocations are lacking. A recent publication tried to solve the problem by conjugating a single-stranded DNA to unfolded protein and then amplify the DNA using PCR after translocation(223). Another study demonstrated the translocation of hyaluronic acid oligosaccharides using high-resolution mass spectrometry(224). Here, we developed a very simple method to provide

evidence of peptide translocation using single molecule fluorescence imaging of samples obtained from the nanopore setup. This method can be easily adapted and employed by other laboratories to validate protein translocation and quantitatively study peptide dynamics through nanopores.

The current blockage signature based on the translocation profile can potentially be used to investigate the length, charge, hydrophobicity, secondary structures and ultimately the amino acid sequences of the peptides. The kinetics of protein folding and unfolding as well as entropic and energetic contributions can be further dissected in the future at the single molecule level.

**CONCLUSIONS**

The portal channel of bacteriophage SPP1 has been reengineered and inserted into lipid bilayer to allow the translocation of peptides with higher ordered structure to produce clear and reproducible electronic signatures. The translocation of peptides observed by single channel conduction assays were verified by optical single molecule fluorescence microscopy assays. The oxidative and the reduced conformation of peptides were clearly differentiated in real-time at single molecule level. When different concentrations of substrates were applied, a similar $\Delta G0$ was obtained, suggesting that the use of SPP1 nanopore for real-time dynamic quantification of peptide folding is feasible.

**AKNOWLEDGEMENTS:**

The research was supported by NIH grant R01 EB012135 to P.G. P.G.'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.

# Chapter 5: Realizing Short Peptide Detection by Enforcing Peptide-Pore Electrostatic Interaction using SPP1 DNA Packaging Motor

This chapter is under preparation for submission. Special thanks to Zhengyi Zhao for help and assistance in preparation of data for figures 5.1 and Dr. Zhi Zhou for help in preparation of data for figure 5.2, 5.3, and 5.4.

## INTRODUCTION

Nanopore technology has recently emerged as a high-throughput and cost-effective single molecule detection method, holding great potential for sensing and biopolymer sequencing(3,4,69). However, one of the critical challenges is the ultra-fast translocation of these biopolymer through nanopore which is far beyond electronic detection limits. To realize DNA sequencing, the first step is to slow down translocation speed. To solve this problem, many attempts have been made, such as controlling the electrolyte temperature, viscosity, the electrical bias voltage, conducting buffer and introducing DNA polymerase(64,66-68). Due to its unique property, currently DNA polymerase has been widely used in DNA sequencing(68,69).

Study the folding of protein and sequencing protein is significant for predicting protein structure and revealing its biological function in many physiology processes. As a counterpart polymer, using nanpore for probing protein structure, folding, detection and sequencing is in early stage because it is still lack of an effective and simple method to slow down the ultra-fast translocation speed. Digesting protein molecule into amino acids and discriminating them with nanopore sequentially may be a possible way toward protein sequencing. One of main obstacles is that short amino acids pass through the protein pore

too fast to be detected with the nowadays available electronics. To slowing down peptide translocation, numerous attempts were made, such adjusting pH (198) and using the AAA+ unfoldase ClpX (225).

SPP1 is a dsDNA bacteriophage that infects *Bacillus subtilis*. The central of DNA packaging motor of the SPP1 is a portal channel, termed connector, which is composed of 12 or 13 copies of protein gp6 (57,186). The connector has an overall diameter of 16 nm and a height of 10.5 nm. The narrowest constriction of the internal channel is ~3 nm (Fig. 5.1)(57,58) which is large compared with other commonly used protein channel, such as α-Hemolysin, MspA, with an internal channel diameter of less than 2 nm(40,73,192). Compare to other larger protein pores such as ClyA(193) and FhuA(194), the uniqueness of the SPP1 channel is it's highly negative charge inside the channel. This provide this channel unique advantage to slow down positive charged peptide translocation.

In this work, the peptides translocation was slowed down in SPP1 DNA packaging motor with low NaCl concentration buffer. The underlying mechanism is that reduced charge screening effect resulted from low salt concentration enhanced electrostatic attractive force between the peptide and pore, which therefore slow down peptide translocation speed. A factor of 500-fold reduction was observed for TAT peptide translocation. Appling such condition in detecting arginine chain peptide allow us to measure peptide as short as two arginine amino acids. Further improve the bandwidth may lead to single amino acid detection and protein sequencing.

## MATERIALS AND METHODS

### *Materials*

The phospholipid 1, 2-diphytanoyl-sn-glycerol-3-phosphocholine (DPhPC) was obtained from Avanti Polar Lipids, Inc. Organic solvents (n-decane and chloroform) were purchased from Fisher Scientific, Inc. and TEDIA, Inc., respectively. TAT peptide and various lengths of poly arginine were custom-ordered from GenScript, Inc. All other reagents were purchased from Sigma or Fisher, if not specified.

***Cloning and Purification of the SPP1 connector channel***

Gene gp6 encoding SPP1 channel was synthesized and cloned into PET3 vector between NhdeI and BamHI by GenScript, Inc. His-tag was inserted into the C-terminal of the SPP1 channel for purification. The plasmid was transformed into BL21 (DE3) for expression and purification. The purification procedure has been described in previous publications(13,174). Briefly, the transformed bacteria were cultured overnight at 37°C in 10 mL LB medium. Then the bacteria were transferred to 1 L of LB medium and cultured until OD600 reaching ~0.5-0.6. Then 5 mM IPTG was added to the medium to induce protein expression. After 3 hours continuous culture, the bacteria were collected by centrifugation. Bacteria was lysed by passing through French press. Lysed cell was separated by centrifugation, and supernatant containing expressed protein was collected and purified with Nickel affinity chromatography (Novagen) (133). His Binding Buffer (15% glycerol, 0.5 M NaCl, 5 mM Imidazole, 10 mM ATP, 50 mM Tris-HCl, pH 8.0), and the cleared lysate was loaded onto a His•Bind® Resin Column and washed with His Washing Buffer (15% glycerol, 0.5 M NaCl, 50 mM Imidazole, 50 mM Tris-HCl, pH 8.0). The His-tagged SPP1 channel was eluted by His Elution Buffer (15% glycerol, 0.5 M NaCl, 0.5 M Imidazole, 50 mM Tris-HCl, pH 8.0). The purified SPP1 channel from chromatography was further purified by a 15-35% glycerol gradient ultracentrifugation at

35K for 2 hrs. Before incorporating into liposomes, the purified SPP1 channel was dialyzed against buffer (0.5 M NaCl, 50 mM Tris-HCl, pH 8.0) to remove excess glycerol.

Insertion of the connector into planar lipid bilayer: The protocol for incorporation of connector channel into lipid bilayer has been published previously (13,56,110,131,174). Briefly, the dehydrated lipid film was formed after the chloroform in 150 μL of 10 mg/mL DPhPC lipids was evaporated in a round bottomed flask, and then rehydrated with 300 μL buffer containing 250 mM sucrose and purified connectors (0.5-1 mg/mL). The suspension was thoroughly vortexed and kept at 45 °C (above the phase transition temperature of DPhPC) followed by extrusion through 400 nm polycarbonate membranes (Avanti Polar Lipids) to produce uniform sized liposome/connector complexes.

Planar bilayer lipid membranes (BLMs) were produced in a BCH-1A horizontal BLM cell (Eastern Scientific). A Teflon partition with a 200 μm aperture was used in the apparatus to separate the BLM cell into two compartments. The aperture was pre-painted with 0.5 μL of 3% (w/v) DPhPC in n-hexane. A conducting buffer was added to both the top and bottom compartments of the BLM cell, and Ag/AgCl electrodes were placed in each compartment. The electrode in the trans-compartment was connected to the headstage of an Axopatch 200B amplifier (Axon Instruments), and the electrode in the top compartment was grounded. A planar lipid bilayer was generated by painting the aperture with 0.5 μL of 3% (w/v) DPhPC in n-decane. 1 μL of the diluted Liposome/connector complex was added to the top compartment directly to generate membrane embedded nanopores.

Electrophysiological measurements: The Axopatch 200B patch clamp amplifier and headstage were connected to a DigiData 1440 analog-digital converter (Axon

Instruments, Inc.) to monitor and record electrochemical currents(13,56,174). The current recordings were low-pass filtered at a frequency of 5 kHz. The sampling frequency was 200 kHz in all experiments, unless otherwise specified. The data were recorded with pClamp 9.1 software (Axon Instruments, Inc.), and analyzed with the Clampfit module of pClamp 9.1 and OriginPro 8.1 (OriginLab Corporation).

***Peptide translocation experiments***

TAT peptide (Cys-Tyr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg) with a final concentration of 23 μg/mL was premixed with the conducting buffer before the insertion of connector channel. For R2, R4, R6, R8, and R10 translocation study, a final concentration of 23 μg/mL was premixed with conducting buffer before the insertion of connector channel.

**RESULTS**

***Charge distribution of SPP1 and α-Hemolysin***

**Fig. 5.1** shows the charge (blue: positive; red: negative) distribution of α-Hemolysin and SPP1 protein nanopore from outside and inter-section view. In α-Hemolysin, both outside and inside of the channel lumen is mainly composed of positive charged amino acids. By contrary, negative charged amino acid residues are clustered outside and the lumen of SPP1 channel, resulting in highly negative charge density. This property of SPP1 has blocked DNA or RNA translocation through the channel even under high voltage in 1 M KCl buffer due to electrostatic repulsing force (data not shown). By screening such force in 2 M LiCl buffer, DNA translocation events has been be observed

due to the higher shielding effect of LiCl than KCl (data not shown). This unique property of SPP1 provides this channel better capability to detect short positive charged molecule.

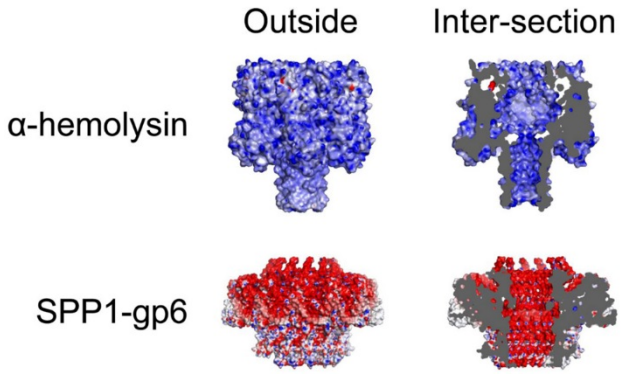



**Figure 5.1** Charge distribution of α-Hemolysin and SPP1 protein nanopore from outside and inter-section view ((blue: positive; red: negative)).

***Slowing down TAT translocation by reducing slat concentration***

To study the effect of electrostatic attractive force on peptide translocation velocity through SPP1, a series NaCl concentrations were used to detect the positive charged TAT peptide (sequence: CYGRKKRRQRRR), of which translocation was studied previously (94). **Fig.5.2** presents typical current trace and TAT translocation events (Left) under 2 M, 1 M, 0.4 M, and 0.2 M NaCl buffer with the same applied voltage and pH. The translocation events under lower NaCl concentration has displayed much longer dwell time than those under higher NaCl concentration, whereas the blockade of TAT peptide under these salt concentration are the same majoring in 55%. **Supplementary Fig.5.1** has shown the dwell time distribution of TAT under these salt concentration. To quantitatively analyze the salt effect on slowing down peptide translocation speed, TAT dwell times under each concentration were derived and plotted in **Fig.5.2**. From the plot, it has clearly shown that dwell time tends to increase when reducing NaCl concentration. The dwell time under 1, 2, 1.5, 0.8, 0.6, 0.4, 0.3,0.2 M NaCl are 1.37±0.07, 0.18±0.01,1±0.09, 4.16±0.12, 8.96±0.38, 12.45±0.08, 63.46±1.05, and 54.30±1.30 ms.

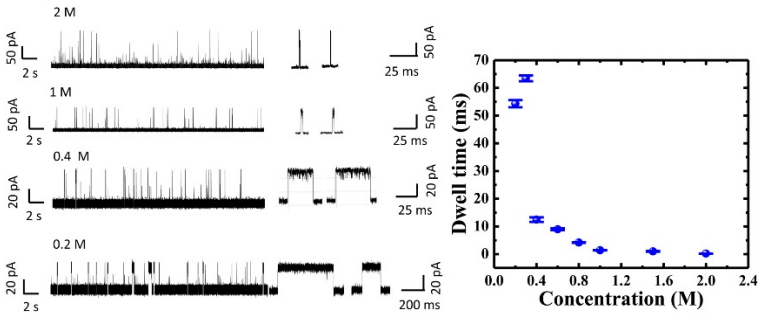
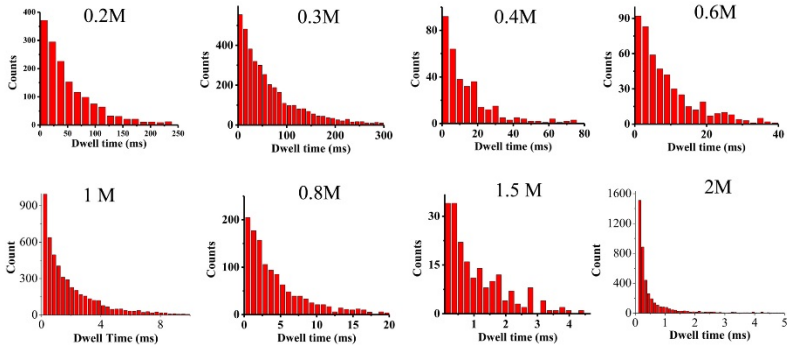


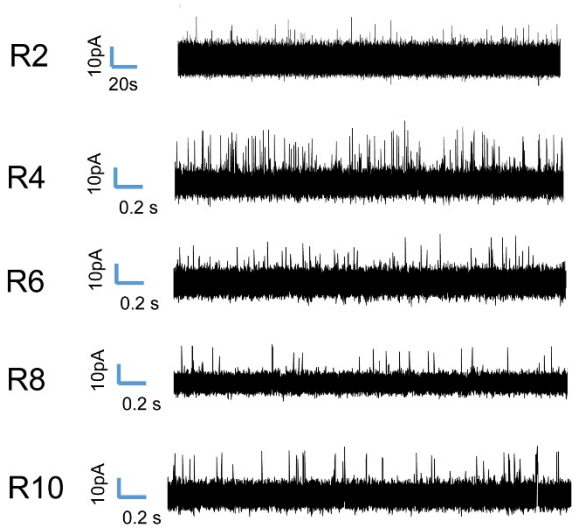

**Figure 5.2 Salt concentration effect on peptide translocation.** Representative current trace of TAT translocation through SPP1 nanopore under 2 M, 1 M, 0.4 M, and 0.2 M NaCl buffer (Left). and TAT translocation events (Left). TAT dwell times versus concentrations (Right).




**Supplementary Figure 5.1** Dwell time distribution of TAT under different NaCl concentration with same pH and 50 mV applied voltage.

***Detect short arginine chain peptide***

Based on the capability of slowing down peptide translocation speed, 0.3 M NaCl buffer (pH 7.9) was chosen to further test whether short poly arginine peptides (R2, R4, R6, R8, and R10), can be detected. **Supplementary Fig.5.2** presents typical current trace of R2, R4, R6, R8, and R10 translocation through SPP1 nanopore. **Fig.5.3** has shown the current blockade of R2, R4, R6, R8, and R10 and the blockades are 23.3±2.2, 35.8±2.6, 29.7±4.3, 32.0±2.3, and 32.7±2.0 respectively. **Supplementary Fig.5.3** has shown the dwell time distribution of R2, R4, R6, R8, and R10 and the dwell time are 0.080±0.002, 0.100±0.001, 0.154±0.005, 0.337±0.009 and 0.531±0.009 respectively. The dwell time for each peptides was derived as previously described (94,220) and plotted in **Fig. 5.4.** To exclude the bandwidth effect, only the events with dwell time longer than 0.2 ms (5 kHz bandwidth) were selected to analyze blockade for each peptides.




**Supplementary Figure 5.2** Representative current trace of f R2, R4, R6, R8, R10 under 0.3M NaCl, 5mM Hepes, pH 8 with 50 mV applied voltage.

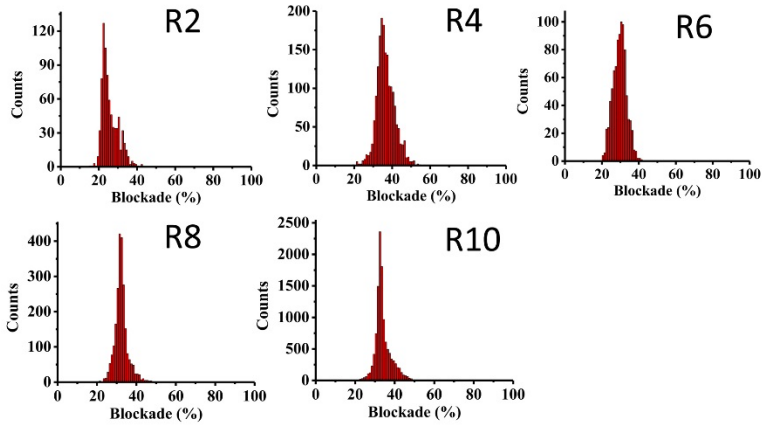
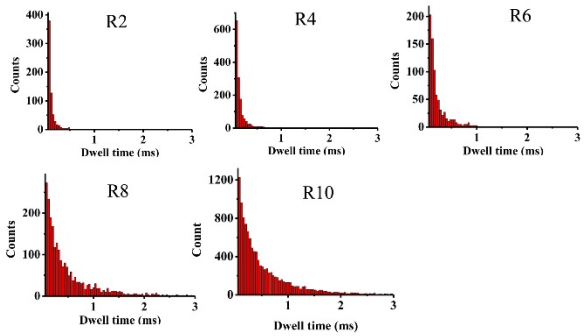



**Figure 5.3** Blockade distribution of R2, R4, R6, R8, R10 under 0.3M NaCl, 5mM Hepes, pH 8 with 50 mV applied voltage.




**Supplementary Fig.5.3** Dwell time distribution of R4, R6, R8, R10 under 0.3M NaCl, 5mM Hepes, pH 8 with 50 mV applied voltage.

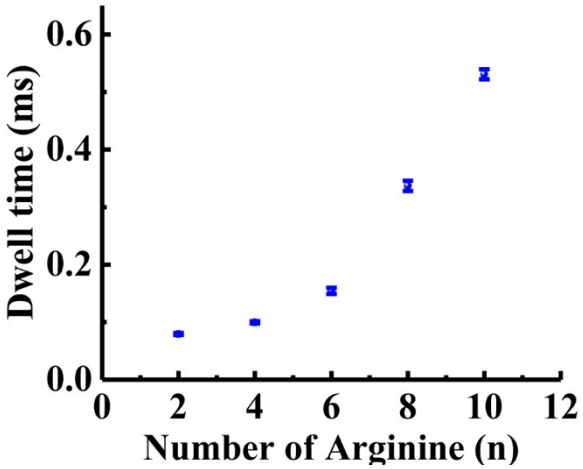



**Figure 5.4** The dwell time versus various length of amino acids under 0.3M NaCl, 5mM Hepes, pH 8 with 50 mV applied voltage.

**DISCUSSION**

For the first-order approximation, Debye length is proportional to 1/C^0.5, and electro-kinetic velocity is on the order of 1/C (226). The red curve in **Fig. 5.2** is a simulation with function y = A/C, where A is a normalized constant and C is NaCl concentration. The fitting result shows similar trend with experimental data. The mismatch of the experiment data and the simulation may arise from several factors: (1) the ignorance of the charge screening of TAT peptide itself; (2) the electro-osmotic effect; (3) possible friction between TAT peptide and SPP1. The counter-trend data at 0.2 M is even smaller than that in 0.3 M NaCl buffer, which may arise from the increased electric driving force due to screened TAT peptide charge.

Compared to R4, R6, R8, and R10, which has shown only one peak, R2 blockade displayed a double-peak distribution. It is speculated that the double peak may arise from two different manners (head-to-tail or shoulder-by-shoulder) of R2 translocation through SPP1, while for longer arginine chain, only head-to-tail manner is allowed.

Recently, Long et al. reported that aerolysin can discriminate Adenine chain from 2 to 10 nucleotides in length by blockade difference which were attributed to the electrostatic interactions between oligonucleotides and aerolysin, and the geometry which does not contain constrict site (39). In this study, 2, 4, 6, 8, 10 arginine amino acids can be differentiated based on dwell time rather than blockade, as shown in Fig. 4. This is because the short peptide shows short dwell time, even though the short peptide events may re-shaped by the built-in filter. The blockade of R2 is about 23%, whereas the blockade of R4, R6, R8, and R10 have similar blockade distribution about 35%. This indicated the depth of the constrict site of SPP1 can probably only accommodate 4 amino acids length

peptide, and any peptide longer than 4 amino acids will not contribute to blockade. This demonstrates that SPP1 nanopore has a high longitudinal spatial-resolution and can potentially be applied for protein sequencing in the future.

## CONCLUSION

In conclusion, TAT and arginine chain peptides were detected with SPP1 DNA packaging motor embedded in lipid bilayer. The dwell time of peptides can be tuned by modulating the electrostatic attractive force between peptide and SPP1 through slat concentration. In 0.3 M NaCl buffer (pH 7.9), TAT demonstrated the highest dwell and lowest translocation velocity. Appling such condition to short arginine chain peptide allow us to sense arginine chain as short as just two arginine amino acids, which is the state-of-art result.

## ACKNOWLEDGEMENTS

The research was supported by NIH grant R01 EB012135 to P.G. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH. P.G's Sylvan G. Frank Endowed Chair position in Pharmaceutics and Drug Delivery is funded by the CM Chen Foundation. P.G. is a consultant of Oxford Nanopore Ltd and RNA Nanobio Ltd.

# Chapter 6: Biomimetic RNA-based Nanopore for Short Peptide Detection

This chapter is under preparation for submission. Special thanks to Hui Li for help and assistance in preparation of data for figures 6.1, 6.2, 6.3, and 6.4.

**INTRODUCTION:**

Nanopore technology has emerged as a cost-effective and high-throughput single molecule detection method, holding great potential for sensing and biopolymer sequencing over the last two decades(3,4,69). To date, many biological nanopore such as α-hemolysin (7), MspA (8), and viral connectors have been successfully used for incredibly versatile applications including detection of small molecules, macromolecule, polymers, polypeptides, as well as DNA and RNA (15-17). All of these nanopore come from biological system with defined channel size and physical property. Creating a new pore *de novo* with the ability to accurately tune the size and its functionality is still of great interest in fundamental and applied science. One of the key challenges in the *de novo* design of membrane channel is, to achieve a defined architecture in atomic scale with defined size, shape and properties (227).

Inspired by biological channels, synthetic nanopore constructed entirely by DNA has been reported previously and serve as cytotoxic agents(228) and drug carriers (229) . One of the designs was constructed with scaffolded DNA origami which was composed of a stem and a barrel-shaped cap (230). Another type was generated based on the archetypal six-helix-bundle structure that enclose a 2-nm-wide lumen with a tunable height (227). In order to insert negatively charged pore into the hydrophobic bilayer membrane,

hydrophobic anchors, such as cholesterol(230), porphyrin(228), and ethane-PPT groups (231) were included in all these artificial pore.

Over the last decade, RNA nanotechnology has advanced rapidly and attracted widespread attention due to the potential for applications in material and medical sciences (232-234). Naturally occurring RNA molecules can fold into unique 3D structures and perform various biological functions *in vitro* and *in vivo*. Compared to DNA, RNA has higher diversity in structure and function. Due to their diversity in structures and functions, RNA molecules were used as building blocks for engineering various *de novo* architectures with unique physical/biochemical properties via bottom-up self-assembly. Varieties of RNA architectures with defined shape, size and stoichiometry have been created for diverse applications, such as therapeutics delivery, cancer targeting, immunomodulation, and sensing(232,233,235).

By introducing polymerase or exonuclease, significant progress has been made in DNA sequencing using nanopore (68,69,72). However, so far there is still lack of effective way for protein sequencing which is critical for understanding its biological function in many physiology processes. One of possible method toward protein sequencing is to digest protein molecule into amino acids and discriminate them using nanopore sequentially. One of main obstacles is that short amino acids pass through the protein pore too fast and are hard to be detected with the nowadays available electronics.

To expand biological nanopores and explore new features of artificial nanopores, here we constructed and assembled RNA nanopores with two different designs. The insertion of RNA nanopore into lipid bilayer and cell membrane were further characterized. Translocation of short amino acids through RNA nanopore was detected and characterized.

Translocation of short amino acids is first time observed in artificial nanopore. This new artificial nanopore has the potential to be used for sensing, disease diagnosis and even protein sequencing.

## MATERIALS AND METHODS:

### *Materials*

The phospholipid 1, 2-diphytanoyl-*sn*-glycerol-3-phosphocholine (DPhPC) was obtained from Avanti Polar Lipids, Inc. Organic solvents (*n*-decane and chloroform) were purchased from Fisher Scientific, Inc. and TEDIA, Inc., respectively. Poly arginine with 2, 4, 6, 8, 10 amino acids peptide was custom-ordered from *GenScript*, Inc. n-octyl-oligo-oxyethylene was purchased from Enzo company. All other reagents were purchased from Sigma or Fisher, if not specified.

### *Design, synthesis, and self-assembly of RNA nanopores*

The NanoEngineer program was used to facilitate the design of the RNA tubes. Unmodified RNA strands were synthesized by *in vitro* T7 transcription and purified by 8 M urea, 8% denaturing PAGE. DNA templates for transcribing the RNA strands were made by polymerase chain reactions (PCR). DNA oligonucleotides for PCR were directly ordered from IDT, Inc. Cholesterol modified RNA strands were ordered from IDT or synthesized by Azco Oligo-800 DNA and RNA synthesizer.

The RNA tubes were self-assembled in a one-pot manner by mixing the synthesized RNA strands in equimolar concentrations in 1×TMS buffer and heated to 95 °C and slowly cool down to room temperature. The step-wise self-assembly of the RNA tubes was

examined by agarose gels. The gels were stained with Ethidium Bromide (EB) and imaged by Typhoon FLA 7000 (GE Healthcare).

***Atomic force microscopy***

The RNA tubes were imaged by atomic force microscopy (AFM), following previously described methods. Briefly, RNA tubes were placed on the APS-modified mica surface and excess samples were washed with DEPC water and dried before imaging. AFM imaging was performed by using the MultiMode AFM NanoScope IV system (Veeco) operated in tapping mode.

***In vitro binding of RNA nanopores to cancer cells***

Prostate cancer LnCap cells were grown on glass cover slides in 24-well plates in RPMI-1640 medium with 10% FBS at 37°C in humidified air containing 5% CO2 overnight. RNA nanopores harboring chosterol and labeled by Cy5 were diluted in optium-MEM medium to 100 nM and incubated with the cells for 2 hrs at 37 °C. RNA nanopores without cholesterol and single strand RNA were used as the negative control. After the incubation, the cells were washed with PBS and fixed by 4% paraformaldehyde. Alexa488 Phalloidin (Life Technologies) was used to stain the cellular actin. The Prolong Gold antifade reagent with DAPI (Life Technologies) were then used to stain the cell nucleuses and mount the cells to the glass slides. The confocal images were recorded by the FluoView FV1000-Filter Confocal Microscope System (Olympus).

***Insertion of the RNA nanopore into planar lipid bilayer***

Briefly, planar bilayer lipid membranes (BLMs) were generated in a BCH-1A horizontal BLM cell (Eastern Scientific). A Teflon partition with a 200 μm aperture was

placed in the apparatus to separate the BLM cell into *cis*- (top) and *trans*- (bottom) compartments. A planar lipid bilayer was formed by painting the aperture with 0.5 μL of 3% (w/v) DPhPC in n-decane. A conducting buffer (1 M KCl, 5 mM HEPES, pH 7.8) was added to both the top and bottom compartments of the BLM cell, and Ag/AgCl electrodes were placed in the buffer of each compartment. The electrode in the trans-compartment was connected to the headstage of an Axopatch 200B amplifier (Axon Instruments, Inc.), and the electrode in the top compartment was grounded.

For direct insertion, a mixture of cholesterol-anchored RNA nanopores and 0.5% n-octyl-oligo-oxyethylene dissolved in conducting buffer was added to the *cis* side of the bilayer to a final concentration of 100 nM nanopores. For pore insertion with liposome, a two-step procedure was employed(13) for this design. DPhPC lipids in chloroform were dehydrated to eliminate solvents first and then rehydrated with buffer containing 250 mM sucrose and purified RNA nanopore with a final concentration 250 nM. The multilamellar lipid-RNA nanopore suspension was then extruded through 400 nm polycarbonate membrane filters to generate uniform unilamellar liposomes with the RNA naopore embedded in the membrane. The resulting liposome-RNA nanopore complex was fused with a planar lipid membrane to generate planar membrane-embedded RNA nanopore.

***Electrophysiological measurements***

The headstage and Axopatch 200B patch clamp amplifier were connected to a DigiData 1440 analog-digital converter (Axon Instruments, Inc.) to monitor and record electrochemical currents through BLMs(13,56,174). The current recordings were low-pass filtered at a frequency of 5 kHz. The sampling frequency was 20 kHz in all experiments, unless otherwise specified. The data were recorded a with pClamp 9.1 software (Axon

Instruments, Inc.). The blockades were calculated with a home-built Matlab program and statistically analyzed with OriginPro 8.1 (OriginLab Corporation). The capture rate was derived by exponentially fitting the time interval between two adjacent events as described previously (236). For peptide translocation experiments, poly arginine was premixed with conducting buffer with a final concentration of 1 μg/mL before the insertion of RNA nanopore.

## RESULTS AND DISCUSSION:

### *Design, synthesis, and self-assembly of RNA nanopores*

RNA tubes were designed by using NanoEngineer software. In the first RNA nanopore design, six RNA strands were used to assemble the tube by utilizing a RNA origami approach similar to DNA origami (**Fig 6.1**). Six RNA strands were designed to fold together to form a hollow channel with the insider diameter of 1.7nm and the outsider diameter of 6.3nm. To assemble the RNA nanopore with cholesterol, two more chemically synthesized RNA strands with cholesterol at the end were added to the structure. For assembling, the synthesized RNA strands were mixed in stoichiometric ratio and annealed in 1×TMS buffer in a one-pot manner. Step-wise self-assembly of the RNA nanopores with and without cholesterol was examined by agarose gels, confirming the formation of the high molecular weight RNA nanopores (**Fig.6.2**).

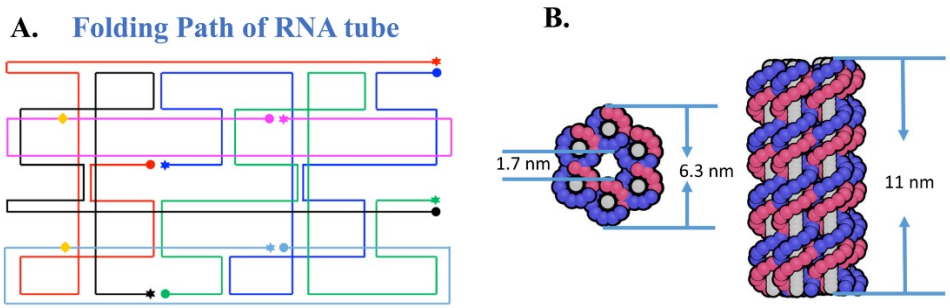



**Figure 6.1 Design and schematic view of origami designed RNA nanopore.** (a) Map of the RNA nanopore composed of 6 RNA strands labeled by six different color. The two yellow squares indicate two cholesterol molecules. (b) Schematic top view (left) and side view (right) of RNA nanopore structure.

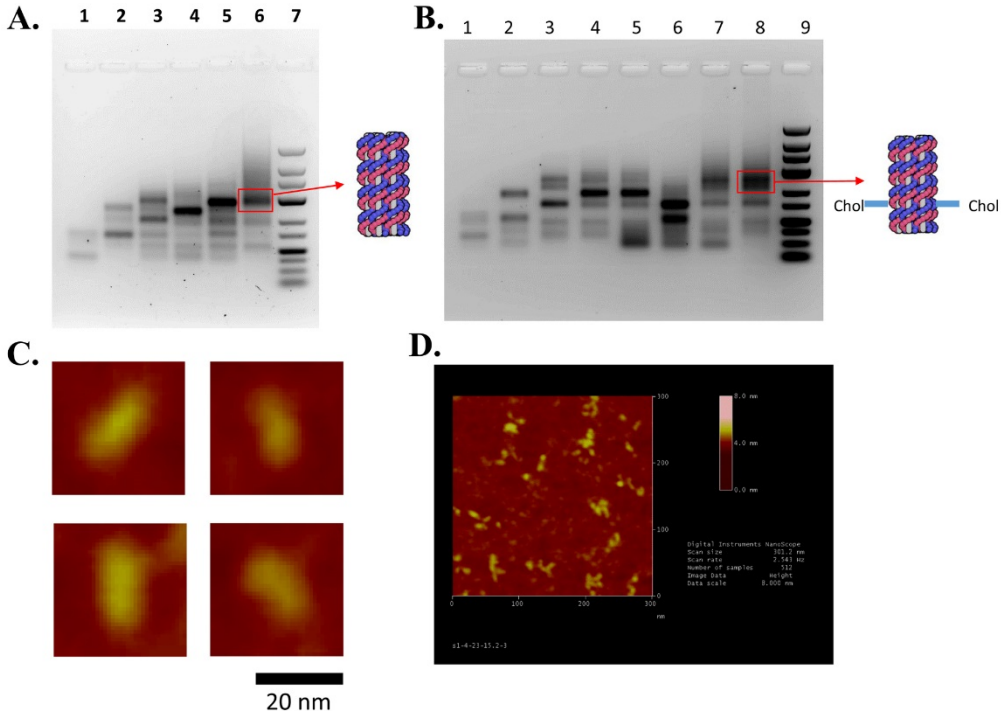



**Figure 6.2 Assemble and AFM of origami designed RNA nanopore.** (a) 4% Agarose gel electrophoresis. Lane 1: Strand 1. Lane 2: Srands 1+2. Lane 3: Srands 1+2+3. Lane 4: Srands 1+2+3+4. Lane 5: Srands 1+2+3+4+5. Lane 6: Srands 1+2+3+4+5+6. Lane 7: DNA Ladder. (b) 4% Agarose gel electrophoresis. Lane 1: Strand 1. Lane 2: Srands 1+2. Lane 3: Srands 1+2+3. Lane 4: Srands 1+2+3+4. Lane 5: Srands 1+2+3+4+5. Lane 6: Srands 1+2+3+4+6. Lane 7: Srands 1+2+3+4+6+3'Chol-RNA. Lane 7: Srands 1+2+3+4+6+3'Chol-RNA+5' Chol-RNA.Lane 9: DNA Ladder. (c) AFM analysis of RNA nanopore: representative individual RNA nanopore (left) and whole view of RNA nanopore (right)

For the second RNA nanopore design with three cholesterol molecules, six RNA strands were designed to be bound together by base paring and each RNA strand form one face of the six-helix channel. Similar channel size with the first RNA nanopore design was predicted, since both designs were composed of six RNA helixes bound together (**Fig 6.3A**). Compared with previous design, this approach has several advantages, such as ease of fabrication, higher yield of pores at higher concentration, and lower synthesis cost. **Fig.6.3B** has shown RNA nanopore without cholesterol can successfully assemble high molecular weight RNA nanopores step-wisely. However, after incorporate the three strands with cholesterol molecules, the assembled structure cannot run into gel and stuck in the loading well. This is possibly due to the formation of aggregated structure resulted from the hydrophobic interaction between cholesterol molecules (**Fig 6.4C**).

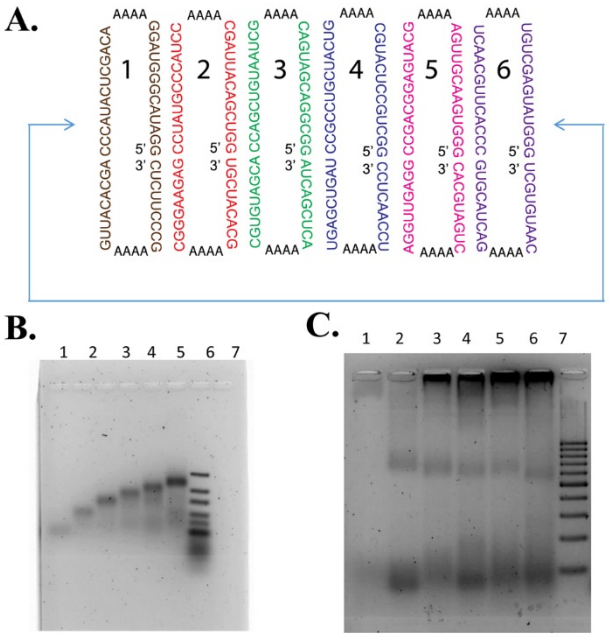



**Figure 6.3 Design and assemble of six bundled RNA nanopore.** (a) Map of the RNA nanopore six RNA strands which bound together by base paring and each RNA strand form one face of the six-helix channel. (b) 4% Agarose gel electrophoresis to confirm assemble of six bundled RNA nanopore. Lane 1: Strand 1. Lane 2: Srands 1+2. Lane 3: Srands 1+2+3. Lane 4: Srands 1+2+3+4. Lane 5: Srands 1+2+3+4+5. Lane 6: Srands 1+2+3+4+5+6. Lane 7: DNA Ladder. (b) 4% Agarose gel electrophoresis to confirm assemble of six bundled RNA nanopore with three cholesterol. Lane 1: Strand 1with cholesterol (chol). Lane 2: Srands 1(Chol)+2. Lane 3: Srands 1(chol)+2+3(chol). Lane 4: Srands 1(chol)+2+3(chol)+4. Lane 5: Srands 1(chol)+2+3(chol)+4+5(chol). Lane 6: Srands 1(chol)+2+3(chol)+4+5(chol)+6. Lane 7: DNA Ladder.

### *Atomic force microscope characterization of RNA nanopores*

Atomic force microscope was used to characterize the RNA nanopores. As shown in the AFM images, individual RNA nanopore can be clearly observed under AFM (**Fig.6.2C, D**). The observed RNA nanopores under atomic force microscope size and shape is in agreement to our design, confirming the successful self-assembly of RNA nanopores.

### *Characterization of RNA nanopore insertion into cellular and lipid bilayer membrane*

To study whether RNA nanopore can interact with cellular membrane, RNA nanopore with and without anchored cholesterol and several other controls, were incubated with LnCap cell. Cy5 fluorophore was conjugated to these structures to track the interaction. The results have shown only cholesterol anchored RNA nanopore group displayed strong cy5 signal in cell membrane indicating cholesterol anchored RNA nanopore group inserted into the cell membrane (**Fig 6.4**). Similar results has also been verified in giant unilamellar vesicles that only RNA nanopore with anchored cholesterol can insert into the lipid membrane of giant unilamellar vesicles (Data not shown).

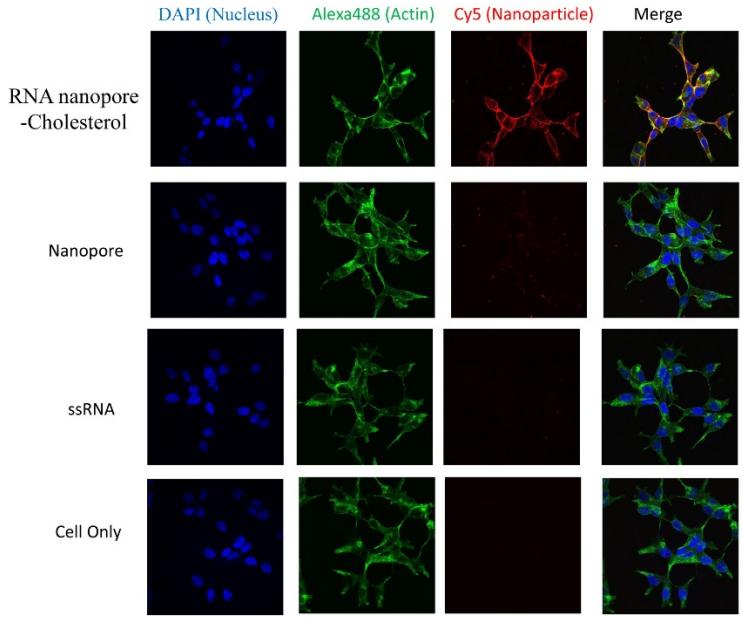



**Figure 6.4 Insertion of RNA nanopores to cellular membrane:** The confocal images of LnCap cell after incubation with RNA nanopore contain cholesterol. RNA nanopores without cholesterol, single strand RNA and cell only used as the negative control as indicated in the image. The DAPI and Alexa 488 were used to stain the cell nucleuses and actin. Cy5 fluorophore conjugated these structure was used to track the interaction.

To incorporate origami designed RNA nanopore into planar lipid membranes, direct insertion of cholesterol-anchored RNA nanopore with detergent n-octyl-oligo-oxyethylene was first tried but no insertion was observed (data not shown). This may be because the hydrophobicity of two cholesterol was not strong enough to overcome the energy barrier of inserting into lipid bilayer. Therefore, a two-step procedure was employed(13) for this design. Briefly, DPhPC lipids in chloroform were dehydrated to eliminate solvents first and then rehydrated with buffer containing 250 mM sucrose and purified origami designed RNA nanopore with a final concentration 250 nM. The multilamellar lipid-RNA nanopore suspension was then extruded through 400 nm polycarbonate membrane filters to generate uniform unilamellar liposomes with the RNA naopore embedded in the membrane. The resulting liposome-RNA nanopore complex was fused with a planar lipid membrane to generate planar membrane-embedded RNA nanopore. Electrophysiological properties of membrane-embedded RNA nanopore was characterized by single channel conductance assay. As revealed in a continuous current trace, the RNA nanopore insertion were observed (**Fig. 6.5A**). The insertion of single RNA nanopore results in various current jump under an applied potential of 50 mV in conducting buffer (1 M KCl, 5 mM HEPES, pH 8). As revealed in **Fig.6.5B**, the conductance distribution of is wide. This wide distribution is attributed to simultaneous insertion of multiple RNA nanopore or multiple RNA nanopore aggregated together.

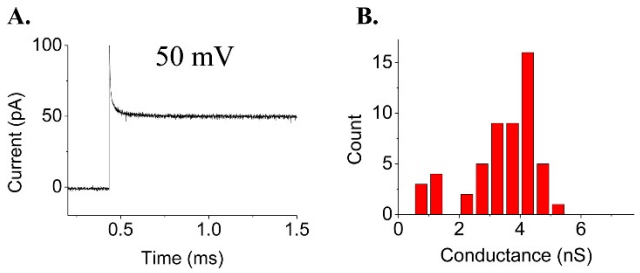



**Figure 6.5 Characterization of origami designed RNA nanopore into lipid bilayer.** (a) Current trace of single channel insertion under 50 mV applied voltage (b) Conductance distribution of origami designed RNA nanopore. Condition: 1M KCl 5mM Hepes, pH 8.

Due to the insertion rate of RNA nanopre is relative low with two anchored cholesterol, another design with three anchored cholesterol and a bundle of six RNA duplexes was constructed. Direct incubation of the three anchored cholesterol RNA nanopore with plannar lipid bilayer has resulted distinguish stepwise current jump (**Fig.6.6B**). Two step prepared liposome-RNA nanopore complex has also resulted in similar stepwise current jump (**Fig.6.6A**). The conductance distribution of this RNA nanopore is also wide under 0.4M KCl, 5mM Hepes, pH8 (**Fig.6.6C,D**). The wide distribution of this RNA nanopore is also possibly due to multiple nanopores aggregated together cholesterol group. This is also revealed in the assemble gel in which the anchored cholesterol RNA nanopore stuck in the well.

***Characterization of peptide translocation through RNA nanopore***

One of possible way toward protein sequencing is to digest protein into amino acids and discriminate them sequentially using nanopore. However, detection of short amino acids by nanopore is still challenging due to the ultra-fast translocation speed. In this study, translocation of poly arginine with 2, 4, 6, 8, 10 amino acids through RNA nanopore were measured and compared.

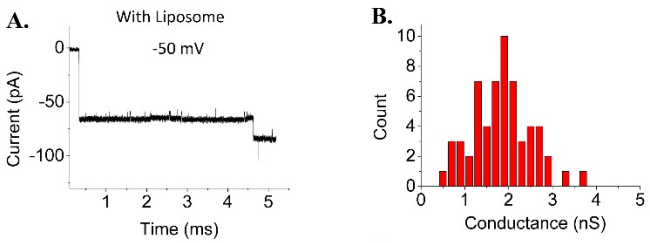
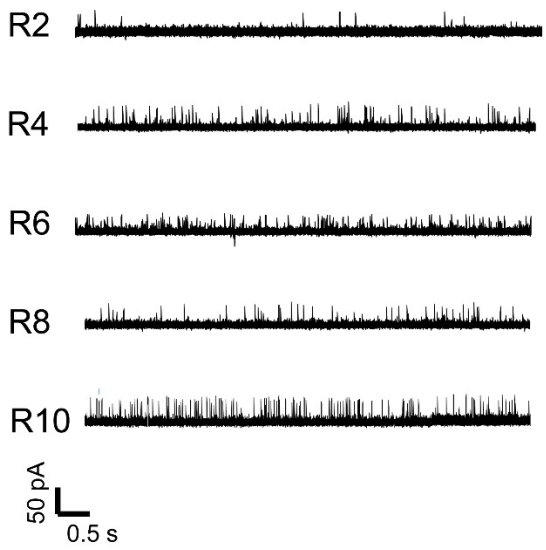
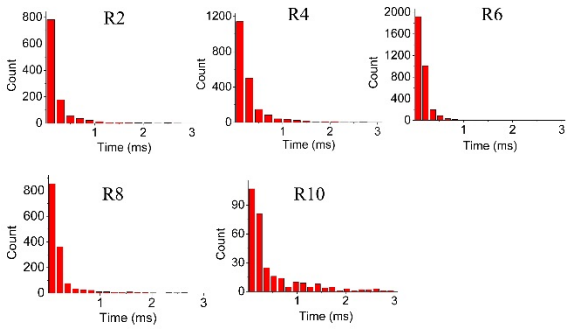



**Figure 6.6 Characterization of six bundled RNA nanopore into lipid bilayer.** (a) Current trace of multiple RNA nanopore/liposome insertion under -50 mV applied voltage (b) Conductance distribution of RNA nanopore/liposome. Condition: 0.4 M KCl 5mM Hepes, pH 8.

**Supp.Fig.6.1** presents typical current trace of R2, R4, R6, R8, and R10 translocation through RNA nanopore with negative 50 mV applied voltage. All of these peptides has displayed similar current blockade. As shown in **Supp.Fig.6.1**, R2 has displayed lower capture rate compared with other peptides, which may be because the translocation speed is too fast for R2 and many events has been filtered. The dwell time for each peptides was derived as previously described (220). The dwell time of R2, R4, R6, R8, and R10 are 0.14±0.01, 0.22±0.01, 0.17±0.01, 0.15±0.01, and 0.28±0.03ms (**Supp.Fig.6.2)**. In this study, we have also observed that only positive or neutral charged peptide are allowed to pass through RNA nanopore, whereas negative charged DNA or peptide are not allowed to pass through RNA nanoproe which may be due to the electrostatic repulsing force between RNA nanopore and negative charged DNA or peptide (data not shown). To study the folding of protein and even sequencing protein, several attempts have been made such as adjusting pH (198) and using the AAA+ unfoldase ClpX to slow down peptide translocation(225). So far, short peptide detection still has not been reported. In this study, short amino acids as short as only two have been successfully detected in our platform, which has potential for applying protein sequencing in the future.




**Supplementary Figure 6.1** Typical current trace of poly arginine R2, R4, R6, R8, and R10 translocation through RNA nanopore. Condition:-50 mV, 0.4M KCl 5mM Hepes, pH 8.




**Supplementary Figure 6.2** Dwell time distributin of poly arginine R2, R4, R6, R8, and R10 translocation through RNA nanopore. Condition:-50 mV, 0.4M KCl 5mM Hepes, pH 8.

## CONCLUSION

In this study, we have successfully designed, constructed and assembled two different biomimetic RNA-based nanopore. The assembled RNA based nanopore anchored with cholesterol have been demonstrated insertion into cellular and lipid membrane. Short amino acids as short as two have been first time detected and characterized. *De novo* designed membrane channel can provide numerous advantages compared protein membrane channel, such as tunable size and its functionality, ease of construction, scale up and modification. Due to RNA unique physical/biochemical properties, biomimetic RNA-based Nanopore have the potential to be applied for sensing and protein sequencing.

## ACKNOWLEDGEMENTS

The research was supported by NIH grant R01 EB012135 to P.G. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH. P.G's Sylvan G. Frank Endowed Chair position in Pharmaceutics and Drug Delivery is funded by the CM Chen Foundation. P.G. is a consultant of Oxford Nanopore Ltd and RNA Nanobio Ltd.

# Chapter 7: Future Direction and Perspectives of the Field

## CONCLUSIONS AND FUTURE DIRECTION:

This thesis work has successfully applied phi29 connector nanopore for antibody detection, developed several new biological nanopore and biomimetic nanopore using RNA nanotechnology, and applied them for short peptide detections. Nanopore technology has emerged as a cost-effective and high-throughput single molecule detection method, holding great potential for sensing and biopolymer sequencing over the last two decades(3,4,69). Chapter 2 demonstrated that EpCAM antibody can be detected and discriminated from the background events by phi29 connector nanopore. Chapter 3 explored more biological nanopore from other bacteriophages and discovered a three step gating property that is common to DNA packaging motors of bacterial viruses. Chapter 4 expanded the applications of portal protein to study the translocation of peptides and the dynamics of peptide oligomeric states in real-time at single molecule level. Chapter 5 has successfully detected and distinguished different lengths of short peptides under optimize condition. Chapter 6 has successfully constructed a new biomimetic nanopore using RNA nanotechnology and applied for short peptides detection.

A wide range of analytes ranging from ions, small molecules, macromolecules, to biopolymers have been detected successfully using nanopore. However, most of these development were still in the laboratory stage. How to improve the sensitivity, stability and specificity especially using clinical sample which is often in the presence of many impurities is still the major bottleneck to transform nanopore for clinical use. Also, more effort should be made to explore detectable capability such as protein and peptide sequencing, protein structure and folding studies.

Continuous improvement and engineering of currently available nanopores and exploration of new channels and pores are still needed. Although many biological channel have been developed and applied for various applications in the last decade, several challenges still remains, such as lacking knowledge on structure-function of available protein channel, low sensitivity and specificity, and limited detectable analytes. In order to obtain high sensitivity and specificity, different analytes requires unique protein nanopore with different shape, size, physical dimensions and hydrophilic/hydrophobic properties.

Compared with traditional DNA sequencing methods, nanopore-based sequencing offers considerable advantages such as ultra-low cost and fast, long readout, label-free, low sample volume and no applications required. Several key challenges that used be the bottlenecks in nanopore-based sequencing has been resolved. First, the ultra-fast translocation of DNA through nanopore which is far beyond the electronic detection limits, have been resolved by introducing enzymes or motors such as DNA polymerase or exonucleases. The second challenge is spatial resolution which is able to detect single bases ideally, however, most of the times in many protein nanopore at least several nucleotides reside within the constrict site at any given point making it difficult to detect a single nucleotide. Noblegen Company employed a strategy that can convert each nucleotide into multiple same nucleotide in the DNA sequence through a process called “cicular DNA conversion” therefore expanding the special resolution (86). Another strategy employed by Oxford Nanopore in R9 pore is that analyzing more than 5 bases at any one time and the sequence is then decoded by a reference map (84,104,237). The third challenge is the low resolution of distinguishing the four bases due to the much closed structure. To solve this problem, Genia Company developed a technology that can modify the four nucleotides

with different sized polymer tag and then be differentiated after cleavage by the nanopore. Another common strategy is the engineering the nanopores to enhance the recognition capability. It is reported that over 700 mutants of R9 have been screened in Oxford Nanopore (82). The challenge for the “exonuclease approach”, which is initially developed in a-HL but no further data revealed by Oxford Nanopore, is the subsequently released nucleotide is not guaranteed to move in to the pore directly or in a strictly sequential manner. One possible approach to resolve this challenge is to fix the exonuclease to the nanopore in specific site and direction ensuring each released nucleotide can move into the pore. Other challenges including the instability of membrane, integrated circuits that enable massively parallel single-molecule DNA sequencing, and software to analyze the data have also been resolved by Oxford Nanopore. The remaining challenge is the accuracy which is still low compared with current sequencing technologies. To enhance accuracy, it still need many works such as protein engineering, large scale screening, exploration new nanopore and improvement of the algorithmic in the future.

REFERENCES


1. Kasianowicz JJ, Brandin E, Branton D, & Deamer DW (1996) Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. *Proc. Natl. Acad. Sci. U. S. A* 93: 13770-13773

2. Pennisi E (2014) Genomics. DNA Sequencers Still Waiting for the Nanopore Revolution. *Science* 343: 829-830

3. Haque F, Shu D, Shu Y, Shlyakhtenko L, Rychahou P, Evers M, & Guo P (2012) Ultrastable Synergistic Tetravalent RNA Nanoparticles for Targeting to Cancers. *Nano Today* 7: 245-257

4. Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Di VM, Garaj S, Hibbs A, Huang X, Jovanovich SB, Krstic PS, Lindsay S, Ling XS, Mastrangelo CH, Meller A, Oliver JS, Pershin YV, Ramsey JM, Riehn R, Soni GV, Tabard-Cossa V, Wanunu M, Wiggin M, & Schloss JA (2008) The Potential and Challenges of Nanopore Sequencing. *Nat. Biotechnol.* 26: 1146-1153

5. Feng Y, Zhang Y, Ying C, Wang D, & Du C (2015) Nanopore-Based Fourth-Generation DNA Sequencing Technology. *Genomics Proteomics Bioinformatics* 13: 4-16

6. Guo BY, Zeng T, & Wu HC (2015) Recent Advances of DNA Sequencing Via Nanopore-Based Technologies. *Science Bulletin* 60: 287-295

7. Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, & Gouaux JE (1996) Structure of Staphylococcal Alpha -Hemolysin, a Heptameric Transmembrane Pore. *Science* 274: 1859-1865

8. Butler TZ, Pavlenok M, Derrington IM, Niederweis M, & Gundlach JH (2008) Single-Molecule DNA Detection With an Engineered MspA Protein Nanopore. *Proceedings of the National Academy of Sciences* 105: 20647-20652

9. Stefureac R, Long YT, Kraatz HB, Howard P, & Lee JS (2006) Transport of Alpha-Helical Peptides Through Alpha-Hemolysin and Aerolysin Pores. *Biochemistry* 45: 9172-9179

10. Mohammad MM, Howard KR, & Movileanu L (2011) Redesign of a Plugged Beta-Barrel Membrane Protein. *J. Biol. Chem.* 286: 8000-8013

11. Chimerel C, Movileanu L, Pezeshki S, Winterhalter M, & Kleinekathofer U (2008) Transport at the Nanoscale: Temperature Dependence of Ion Conductance. *Eur. Biophys. J.* 38: 121-125

12. Zhuang T & Tamm LK (2014) Control of the Conductance of Engineered Protein Nanopores Through Concerted Loop Motions. *Angew. Chem. Int. Ed Engl.* 53: 5897-5902

13. Wendell D, Jing P, Geng J, Subramaniam V, Lee TJ, Montemagno C, & Guo P (2009) Translocation of Double-Stranded DNA Through Membrane-Adapted Phi29 Motor Protein Nanopores. *Nat. Nanotechnol.* 4: 765-772

14. Wang S, Ji Z, Yan E, Haque F, & Guo P (2016) Three-Step Channel Conformational Changes Common to DNA Packaging Motors of Bacterial Viruses T3, T4, SPP1, and Phi29. *Virology*

15. Haque F, Lunn J, Fang H, Smithrud D, & Guo P (2012) Real-Time Sensing and Discrimination of Single Chemicals Using the Channel of Phi29 DNA Packaging Nanomotor. *ACS Nano* 6: 3251-3261

16. Howorka S & Siwy Z (2009) Nanopore Analytics: Sensing of Single Molecules. *Chem. Soc. Rev.* 38: 2360-2384

17. Healy K (2007) Nanopore-Based Single-Molecule DNA Analysis. *Nanomedicine* 2: 459-481

18. Bayley H & Jayasinghe L (2004) Functional Engineered Channels and Pores (Review). *Molecular Membrane Biology* 21: 209-220

19. Brannigan JA & Wilkinson AJ (2002) Protein Engineering 20 Years on. *Nat. Rev Mol. Cell Biol.* 3: 964-970

20. Buchsbaum SF, Mitchell N, Martin H, Wiggin M, Marziali A, Coveney PV, Siwy Z, & Howorka S (2013) Disentangling Steric and Electrostatic Factors in Nanoscale Transport Through Confined Space. *Nano Lett.* 13: 3890-3896

21. Langecker M, Arnaut V, Martin TG, List J, Renner S, Mayer M, Dietz H, & Simmel FC (2012) Synthetic Lipid Membrane Channels Formed by Designed DNA Nanostructures. *Science* 338: 932-936

22. Akeson M, Branton D, Kasianowicz JJ, Brandin E, & Deamer DW (1999) Microsecond Time-Scale Discrimination Among Polycytidylic Acid, Polyadenylic Acid, and Polyuridylic Acid As Homopolymers or As Segments Within Single RNA Molecules. *Biophys. J.* 77: 3227-3233

23. Hernandez-Ainsa S & Keyser UF (2014) DNA Origami Nanopores: Developments, Challenges and Perspectives. *Nanoscale.* 6: 14121-14132

24. Gilboa T & Meller A (2015) Optical Sensing and Analyte Manipulation in Solid-State Nanopores. *Analyst* 140: 4733-4747

25. Kudr J, Skalickova S, Nejdl L, Moulick A, Ruttkay-Nedecky B, Adam V, & Kizek R (2015) Fabrication of Solid-State Nanopores and Its Perspectives. *Electrophoresis* 36: 2367-2379

26. Cuervo A & Carrascosa JL (2012) Viral Connectors for DNA Encapsulation. *Curr. Opin. Biotechnol.* 23: 529-536

27. Gu LQ & Shim JW (2010) Single Molecule Sensing by Nanopores and Nanopore Devices. *Analyst* 135: 441-451

28. Stoddart D, Heron AJ, Mikhailova E, Maglia G, & Bayley H (2009) Single-Nucleotide Discrimination in Immobilized DNA Oligonucleotides With a Biological Nanopore. *Proceedings of the National Academy of Sciences* 106: 7702-7707

29. Kang XF, Gu LQ, Cheley S, & Bayley H (2005) Single Protein Pores Containing Molecular Adapters at High Temperatures. *Angew. Chem. Int. Ed Engl.* 44: 1495-1499

30. Vercoutere W, Winters-Hilt S, Olsen H, Deamer D, Haussler D, & Akeson M (2001) Rapid Discrimination Among Individual DNA Hairpin Molecules at Single-Nucleotide Resolution Using an Ion Channel. *Nat Biotech* 19: 248-252

31. Heinz C, Roth E, & Niederweis M (2003) Purification of Porins From Mycobacterium Smegmatis. *Methods Mol. Biol.* 228: 139-150

32. Faller M, Niederweis M, & Schulz GE (2004) The Structure of a Mycobacterial Outer-Membrane Channel. *Science* 303: 1189-1192

33. Howorka S & Bayley H (2002) Probing Distance and Electrical Potential Within a Protein Pore With Tethered DNA. *Biophys. J.* 83: 3202-3210

34. Heinz C, Engelhardt H, & Niederweis M (2003) The Core of the Tetrameric Mycobacterial Porin MspA Is an Extremely Stable Beta-Sheet Domain. *J. Biol. Chem.* 278: 8678-8685

35. Pavlenok M, Derrington IM, Gundlach JH, & Niederweis M (2012) MspA Nanopores From Subunit Dimers. *PLoS ONE* 7: e38726

36. Merstorf C, Cressiot B, Pastoriza-Gallego M, Oukhaled A, Betton JM, Auvray L, & Pelta J (2012) Wild Type, Mutant Protein Unfolding and Phase Transition Detected by Single-Nanopore Recording. *ACS Chem Biol* 7: 652-658

37. Cabiaux V, Buckley JT, Wattiez R, Ruysschaert JM, Parker MW, & van der Goot FG (1997) Conformational Changes in Aerolysin During the Transition From the Water-Soluble Protoxin to the Membrane Channel. *Biochemistry* 36: 15224-15232

38. Chakraborty T, Schmid A, Notermans S, & Benz R (1990) Aerolysin of Aeromonas Sobria: Evidence for Formation of Ion-Permeable Channels and Comparison With Alpha-Toxin of Staphylococcus Aureus. *Infect. Immun.* 58: 2127-2132

39. Cao C, Ying YL, Hu ZL, Liao DF, Tian H, & Long YT (2016) Discrimination of Oligonucleotides of Different Lengths With a Wild-Type Aerolysin Nanopore. *Nat. Nanotechnol.* 11: 713-718

40. Pastoriza-Gallego M, Rabah L, Gibrat G, Thiebot B, van der Goot FG, Auvray L, Betton JM, & Pelta J (2011) Dynamics of Unfolded Protein Transport Through an Aerolysin Pore. *J. Am. Chem. Soc.* 133: 2923-2931

41. Goyal P, Krasteva PV, Van GN, Gubellini F, Van dB, I, Troupiotis-Tsailaki A, Jonckheere W, Pehau-Arnaudet G, Pinkner JS, Chapman MR, Hultgren SJ, Howorka S, Fronzes R, & Remaut H (2014) Structural and Mechanistic Insights into the Bacterial Amyloid Secretion Channel CsgG. *Nature* 516: 250-253

42. Chen M, Khalid S, Sansom MS, & Bayley H (2008) Outer Membrane Protein G: Engineering a Quiet Pore for Biosensing. *Proc. Natl. Acad. Sci. U. S. A* 105: 6272-6277

43. Yildiz O, Vinothkumar KR, Goswami P, & Kuhlbrandt W (2006) Structure of the Monomeric Outer-Membrane Porin OmpG in the Open and Closed Conformation. *EMBO J.* 25: 3702-3713

44. Conlan S, Zhang Y, Cheley S, & Bayley H (2000) Biochemical and Biophysical Characterization of OmpG: A Monomeric Porin. *Biochemistry* 39: 11845-11854

45. Bonhivers M, Desmadril M, Moeck GS, Boulanger P, Colomer-Pallas A, & Letellier L (2001) Stability Studies of FhuA, a Two-Domain Outer Membrane Protein From Escherichia Coli. *Biochemistry* 40: 2606-2613

46. Krewinkel M, Dworeck T, & Fioroni M (2011) Engineering of an E. Coli Outer Membrane Protein FhuA With Increased Channel Diameter. *J. Nanobiotechnology.* 9: 33

47. Niedzwiecki DJ, Mohammad MM, & Movileanu L (2012) Inspection of the Engineered FhuA DeltaC/Delta4L Protein Nanopore by Polymer Exclusion. *Biophys. J.* 103: 2115-2124

48. Mueller M, Grauschopf U, Maier T, Glockshuber R, & Ban N (2009) The Structure of a Cytolytic Alpha-Helical Toxin Pore Reveals Its Assembly Mechanism. *Nature* 459: 726-730

49. Soskine M, Biesemans A, Moeyaert B, Cheley S, Bayley H, & Maglia G (2012) An Engineered ClyA Nanopore Detects Folded Target Proteins by Selective External Association and Pore Entry. *Nano Lett* 12: 4895-4900

50. Klimpel KR, Molloy SS, Thomas G, & Leppla SH (1992) Anthrax Toxin Protective Antigen Is Activated by a Cell Surface Protease With the Sequence Specificity and Catalytic Properties of Furin. *Proc Natl Acad Sci U S A* 89: 10277-10281

51. Petosa C, Collier RJ, Klimpel KR, Leppla SH, & Liddington RC (1997) Crystal Structure of the Anthrax Toxin Protective Antigen. *Nature* 385: 833-838

52. Nestorovich EM, Karginov VA, Berezhkovskii AM, & Bezrukov SM (2010) Blockage of Anthrax PA63 Pore by a Multicharged High-Affinity Toxin Inhibitor. *Biophys. J.* 99: 134-143

53. Blaustein RO, Koehler TM, Collier RJ, & Finkelstein A (1989) Anthrax Toxin: Channel-Forming Activity of Protective Antigen in Planar Phospholipid Bilayers. *Proc Natl Acad Sci U S A* 86: 2209-2213

54. Blaustein RO & Finkelstein A (1990) Voltage-Dependent Block of Anthrax Toxin Channels in Planar Phospholipid Bilayer Membranes by Symmetric Tetraalkylammonium Ions. Effects on Macroscopic Conductance. *J. Gen. Physiol* 96: 905-919

55. Nestorovich EM & Bezrukov SM (2012) Obstructing Toxin Pathways by Targeted Pore Blockage. *Chem. Rev* 112: 6388-6430

56. Jing P, Haque F, Vonderheide A, Montemagno C, & Guo P (2010) Robust Properties of Membrane-Embedded Connector Channel of Bacterial Virus Phi29 DNA Packaging Motor. *Mol. Biosyst.* 6: 1844-1852

57. Lebedev AA, Krause MH, Isidro AL, Vagin AA, Orlova EV, Turner J, Dodson EJ, Tavares P, & Antson AA (2007) Structural Framework for DNA Translocation *Via* the Viral Portal Protein. *EMBO J.* 26: 1984-1994

58. Lhuillier S, Gallopin M, Gilquin B, Brasiles S, Lancelot N, Letellier G, Gilles M, Dethan G, Orlova EV, Couprie J, Tavares P, & Zinn-Justin S (2009) Structure of Bacteriophage SPP1 Head-to-Tail Connection Reveals Mechanism for Viral DNA Gating. *Proc. Natl. Acad. Sci. U. S. A* 106: 8507-8512

59. Valpuesta JM, Sousa N, barthelemy I, Fernandez JJ, Fujisawa H, Ibarra B, & Carrascosa JL (2000) Structural Analysis of the Bacteriophage T3 Head-to-Tail Connector. *J. Struct. Biol.* 131: 146-155

60. Sun L, Zhang X, Gao S, Rao PA, Padilla-Sanchez V, Chen Z, Sun S, Xiang Y, Subramaniam S, Rao VB, & Rossmann MG (2015) Cryo-EM Structure of the Bacteriophage T4 Portal Protein Assembly at Near-Atomic Resolution. *Nat. Commun.* 6: 7548

61. Meller A, Nivon L, Brandin E, Golovchenko J, & Branton D (2000) Rapid Nanopore Discrimination Between Single Polynucleotide Molecules. *Proc. Natl. Acad. Sci. U. S. A* 97: 1079-1084

62. Stoddart D, Heron AJ, Klingelhoefer J, Mikhailova E, Maglia G, & Bayley H (2010) Nucleobase Recognition in SsDNA at the Central Constriction of the Alpha-Hemolysin Pore. *Nano Lett.* 10: 3633-3637

63. Wallace EV, Stoddart D, Heron AJ, Mikhailova E, Maglia G, Donohoe TJ, & Bayley H (2010) Identification of Epigenetic DNA Modifications With a Protein Nanopore. *Chem. Commun. (Camb. )* 46: 8195-8197

64. Kawano R, Schibel AEP, Cauley C, & White HS (2008) Controlling the Translocation of Single-Stranded DNA Through +?Hemolysin Ion Channels Using Viscosity. L*angmuir* 25: 1233-1237
65. Meller A & Branton D (2002) Single Molecule Measurements of DNA Transport Through a Nanopore. E*lectrophoresis* 23: 2583-2591

66. de Zoysa RS, Jayawardhana DA, Zhao Q, Wang D, Armstrong DW, & Guan X (2009) Slowing DNA Translocation Through Nanopores Using a Solution Containing Organic Salts. J. *Phys. Chem. B* 113: 13332-13336

67. Howorka S, Cheley S, & Bayley H (2001) Sequence-Specific Detection of Individual DNA Strands Using Engineered Nanopores. N*at. Biotechnol.* 19: 636-639

68. Olasagasti F, Lieberman KR, Benner S, Cherf GM, Dahl JM, Deamer DW, & Akeson M (2010) Replication of Individual DNA Molecules Under Electronic Control Using a Protein Nanopore. N*at Nanotechnol.* 5: 798-806

69. Cherf GM, Lieberman KR, Rashid H, Lam CE, Karplus K, & Akeson M (2012) Automated Forward and Reverse Ratcheting of DNA in a Nanopore at 5-A Precision. N*at Biotechnol.* 30: 344-348

70. Ayub M & Bayley H (2012) Individual RNA Base Recognition in Immobilized Oligonucleotides Using a Protein Nanopore. N*ano Lett.* 12: 5637-5643

71. Ayub M, Hardwick SW, Luisi BF, & Bayley H (2013) Nanopore-Based Identification of Individual Nucleotides for Direct RNA Sequencing. N*ano Lett.* 13: 6144-6150

72. Fuller CW, Kumar S, Porel M, Chien M, Bibillo A, Stranges PB, Dorwart M, Tao C, Li Z, Guo W, Shi S, Korenblum D, Trans A, Aguirre A, Liu E, Harada ET, Pollard J, Bhat A, Cech C, Yang A, Arnold C, Palla M, Hovis J, Chen R, Morozova I, Kalachikov S, Russo JJ, Kasianowicz JJ, Davis R, Roever S, Church GM, & Ju J (2016) Real-Time Single-Molecule Electronic DNA Sequencing by Synthesis Using Polymer-Tagged Nucleotides on a Nanopore Array. P*roc Natl Acad Sci U S A* 113: 5233-5238

73. Manrao EA, Derrington IM, Laszlo AH, Langford KW, Hopper MK, Gillgren N, Pavlenok M, Niederweis M, & Gundlach JH (2012) Reading DNA at Single-Nucleotide Resolution With a Mutant MspA Nanopore and Phi29 DNA Polymerase. N*at Biotechnol.* 30: 349-353

74. Laszlo AH, Derrington IM, Ross BC, Brinkerhoff H, Adey A, Nova IC, Craig JM, Langford KW, Samson JM, Daza R, Doering K, Shendure J, & Gundlach JH (2014) Decoding Long Nanopore Sequencing Reads of Natural DNA. N*at. Biotechnol.* 32: 829-833

75. Ayub M, Stoddart D, & Bayley H (2015) Nucleobase Recognition by Truncated Alpha-Hemolysin Pores. A*CS Nano* 9: 7895-7903

76. Clarke J, Wu HC, Jayasinghe L, Patel A, Reid S, & Bayley H (2009) Continuous Base Identification for Single-Molecule Nanopore DNA Sequencing. N*at. Nanotechnol.* 4: 265-270

77. Derrington IM, Butler TZ, Collins MD, Manrao E, Pavlenok M, Niederweis M, & Gundlach JH (2010) Nanopore DNA Sequencing With MspA. P*roc. Natl. Acad. Sci. U. S. A* 107: 16060-16065

78. Laszlo AH, Derrington IM, Brinkerhoff H, Langford KW, Nova IC, Samson JM, Bartlett JJ, Pavlenok M, & Gundlach JH (2013) Detection and Mapping of 5-Methylcytosine and 5-Hydroxymethylcytosine With Nanopore MspA. P*roc Natl Acad Sci U S A* 110: 18904-18909

79. Schreiber J, Wescoe ZL, bu-Shumays R, Vivian JT, Baatar B, Karplus K, & Akeson M (2013) Error Rates for Nanopore Discrimination Among Cytosine, Methylcytosine, and Hydroxymethylcytosine Along Individual DNA Strands. P*roc Natl Acad Sci U S A* 110: 18910-18915

80. Bhattacharya S, Derrington IM, Pavlenok M, Niederweis M, Gundlach JH, & Aksimentiev A (2012) Molecular Dynamics Study of MspA Arginine Mutants Predicts Slow DNA Translocations and Ion Current Blockades Indicative of DNA Sequence. A*CS Nano* 6: 6960-6968

81. Maglia G, Restrepo MR, Mikhailova E, & Bayley H (2008) Enhanced Translocation of Single DNA Molecules Through Alpha-Hemolysin Nanopores by Manipulation of Internal Charge. P*roc Natl Acad Sci U S A* 105: 19720-19725

82. (2016) No Thanks, I've already got one.

83. Loman NJ & Watson M (2015) Successful Test Launch for Nanopore Sequencing. N*at. Methods* 12: 303-304

84. Jain M, Fiddes IT, Miga KH, Olsen HE, Paten B, & Akeson M (2015) Improved Data Analysis for the MinION Nanopore Sequencer. N*at. Methods* 12: 351-356

85. Kumar S, Tao C, Chien M, Hellner B, Balijepalli A, Robertson JW, Li Z, Russo JJ, Reiner JE, Kasianowicz JJ, & Ju J (2012) PEG-Labeled Nucleotides and Nanopore Detection for Single Molecule DNA Sequencing by Synthesis. S*ci. Rep.* 2: 684

86. McNally B, Singer A, Yu Z, Sun Y, Weng Z, & Meller A (2010) Optical Recognition of Converted DNA Nucleotides for Single-Molecule DNA Sequencing Using Nanopore Arrays. N*ano Lett.* 10: 2237-2244

87. Bayley H & Cremer PS (2001) Stochastic Sensors Inspired by Biology. N*ature* 413: 226-230

88. Thomson K, Amin I, Morales E, & Winters-Hilt S (2007) Preliminary Nanopore Cheminformatics Analysis of Aptamer-Target Binding Strength. B*MC. Bioinformatics.* 8 Suppl 7: S11

89. Braha O, Walker B, Cheley S, Kasianowicz JJ, Song L, Gouaux JE, & Bayley H (1997) Designed Protein Pores As Components for Biosensors. C*hem. & Biol.* 4: 497-505

90. Braha O, Gu LQ, Zhou L, Lu X, Cheley S, & Bayley H (2000) Simultaneous Stochastic Sensing of Divalent Metal Ions. N*at Biotech* 18: 1005-1007

91. Guan X, Gu LQ, Cheley S, Braha O, & Bayley H (2005) Stochastic Sensing of TNT With a Genetically Engineered Pore. C*hembiochem* 6: 1875-1881

92. Geng J & Guo P (2011) Membrane-Embedded Channel of Bacteriophage Phi29 DNA Packaging Motor for Single Molecule Sensing and Nanomedicine. C*hinese Bulletin of Life Sciences* 23: 1114-1129

93. Ji Z, Wang S, Zhao Z, Zhou Z, Haque F, & Guo P (2016) Fingerprinting of Peptides With a Large Channel of Bacteriophage Phi29 DNA Packaging Motor. S*mall*

94. Zhou Z, Ji Z, Wang S, Haque F, & Guo P (2016) Oriented Single Directional Insertion of Nanochannel of Bacteriophage SPP1 DNA Packaging Motor into Lipid Bilayer Via Polar Hydrophobicity. B*iomaterials* 105: 222-227

95. Movileanu L, Howorka S, Braha O, & Bayley H (2000) Detecting Protein Analytes That Modulate Transmembrane Movement of a Polymer Chain Within a Single Protein Pore. N*ature Biotechnology* 18: 1091-1095

96. Rotem D, Jayasinghe L, Salichou M, & Bayley H (2012) Protein Detection by Nanopores Equipped With Aptamers. J. *Am. Chem. Soc.* 134: 2781-2787

97. Howorka S, Nam J, Bayley H, & Kahne D (2004) Stochastic Detection of Monovalent and Bivalent Protein-Ligand Interactions. A*ngew. Chem. Int. Ed Engl.* 43: 842-846

98. Xie H, Braha O, Gu LQ, Cheley S, & Bayley H (2005) Single-Molecule Observation of the Catalytic Subunit of CAMP-Dependent Protein Kinase Binding to an Inhibitor Peptide. C*hem. Biol* 12: 109-120

99. Luchian T, Shin SH, & Bayley H (2003) Single-Molecule Covalent Chemistry With Spatially Separated Reactants. A*ngew. Chem. Int. Ed Engl.* 42: 3766-3771

100. Loudwig S & Bayley H (2006) Photoisomerization of an Individual Azobenzene Molecule in Water: an on-Off Switch Triggered by Light at a Fixed Wavelength. J. *Am. Chem. Soc.* 128: 12404-12405

101. Shin SH, Luchian T, Cheley S, Braha O, & Bayley H (2002) Kinetics of a Reversible Covalent-Bond-Forming Reaction Observed at the Single-Molecule Level. A*ngew. Chem. Int. Ed Engl.* 41: 3707-3709

102. Cheley S, Xie H, & Bayley H (2006) A Genetically Encoded Pore for the Stochastic Detection of a Protein Kinase. C*hembiochem* 7: 1923-1927

103. Gu LQ, Cheley S, & Bayley H (2001) Prolonged Residence Time of a Noncovalent Molecular Adapter, Beta-Cyclodextrin, Within the Lumen of Mutant Alpha-Hemolysin Pores. J. *Gen. Physiol* 118: 481-494

104. Ho CW, Van M, V, Tsai KC, De Temmerman PJ, Mast J, & Maglia G (2015) Engineering a Nanopore With Co-Chaperonin Function. S*ci Adv.* 1: e1500905

105. Wang S, Haque F, Rychahou PG, Evers BM, & Guo P (2013) Engineered Nanopore of Phi29 DNA-Packaging Motor for Real-Time Detection of Single Colon Cancer Specific Antibody in Serum. A*CS Nano* 7: 9814-9822

106. Gu LQ, la Serra M, Vincent JB, Vigh G, Cheley S, Braha O, & Bayley H (2000) Reversal of Charge Selectivity in Transmembrane Protein Pores by Using Noncovalent Molecular Adapters. P*roc Natl Acad Sci USA* 97: 3959-3964

107. Gu LQ, Braha O, Conlan S, Cheley S, & Bayley H (1999) Stochastic Sensing of Organic Analytes by a Pore-Forming Protein Containing a Molecular Adapter. N*ature* 398: 686-690

108. Astier Y, Braha O, & Bayley H (2006) Toward Single Molecule DNA Sequencing: Direct Identification of Ribonucleoside and Deoxyribonucleoside 5'-Monophosphates by Using an Engineered Protein Nanopore Equipped With a Molecular Adapter. J. *Am. Chem. Soc.* 128: 1705-1710

109. Wu HC, Astier Y, Maglia G, Mikhailova E, & Bayley H (2007) Protein Nanopores With Covalently Attached Molecular Adapters. J. *Am. Chem. Soc.* 129: 16142-16148

110. Jing P, Haque F, Shu D, Montemagno C, & Guo P (2010) One-Way Traffic of a Viral Motor Channel for Double-Stranded DNA Translocation. N*ano Lett.* 10: 3620-3627

111. Hammerstein AF, Jayasinghe L, & Bayley H (2011) Subunit Dimers of Alpha-Hemolysin Expand the Engineering Toolbox for Protein Nanopores. J. *Biol. Chem.* 286: 14324-14334

112. Geng J, Wang S, Fang H, & Guo P (2013) Channel Size Conversion of Phi29 DNA-Packaging Nanomotor for Discrimination of Single- and Double-Stranded Nucleic Acids. A*CS Nano* 7: 3315-3323

113. Rao VB & Feiss M (2008) The Bacteriophage DNA Packaging Motor. A*nnu. Rev. Genet.* 42: 647-681

114. Cingolani G, Moore SD, Prevelige PE, Jr., & Johnson JE (2002) Preliminary Crystallographic Analysis of the Bacteriophage P22 Portal Protein. J. *Struct. Biol.* 139: 46-54

115. Dube P, Tavares P, Lurz R, & van HM (1993) The Portal Protein of Bacteriophage SPP1: a DNA Pump With 13-Fold Symmetry. E*MBO J* 12: 1303-1309

116. Kocsis E, Cerritelli ME, Trus BL, Cheng N, & Steven AC (1995) Improved Methods for Determination of Rotational Symmetries in Macromolecules. U*ltramicroscopy* 60: 219-228

117. Trus BL, Cheng N, Newcomb WW, Homa FL, Brown JC, & Steven AC (2004) Structure and Polymorphism of the UL6 Portal Protein of Herpes Simplex Virus Type 1. J. *Virol.* 78: 12668-12671

118. Neher E & Stevens CF (1977) Conductance Fluctuations and Ionic Pores in Membranes. A*nnu. Rev. Biophys. Bioeng.* 6: 345-381

119. Doyle DA, Morais CJ, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, Chait BT, & MacKinnon R (1998) The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. S*cience* 280: 69-77

120. Hanson PI & Whiteheart SW (2005) AAA+ Proteins: Have Engine, Will Work. N*at. Rev. Mol Cell Biol.* 6: 519-529

121. Zhao Z, Khisamutdinov E, Schwartz C, & Guo P (2013) Mechanism of One-Way Traffic of Hexameric Phi29 DNA Packaging Motor With Four Electropositive Relaying Layers Facilitating Anti-Parallel Revolution. A*CS Nano* 7: 4082-4092

122. Schwartz C, De Donatis GM, Fang H, & Guo P (2013) The ATPase of the Phi29 DNA-Packaging Motor Is a Member of the Hexameric AAA+ Superfamily. V*irology* 443: 20-27

123. Guo PX & Lee TJ (2007) Viral Nanomotors for Packaging of DsDNA and DsRNA. M*ol. Microbiol.* 64: 886-903

124. Sun S, Rao VB, & Rossmann MG (2010) Genome Packaging in Viruses. C*urr. Opin. Struct. Biol.* 20: 114-120

125. Zhang F, Lemieux S, Wu X, St.-Arnaud S, McMurray CT, Major F, & Anderson D (1998) Function of Hexameric RNA in Packaging of Bacteriophage Phi29 DNA in Vitro. M*ol. Cell.* 2: 141-147

126. Guo P, Erickson S, & Anderson D (1987) A Small Viral RNA Is Required for i*n Vitro* Packaging of Bacteriophage Phi29 DNA. S*cience* 236: 690-694

127. Guo P, Zhang C, Chen C, Trottier M, & Garver K (1998) Inter-RNA Interaction of Phage Phi29 PRNA to Form a Hexameric Complex for Viral DNA Transportation. M*ol. Cell.* 2: 149-155

128. Shu D, Zhang H, Jin J, & Guo P (2007) Counting of Six PRNAs of Phi29 DNA-Packaging Motor With Customized Single Molecule Dual-View System. E*MBO J.* 26: 527-537

129. Simpson AA, Leiman PG, Tao Y, He Y, Badasso MO, Jardine PJ, Anderson DL, & Rossman MG (2001) Structure Determination of the Head-Tail Connector of Bacteriophage Phi29. A*cta Cryst* D57: 1260-1269

130. Guasch A, Pous J, Ibarra B, Gomis-Ruth FX, Valpuesta JM, Sousa N, Carrascosa JL, & Coll M (2002) Detailed Architecture of a DNA Translocating Machine: the High-Resolution Structure of the Bacteriophage Phi29 Connector Particle. J. *Mol. Biol.* 315: 663-676

131. Geng J, Fang H, Haque F, Zhang L, & Guo P (2011) Three Reversible and Controllable Discrete Steps of Channel Gating of a Viral DNA Packaging Motor. B*iomaterials* 32: 8234-8242

132. Ibanez C, Garcia JA, Carrascosa JL, & Salas M (1984) Overproduction and Purification of the Connector Protein of B*acillus Subtilis* Phage F29. N*ucleic Acids Res.* 12: 2351-2365

133. Robinson MA, Wood JP, Capaldi SA, Baron AJ, Gell C, Smith DA, & Stonehouse NJ (2006) Affinity of Molecular Interactions in the Bacteriophage Phi29 DNA Packaging Motor. N*ucleic Acids Res.* 34: 2698-2709

134. Guo Y, Blocker F, Xiao F, & Guo P (2005) Construction and 3-D Computer Modeling of Connector Arrays With Tetragonal to Decagonal Transition Induced by PRNA of Phi29 DNA-Packaging Motor. J. *Nanosci. Nanotechnol.* 5: 856-863

135. Xiao F, Sun J, Coban O, Schoen P, Wang JC, Cheng RH, & Guo P (2009) Fabrication of Massive Sheets of Single Layer Patterned Arrays Using Lipid Directed Reengineered Phi29 Motor Dodecamer. A*CS Nano* 3: 100-107

136. Cai Y, Xiao F, & Guo P (2008) The Effect of N- or C-Terminal Alterations of the Connector of Bacteriophage Phi29 DNA Packaging Motor on Procapsid Assembly, PRNA Binding, and DNA Packaging. N*anomedicine* 4: 8-18

137. Haque F, Li J, Wu H-C, Liang X-J, & Guo P (2013) Solid-State and Biological Nanopore for Real-Time Sensing of Single Chemical and Sequencing of DNA. N*ano Today* 8: 56-74

138. Venkatesan BM & Bashir R (2011) Nanopore Sensors for Nucleic Acid Analysis. N*ature Nanotechnology* 6: 615-624

139. Majd S, Yusko EC, Billeh YN, Macrae MX, Yang J, & Mayer M (2010) Applications of Biological Pores in Nanomedicine, Sensing, and Nanoelectronics. C*urrent Opinion in Biotechnology* 21: 439-476

140. (2011) N*anopores: Sensing and Fundamental Biological Interactions,* Springer London, Limited,

141. Raj A & van OA (2009) Single-Molecule Approaches to Stochastic Gene Expression. A*nnu. Rev. Biophys.* 38: 255-270

142. Perkins TJ & Swain PS (2009) Strategies for Cellular Decision-Making. M*ol. Syst. Biol.* 5: 326

143. Bayley H & Martin CR (2000) Resistive-Pulse Sensing-From Microbes to Molecules. C*hem. Rev.* 100: 2575-2594

144. Sauer-Budge AF, Nyamwanda JA, Lubensky DK, & Branton D (2003) Unzipping Kinetics of Double-Stranded DNA in a Nanopore. P*hys. Rev. Lett.* 90: 238101

145. Balzar M, Winter MJ, de Boer CJ, & Litvinov SV (1999) The Biology of the 17-1A Antigen (Ep-CAM). J. *Mol Med. (Berl)* 77: 699-712

146. de Boer CJ, van Krieken JH, Janssen-van Rhijn CM, & Litvinov SV (1999) Expression of Ep-CAM in Normal, Regenerating, Metaplastic, and Neoplastic Liver. J. *Pathol.* 188: 201-206

147. High AS, Robinson PA, & Klein CE (1996) Increased Expression of a 38kd Cell-Surface Glycoprotein MH99 (KS 1/4) in Oral Mucosal Dysplasias. J. *Oral Pathol. Med.* 25: 10-13

148. Pauli C, Munz M, Kieu C, Mack B, Breinl P, Wollenberg B, Lang S, Zeidler R, & Gires O (2003) Tumor-Specific Glycosylation of the Carcinoma-Associated Epithelial Cell Adhesion Molecule EpCAM in Head and Neck Carcinomas. C*ancer Lett.* 193: 25-32

149. Poczatek RB, Myers RB, Manne U, Oelschlager DK, Weiss HL, Bostwick DG, & Grizzle WE (1999) Ep-Cam Levels in Prostatic Adenocarcinoma and Prostatic Intraepithelial Neoplasia. J. *Urol.* 162: 1462-1466

150. Shetye J, Christensson B, Rubio C, Rodensjo M, Biberfeld P, & Mellstedt H (1989) The Tumor-Associated Antigens BR55-2, GA73-3 and GICA 19-9 in Normal and Corresponding Neoplastic Human Tissues, Especially Gastrointestinal Tissues. A*nticancer Res* 9: 395-404

151. Spizzo G, Gastl G, Wolf D, Gunsilius E, Steurer M, Fong D, Amberger A, Margreiter R, & Obrist P (2003) Correlation of COX-2 and Ep-CAM Overexpression in Human Invasive Breast Cancer and Its Impact on Survival. B*r. J. Cancer* 88: 574-578

152. Sarma VR, Silverton EW, Davies DR, & Terry WD (1971) The Three-Dimensional Structure at 6 A Resolution of a Human Gamma Gl Immunoglobulin Molecule. J. *Biol Chem.* 246: 3753-3759

153. Li W, Bell NA, Hernandez-Ainsa S, Thacker VV, Thackray AM, Bujdoso R, & Keyser UF (2013) Single Protein Molecule Detection by Glass Nanopores. A*CS Nano* 7(5):4129-34

154. Wei R, Gatterdam V, Wieneke R, Tampe R, & Rant U (2012) Stochastic Sensing of Proteins With Receptor-Modified Solid-State Nanopores. N*at Nanotechnol.* 7: 257-263

155. Myszka DG (1997) Kinetic Analysis of Macromolecular Interactions Using Surface Plasmon Resonance Biosensors. C*urr. Opin. Biotechnol.* 8: 50-57

156. Abe H, Kuroki M, Imakiire T, Yamauchi Y, Yamada H, Arakawa F, & Kuroki M (2002) Preparation of Recombinant MK-1/Ep-CAM and Establishment of an ELISA System for Determining Soluble MK-1/Ep-CAM Levels in Sera of Cancer Patients. J. *Immunol. Methods* 270: 227-233

157. Sterzynska K, Kempisty B, Zawierucha P, & Zabel M (2012) Analysis of the Specificity and Selectivity of Anti-EpCAM Antibodies in Breast Cancer Cell Lines. F*olia Histochem. Cytobiol.* 50: 534-541

158. Guo P, Noji H, Yengo CM, Zhao Z, & Grainge I (2016) Biological Nanomotors With Revolution, Linear, or Rotation Motion Mechanism. M*icrobiology and Molecular Biology Reviews* 80: 161-186

159. Guo P, Peterson C, & Anderson D (1987) Prohead and DNA-Gp3-Dependent ATPase Activity of the DNA Packaging Protein Gp16 of Bacteriophage Phi29. J *Mol Biol* 197: 229-236

160. Chaban Y, Lurz R, Brasiles S, Cornilleau C, Karreman M, Zinn-Justin S, Tavares P, & Orlova EV (2015) Structural Rearrangements in the Phage Head-to-Tail Interface During Assembly and Infection. P*roc. Natl. Acad. Sci. U. S A* 112: 7009-7014

161. Zhang H, Schwartz C, De Donatis GM, & Guo P (2012) "Push Through One-Way Valve" Mechanism of Viral DNA Packaging. A*dv. Virus Res* 83: 415-465

162. Schwartz C, De Donatis GM, Zhang H, Fang H, & Guo P (2013) Revolution Rather Than Rotation of AAA+ Hexameric Phi29 Nanomotor for Viral DsDNA Packaging Without Coiling. V*irology* 443: 28-39

163. Guo P (2014) Biophysical Studies Reveal New Evidence for One-Way Revolution Mechanism of Bacteriophage Phi29 DNA Packaging Motor. B*iophysical Journal* 106: 1837-1838

164. De-Donatis G, Zhao Z, Wang S, Huang PL, Schwartz C, Tsodikov VO, Zhang H, Haque F, & Guo P (2014) Finding of Widespread Viral and Bacterial Revolution DsDNA Translocation Motors Distinct From Rotation Motors by Channel Chirality and Size. C*ell Biosci* 4: 30

165. Ray K, Sabanayagam CR, Lakowicz JR, & Black LW (2010) DNA Crunching by a Viral Packaging Motor: Compression of a Procapsid-Portal Stalled Y-DNA Substrate. V*irology* 398: 224-232

166. Dixit AB, Ray K, & Black LW (2012) Compression of the DNA Substrate by a Viral Packaging Motor Is Supported by Removal of Intercalating Dye During Translocation. P*roc. Natl. Acad. Sci. U. S. A* 109: 20419-20424

167. Harvey SC (2015) The Scrunchworm Hypothesis: Transitions Between A-DNA and B-DNA Provide the Driving Force for Genome Packaging in Double-Stranded DNA Bacteriophages. J *Struct Biol* 189: 1-8

168. Urbaneja MA, Rivqas S, Carrascosa JL, & Valpuesta JM (1994) An Intrinsic-Tryptophan-Fluorescence Study of Phage Phi29 Connector/Nucleic Acid Interactions. E*ur. J. Biochem.* 225: 747-753

169. Tolley AC & Stonehouse NJ (2008) Conformational Changes in the Connector Protein Complex of the Bacteriophage Phi29 DNA Packaging Motor. C*omputational and Mathematical Methods in Medicine* 9: 327-337

170. Grimes S, Ma S, Gao J, Atz R, & Jardine PJ (2011) Role of Phi29 Connector Channel Loops in Late-Stage DNA Packaging. J. *Mol. Biol.* 410: 50-59

171. Tang JH, Olson N, Jardine PJ, Girimes S, Anderson DL, & Baker TS (2008) DNA Poised for Release in Bacteriophage Phi29. S*tructure* 16: 935-943

172. Fang H, Jing P, Haque F, & Guo P (2012) Role of Channel Lysines and "Push Through a One-Way Valve" Mechanism of Viral DNA Packaging Motor. B*iophysical Journal* 102: 127-135

173. Hu B, Margolin W, Molineux IJ, & Liu J (2013) The Bacteriophage T7 Virion Undergoes Extensive Structural Remodeling During Infection. S*cience* 339: 576-579

174. Haque F, Geng J, Montemagno C, & Guo P (2013) Incorporation of Viral DNA Packaging Motor Channel in Lipid Bilayers for Real-Time, Single-Molecule Sensing of Chemicals and Double-Stranded DNA. N*at. Protoc.* 8: 373-392

175. Quinten TA & Kuhn A (2012) Membrane Interaction of the Portal Protein Gp20 of Bacteriophage T4. J *Virol.* 86: 11107-11114

176. Coulter, W. H. (1953) U.S. Patent 2,656,508.

177. Camacho AG, Gual A, Lurz R, Tavares P, & Alonso JC (2003) Bacillus Subtilis Bacteriophage SPP1 DNA Packaging Motor Requires Terminase and Portal Proteins. J. *Biol. Chem.* 278: 23251-23259

178. Tsuprun V, Anderson D, & Egelman EH (1994) The Bacteriophage F29 Head-Tail Connector Shows 13-Fold Symmetry in Both Hexagonally Packed Arrays and As Single Particles. B*iophys. J.* 66: 2139-2150

179. Orlova EV, Gowen B, Droge A, Stiege A, Weise F, Lurz R, van HM, & Tavares P (2003) Structure of a Viral DNA Gatekeeper at 10 A Resolution by Cryo-Electron Microscopy. E*MBO J.* 22: 1255-1262

180. Serwer P, Wright ET, Liu Z, & Jiang W (2014) Length Quantization of DNA Partially Expelled From Heads of a Bacteriophage T3 Mutant. V*irology* 456-457: 157-170

181. Bjornsti MA, Reilly BE, & Anderson DL (1983) Morphogenesis of Bacteriophage F29 of B*acillus Subtilis:* Oriented and Quantized i*n Vitro* Packaging of DNA Protein Gp3. J. *Virol.* 45: 383-396

182. Pi F, Vieweger M, Zhao Z, Wang S, & Guo P (2015) Discovery of a New Method for Potent Drug Development Using Power Function of Stoichiometry of Homomeric Biocomplexes or Biological Nanomotors. E*xpert Opin. Drug Deliv.* 1-14

183. Guo P, Zhao Z, Haak J, Wang S, Wu D, Meng B, & Weitao T (2014) Common Mechanisms of DNA Translocation Motors in Bacteria and Viruses Using One-Way Revolution Mechanism Without Rotation. B*iotechnology Advances* 32: 853-872

184. Cuervo A & Carrascosa JL (2011) Viral Connectors for DNA Encapsulation. C*urr. Opin. Biotechnol.*

185. Hu B, Margolin W, Molineux IJ, & Liu J (2015) Structural Remodeling of Bacteriophage T4 and Host Membranes During Infection Initiation. P*roc. Natl. Acad. Sci. U. S A* 112: E4919-E4928

186. Lurz R, Orlova EV, Gunther D, Dube P, Droge A, Weise F, van Heel M, & Tavares P (2001) Structural Organisation of the Head-to-Tail Interface of a Bacterial Virus 1. J *Mol Biol* 310: 1027-1037

187. Kasianowicz JJ, Robertson JW, Chan ER, Reiner JE, & Stanford VM (2008) Nanoscopic Porous Sensors. A*nnu. Rev. Anal. Chem. (Palo. Alto. Calif. )* 1: 737-766

188. Reiner JE, Balijepalli A, Robertson JW, Campbell J, Suehle J, & Kasianowicz JJ (2012) Disease Detection and Management Via Single Nanopore-Based Sensors. C*hem. Rev.* 112: 6431-6451

189. Geng J, Kim K, Zhang J, Escalada A, Tunuguntla R, Comolli LR, Allen FI, Shnyrova AV, Cho KR, Munoz D, Wang YM, Grigoropoulos CP, jo-Franklin CM, Frolov VA, & Noy A (2014) Stochastic Transport Through Carbon Nanotubes in Lipid Bilayers and Live Cell Membranes. N*ature* 514: 612-615

190. Chang H, Venkatesan BM, Iqbal SM, Andreadakis G, Kosari F, Vasmatzis G, Peroulis D, & Bashir R (2006) DNA Counterion Current and Saturation Examined by a MEMS-Based Solid State Nanopore Sensor. B*iomed. Microdevices.* 8: 263-269

191. van den HM, Hall AR, Wu MY, Zandbergen HW, Dekker C, & Dekker NH (2010) Controlling Nanopore Size, Shape and Stability. n*anotechnology* 21: 115304

192. Shahzad A, Knapp M, Lang I, & Kohler G (2011) The Use of Fluorescence Correlation Spectroscopy (FCS) As an Alternative Biomarker Detection Technique: a Preliminary Study. J *Cell Mol Med.* 15: 2706-2711

193. Soskine M, Biesemans A, De Maeyer M, & Maglia G (2013) Tuning the Size and Properties of ClyA Nanopores Assisted by Directed Evolution. J. *Am. Chem. Soc.* 135: 13456-13463

194. Mohammad MM, Iyer R, Howard KR, McPike MP, Borer PN, & Movileanu L (2012) Engineering a Rigid Protein Tunnel for Biomolecular Detection. J. *Am. Chem. Soc.*

195. Haque F, Wang S, Stites C, Chen L, Wang C, & Guo P (2015) Single Pore Translocation of Folded, Double-Stranded, and Tetra-Stranded DNA Through Channel of Bacteriophage Phi29 DNA Packaging Motor. B*iomaterials* 53: 744-752

196. Wen S, Zeng T, Liu L, Zhao K, Zhao Y, Liu X, & Wu HC (2011) Highly Sensitive and Selective DNA-Based Detection of Mercury(II) With Alpha-Hemolysin Nanopore. J. *Am. Chem. Soc.* 133: 18312-18317

197. Wang HY, Gu Z, Cao C, Wang J, & Long YT (2013) Analysis of a Single Alpha-Synuclein Fibrillation by the Interaction With a Protein Nanopore. A*nal. Chem.* 85: 8254-8261

198. Mereuta L, Roy M, Asandei A, Lee JK, Park Y, Andricioaei I, & Luchian T (2014) Slowing Down Single-Molecule Trafficking Through a Protein Nanopore Reveals Intermediates for Peptide Translocation. S*cientific reports* 4:

199. Ying YL, Li DW, Liu Y, Dey SK, Kraatz HB, & Long YT (2012) Recognizing the Translocation Signals of Individual PeptideG 囚 Oligonucleotide Conjugates Using an +?Hemolysin Nanopore. Che*mical Communications 48*: 8784-8786

200. Mereuta L, Asandei A, Seo CH, Park Y, & Luchian T (2014) Quantitative Understanding of PH-and Salt-Mediated Conformational Folding of Histidine-Containing, +?Hairpin-Like Peptides, Through Single-Molecule Probing With Protein Nanopores. ACS *applied materials & interfaces 6:* 13242-13256

201. Sutherland TC, Long YT, Stefureac RI, Bediako-Amoa I, Kraatz HB, & Lee JS (2004) Structure of Peptides Investigated by Nanopore Analysis. Nano *Letters 4:* 1273-1277

202. Uhlen M & Ponten F (2005) Antibody-Based Proteomics for Human Tissue Profiling. Mol *Cell Proteomics 4:* 384-393

203. Barnham KJ, Kenche VB, Ciccotosto GD, Smith DP, Tew DJ, Liu X, Perez K, Cranston GA, Johanssen TJ, Volitakis I, Bush AI, Masters CL, White AR, Smith JP, Cherny RA, & Cappai R (2008) Platinum-Based Inhibitors of Amyloid-Beta As Therapeutic Agents for Alzheimer's Disease. Proc. *Natl. Acad. Sci. U. S A 105*: 6813-6818

204. Wei R, Gatterdam V, Wieneke R, Tampe R, & Rant U (2012) Stochastic Sensing of Proteins With Receptor-Modified Solid-State Nanopores. Nat. *Nanotechnol. 7:* 257-263

205. Plesa C, Kowalczyk SW, Zinsmeester R, Grosberg AY, Rabin Y, & Dekker C (2013) Fast Translocation of Proteins Through Solid State Nanopores. Nano *Lett. 13:* 658-663

206. Fologea D, Ledden B, McNabb DS, & Li J (2007) Electrical Characterization of Protein Molecules by a Solid-State Nanopore. Appl. *Phys Lett. 91:* 539011-539013

207. Wang HY, Ying YL, Li Y, Kraatz HB, & Long YT (2011) Nanopore Analysis of Beta-Amyloid Peptide Aggregation Transition Induced by Small Molecules. Anal. *Chem. 83:* 1746-1752

208. Movileanu L (2008) Squeezing a Single Polypeptide Through a Nanopore. Soft *Matter 4:* 925-931

209. Wang Y, Montana V, Grubisic V, Stout RF, Jr., Parpura V, & Gu LQ (2015) Nanopore Sensing of Botulinum Toxin Type B by Discriminating an Enzymatically Cleaved Peptide From a Synaptic Protein Synaptobrevin 2 Derivative. ACS *Appl. Mater. Interfaces 7:* 184-192

210. Li W, Bell NA, Hernandez-Ainsa S, Thacker VV, Thackray AM, Bujdoso R, & Keyser UF (2013) Single Protein Molecule Detection by Glass Nanopores. ACS *Nano 7:* 4129-4134

211. Siwy Z, Trofin L, Kohli P, Baker LA, Trautmann C, & Martin CR (2005) Protein Biosensors Based on Biofunctionalized Conical Gold Nanotubes. J Am. *Chem Soc 127*: 5000-5001

212. Han A, Schmann G, Mondin G, Bitterli RA, Hegelbach NG, de Rooij NF, & Staufer U (2006) Sensing Protein Molecules Using Nanofabricated Pores. Appli*ed Physics Letters 88(9*): 3901

213. Steinbock LJ, Krishnan S, Bulushev RD, Borgeaud S, Blokesch M, Feletti L, & Radenovic A (2014) Probing the Size of Proteins With Glass Nanopores. Nanos*cale 6: 1*4380-14387

214. Nivala J, Mulroney L, Li G, Schreiber J, & Akeson M (2014) Discrimination Among Protein Variants Using an Unfoldase-Coupled Nanopore. ACS N*ano 8: 1*2365-12375

215. Payet L, Martinho M, Pastoriza-Gallego M, Betton JM, Auvray L, Pelta J, & Mathe J (2012) Thermal Unfolding of Proteins Probed at the Single Molecule Level Using Nanopores. Anal. *Chem 84:* 4071-4076

216. Preus S, Noer SL, Hildebrandt LL, Gudnason D, & Birkedal V (2015) ISMS: Single-Molecule FRET Microscopy Software. Natur*e Methods 12:* 593-594

217. Wang S, Ji Z, Yan E, Haque F, & Guo P (2016) Three**-Step Channel Conformational Changes Common to DNA Packaging Motors of Bacterial Viruses T3, T4, SPP1, and Phi29 . Vir**ol*ogy*

*218.* Maupetit J, Derreumaux P, & Tuffery P (2009) PEP-FOLD: an Online Resource for De Novo Peptide Structure Prediction. Nucle*ic Acids Res 37:* W498-W503

219. Wanunu M, Morrison W, Rabin Y, Grosberg AY, & Meller A (2010) Electrostatic Focusing of Unlabelled DNA into Nanoscale Pores Using a Salt Gradient. Natur*e Nanotechnology 5: 1*60-165

220. Zhou Z, Hu Y, Shan X, Li W, Bai X, Wang P, & Lu X (2015) Revealing Three Stages of DNA-Cisplatin Reaction by a Solid-State Nanopore. Sci. *Rep. 5: 1*1868

221. David C, Foley S, & Enescu M (2009) Protein S-S Bridge Reduction: a Raman and Computational Study of Lysozyme Interaction With TCEP. Phys *Chem. Chem. Phys 11:* 2532-2542

222. Kasianowicz JJ, Brandin E, Branton D, & Deamer DW (1996) Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. Proc. *Natl. Acad. Sci. U. S A 93:* 13770-13773

223. Pastoriza-Gallego M, Breton MF, Discala F, Auvray L, Betton JM, & Pelta J (2014) Evidence of Unfolded Protein Translocation Through a Protein Nanopore. ACS N*ano 8: 1*1350-11360

224. Fennouri A, Przybylski C, Pastoriza-Gallego M, Bacri L, Auvray L, & Daniel R (2012) Single Molecule Detection of Glycosaminoglycan Hyaluronic Acid Oligosaccharides and Depolymerization Enzyme Activity Using a Protein Nanopore. ACS N*ano. 6: 9*672-9678

225. Nivala J, Marks DB, & Akeson M (2013) Unfoldase-Mediated Protein Translocation Through an Alpha-Hemolysin Nanopore. Nat B*iotechnol 31:* 247-250

226. Schoch R, Han J, & Renaud P (2008) Transport Phenomena in Nanofluidics. Revie*ws of Modern Physics 80:* 839

227. Burns JR, Seifert A, Fertig N, & Howorka S (2016) A Biomimetic DNA-Based Channel for the Ligand-Controlled Transport of Charged Molecular Cargo Across a Biological Membrane. Nat. *Nanotechnol. 11:* 152-156

228. Burns JR, Al-Juffali N, Janes SM, & Howorka S (2014) Membrane-Spanning DNA Nanopores With Cytotoxic Effect. Angew. *Chem. Int. Ed Engl. 53:* 12466-12470

229. Hernandez-Ainsa S & Keyser UF (2013) DNA Origami Nanopores: an Emerging Tool in Biomedicine. Nanom*edicine. (Lond) 8: 1*551-1554

230. Langecker M, Arnaut V, Martin TG, List J, Renner S, Mayer M, Dietz H, & Simmel FC (2012) Synthetic Lipid Membrane Channels Formed by Designed DNA Nanostructures. Scien*ce 338:* 932-936

231. Gopfrich K, Li CY, Mames I, Bhamidimarri SP, Ricci M, Yoo J, Mames A, Ohmann A, Winterhalter M, Stulz E, Aksimentiev A, & Keyser UF (2016) Ion Channels Made From a Single Membrane-Spanning DNA Duplex. Nano *Lett. 16:* 4665-4669

232. Guo P (2010) The Emerging Field of RNA Nanotechnology. Natur*e Nanotechnology 5: 8*33-842

233. Li H, Lee T, Dziubla T, Pi F, Guo S, Xu J, Li C, Haque F, Liang X, & Guo P (2015) RNA As a Stable Polymer to Build Controllable and Defined Nanostructures for Material and Biomedical Applications. Nano *Today 10:* 631-655

234. Shukla GC, Haque F, Tor Y, Wilhelmsson LM, Toulme JJ, Isambert H, Guo P, Rossi JJ, Tenenbaum SA, & Shapiro BA (2011) A Boost for the Emerging Field of RNA Nanotechnology. ACS N*ano 5: 3*405-3418

235. Shu Y, Pi F, Sharma A, Rajabi M, Haque F, Shu D, Leggas M, Evers BM, & Guo P (2014) Stable RNA Nanoparticles As Potential New Generation Drugs for Cancer Therapy. Adv. *Drug Deliv. Rev. 66C:* 74-89

236. Zhou Z, Hu Y, Shan X, Li W, Bai X, Wang P, & Lu X (2015) Revealing Three Stages of DNA-Cisplatin Reaction by a Solid-State Nanopore. Sci. *Rep. 5: 1*1868

237. Mhlanga MM & Tyagi S (2006) Using TRNA-Linked Molecular Beacons to Image Cytoplasmic MRNAs in Live Cells. Nat. *Protoc. 1: 1*392-1398

# Vita

**Shaoying Wang**

## Educational Institutions

| | | | |
|---|---|---|---|
| China Pharmaceutical University | 2004 – 2008 | Bachelors | Biotechnology |
| University of Kentucky | 2012 - 2016 | PhD Candidate | Pharmaceutical Sciences |

## Professional Publications

1. **Wang S**, Zhou Z, Zhao,Z, Zhang H, Haque F and Guo P. Channel of SPP1 DNA Packaging Motor for Real Time Kinetic Analysis of Peptide States. *Submitted to Scientific Reports*.
2. Ji Z, **Wang S**, Zhao Z, Zhou Z, Haque F and Guo P. Fingerprinting of Peptides with a Large Channel of Bacteriophage Phi29 DNA Packaging Motor. ***Small***. 20 JUL 2016
3. Zhao Z, Ji Z, **Wang S**, Zhao Z, Haque F and Guo P. Oriented Single Directional Insertion of Nanochannel of Bacteriophage SPP1 DNA Packaging Motor into Lipid Bilayer via Polar Hydrophobicity. ***Biomaterial***.
4. **Wang S**, Ji Z, Yan E, Haque F and Guo P. Three-step Channel Conformational Changes Common to DNA Packaging Motors of Bacterial Viruses T3, T4, SPP1, and Phi29. ***Virology***, 2016.
5. Pi F, Vieweger M, Zhao Z, **Wang S** and Guo P. Discovery of a new method for potent drug development using power function of stoichiometry of homomeric biocomplexes or biological nanomotors. ***Expert Opin Drug Deliv***. 2015 Aug 26
6. Haque F, **Wang S**, Stites C, Chen L, Wang C and Guo P. Single pore translocation of folded, double-stranded, and tetra-stranded DNA through channel of bacteriophage phi29 DNA packaging motor. ***Biomaterials***. 2015 Jun;53:744-52.
7. Li J, Tang Z, Hu R, Fu Q, Yan E, **Wang S**, Guo P, Zhao Q. Probing surface hydrophobicity of individual protein at single-molecule resolution using solid-state nanopores. ***Science China Materials.*** June 2015
8. De-Donatis GM, Zhao Z, **Wang S**, Huang LP, Schwartz C, Tsodikov OV, Zhang H, Haque F and Guo P. Finding of widespread viral and bacterial revolution dsDNA translocation motors distinct from rotation motors by channel chirality and size. ***Cell Biosci***. 2014 Jun 1;4:30.
9. Guo P, Zhao Z, Haak J, **Wang S**, Wu D, Meng B and Weitao T. Common mechanisms of DNA translocation motors in bacteria and viruses using one-way revolution mechanism without rotation. ***Biotechnol Adv***. 2014 Jul-Aug;32(4):853-72
10. Guo P, Haque F, Hallahan B, Reif R, Li H, and **Wang S.** Uniqueness, Advantages, Challenges, Solutions, and Perspectives in Therapeutics Applying RNA Nanotechnology. ***Book chapter in RNA Nanotechnology and Therapeutics***. July 9, 2013 by CRC press.
11. **Wang S**, Haque F, Rychahou P, Evers M and Guo P. Engineered nanopore of Phi29 DNA-packaging motor for real-time detection of single colon cancer specific antibody in serum. ***ACS Nano***. 2013 Nov 26;7(11):9814-22.
12. Geng J, **Wang S**, Fang H and Guo P. Channel size conversion of Phi29 DNA-packaging nanomotor for discrimination of single- and double-stranded nucleic acids. ***ACS Nano***. 2013 Apr 23;7(4):3315-23
13. **Wang S**, Qin B, and Cheng K. Delivery of Nucleic Acids. ***Book Chapter 14 in***

***Advanced Drug Delivery***. ISBN: 978-1-118-02266-5. November 2013.
14. Yang X, Yuan X, Cai D, **Wang S** and Zong L. Low molecular weight chitosan in DNA vaccine delivery via mucosa. ***Int J Pharm.*** Jun 22, 2009.

## Scholastic and Professional Honors

| | |
|---|---|
| Oral Presentation at RNA Nanotechnology Conference, London, UK | 2016 |
| First Place of Research Day Award at College of Pharmacy Research Day of The Ohio State University | 2016 |
| Second place of poster presentation at the Seventh Markey Cancer Center Research Day | 2014 |
| Oral Presentation at Annual Scientific Meeting of American Society for Virology | 2014 |