Published in final edited form as:

Curr Opin Biotechnol. 2018 June ; 51: 80–89. doi:10.1016/j.copbio.2017.11.006.

Engineering of Protein Nanopores for Sequencing, Chemical or Protein Sensing and Disease Diagnosis

Shaoying Wang¹, Zhengyi Zhao², Farzin Haque^{1,*}, and Peixuan Guo^{3,*}

¹P&Z Biological Technology, Newark, NJ, USA.

²Nanobio Delivery Pharmaceutical Co. Ltd, Columbus, OH, USA.

³College of Pharmacy, Division of Pharmaceutics & Pharmaceutical Chemistry; College of Medicine, Comprehensive Cancer Center; and Center for RNA Nanobiotechnology and Nanomedicine; The Ohio State University, Columbus, OH, USA.

Abstract

Biological systems contain highly-ordered structures performing diverse functions. The elegant structures of biomachines have inspired the development of nanopores as single molecule sensors. Over the years, the utility of nanopores for detecting a wide variety of analytes have rapidly emerged for sensing, sequencing and diagnostic applications. Several protein channels with diverse shapes and sizes, such as motor channels from bacteriophage Phi29, SPP1, T3, and T4, as well as α -hemolysin, MspA, aerolysin, FluA, OmpF/G, CsgG, ClyA, have been continually investigated and developed as nanopores. Herein, we focus on advances in biological nanopores for single molecule sensing and DNA sequencing from a protein engineering standpoint for changing pore sizes, altering charge distributions, enhancing sensitivity, improving stability, and imparting new detection capabilities.

GRAPHICAL ABSTRACT

INTRODUCTION

The intriguing design and elegant architectures of biological machineries have inspired the development of nanopores. The proof-of-concept was established in 1996 using a

*Address correspondence to: Peixuan Guo, Ph.D., Sylvan G. Frank Endowed Chair in Pharmaceutics and Drug Delivery, The Ohio State University, 460 W 12th Avenue, 418 Biomedical Research Tower, Columbus, OH 43210, USA, guo.1091@osu.edu; Phone: (614) 293-2114, Farzin Haque, Ph.D., Director of Research, P&Z Biological Technology, 211 Warren Street, Suite 503, Newark, NJ-07103, USA, farzin.haque@pzbiologyusa.com; Phone: (920) 284-4749.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest

membrane-embedded α -hemolysin channel[1]. Single-stranded DNA was electrically driven through α -hemolysin and the resulting translocation was verified by qRT-PCR. Over the last decades, nanopore has demonstrated potentials for sensing a wide range of analytes, and is well positioned to bring a revolution to medical diagnostics and DNA/RNA sequencing. Compared with other technologies, nanopore offers[2;3]: single molecule detection with high specificity and sensitivity; is label-free; is amplification-free; offers real time recognition; requires low sample volumes and minimal sample processing; rapid electrical detection method; high-throughput in nature; multiplexing capabilities; and requires no special expertise for operation.

To date, several protein channels, such as α -hemolysin[4], MspA[5], aerolysin[6], FluA[7], Omp F/G[8;9], CsgG[10], ClyA[11], PA₆₃[12] and, viral connectors phi29[13], SPP1[14], T3[14] and T4[14] (Fig. 1) have been successfully developed as nanopore for detection of small molecules, polymers, polypeptides, and DNA/RNA[2;15]. Protein engineering, such as site directed mutagenesis, insertion, and deletion of amino acids, and introduction of functional modules have been extensively employed to tune nanopore properties[16]. Different analytes require unique nanopores with different shape, size, and hydrophilic/hydrophobic properties in order to be detected with high sensitivity and specificity. Herein, we summarize progress in engineering protein nanopores for sensing and sequencing applications. Due to space limitations, we omitted discussion on DNA-based biological pores and synthetic pores, which have been reviewed elsewhere[17;18].

GENERAL STRATEGIES FOR ENGINEERING PROTEIN NANOPORE OR CHANNELS

Among various channel forming proteins[16;19], only a handful are fit to serve as nanopores (Fig. 1). A major limitation is the lack of crystal structures to guide engineering at specific sites. We also have limited knowledge of how the structure of the channel relates to its function. Generally, to be a good candidate, wild-type or engineered channel needs to meet at least four criteria: (1) robust with channel forming capability after mutagenesis; (2) appropriate geometry including the entrance, vestibule, and constriction; (3) proper charge and hydrophilic/hydrophobic surface; and, (4) stable current with controllable voltage gating.

Several engineering approaches have been employed to change the intrinsic properties of nanopores for increasing the stability, changing the oligomeric state, introducing recognition sites, enhancing sensitivity, facilitate membrane insertion, and imparting selectivity for target analyte (Fig. 2). The first approach is based on site-directed mutagenesis that substitutes specific amino acids at certain locations in the primary sequence. The availability of a wide range of unnatural amino acids with unique side chains and functional properties makes it possible to expand the scope of nanopore engineering. Furthermore, mutations can result in conformational changes of the channel, making it possible to tune the channel size and to enhance the sensing scope and sensitivity. The second approach is to attach specific tags and enzymes (such as, His-tag, Strep tag, or TAT as well as proteins like DNA polymerase, helicase and exonuclease) either through fusion protein expression or in vitro

conjugation to the anticipated sites of the pore. The third approach is to introduce adapter ligands either covalently or non-covalently into the pore for substrate binding. The fourth approach involves more significant truncation or insertion of several amino acids to alter the chain segment aggressively.

ENGINEERING PROTEIN NANOPORES OR CHANNELS FOR DNA AND RNA SEQUENCING

α -hemolysin

The idea of using nanopores for sequencing of an intact DNA strand was proposed three decades ago by David Deamer and Daniel Branton. Since the publication of the landmark paper[1] showing translocation of ssDNA through α -hemolysin, several follow-up studies showed that DNA/RNA homo-polymers or co-polymers can be distinguished[20;21]. In addition, through immobilizing ssDNA *via* biotin-streptavidin conjugation on the channel entrance, three recognition sites were identified that are located near the constriction, in the middle of the barrel, and the trans-entrance[22]. With modification of three recognition sites E111N/K147N/M113Y, epigenetic DNA modification, 5-methylcytosine and 5-hydroxymethylcytosine could be identified[23]. The next challenge was to slow the translocation of DNA. Several passive approaches have been tested including increasing viscosity[24], decreasing temperature[25], adding organic salts[26], as well as more active approaches by attaching complementary DNA sequences[27], and enzymes[28]. Incorporation of enzyme turned out to be an efficient way to slow DNA translocation enough to enable DNA sequencing. Initial success was achieved by introducing T7 DNA polymerase into α -hemolysin. The blockage current caused by sequential nucleotide incorporation were recorded when DNA-polymerase complex was held at the entrance of the nanopore[28]. However, three sequential 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[29]. Later, it was shown that phi29 DNA polymerase could overcome these challenges because it can maintain normal elongation and remain bound to the DNA template while captured in the nanopore[29]. With phi29 DNA polymerase, α -hemolysin is able to ratchet DNA through the pore at 5-Å precision[29] (Fig. 3A).

By using an exonuclease approach (Fig. 3B), a DNA strand was digested and the released nucleotides were then captured and sensed with non-covalent adapter[3]. Individual RNA bases and even modified bases can be distinguished by immobilizing ssRNA with α -hemolysin mutant E111N/K147N/M113Y[30]. The feasibility of direct RNA sequencing was demonstrated by introducing derivatives of β -cyclodextrin as adaptors and polynucleotide phosphorylase which cut ssRNA in the 3'→5' direction processively and liberated rNDPs[31] (Fig. 3C). Kinetic parameters for rNDP identification was thoroughly compared in different mutant systems to identify RNA nucleotides[31].

Another approach for DNA sequencing is the combination of sequencing and synthesis using nanopore[32] (Fig. 3D). As DNA polymerase elongates the template, the incoming complementary-tagged nucleotide is cleaved and the nucleotide can be differentiated by

different-sized polymer tags[32]. Herein, a single polymerase is covalently attached to a single subunit of α -hemolysin at an optimized distance that permits fast capture of the tag.

Phi29 and other channels of viral DNA packaging motors

The membrane-embedded phi29 nanopore is the first example of a protein channel that is neither an ion channel nor a membrane protein[13;33]. Compared to other protein channels, phi29 channel has a large diameter of 3.6 nm and thus suitable for conjugating various adapters in the lumen for sensing capabilities and for protein engineering to create a sharper detection region for sequencing capabilities. The channel can translocate ssDNA/ssRNA, dsDNA/dsRNA, peptides and polymers[13;34–37]. The larger channel makes it possible to detect larger molecules. The connector can be functionalized with various probes that can bind individual targets with high selectivity and sensitivity[38;39]. The connector shrinks and exhibits discrete step-wise conformational changes in response to ligand binding to any of its twelve subunits[39]. The distinctive current signatures resulting from ligand triggered defined channel shrinking can clearly discriminate specific and non-specific binding at single molecule level with high accuracy. Dynamic interactions between the analyte and the binding sites can also be studied in real-time at single molecule level[38;39].

In native phi29 DNA packaging motor, the speed of dsDNA translocation is about 100-200 base pairs per second, which is well within the current detection technologies[40]. Using one ATP, the motor can package 1.75 base pairs of DNA[37;41;42]. The initial packaging rate is 6 ms/bp, which is 1000-fold slower than the speed of DNA electrophoretically driven through nanopores. Constructing an artificial phi29 DNA packaging motor embedded in a membrane offers the possibility of developing an active ATP-driven DNA-sequencing apparatus by combining the advantages of viral motors and nanopore technology.

MspA

Wild-type MspA exhibited spontaneous blockades without DNA under potentials >60 mV. To circumvent this problem, a mutant channel was generated whereby negatively charged aspartates were replaced with asparagines in the constriction zone resulting in D90N/D91N/D93N mutant[43;44]. Radical protein engineering, such as truncated barrel mutant (TBM 6), Met113→Gly, and replacement of negatively charged residues around the entrance and in the vestibule with positively charged residues (D90N/D91N/D93N/D118R/D134R/E139K) generated a sharper reading head and improved the resolution for base identification[45]. Positively charged β -cyclodextrin derivatives can also reside in the interior of the mutant, thereby allowing continuous recognition of mononucleotides, which is a prerequisite for exonuclease sequencing.

By combining reengineered MspA with phi29 DNA polymerase, the nanopore was able to read DNA at single-nucleotide resolution[43] (Fig. 3E) as well as detect 5-methylcytosine and 5-hydroxymethylcytosine, which are common DNA lesions[46;47]. For long reads of DNA sequence, the platform was modified by introducing two asymmetric adaptors: a cholesterol-tailed adaptor and a nicked, hairpin adaptor to the ends of linearized phiX174 dsDNA. Since each four-nucleotide combination has a unique current fingerprint, a current blockade map of all 256 four-nucleotide combinations was generated. Using this map as a

reference, MspA is able to read 4500-bp sequence corresponding to natural DNA sequences[44]. However, errors can result from shifts in the positioning of the DNA within the pore's constriction and nucleotides outside of the quadromer.

The speed of DNA translocation in MspA is faster than α -hemolysin[5] due to differences between the widths of the channel regions flanking the narrowest constrictions. In α -hemolysin, about 10–20 bases are confined in the constriction site, whereas in MspA, there are only 2–4 bases. Another reason is the charge distribution within the constriction site. In α -hemolysin, this region is formed by E111, K147, and M113, whereas there are no charged residues in the constriction of MspA mutant, which may be responsible for the faster translocation speed[5]. These results demonstrate that a better understanding of nanopore geometry and charge distribution is critical for sequencing and sensing.

Commercial ventures

Oxford Nanopore Technologies was the first company to release a DNA sequencing prototype in the market starting with the *MinION*, a portable hand-held device operated through the USB port of a laptop. The detection platform is based on a mutant CsgG pore. The chip contains 4 segments each housing 512 channels and needs ~10 pg DNA without PCR amplification. Read length can be up to 200 kbp. By analyzing more than 5 bases at any one time, the sequence can be decoded by a reference map[48]. The accuracy in 1D reads (linear reads) and 2D (read both the template and complement strands) reads has been reported to be 90% and 95% respectively[49]. Other products under development include a higher-throughput *PromethION* instrument with increasing flow cells and a smaller smart-phone based *SmidgION* for field analyses. *Genia Technologies* (acquired by Roche) employs a Nanotag sequencing approach[50] and sequencing by synthesis with α -hemolysin[32]. *NobleGen Biosciences* coupled color-coded 'beacon and nanopores for sequencing[51]. Finally, *Illumina* licensed the MspA-nanopore based patents but no developments have been reported so far.

ENGINEERING PROTEIN NANOPORES FOR SINGLE CHEMICAL OR MACRO-MOLECULE SENSING

Another burgeoning area is single molecule detection of ions, small molecules, macromolecules, and biopolymers using nanopores. Various approaches have been used to introduce probes, including site directed mutagenesis, fused expression, and covalent or non-covalent attachment of adaptors (Fig. 4). Compared with solid state pores, protein pores are amenable for selective modification with various probes that can interact with individual targets with high sensitivity and selectivity.

Sensing directly using site directed mutagenesis

Introducing artificial molecular recognition sites can greatly broaden the sensing capabilities of nanopores. Four histidines were introduced into α -hemolysin subunit which was then incorporated into a heteromeric pore composed of six wild-type and one mutant α -hemolysin subunits (WT₆H₁). The resulting pore was able to detect two or more divalent metal ions simultaneously[52] (Fig. 4A). TNT can be detected by three aromatic

(Phe, Trp, Tyr) Met113 α -hemolysin mutants[53]. In phi29 nanopore, ethane, thymine and benzene with reactive thioester moieties were clearly distinguished upon binding with cysteine mutated from K234[38] (Fig. 4D–E). Fingerprints of various peptides and oligomeric states of peptides can be detected by phi29 channel based on distinctive current blockage signatures[54]. One-way traffic of peptide translocation has been observed in both SPP1 and phi29 motor channels[55]. Furthermore, the orientation of channels inserted into membrane was governed by the hydrophobicity/hydrophilicity of the terminal ends of the protein.

Sensing via probes introduced through covalent linker

One end of a single 3.4 kDa polyethylene glycol chain end-labeled with biotin was attached to S106C α -hemolysin mutant to detect streptavidin[56] (Fig. 4A). By hybridization of DNA aptamer to N17C α -hemolysin mutant, thrombin can be detected[57] (Fig. 4C). Similarly, engineered ClyA nanopore containing DNA aptamers with activated thiol group can distinguish human and bovine thrombin despite 86% sequence similarities[11]. In addition, the kinetics of binding can be studied as demonstrated by binding of tetravalent lectin to a Gal- β -1,3-GalNAc derived ligand engineered α -hemolysin pore[58]. Similarly, a ligand, PKIP5-24, was tethered to *trans*-entrance of α -hemolysin at position C129 to study the binding interaction of cAMP-dependent protein kinase to the ligand[59] (Fig. 4B). In another study, a single photolabile carbamate group was tethered to the interior wall at Thr117Cys of α -hemolysin. Upon UV illumination, the intermediates of the decomposition can be detected because the current depends on the shape, size, and polarity of the reactants tethered within the pore[60]. Similarly, azobenzene can be attached *via* disulfide bond to Thr117Cys α -hemolysin to study light induced isomerisation at single molecule level[61].

Sensing via probes introduced through fusion protein expression

Introducing new functions through fused expression of the encoded genetic sequence of proteins is challenging. One method is combining evolutionarily preserved regions of proteins with generally independent protein domains. In one study, functional elements of co-chaperonin GroES was introduced to α -hemolysin to study the interaction with GroEL at the single molecule level. The fused flexible loops of GroES showed correct folding with enzymatic activity mediating GroEL-assisted protein folding efficiently as native GroES[62]. In another study, epithelial cell adhesion molecule (EpCAM) peptide was co-expressed with phi29 connector at the C-terminus[39]. Binding of EpCAM antibodies sequentially to each peptide probe induced stepwise blocks in current (Fig. 4F–G). The kinetics of probe-antibody docking can be analyzed in real time at single molecule level by the distinctive current signatures. Importantly, EpCAM antibody can be distinguished in presence of serum or nonspecific antibody (Fig. 4G).

Sensing with non-covalent adaptors

Lodging of an adapter in the lumen of protein nanopores can provide host-guest interactions. One class of adapters for binding small analytes are ring-shaped molecules, such as β -cyclodextrin, which has a hydrophobic interior cavity and a hydrophilic exterior. The possible lodging sites of β -cyclodextrins at the internal end of the β -barrel were enhanced by site directed mutagenesis at Met113, Lys147 and Glu111. The binding affinity was increased

10^4 fold compared to wild type, resulting in prolonged residence time of about tens of seconds[63] (Fig. 4B). Organic ligands or drugs reversibly bound to the adapter and resulted in analyte characteristic blockade fingerprints[63].

Sensing with covalent adaptors

β -cyclodextrin can also be covalently attached to Met113Cys α -hemolysin through disulfide bond[64]. This is especially useful for single molecule exonuclease DNA sequencing which requires continuous presence of a molecular adapter. The orientation of the adapter can also be fixed which will allow analytes to bind through only one of the entrances to the cyclodextrin cavity[64]. For instance, unlabeled nucleoside 5'-monophosphate and methylated cytosine can be identified by a covalently attached adapter with accuracies approaching 99.8% [65].

Sensing via conformational changes on the channel

Recently a novel concept of sensing via conformational change was demonstrated by phi29 connector. The underlying principle is based on step-wise conformational changes of the channel induced by binding of ligands to probes immobilized on the connector. The DNA packaging motor of bacteriophage SPPI, T3, T4 and phi29 all exercise three discrete steps of gating (shrinking) during DNA translocation, with each step reducing 32% of channel dimensions. This is due to the unique biological property of channel conformational changes via chirality transition of the DNA packaging motor. During DNA packaging, the motor uses the left-handed channel to translocate the dsDNA, and subsequent ejection of genomic dsDNA for host infection requires the change to the right-handed channel to control the one-way traffic of dsDNA[14]. As proof-of-concept demonstrating the use of the proposed method for the detection of analytes with high sensitivity and specificity, an Epithelial Cell Adhesion Molecule (EpCAM) peptide was incorporated into the C-terminus of the connector as a probe to specifically detect EpCAM antibody (Ab) in the serum at nM concentration[39]. The specific signal of EpCAM Ab can be clearly distinguished from non-specific background signals arising from non-specific Ab or serum proteins (Fig. 4F–G).

Changing oligomeric state of channel

Control of subunit stoichiometry is critical to tailor protein nanopores for various applications since it can change the constriction zone size and physical properties. In α -hemolysin both homo-heptameric and hetero-heptameric forms have been achieved. Since both N-terminus and C-terminus of a single subunit of α -hemolysin are on the same side of the lipid membrane, the N-terminus and C-terminus of two subunits can be linked to form dimers to generate an even number of subunits of α -hemolysin. The dimer still can assemble into a heptamer form indicating that one or more subunits are expelled from the central pore[66]. In MspA, the oligomer state was changed by connecting two mutated MspA subunits through a 17–62 amino acid linker and the resulting mutant formed functional channels[67]. In phi29 connector, deleting a loop in the central part of the channel reduced the conductance by 40% compared to wild-type indicating change in the oligomeric state. The resulting mutant can distinguish ssRNA and ssDNA[68]. Although the connectors from different phages exist mostly as dodecamers[69], the stoichiometry can vary from 11- to 14-mer *in vitro* following ectopic expression and assembly[70–72].

PERSPECTIVES

In the last two decades, significant progress has been made in applying biological nanopores for DNA sequencing and single molecule detection. Now nanopores are well positioned to bring a revolution to medical diagnostics and DNA sequencing market. To date, several protein channels such as α -hemolysin, MspA, aerolysin, FlpA, Omp G, and viral motor channels Phi29, SPP1, T3 and T4 have been successfully used for incredibly versatile applications. Several challenges still remain, such as limited knowledge of structure-function of protein channels. Analytes that can be detected by nanopores are still limited because different analytes require unique protein nanopore with different shape, size, physical dimensions and hydrophilic/hydrophobic properties. Therefore, it still requires continuous improvement of existing nanopores and exploration of new channels. How to improve the sensitivity, stability and specificity especially using clinical samples, which is often in the presence of many impurities, is still the major bottleneck to transition the technology to the clinic for diagnostic use.

One intriguing application is to incorporate self-assembled protein channels in live cells as a passageway to secrete DNA from the cytoplasm for instance to the periplasm of bacteria. Such a system would perhaps enable a novel type of communication in bacteria based on the presence of DNA in outer-membrane vesicles. Furthermore, transport of DNA/RNA across the membranes, which is a common process in living cells, could have important applications in synthetic biology, for instance to edit the genome or for cell-to-cell communication or overcome drug resistance. A proof of concept was recently demonstrated using Carbon Nanotube-based porins which can spontaneously insert into the membranes of live cells and display stable single-channel activity [76].

For DNA /RNA sequencing, several key challenges that used to be the bottlenecks in nanopore-based sequencing has now been somewhat addressed, such as procedures to overcome the fast translocation speed, spatial resolution and platform stability. It is reported that over 700 mutants of R9 have been screened by Oxford Nanopore[49]. However, the accuracy is still low compared with current sequencing technologies. Additional efforts are still needed to enhance nanopore signal detecting capabilities, optimize the platform through protein engineering, large scale screening, new nanopore exploration and improving the algorithm in the future.

Acknowledgments

The work was supported by NIH grants R01EB012135 to P.G. Funding to PG's Endowed Chair position is by CM Chen Foundation. P.G. is a consultant of Oxford Nanopore Technologies and Nanobio Delivery Pharmaceutical Co. Ltd. He is also the co-founder of P&Z Biological Technology.

References

1. Kasianowicz JJ, Brandin E, Branton D, Deamer DW. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. U. S. A.* 1996; 93:13770–13773. • This work is the first experimental evidence demonstrating that a nanopore could be used for nucleic acids detection using a membrane-embedded α -hemolysin channel. [PubMed: 8943010]
2. Haque F, Li J, Wu H-C, Liang X-J, Guo P. Solid-state and biological nanopore for real-time sensing of single chemical and sequencing of DNA. *Nano Today.* 2013; 8:56–74. [PubMed: 23504223]

3. 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. The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* 2008; 26:1146–1153. [PubMed: 18846088]
4. Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, Gouaux JE. Structure of Staphylococcal alpha -Hemolysin, a Heptameric Transmembrane Pore. *Science.* 1996; 274:1859–1865. [PubMed: 8943190]
5. Butler TZ, Pavlenok M, Derrington IM, Niederweis M, Gundlach JH. Single-molecule DNA detection with an engineered MspA protein nanopore. *Proceedings of the National Academy of Sciences.* 2008; 105:20647–20652. • By mutagenesis of the negative charge of MspA in the constriction part, this work has first time shown that MspA nanopore has the capability to detect ssDNA which built the foundation for future DNA sequencing studies.
6. Stefureac R, Long YT, Kraatz HB, Howard P, Lee JS. Transport of alpha-helical peptides through alpha-hemolysin and aerolysin pores. *Biochemistry.* 2006; 45:9172–9179. [PubMed: 16866363]
7. Mohammad MM, Howard KR, Movileanu L. Redesign of a plugged beta-barrel membrane protein. *J. Biol. Chem.* 2011; 286:8000–8013. [PubMed: 21189254]
8. Chimere C, Movileanu L, Pezeshki S, Winterhalter M, Kleinekathofer U. Transport at the nanoscale: temperature dependence of ion conductance. *Eur. Biophys. J.* 2008; 38:121–125. [PubMed: 18726094]
9. Zhuang T, Tamm LK. Control of the conductance of engineered protein nanopores through concerted loop motions. *Angew. Chem. Int. Ed Engl.* 2014; 53:5897–5902. [PubMed: 24777684]
10. 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. Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature.* 2014; 516:250–253. [PubMed: 25219853]
11. Soskine M, Biesemans A, Moeyaert B, Cheley S, Bayley H, Maglia G. An engineered ClyA nanopore detects folded target proteins by selective external association and pore entry. *Nano Lett.* 2012; 12:4895–4900. [PubMed: 22849517]
12. Nestorovich EM, Karginov VA, Berezhkovskii AM, Bezrukov SM. Blockage of anthrax PA63 pore by a multicharged high-affinity toxin inhibitor. *Biophys. J.* 2010; 99:134–143. [PubMed: 20655841]
13. Wendell D, Jing P, Geng J, Subramaniam V, Lee TJ, Montemagno C, Guo P. Translocation of double-stranded DNA through membrane-adapted phi29 motor protein nanopores. *Nat. Nanotechnol.* 2009; 4:765–772. • Pioneering paper demonstrating that a motor channel from bacteriophage phi29 is able to insert into lipid bilayer and detect double stranded DNA translocation. [PubMed: 19893523]
14. Wang S, Ji Z, Yan E, Haque F, Guo P. Three-step channel conformational changes common to DNA packaging motors of bacterial viruses T3, T4, SPP1, and Phi29. *Virology.* 2016
15. Howorka S, Siwy Z. Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev.* 2009; 38:2360–2384. [PubMed: 19623355]
16. Bayley H, Jayasinghe L. Functional engineered channels and pores (Review). *Molecular Membrane Biology.* 2004; 21:209–220. [PubMed: 15371010]
17. Hernandez-Ainsa S, Keyser UF. DNA origami nanopores: developments, challenges and perspectives. *Nanoscale.* 2014; 6:14121–14132. [PubMed: 25325422]
18. Kudr J, Skalickova S, Nejd L, Moulick A, Ruttkay-Nedecky B, Adam V, Kizek R. Fabrication of solid-state nanopores and its perspectives. *Electrophoresis.* 2015; 36:2367–2379. [PubMed: 26046318]
19. Cuervo A, Carrascosa JL. Viral connectors for DNA encapsulation. *Curr. Opin. Biotechnol.* 2012; 23:529–536. [PubMed: 22186221]
20. Meller A, Nivon L, Brandin E, Golovchenko J, Branton D. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Acad. Sci. U. S. A.* 2000; 97:1079–1084. [PubMed: 10655487]

21. Akeson M, Branton D, Kasianowicz JJ, Brandin E, Deamer DW. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.* 1999; 77:3227–3233. [PubMed: 10585944]
22. Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proceedings of the National Academy of Sciences.* 2009; 106:7702–7707.
23. Wallace EV, wallace D, Heron AJ, Mikhailova E, Maglia G, Donohoe TJ, Bayley H. Identification of epigenetic DNA modifications with a protein nanopore. *Chem. Commun. (Camb.).* 2010; 46:8195–8197. [PubMed: 20927439]
24. Kawano R, Schibel AEP, Cauley C, White HS. Controlling the Translocation of Single-Stranded DNA through alpha-Hemolysin Ion Channels Using Viscosity. *Langmuir.* 2008; 25:1233–1237.
25. Meller A, Branton D. Single molecule measurements of DNA transport through a nanopore. *Electrophoresis.* 2002; 23:2583–2591. [PubMed: 12210161]
26. de Zoysa RS, Jayawardhana DA, Zhao Q, Wang D, Armstrong DW, Guan X. Slowing DNA translocation through nanopores using a solution containing organic salts. *J. Phys. Chem. B.* 2009; 113:13332–13336. [PubMed: 19736966]
27. Howorka S, Cheley S, Bayley H. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nat. Biotechnol.* 2001; 19:636–639. [PubMed: 11433274]
28. Olasagasti F, Lieberman KR, Benner S, Cherf GM, Dahl JM, Deamer DW, Akeson M. Replication of individual DNA molecules under electronic control using a protein nanopore. *Nat Nanotechnol.* 2010; 5:798–806. [PubMed: 20871614]
29. Cherf GM, Lieberman KR, Rashid H, Lam CE, Karplus K, Akeson M. Automated forward and reverse ratcheting of DNA in a nanopore at 5-A precision. *Nat Biotechnol.* 2012; 30:344–348. • In this work, phi29 DNA polymerase was first time incorporated into α -hemolysin nanopore system to control the speed and direction of DNA which enable α -hemolysin to detect the translocation of DNA base-by-base. [PubMed: 22334048]
30. Ayub M, Bayley H. Individual RNA base recognition in immobilized oligonucleotides using a protein nanopore. *Nano Lett.* 2012; 12:5637–5643. [PubMed: 23043363]
31. Ayub M, Hardwick SW, Luisi BF, Bayley H. Nanopore-based identification of individual nucleotides for direct RNA sequencing. *Nano Lett.* 2013; 13:6144–6150. [PubMed: 24171554]
32. 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. Real-time single-molecule electronic DNA sequencing by synthesis using polymer-tagged nucleotides on a nanopore array. *Proc Natl Acad Sci U S A.* 2016; 113:5233–5238. [PubMed: 27091962]
33. Haque F, Geng J, Montemagno C, Guo P. Incorporation of Viral DNA Packaging Motor Channel in Lipid Bilayers for Real-Time, Single-Molecule Sensing of Chemicals and Double-Stranded DNA. *Nat. Protoc.* 2013; 8:373–392. [PubMed: 23348364]
34. Jing P, Haque F, Shu D, Montemagno C, Guo P. One-Way Traffic of a Viral Motor Channel for Double-Stranded DNA Translocation. *Nano Lett.* 2010; 10:3620–3627. [PubMed: 20722407]
35. Jing P, Haque F, Vonderheide A, Montemagno C, Guo P. Robust Properties of Membrane-Embedded Connector Channel of Bacterial Virus Phi29 DNA Packaging Motor. *Mol. Biosyst.* 2010; 6:1844–1852. [PubMed: 20523933]
36. Geng J, Fang H, Haque F, Zhang L, Guo P. Three reversible and controllable discrete steps of channel gating of a viral DNA packaging motor. *Biomaterials.* 2011; 32:8234–8242. [PubMed: 21807410]
37. Fang H, Jing P, Haque F, Guo P. Role of channel Lysines and "Push Through a One-way Valve" Mechanism of Viral DNA packaging Motor. *Biophysical Journal.* 2012; 102:127–135. [PubMed: 2225806]
38. Haque F, Lunn J, Fang H, Smithrud D, Guo P. Real-Time Sensing and Discrimination of Single Chemicals Using the Channel of Phi29 DNA Packaging Nanomotor. *ACS Nano.* 2012; 6:3251–3261. [PubMed: 22458779]

39. Wang S, Haque F, Rychahou PG, Evers BM, Guo P. Engineered Nanopore of Phi29 DNA-Packaging Motor for Real-Time Detection of Single Colon Cancer Specific Antibody in Serum. *ACS Nano*. 2013; 7:9814–9822. [PubMed: 24152066]
40. Zhang H, Schwartz C, De Donatis GM, Guo P. "Push Through One-Way Valve" Mechanism of Viral DNA Packaging. *Adv. Virus Res.* 2012; 83:415–465. [PubMed: 22748815]
41. Guo P, Peterson C, Anderson D. Prohead and DNA-gp3-dependent ATPase activity of the DNA packaging protein gp16 of bacteriophage phi29. *J Mol Biol.* 1987; 197:229–236. [PubMed: 2960820]
42. Schwartz C, Fang H, Huang L, Guo P. Sequential action of ATPase ATP, ADP, Pi and dsDNA in procapsid-free system to enlighten mechanism in viral dsDNA packaging. *Nucleic Acids Res.* 2012; 40:2577–2586. [PubMed: 22110031]
43. Manrao EA, Derrington IM, Laszlo AH, Langford KW, Hopper MK, Gillgren N, Pavlenok M, Niederweis M, Gundlach JH. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat Biotechnol.* 2012; 30:349–353. • By combining a mutant of MspA and phi29 DNA polymerase, this work demonstrates that MspA system can resolve changes in current corresponding to a known DNA sequence. [PubMed: 22446694]
44. 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. Decoding long nanopore sequencing reads of natural DNA. *Nat. Biotechnol.* 2014; 32:829–833. [PubMed: 24964173]
45. Ayub M, Stoddart D, Bayley H. Nucleobase Recognition by Truncated alpha-Hemolysin Pores. *ACS Nano*. 2015; 9:7895–7903. [PubMed: 26114210]
46. Laszlo AH, Derrington IM, Brinkerhoff H, Langford KW, Nova IC, Samson JM, Bartlett JJ, Pavlenok M, Gundlach JH. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc Natl Acad Sci U S A.* 2013; 110:18904–18909. [PubMed: 24167255]
47. Schreiber J, Wescoe ZL, bu-Shumays R, Vivian JT, Baatar B, Karplus K, Akeson M. Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proc Natl Acad Sci U S A.* 2013; 110:18910–18915. [PubMed: 24167260]
48. Jain M, Fiddes IT, Miga KH, Olsen HE, Paten B, Akeson M. Improved data analysis for the MinION nanopore sequencer. *Nat. Methods.* 2015; 12:351–356. • This work evaluated and optimized the performance of commercial nanopore sequencer MinION, demonstrating that MinION is able to sequence long DNA with high accuracy. [PubMed: 25686389]
49. No Thanks, I've already got one. 2016. Technology focus from Clive G Brown, CTO of Oxford Nanopore Technologies. (<https://www.youtube.com/watch?v=nizGyutn6v4>). • This work reported development of Oxford Nanopore Technologies for DNA Sequencing.
50. Kumar S, Tao C, Chien M, Hellner B, Balijepalli A, Robertson JW, Li Z, Russo JJ, Reiner JE, Kasianowicz JJ, Ju J. PEG-labeled nucleotides and nanopore detection for single molecule DNA sequencing by synthesis. *Sci. Rep.* 2012; 2:684. [PubMed: 23002425]
51. McNally B, Singer A, Yu Z, Sun Y, Weng Z, Meller A. Optical recognition of converted DNA nucleotides for single-molecule DNA sequencing using nanopore arrays. *Nano Lett.* 2010; 10:2237–2244. [PubMed: 20459065]
52. Braha O, Gu LQ, Zhou L, Lu X, Cheley S, Bayley H. Simultaneous stochastic sensing of divalent metal ions. *Nat Biotech.* 2000; 18:1005–1007.
53. Guan X, Gu LQ, Cheley S, Braha O, Bayley H. Stochastic sensing of TNT with a genetically engineered pore. *Chembiochem.* 2005; 6:1875–1881. [PubMed: 16118820]
54. Ji Z, Wang S, Zhao Z, Zhou Z, Haque F, Guo P. Fingerprinting of Peptides with a Large Channel of Bacteriophage Phi29 DNA Packaging Motor. *Small.* 2016
55. Zhou Z, Ji Z, Wang S, Haque F, Guo P. Oriented single directional insertion of nanochannel of bacteriophage SPP1 DNA packaging motor into lipid bilayer via polar hydrophobicity. *Biomaterials.* 2016; 105:222–227. [PubMed: 27529454]
56. Movileanu L, Howorka S, Braha O, Bayley H. Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nature Biotechnology.* 2000; 18:1091–1095.

57. Rotem D, Jayasinghe L, Salichou M, Bayley H. Protein detection by nanopores equipped with aptamers. *J. Am. Chem. Soc.* 2012; 134:2781–2787. [PubMed: 22229655]
58. Howorka S, Nam J, Bayley H, Kahne D. Stochastic detection of monovalent and bivalent protein-ligand interactions. *Angew. Chem. Int. Ed Engl.* 2004; 43:842–846. [PubMed: 14767954]
59. Xie H, Braha O, Gu LQ, Cheley S, Bayley H. Single-molecule observation of the catalytic subunit of cAMP-dependent protein kinase binding to an inhibitor peptide. *Chem. Biol.* 2005; 12:109–120. [PubMed: 15664520]
60. Luchian T, Shin SH, Bayley H. Single-molecule covalent chemistry with spatially separated reactants. *Angew. Chem. Int. Ed Engl.* 2003; 42:3766–3771. [PubMed: 12923839]
61. Ludwig S, Bayley H. Photoisomerization of an individual azobenzene molecule in water: an on-off switch triggered by light at a fixed wavelength. *J. Am. Chem. Soc.* 2006; 128:12404–12405. [PubMed: 16984176]
62. Ho CW, Van MV, Tsai KC, De Temmerman PJ, Mast J, Maglia G. Engineering a nanopore with co-chaperonin function. *Sci Adv.* 2015; 1:e1500905. [PubMed: 26824063]
63. Gu LQ, Braha O, Conlan S, Cheley S, Bayley H. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature.* 1999; 398:686–690. [PubMed: 10227291]
64. Wu HC, Astier Y, Maglia G, Mikhailova E, Bayley H. Protein nanopores with covalently attached molecular adapters. *J. Am. Chem. Soc.* 2007; 129:16142–16148. [PubMed: 18047341]
65. Clarke J, Wu HC, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* 2009; 4:265–270. [PubMed: 19350039]
66. Hammerstein AF, Jayasinghe L, Bayley H. Subunit dimers of alpha-hemolysin expand the engineering toolbox for protein nanopores. *J. Biol. Chem.* 2011; 286:14324–14334. [PubMed: 21324910]
67. Pavlenok M, Derrington IM, Gundlach JH, Niederweis M. MspA nanopores from subunit dimers. *PLoS ONE.* 2012; 7:e38726. [PubMed: 22719928]
68. Geng J, Wang S, Fang H, Guo P. Channel size conversion of Phi29 DNA-packaging nanomotor for discrimination of single- and double-stranded nucleic acids. *ACS Nano.* 2013; 7:3315–3323. [PubMed: 23488809]
69. Rao VB, Feiss M. The bacteriophage DNA packaging motor. *Annu. Rev. Genet.* 2008; 42:647–681. [PubMed: 18687036]
70. Cingolani G, Moore SD, Prevelige PE Jr, Johnson JE. Preliminary crystallographic analysis of the bacteriophage P22 portal protein. *J. Struct. Biol.* 2002; 139:46–54. [PubMed: 12372319]
71. Dube P, Tavares P, Lurz R, van HM. The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. *EMBO J.* 1993; 12:1303–1309. [PubMed: 8467790]
72. Valpuesta JM, Sousa N, barthelemy I, Fernandez JJ, Fujisawa H, Ibarra B, Carrascosa JL. Structural analysis of the bacteriophage T3 head-to-tail connector. *J. Struct. Biol.* 2000; 131:146–155. [PubMed: 11042085]
73. Bayley H, Cremer PS. Stochastic sensors inspired by biology. *Nature.* 2001; 413:226–230. [PubMed: 11557992]
74. Cheley S, Xie H, Bayley H. A genetically encoded pore for the stochastic detection of a protein kinase. *Chembiochem.* 2006; 7:1923–1927. [PubMed: 17068836]
75. Clarke J, Wu H-C, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology.* 2009; 4:265–270.
76. Geng J, Kim K, Zhang J, Escalada A, Tunuguntla R, Comolli LR, Allen FI, Shnyrova AV, Cho KR, Munoz D, Wang YM, Grigoropoulos CP, Ajo-Franklin CM, Frolov VA, Noy A. Stochastic transport through carbon nanotubes in lipid bilayers and live cell membranes. *Nature.* 2014; 514:612–5. [PubMed: 25355362]

HIGHLIGHTS

- Nanopore offers single molecule detection with high specificity and sensitivity.
- Protein pores have emerged as a versatile platform for single molecule sensing.
- Protein pores can be engineered and conjugated with ligands at specific locations.
- Amenability and homogeneity of protein pore ensure reproducibility of sensing data.
- Nanopores are set to bring a revolution to medical diagnostics and DNA sequencing.

Figure 1. Structure of commonly used biological nanopores

Bottom, cross-section and side views of: (A) phi29 connector (PDB: 1H5W) (B) α -hemolysin (PDB: 3ANZ), (C) MspA (PDB: 1UUN), (D) CsgG (PDB: 4UV3), (E) PA₆₃ (PDB: 1V36), (F) ClyA (PDB: 2WCD), (G) FhuA (PDB: 1BY5), and (H) SPP1 connector (PDB: 2JES).

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

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

Figure 2. Engineering of protein nanopores

Cross-sectional view of α-hemolysin (top) and MspA (bottom) showing common mutation sites and effect of the mutations [5,22,23,53,64].

Figure 3. Strategies for sequencing DNA and RNA at single-nucleotide resolution using DNA polymerase or exonuclease

(A) Forward and reverse ratcheting of DNA through α -hemolysin equipped with phi29 DNA polymerase[29]. (B) Left: Proposed exonuclease-based sequencing model[3]. An exonuclease (pale blue) is attached to the top of a nanopore by chemical linker or genetic fusion. After cleavage by exonuclease, the identity of individual nucleotides is determined sequentially by the level of the current blockades. Right and bottom: Single-channel recording from α -hemolysin with adapter and residual current histogram of nucleotide binding events showing discrimination of dGMP, dTMP, dAMP and dCMP[65]. (C)

Principle of nanopore DNA sequencing by syntheses[32]. A DNA polymerase is linked to α -hemolysin. Template DNA and primer along with the tagged nucleotide form a ternary complex with the polymerase. The tagged nucleotide is captured and detected sequentially with specific current blockades as they are incorporated into the growing DNA strand. **(D)** Nanopore RNA sequencing. A ssRNA oligonucleotide (circles) is cleaved by polynucleotide phosphorylase (green) sequentially. The released nucleotides are detected by mutant α -hemolysin equipped with a cyclodextrin adapter. **(E)** MspA nanopore equipped with phi29 DNA polymerase[43]. Figures reproduced with permission from Ref. [29, 3, 65, 32 and 43].