Biophysical *Journal*
Article

Biophysical Society

Control of subunit stoichiometry in single-chain MspA nanopores

Mikhail Pavlenok,[1] Luning Yu,[2] Dominik Herrmann,[1] Meni Wanunu,[2,3,4] and Michael Niederweis[1,*]
[1]Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama; [2]Department of Physics, Northeastern University, Boston, Massachusetts; [3]Department of Bioengineering, Northeastern University, Boston, Massachusetts; and [4]Department of Chemistry and Chemical Biology, Northeastern University, Boston, Massachusetts

ABSTRACT Transmembrane protein channels enable fast and highly sensitive detection of single molecules. Nanopore sequencing of DNA was achieved using an engineered *Mycobacterium smegmatis* porin A (MspA) in combination with a motor enzyme. Due to its favorable channel geometry, the octameric MspA pore exhibits the highest current level compared with other pore proteins. To date, MspA is the only protein nanopore with a published record of DNA sequencing. While widely used in commercial devices, nanopore sequencing of DNA suffers from significant base-calling errors due to stochastic events of the complex DNA-motor-pore combination and the contribution of up to five nucleotides to the signal at each position. Different mutations in specific subunits of a pore protein offer an enormous potential to improve nucleotide resolution and sequencing accuracy. However, individual subunits of MspA and other oligomeric protein pores are randomly assembled in vivo and in vitro, preventing the efficient production of designed pores with different subunit mutations. In this study, we converted octameric MspA into a single-chain pore by connecting eight subunits using peptide linkers. Lipid bilayer experiments demonstrated that single-chain MspA formed membrane-spanning channels and discriminated all four nucleotides identical to MspA produced from monomers in DNA hairpin experiments. Single-chain constructs comprising three, five, six, and seven connected subunits assembled to functional channels, demonstrating a remarkable plasticity of MspA to different subunit stoichiometries. Thus, single-chain MspA constitutes a new milestone in the optimization of MspA as a biosensor for DNA sequencing and many other applications by enabling the production of pores with distinct subunit mutations and pore diameters.

SIGNIFICANCE Nanopore sequencing of DNA is a fast and cheap technology that uniquely delivers multi-kilobase reads. It is currently used worldwide in many applications such as genome sequencing, epigenetics, and surveillance of viral and bacterial pathogens and has started to revolutionize human lives in medicine, agriculture, and environmental studies. However, the high base-calling error rates prevent nanopore DNA sequencing from reaching its full potential. In this study, we converted octameric MspA into a single-chain pore, enabling different mutations in each subunit to fine-tune the pore geometry and chemistry and optimize MspA as a biosensor for DNA sequencing and many other applications.

INTRODUCTION

Protein pores have attracted enormous interest as biosensors, as they enable fast and highly sensitive detection of analytes by recording current traces. This analytical approach relies on the modulation of the transmembrane conductance of a pore protein by random and transient associations with an analyte and was termed "stochastic sensing" (1–3). Stochastic sensing was pioneered using bacterial and fungal pores (4,5) and has been applied to a wide variety of small molecules and polymers (6). In particular, nucleotide sensing by pore proteins has been developed into a promising DNA-sequencing technology due to low costs, fast processing times, and minimal sample requirements by direct sequencing of single DNA molecules (7). In this method, single-stranded DNA is driven by an electrical field through a membrane-embedded protein pore. The ion current flowing in the pore is reduced by the translocating single-stranded DNA in a nucleotide-specific manner, enabling DNA sequencing by current measurements (8). While the principle of nanopore DNA sequencing was proposed on the basis of work with the staphylococcal α-hemolysin pore (9), it was first demonstrated to be a feasible technology by using the *Mycobacterium smegmatis* porin A (MspA) (10). MspA is a homo-octameric porin (11) located

Submitted September 19, 2021, and accepted for publication January 25, 2022.

*Correspondence: mnieder@uab.edu

Editor: Michael Brown.
https://doi.org/10.1016/j.bpj.2022.01.022

in the outer membrane of *M. smegmatis* (12) and constitutes a major uptake pathway for many hydrophilic nutrient molecules and antibiotics (13–17). MspA has a goblet-like structure with a short and narrow constriction zone (11) that is suitable for sensing nucleotides as they pass through the pore. In addition, MspA is the most stable of the known pore proteins, resisting denaturation by 2% SDS at 100°C (18), and is functional in many different membrane-like environments (19–21), an important feature for technical applications (6). Based on these beneficial characteristics, we first engineered MspA to translocate DNA (22) and then showed that MspA distinguishes the four nucleobases (23). Enzymatic control of the DNA translocation rate using a motor enzyme increased the signal-to-noise ratio of nucleotide sensing by MspA dramatically (24), enabling for the first time nanopore sequencing of DNA (25). Several other channel proteins have also been used for nucleotide sensing, including α-hemolysin (26), CsgG (27), ClyA (28), FraC (29), and aerolysin (30). However, MspA exhibits the highest current level compared with α-hemolysin, CsgG, and aerolysin, likely due to its short channel constriction and its wide vestibule at the pore entrance (31). Mutants of the CsgG nanopore are apparently being used in commercially available nanopore sequencing devices (32), but to our knowledge no molecular details of the performance of CsgG in DNA-sequencing experiments have been published in peer-reviewed journals. This is in contrast to MspA, which has been extensively characterized in nanopore sequencing experiments (25,33–36).

One of the main problems of nanopore sequencing is the very fast translocation of DNA (37). Control of the DNA translocation rate is currently achieved by DNA-processing enzymes, but this adds complexity and stochastic signals from the motor protein (25,38). In addition, the residual current of single-stranded DNA passing through the pore is determined by four to five nucleotides at each position (25). These and other limitations result in raw base-calling errors of up to 12% for MspA (39). To our knowledge, raw base-calling errors are not published for any other nanopore. Many of these challenges could, in principle, be addressed by protein engineering. However, the oligomeric nature of MspA and other protein nanopores severely limits these efforts (40). Distinct mutations in different subunits are required to efficiently improve nucleotide recognition, to control DNA capture and translocation rates, and to control the path of DNA translocation. We call these mutations "asymmetric" because they disrupt the octameric symmetry of the MspA pore. These asymmetric mutations open numerous avenues to improve the performance of the MspA pore in DNA sequencing. While an attempt in this direction has been made by electrophoretically separating permutations of a mixture of two distinct hemolysin subunits after chemical modifications (41), this process is tedious and difficult to scale, and is not usable for a combination of mutations in different subunits. Asymmetric oligomeric pores can only be efficiently produced by synthesizing one gene which encodes all subunits. In a first step, we constructed an MspA dimer with two covalently connected subunits and showed that the MspA dimer indeed assembles to functional MspA pores (40), demonstrating that connecting MspA subunits by linkers is feasible. However, obtaining pure and active single-chain MspA, in which all eight subunits are linked (Fig. 1 *B*), has been elusive.

In this study, we show that it is feasible to convert octameric MspA into a monomeric pore in which all eight subunits are linked. We also constructed and purified single-chain MspA pores composed of three, five, six, and seven linked subunits and demonstrated the channel activity of these proteins in lipid bilayer experiments. These results highlight the potential of single-chain MspA and pave the way to improve its capabilities as a nanosensor for DNA sequencing.

MATERIALS AND METHODS

Chemicals and enzymes

Chemicals were of the highest purity available from Sigma-Aldrich (St. Louis, MO), Merck (Kenilworth, NJ), Invitrogen (Waltham, MA), or Thermo Fisher Scientific (Waltham, MA) unless otherwise noted. The lipid 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC) was purchased from Avanti Polar Lipids (Alabaster, AL). The detergent *n*-octyl-polyoxyethylene (OPOE) was purchased from Santa Cruz Biotechnology (Dallas, TX). Amphiphilic block-copolymer poly(1,2-butadiene)-*b*-poly(ethylene oxide) (PBD-PEO) was obtained from Polymer Source (Montreal, Canada). Middlebrook 7H9 broth and Middlebrook 7H10 agar were from Difco (Franklin Lakes, NJ). Restriction enzymes and other molecular biology reagents were purchased from New England Biolabs (Ipswich, MA). Genes were synthesized by GenScript (Piscataway, NJ). The oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA).

Bacterial strains and growth conditions

M. smegmatis ML712, which lacks the porin genes *mspA*, *mspB*, *mspC*, and *mspD* (42), was used for purification of octameric MspA proteins (wild-type (wt) MspA, MspA M1, and MspA M2) and grown at 37°C in Middlebrook 7H9 liquid medium (BD Biosciences, Franklin Lakes, NJ) supplemented with 0.2% glycerol and 0.05% Tween 80 or on Middlebrook 7H10 agar (BD Biosciences) supplemented with 0.2% glycerol. Hygromycin was used in concentrations of 50 μg/mL for *M. smegmatis* ML712. *Escherichia coli* DH5α was used for cloning experiments and was routinely grown in Luria-Bertani (LB) broth at 37°C. For single-chain MspA production and purification, *E. coli* Omp8 (Table S2) (43) was grown in LB broth. Ampicillin and streptomycin were used in concentrations of 100 μg/mL and 50 μg/mL, respectively, for *E. coli* DH5α and Omp8.

Plasmid construction

Full-length single-chain *mspA* genes and *mspA m2-1* gene encoding a histidine$_8$ tag and the *mspA m2-8* encoding TwinStrepII tag were ordered from GenScript. The resulting plasmids used in this study (Figs. S1 and S2) were constructed using standard molecular biology methods. Plasmids were analyzed by digestion with restriction enzymes and by sequencing to verify introduced genes and/or mutations.

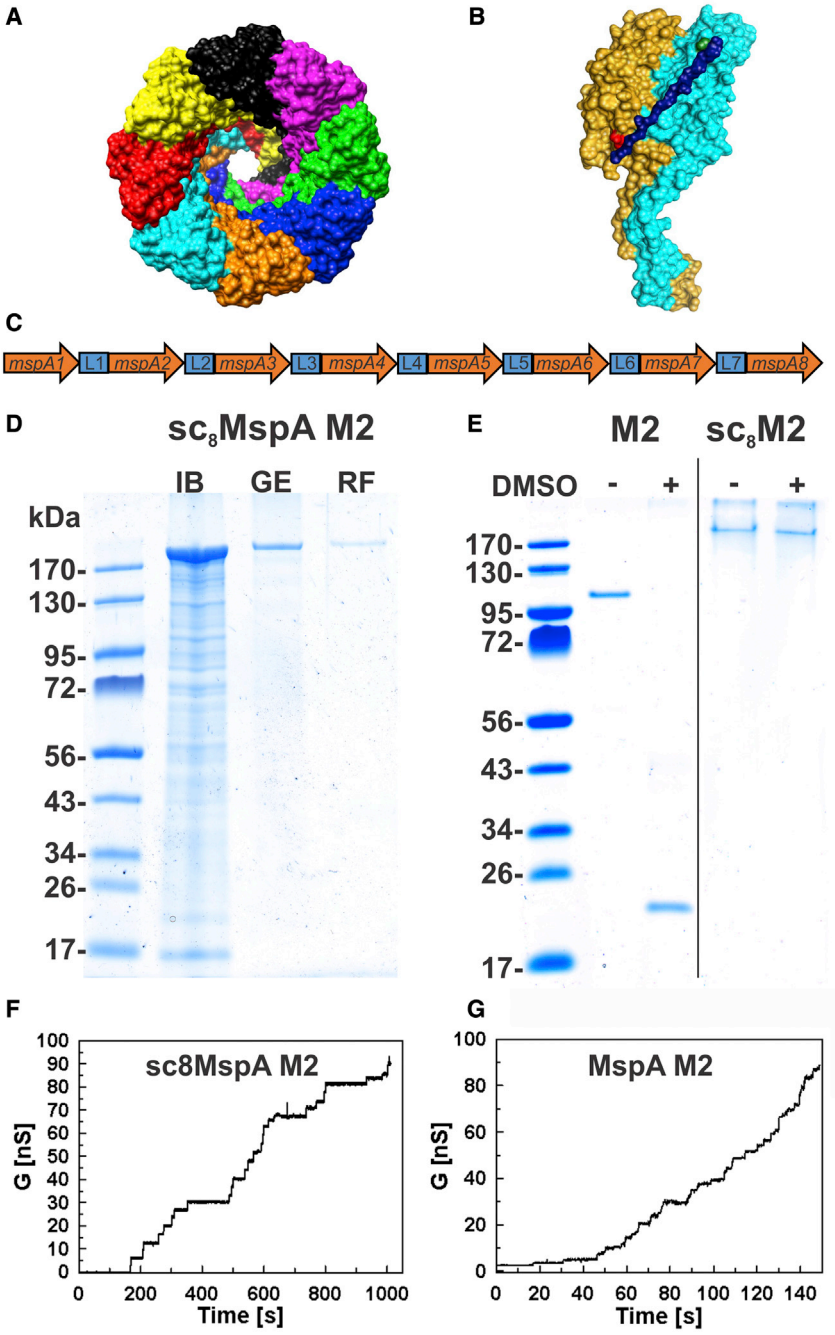

FIGURE 1 Design, production, and purification of single-chain MspA. (*A*) Structure of wt MspA (PDB: 1UUN) with eight monomers in different colors (*top view*). (*B*) Model of the covalent peptide linker between two MspA subunits in single-chain MspA (*side view*). The linker connects the C-terminus (*red*) of one subunit to the N-terminus (*blue*) of the adjacent subunit. Models in (*A*) and (*B*) were prepared using Chimera 11. (*C*) The single-chain *mspA* gene. Numbered arrows represent translationally fused *mspA m2* genes. Blue boxes with numbers show the linker regions connecting adjacent *mspA m2* genes. The scheme is not to scale. (*D*) Purification of sc_8MspA M2. Samples were loaded onto 10% polyacrylamide gel followed by staining with Coomassie. Lanes: IB, inclusion bodies purified from Omp8 *E. coli* solubilized in 8 M urea; GE, sample after gel-extraction procedure; RF, refolded sample after dialysis. Numbers on the left are molecular weights of the marker in kDa. (*E*) Denaturation of MspA M2 and sc_8MspA M2. Coomassie-stained 8% polyacrylamide gel of untreated (−) and samples boiled in 80% (v/v) dimethyl sulfoxide (DMSO) (+) to denature proteins. Lanes: M2, octameric MspA M2; sc_8M2, sc_8MspA M2 with eight covalently linked subunits. One microgram of protein was loaded on each lane. (*F* and *G*) Current trace of sc_8MspA M2 (*F*) and octameric MspA M2 (*G*) in a diphytanoyl phosphatidylcholine (DPhPC) membrane with an aperture diameter of 1 mm in a Montal-Mueller system. The applied potential was −10 mV, the electrolyte was 1 M KCl, 10 mM HEPES (pH 7.5).

Purification of octameric MspA proteins

Purification of MspA M2 was performed as described previously (12,44) with slight modifications. In brief, *M. smegmatis* ML712 harboring the pML844 plasmid was grown for 2 days at 37°C. Cell pellets were collected, washed, and resuspended in extraction buffer (100 mM NaCl, 0.1 mM Na_2EDTA, 150 mM NaH_2PO_4/Na_2HPO_4, 0.5% (v/v) OPOE, pH 7.0) followed by boiling for 30 min. The protein extract was precipitated with ice-cold acetone and incubated on ice overnight. The precipitated protein was resuspended in OPOE buffer (25 mM HEPES, 150 mM NaCl, 0.5% (v/v) OPOE, pH 7.5) and applied onto a Superdex S200 HiLoad 26/60 column for gel filtration. Fractions containing apparently pure MspA M2 were pooled and used for the experiments described here.

Production and purification of single-chain MspA M2 and single-chain MspA PN1 proteins

The *E. coli* strain Omp8 transformed with plasmids encoding different single-chain MspA PN1 constructs (pML3216, pML3222, pML 3231, pML3232, pML3233, pML3234 (Table S3)) was used for protein production and purification. Cells were grown overnight at 37°C in LB medium containing 100 μg/mL ampicillin. Overnight cultures were then diluted into 1 L of fresh LB medium to an OD_{600} of approximately 0.1 and incubated at 37°C. At OD_{600} of 0.6, expression was induced with 1.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After the induction, cells were grown for 2 h at 37°C followed by purification of inclusion bodies as described previously (45). In brief, after sonication cells were centrifuged

at 1500 × *g* for 10 min at 4°C to remove cell debris. Triton X-100 (1% v/v, final) was added to solubilize membrane proteins, and the mixture was incubated on ice for 10 min. The sample was then centrifuged at 7000 × *g* for 20 min at 4°C to collect an insoluble pellet. The pellet was washed three more times with lysis buffer to remove Triton X-100. The resulting pellet containing inclusion bodies was resuspended in 8 M urea and incubated overnight at room temperature. Next, inclusion bodies were separated on an 8% SDS-polyacrylamide gel followed by Coomassie blue staining (Simply Blue Safe Stain (Invitrogen)). The single-chain MspA PN1 proteins were then eluted from the gel as follows. The bands corresponding to the theoretical molecular weight of single-chain MspA were excised from a gel with a clean razor. The gel bands were crushed and protein elution buffer (25 mM HEPES, 150 mM NaCl, 0.5% (v/v) OPOE, pH 7.5) was added, followed by brief sonication on ice to further disperse gel particles. The ratio of the gel volume to protein elution buffer was 1:3. The mixture was placed on a rotary shaker and incubated overnight at 30°C. After incubation, the sample was centrifuged at 16,000 × *g* for 5 min to precipitate the polyacrylamide gel pieces. The supernatant containing the eluted protein was used for refolding by dialysis against 2 L of refolding buffer (150 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4, 200 mM L-arginine, 800 mM urea, 0.5% OPOE, 1 mM phenylmethylsulfonyl fluoride (PMSF), Complete Protease Inhibitor cocktail, 0.02% sodium azide, pH 8.0). The concentration of the refolded protein was measured using bicinchoninic acid (BCA kit, Thermo Fisher Scientific). The refolded protein was used immediately or frozen at −20°C with glycerol (50% v/v) for storage.

Expression and purification of single-chain double-tagged MspA M2

E. coli strain Omp8 transformed with pML4170 was grown overnight at 37°C in LB medium containing 50 μg/mL streptomycin. The cells were then diluted into 1–2 L of fresh medium to give OD_{600} of 0.1. When OD_{600} reached 0.6, the cells were induced with 1 mM IPTG (final). The cultures were then transferred to 18°C and grown for 14 h. Harvested cells were washed in PBS and resuspended in lysis buffer at a ratio of 1:5 (150 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4, pH 7.4), supplemented with Benzonase (Novagen, Madison, WI) and Complete Protease Inhibitor cocktail (Roche, San Francisco, CA). The cells were sonicated on a Misonix sonicator (Farmingdale, NY) for 20 min on ice (30 s on/off cycle, 50 W). The lysate was then used for inclusion body purification as described in the previous paragraph. Inclusion bodies in 8 M urea were loaded on Ni-NTA agarose resin (Qiagen, Germantown, MD) to bind single-chain MspA. Ni-affinity purification was performed in denaturing conditions with buffer composition of 150 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4, and 6 M urea, pH 7.4. scMspA was eluted with denaturing buffer containing 700 mM imidazole. Elution fractions were pooled, concentrated on an Amicon spin column (Millipore, Burlington, MA) with 50 kDa cutoff, and loaded onto a Superdex S200 26/60 HiLoad column (GE Life Sciences, Piscataway, NJ) for gel filtration in denaturing conditions (150 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4, 6 M urea, 0.2% (w/v) SDS, pH 7.5). Fractions containing scMspA were combined, concentrated, and used for StrepII tag affinity purification on Strep-Tactin XT resin (IBA, Göttingen, Germany). Biotin (50 mM in 150 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4, 6 M urea, pH 8.0 buffer) was used to elute scMspA protein. The elution fractions were combined, and OPOE (0.5% v/v, final) was added prior to refolding by dialysis overnight at room temperature against 2 L of refolding buffer (150 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4, 200 mM L-arginine, 800 mM urea, 0.5% OPOE, 1 mM PMSF, Complete Protease Inhibitor cocktail, 0.02% sodium azide, pH 8.0). This sample is referred to as refolded sample. As the final step to remove contaminants and refolding buffer components, refolded sample was concentrated on an Amicon spin column (Millipore) with 50 kDa cutoff, loaded onto a Supredex S200 Increase 10/300 GL column (GE Healthcare), and eluted with 150 mM NaCl, 50 mM NaH_2PO_4/Na_2HPO_4, 0.5% OPOE, pH 7.4 buffer. Individual fractions were used for lipid bilayer experiments and downstream analysis.

Denaturation of single-chain MspA

For all experiments, equal amounts of protein (typically 2 μg) were used. Octameric MspA M2 purified from *M. smegmatis* ML712 or refolded single-chain MspA proteins were mixed with dimethyl sulfoxide (DMSO) (80% v/v, final). The mixture was incubated for 15 min at 99°C followed by addition of 10 volumes of ice-cold acetone to precipitate the protein, and incubated on ice for 15 min. Precipitated samples were centrifuged at 16,000 × *g* for 15 min at 4°C. The protein pellet was dried under vacuum to remove acetone. After drying, the samples were resuspended in the initial volume, mixed with loading dye, and separated on polyacrylamide gel.

Lipid bilayer measurements

Lipid bilayer experiments were performed in a custom-made lipid bilayer apparatus as previously described (12). In brief, a Teflon cuvette with 10 mL volume was separated into two compartments (*cis* and *trans*) by a wall with an aperture of approximate diameter of 1 mm. Ag/AgCl electrodes were bathed in a 1 M KCl, 10 mM HEPES, pH 7.4 electrolyte solution. The cuvette was primed on both sides of the aperture with 2% DPhPC (Avanti Polar Lipids) in chloroform. Lipid membranes were painted across aperture from a solution of 1% DPhPC in *n*-decane. The samples were added to both sides of the cuvette. Baseline and detergent-containing buffers were examined to exclude contamination and detergent interference. Single-channel conductances for more than 100 pores for a protein sample were recorded. Recordings were performed at −10 mV potential. Current was recorded using Keithley (Cleveland, OH) 428 Current Amplifier with a filter rise time of 30 ms, and digitized by a computer equipped with a Keithley Metrabyte STA 1800 U interface. The data were recorded with Test Point 4.0 software (Keithley). The raw data were analyzed using IGOR Pro 5.03 (WaveMetrics, Portland, OR) using a macro provided by Dr. Harald Engelhardt. The data were further analyzed in SigmaPlot 11.0 (Systat Software, San Jose, CA) to generate the graphs shown here.

Single-channel DNA hairpin measurements

A chip with a wedge-on-pillar aperture of 100 μm diameter made from the hydrophobic SU-8 polymer (46) was glued and sealed to a custom-designed fluidic cell, separating *cis* and *trans* chambers. The aperture was pretreated by 4 mg/mL PBD_{11}-PEO_8 block-copolymer (Polymer Source) dissolved in hexane. After evaporation of the hexane solvent, a dry layer of polymer was formed on the aperture edge. The *cis* and *trans* chambers of the measuring cuvette were filled with electrolyte buffer, and a pair of Ag/AgCl electrodes connected to an Axon 200B patch-clamp amplifier (Molecular Devices, LLC, San Jose, CA) were inserted into the cuvette. The polymer membrane was painted across the pretreated aperture using 8 mg/mL polymer dissolved in n-decane. More than 60 min waiting time is needed for the polymer membrane to thin down until it forms a bilayer (membrane capacitance range: 60–80 pF). 0.063 nM octameric MspA M2 or 4.2 nM sc_8MspA_{dt} M2 was added to the *cis* chamber. After insertion of a single pore, DNA hairpins were added to the *cis* chamber. The data were collected at 250 kHz sampling rate using a 10 kHz low-pass filter. All oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA)).

RESULTS

Purification and characterization of single-chain MspA

Eight MspA subunits assemble in the outer membrane of *M. smegmatis* to produce a central water-filled channel (Fig. 1 *A*). We have shown previously that the C-terminus of a monomer can be connected to the N-terminus of a

neighboring subunit by a 17-mer peptide linker due to the close proximity of the termini (40). Here, we used the same approach to synthesize a single-chain *mspA m2* (*scmspA m2*) gene with eight *mspA m2* genes encoding the mutations D90N/D91N/D93N/D118R/D134K/D139R, which enable efficient DNA capture and translocation (22), connected by DNA fragments encoding $(GGGGS)_3$ peptide linkers (Fig. 1 *B* and Table S1). Each gene is flanked by unique restriction sites to enable specific modifications of each MspA subunit (Table S1). We first tried to produce scMspA M2 protein in the porin quadruple mutant *M. smegmatis* ML712, which lacks all four *msp* genes (42). However, all *scmspA m2* genes in clones producing MspA protein were scrambled and protein quantities were very low (not shown). We then used *E. coli* to produce scMspA M2 protein. To reduce the probability of homologous recombination between the individual *mspA* genes, we altered every second codon in the *mspA* subunit genes 2 through 8 to generate DNA sequence differences without altering the amino acid sequence. In addition, the *scmspA m2* DNA sequence was altered for optimal expression in *E. coli*, and the signal peptide was removed for cytoplasmic expression of MspA. We expressed the synthetic *scmspA m2* gene under the control of the T7 promoter using the plasmid pML3216 (Fig. S1) in the BL21(DE3)Omp8 strain lacking the three major porins of *E. coli* (43) to avoid contamination with endogenous porins. MspA protein (166 kDa theoretical mass) was purified from inclusion bodies by gel electrophoresis and excision of a protein band of 170 kDa and refolded in a buffer containing the detergent OPOE (Fig. 1 *D*). The protein yield was 12 μg of purified sc_8MspA M2 per single preparation from four polyacrylamide gels (approximately 190 μg per liter of *E. coli* culture). It should be noted that all subunits of sc_8MspA M2 are identical, i.e., they do not contain different mutations. Octameric MspA is an extremely stable protein that does not denature after boiling for 10 min in 2% SDS and other harsh denaturing conditions (18). To dissociate octameric MspA into its subunits, we boiled the purified proteins in 80% (v/v) DMSO (40). Only octameric MspA M2 dissociated into monomers while sc_8MspA M2 was stable, demonstrating that all eight subunits are covalently linked and that we purified full-length scMspA (Fig. 1 *E*). To examine whether sc_8MspA M2 forms functional channels, we performed lipid bilayer experiments in DPhPC membranes of 1 mm diameter in a Montal-Mueller setup (40). Addition of both octameric MspA M2 purified from *M. smegmatis* (Fig. 1 *G*) and recombinant sc_8MspA M2 protein to the membranes resulted in a stepwise current increase indicative of channel insertions (Fig. 1 *F*), while no channel activity was observed when only detergent-containing buffer was added (not shown). Analysis of the single-channel insertions revealed similar conductance values ranging from 0.5 to 6 nS for octameric MspA M2 and from 0.2 to 4.5 nS for single-chain MspA (Fig. S4). These wide ranges of single-channel conductances were also observed for other octameric MspA proteins such as wt MspA and MspA M1 (Fig. S4) and seem to be an intrinsic feature of the MspA pore. Taken together, these experiments demonstrate that sc_8MspA M2 forms functional pores and show that it is feasible to convert an oligomeric pore protein into a functional monomeric protein, enabling in the future asymmetric mutations of MspA to optimize its channel properties for specific applications.

Single-chain MspA variants with altered subunit stoichiometries form functional channels

To highlight the advantages of scMspA in tailoring the pore for specific applications, we aimed at changing its channel diameter as one of the most important properties determining the interactions of the pore constriction with the analyte. Previously, the constriction diameter could only be altered by mutations of individual amino acids in the protein monomer, which also changes the chemical nature of the interactions with the analyte. A monomeric pore consisting of identical subunits should enable the design of pores with different channel diameters by altering the subunit stoichiometry. To examine whether MspA is capable of forming functional pores with a stoichiometry different from 8, we constructed scMspA variants with three, five, six, seven, and, as a control, eight subunits based on MspA M2 with an additional P97F mutation (Table S5). This additional mutation was chosen because octameric MspA PN1 has a well-defined channel distribution with a peak at 1.2 nS (Fig. S5 *A*), in contrast to all other examined octameric MspA pores (Fig. S4). All PN1 constructs contain the MspA M2 mutations (D90N/D91N/D93N/D118R/D134K/D139R) as described in (22) in addition to the P97F mutation in all subunits. The corresponding proteins were named sc_xMspA PN1 with "x" numbers of covalently linked subunits and were purified from *E. coli* BL21(DE3)Omp8 as described above. The target proteins accounted for approximately 10% of proteins in the whole-cell lysate as determined by quantitative image analysis of protein gels (Fig. S3). The apparent molecular masses for sc_3MspA PN1, sc_5MspA PN1, sc_6MspA PN1, and sc_7MspA PN1 proteins were 60 kDa, 110 kDa, 120 kDa, and 140 kDa, respectively (Fig. 2 *A*), in good agreement with the theoretical molecular masses. To examine whether the scMspA PN1 subunits in these constructs are covalently linked, we denatured the proteins by boiling in 80% (v/v) DMSO. As expected, DMSO treatment dissociated octameric MspA M2 protein into its 20 kDa monomers while all scMspA PN1 proteins were stable under these conditions (Fig. 2 *A*), demonstrating that the subunits of these proteins are indeed covalently linked. To examine the channel-forming properties of scMspA PN1 pores, we performed planar lipid bilayer experiments as previously described (42). Surprisingly, in these experiments all single-chain constructs including those with an uneven number of connected

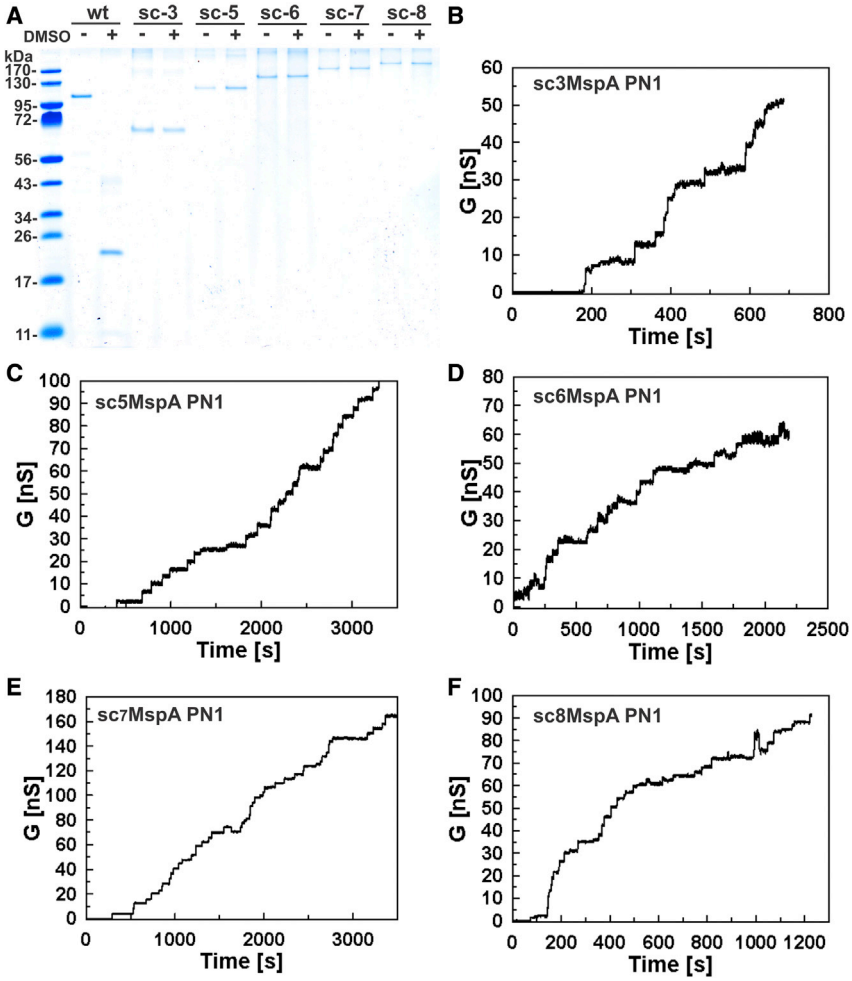

FIGURE 2 Channel activity of MspA pores with different subunit stoichiometries. (*A*) Coomassie-stained 8% polyacrylamide gel with different single-chain MspA PN1 constructs. Lanes: wt, octameric MspA M2; sc-3, sc-5, sc-6, sc-7, sc-8, single-chain MspA PN1 with three, five, six, seven, eight covalently linked subunits. Octameric MspA M2 and MspA proteins with different subunit compositions were boiled in 80% DMSO (+) and compared with untreated protein (−). Note that the 20 kDa band is the monomeric subunit of MspA. Equal amounts (2 μg) of each protein were loaded on the gel. (*B–F*) Current traces of single-channel conductances of scMspA PN1 constructs. Membrane currents were recorded in an electrolyte containing 1 M KCl, 10 mM HEPES (pH 7.4) at a voltage of −10 mV. The diameter of the diphytanoyl phosphatidylcholine (DPhPC) bilayer was approximately 1 mm. (*B*) sc_3MspA PN1: 95 insertion events from 4 membranes; (*C*) sc_5MsA PN1: 221 insertions from 7 membranes; (*D*) sc_6MspA PN1: 463 insertion events from 7 different membranes; (*E*) sc_7MspA PN1: 62 insertions from 5 membranes; (*F*) sc_8MspA PN1: 150 insertions from 6 membranes. In these experiments 70–840 ng of refolded proteins was added to the membranes.

subunits produced a stepwise current increase indicative of channel insertions into lipid membranes (Fig. 2, *B–F*). The channel insertion frequencies were similar for all scMspA PN1 proteins. The scMspA PN1 pores showed conductance values ranging between 0.5 and 5 nS (Fig. S5), similar to sc_8MspA M2 and octameric MspA proteins (Fig. S4). These results demonstrated that functional MspA pores are formed with different subunit numbers. The single-channel conductances for the sc_3MspA PN1 and sc_5MspA PN1 variants shifted to values larger than that of sc_8MspA PN1 (Fig. S5), indicating the presence of pores with a larger number of subunits. This is likely due to the assembly of two or more sc_3MspA and sc_5MspA constructs. Similar results were obtained for the sc_6MspA PN1 and sc_7MspA PN1 variants (Fig. S5). Taken together, these experiments demonstrate that it is feasible to construct MspA pores with altered stoichiometries using covalently linked subunits. This remarkable plasticity enables the adaptation of the sensing capabilities of MspA for different substrates not only by making asymmetric mutations but also by designing pores with altered constriction zone diameters.

Improvement of the purification of recombinant single-chain MspA protein

While the above experiments demonstrated that single-chain MspA produces functional pores, the purification based on gel extraction is labor intensive and inefficient, resulting in low yields of approximately 10 μg per preparation. We estimated that ∼98% of the initial amount of scMspA proteins in the inclusion bodies was lost during this process. One of the main challenges was to separate full-length sc_8MspA M2 from its many degradation products which are only marginally smaller, i.e., these degradation products may have cleaved linker peptides that appear to be very susceptible to proteolysis (Fig. S6). To address this issue, we exploited the fact that both termini of MspA are accessible on the outside of the MspA channel (Fig. 1 *B*). Thus, we added a His_8 tag at the N-terminus and a Twin-StrepII tag at the C-terminus to generate double-tagged single-chain MspA M2 (sc_8MspA_{dt} M2) in which all eight subunits are identical, containing the D90N/D91N/D93N/D118R/D134K/D139R mutations (Fig. 3 *A*).

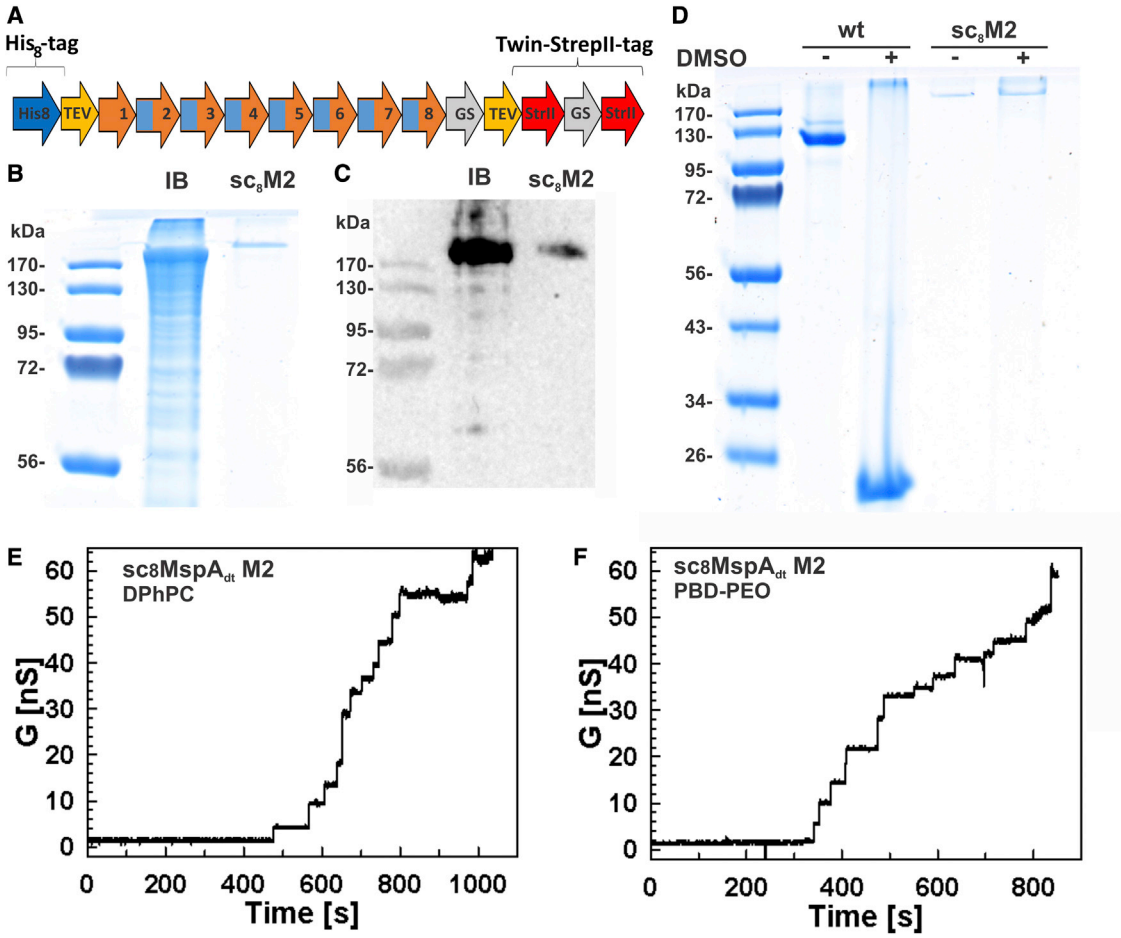

FIGURE 3 Purification and channel activity of single-chain MspA. (*A*) Scheme of the *scmspA* gene in the expression vector for protein (sc_8MspA_{dt} M2) purification from *E. coli*. Eight identical *mspA m2* genes encoding MspA M2 (22) with the mutations D90N/D91N/D93N/D118R/D134K/D139R (*orange arrows 1–8*) are flanked by an N-terminal His_8 tag and a C-terminal Twin-StrepII tag (StrII). Peptide linkers (*blue boxes*) connect the individual MspA subunits (Table S1). GS, glycine-serine linker; TEV, tobacco etch virus protease recognition site. Scheme is not to scale. (*B*) Coomassie-stained 8% polyacrylamide gel with inclusion bodies (IB) from *E. coli* BL21(DE3)Omp8 and the purified sc_8MspA_{dt} M2 protein. (*C*) Western blot of the samples shown in (*B*). Proteins were transferred onto polyvinylidene fluoride membrane and stained with anti-StrepII-tag HRP-conjugated antibodies. (*D*) Denaturation of MspA M2 and sc_8MspA_{dt} M2. Coomassie-stained 8% polyacrylamide gel of DMSO treated (+) and untreated (−) samples. Proteins were boiled in 80% (v/v) DMSO to induce denaturing. Note that the 20 kDa band is a monomeric subunit of MspA that is visible after DMSO treatment. Lanes: wt, octameric MspA M2; sc_8M2, sc_8MspA_{dt} M2 with eight covalently linked subunits. Three micrograms of MspA M2 and 1 μg of sc_8MspA_{dt} M2 were loaded. (*E*) Current trace of sc_8MspA_{dt} M2 in 1 M KCl, 10 mM HEPES (pH 7.4) electrolyte at −10 mV applied potential in diphytanoyl phosphatidylcholine (DPhPC) bilayer with an aperture of approximately 1 mm in diameter. The concentration of the protein in the cuvette was 16 ng/mL. (*F*) Current trace of sc_8MspA_{dt} M2 in 1 M KCl, 10 mM HEPES (pH 7.4) electrolyte at −10 mV applied potential in a poly(1,2-butadiene)-*b*-poly(ethylene oxide) (PBD-PEO) polymer bilayer with an aperture of approximately 1 mm in diameter. The concentration of the protein in the cuvette was 16 ng/mL.

The plasmid pML4170 carrying the sc_8mspA_{dt} *m2* gene was transformed into *E. coli* BL21(DE3)Omp8 for protein production and purification. Sc_8MspA_{dt} M2 protein was purified under denaturing conditions using subsequent Ni(II)- and streptavidin-affinity chromatography to isolate full-length protein (Figs. 3, *B* and *C*; Fig. S7). Refolding in the detergent OPOE and a final size exclusion (Fig. S7) yielded 0.2 mg of apparently pure protein per liter of culture (Fig. 3 *B*). Importantly, we did not observe scMspA degradation products as shown in a Western blot of purified sc_8MspA_{dt} M2 (Fig. 3 *C*), in contrast to previous preparations using gel excision (Fig. S6). As a quality control we performed denaturing experiments with 80% DMSO. As expected, only octameric MspA M2 dissociated into its monomeric subunits of 20 kDa while sc_8MspA_{dt} M2 was stable, indicating that all eight subunits are covalently linked in the purified refolded protein (Fig. 3 *D*). Thus, the combination of N- and C-terminal affinity tags enabled us to purify full-length scMspA protein with a 20-fold increase in protein yield in comparison with the previous gel-excision method.

Channel-forming properties of single-chain MspA in lipid bilayer experiments

To examine whether sc_8MspA_{dt} M2 forms functional channels, we performed lipid bilayer experiments in a

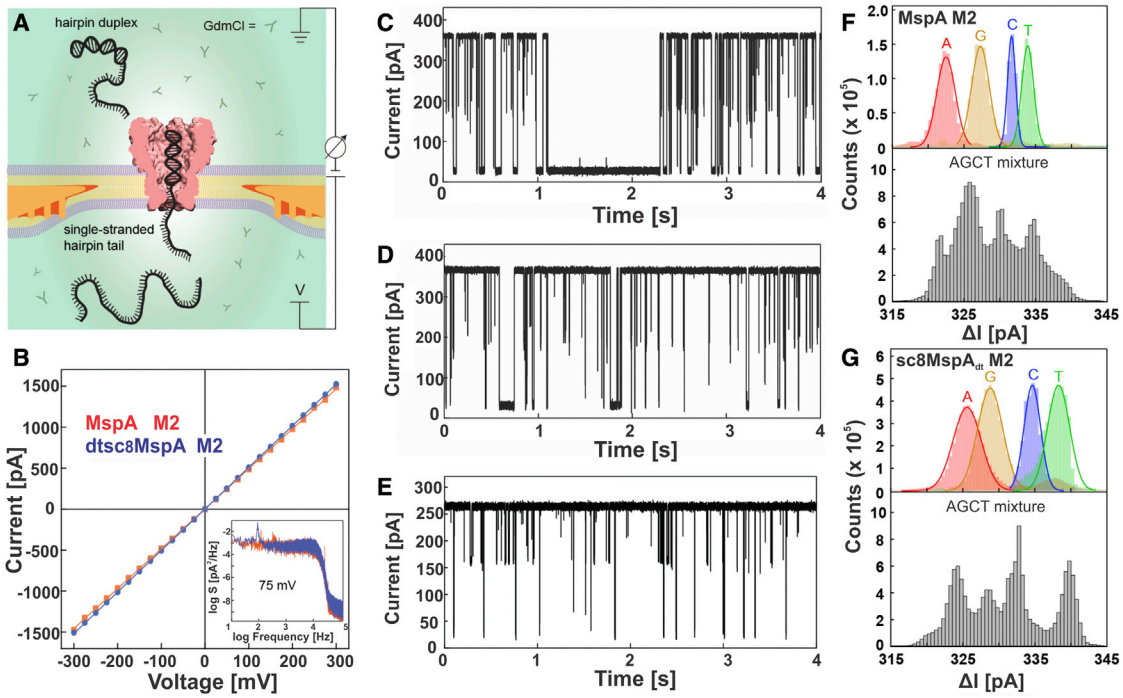

FIGURE 4 Nucleotide recognition by single-chain MspA in DNA hairpin experiments. (*A*) A free-standing PBD-PEO polymer bilayer membrane is supported by wedge-on-pillar 100 μm aperture. A single MspA nanopore is shown integrated into the polymer bilayer bathing in an electrolyte containing 2 M guanidinium chloride (GdmCl). A patch-clamp amplifier is connected to Ag/AgCl electrodes to measure the ion current through the *cis* (*top*) and *trans* (*bottom*) chambers. (*B*) Current/voltage curves of sc$_8$MspA$_{dt}$ M2 and MspA M2 bathing in an electrolyte containing 1 M KCl, 2 M GdmCl, 10 mM Tris (pH 7.5) were recorded from −300 to +300 mV. The power spectral density plots were recorded at 75 mV. (*C* and *D*) Continuous current traces show translocation events for poly-dT DNA hairpins through MspA M2 (*C*) and sc$_8$MspA$_{dt}$ M2 (*D*) nanopores in an electrolyte containing 1 M KCl, 2 M GdmCl, 10 mM Tris (pH 7.5) at a voltage of 75 mV. (*E*) Continuous current trace for sc$_8$MspA$_{dt}$ M2 in the same electrolyte as in (*D*) without GdmCl at a voltage of 140 mV. (*F* and *G*) Histograms of current change (ΔI) caused by the blockade of the pores by individual DNA hairpins (*top*) and by DNA hairpin mixtures (*bottom*) for MspA M2 (*F*) and sc$_8$MspA$_{dt}$ M2 (*G*). The experiments were done in an electrolyte containing 1 M KCl, 2 M GdmCl, 10 mM Tris (pH 7.5) at a voltage of 75 mV. The concentration of each DNA hairpin (poly-dA, poly-dC, poly-dG, and poly-dT) was 0.3 μM.

Montal-Mueller setup using cuvettes with an aperture diameter of 1 mm. As expected, refolded sc$_8$MspA$_{dt}$ M2 had channel-forming activity in DPhPC membranes as shown by the stepwise current increase after addition of the protein (Fig. 3 *E*). Analysis of 432 insertions from eight different membranes showed predominant conductance peaks of 1.3 nS, 1.9 nS, and 3.0 nS (Fig. S4 *E*). Notably, sc$_8$MspA$_{dt}$ M2 had a broader conductance distribution than sc$_8$MspA PN1 (Fig. S5 *F*). These differences may be the result of different purification methods, the absence of the P97F mutation near the constriction zone of sc$_8$MspA$_{dt}$ M2 compared with sc$_8$MspA PN1, and/or the presence of purification tags on the termini of sc$_8$MspA$_{dt}$ M2.

We recently demonstrated that octameric MspA M2 forms channels in bilayers composed of PBD-PEO polymer (21). PBD-PEO bilayers have a two- to three-fold increased lifetime and are more robust toward chemicals and high voltages than membranes made from biological lipids, and can be used in nanopore sequencing experiments (21). Thus, we also examined the channel properties of scMspA insertions in PBD-PEO bilayers. Similarly to DPhPC bilayers, we observed insertions of sc$_8$MspA$_{dt}$ M2 into PBD-PEO bilayers (Fig. 3 *F*). Analysis of 207 insertions from 28 membranes showed broad distribution of single-channel conductances (Fig. S4 *F*) in PBD-PEO bilayers similar to those obtained for sc$_8$MspA$_{dt}$ M2 in DPhPC bilayers. It should be noted that insertion frequency of sc$_8$MspA$_{dt}$ M2 was reduced in PBD-PEO bilayers in comparison with DPhPC membranes. Taken together, these experiments showed that sc$_8$MspA$_{dt}$ M2 forms functional channels in both DPhPC and PBD-PEO bilayers.

DNA hairpin translocation through the single-chain MspA pore

To examine whether the nucleotide recognition capability of MspA is preserved in scMspA, we used sc$_8$MspA$_{dt}$ M2 in DNA hairpin experiments as described previously (23). This assay is based on distinct residual currents when the single-stranded homopolymer tail is located inside the constriction zone (Table S4), while the double-stranded region of the DNA hairpins is stalled in the lumen of the MspA pore and temporarily prevents translocation (Fig. 4 *A*). Single sc$_8$MspA$_{dt}$ M2 pores were inserted into PBD-PEO block-copolymer bilayer membranes with a diameter of 100 μm and bathed in a buffer containing 10 mM Tris (pH 7.5), 1 M

KCl, and 2 M guanidinium chloride (GdmCl). Based on previous data (21), we used GdmCl in the electrolyte solution because it induces a semimelting state of DNA hairpins that enables smooth DNA passage through the nanopore at low applied voltage, resulting in increased DNA capture rate and decreased trapping time.

The current/voltage (I/V) curves and power spectral density plots reveal that the conductance and trace noise are almost identical for sc_8MspA_{dt} M2 and octameric MspA M2 (Fig. 4 *B*). Both proteins exhibited gating when negative voltage was applied (Fig. S8 *A*). Since the gating conductance varies for the same MspA pore even at the same voltage, we used the ungated current values for the I/V curves, resulting in identical symmetric I/V curves for both sc_8MspA_{dt} M2 and MspA M2 (Fig. 4 *B*). Uncorrected I/V curves for the same raw data are shown in Fig. S8 *B*. Interestingly, the protein concentration required to observe channel insertions was approximately 60-fold higher for sc_8MspA_{dt} M2 (4.2 nM) than for octameric MspA M2 (0.063 nM). Consistent with previous results in large aperture bilayers (Fig. S4 *F*), a wide conductance distribution was observed for individual sc_8MspA_{dt} M2 channel ranging from 1 to 5.5 nS. However, only the pore with conductance of around 4.9 nS resulted in clean and deep DNA translocation signals. This appears to be a general property of the MspA pore, as populations of pores with different conductance levels were previously observed in DNA hairpin experiments with octameric MspA isolated from *M. smegmatis* (40).

The current traces after addition of the poly-dT hairpin to sc_8MspA_{dt} M2 and MspA M2 show current blockades resulting from translocation of the DNA hairpin through the pore in GdmCl-containing buffer (Fig. 4, *C* and *D*). Similar current blockades were observed with the poly-dT hairpin and sc_8MspA_{dt} M2 when no GdmCl was present in the electrolyte buffer (Fig. 4 *E*). However, a higher voltage of 140 mV instead of 75 mV was necessary to overcome the energy barrier for DNA translocation, in accordance with previous hairpin experiments with octameric MspA in which 140–180 mV was used for hairpin DNA translocation (23). Our data further indicate that GdmCl improves both capture and translocation of hairpin DNA through the pore.

Nucleotide recognition by single-chain MspA in DNA hairpin experiments

To examine nucleotide recognition by sc_8MspA_{dt} M2, we used DNA hairpins with a duplex region of 14 nucleotides and a homopolymer tail of 50 nt (hp14-dT50 (poly-dT), hp14-dA50 (poly-dA), hp14-dC27 (poly-dC), and hp14-dG3dA47 (poly-dG)) (Table S4 and Fig. S9). It should be noted that we used only three dG nucleotides in the "poly"-dG hairpin tail in otherwise dA background to avoid G-tetrad formation as in our previous experiments (23). Histograms of the current change ΔI (current I in the presence of a hairpin minus the baseline current level I_0) are shown for the four individual DNA hairpins and mixture of four hairpins with octameric MspA M2 (Fig. 4 *F*) and sc_8MspA_{dt} M2 (Fig. 4 *G*), respectively. Fitted Gaussians for each individual hairpin with a homopolymer tail were well resolved and separated, and looked almost identical for both MspA M2 and sc_8MspA_{dt} M2 (Fig. 4, *F* and *G*, *top*). The partial overlap between purines was more pronounced for sc_8MspA_{dt} M2; nonetheless, individual peaks for poly-dA and poly-dG were still distinguished (Fig. 4 *G*, *top*). Remarkably, the histograms of DNA hairpin mixture clearly show four separate peaks representing the four different nucleotides, which match the current blockades for individual hairpin peaks (Fig. 4, *F* and *G*, *bottom*). In this experiment sc_8MspA_{dt} M2 had better resolving properties than octameric MspA M2. In conclusion, our DNA hairpin experiments showed that sc_8MspA_{dt} M2 and octameric MspA M2 have very similar capabilities in distinguishing all four DNA nucleotides. These results indicate that covalent linking of all eight MspA subunits preserves the nucleotide recognition properties of the MspA pore and open many avenues to tailor the channel properties of the MspA pore for specific sensing applications.

DISCUSSION

Properties and potential applications of single-chain MspA nanopores

In this study we showed that it is feasible to convert multimeric MspA into a functional monomeric nanopore with identical electrophysiological properties. This achievement is a major improvement over previous approaches based on self-assembly of wt and mutated subunits, which required the physical separation of pores with different subunit combinations lowering the yield dramatically and suffered from different permutations with identical physical properties (6). By contrast, single-chain MspA as produced in this study is a protein with all subunits connected in a single polypeptide. ScMspA enables the design of pores with different mutations in different subunits, in contrast to octameric MspA in which the same mutations are present eight times. These "asymmetric" mutations open numerous avenues to improve the performance of the MspA pore in DNA sequencing. 1) Distinct amino acids in the MspA constriction zone will enable different chemical interactions with the DNA nucleobases. Previously, mutations that enable DNA translocation were present in all eight MspA monomers (22), making asymmetric interactions with DNA impossible. Asymmetric mutations open numerous avenues to increase the specificity of the current blockade for each nucleotide and to increase the dwell time in the construction zone. Both of these effects are likely to reduce the contribution of neighboring nucleotides to the current blockade, which is currently 4–5 nucleotides (47), and concomitantly reduce the high raw data error rates in nanopore sequencing (48). 2) Single-chain MspA can

be used to slightly alter the diameter of the pore to modulate the interactions between amino acids in the constriction zone and nucleotides. This can be achieved by changing the subunit stoichiometry as shown in this study (Fig. 2). In a simple geometrical model, adding or removing a subunit alters the diameter of the MspA pore by one-eighth, which is equivalent to approximately 1.5 Å, a reasonable range for significantly changing chemical interactions. Both the hexameric and heptameric pores are functional, demonstrating that this approach is feasible. 3) Motor proteins such as DNA polymerases employed to control the translocation rate of DNA through the pore have been instrumental for sequencing (25) and have become standard in nanopore sequencing devices (49). However, the long distance between the motor proteins and the MspA constriction zone decreases the positional precision and increases error due to thermal motions of the flexible DNA strand (50) and the motor protein. Random positioning of the motor protein on top of the MspA pore might enhance this error. A single cysteine in one of the surface loops of single-chain MspA would enable the covalent attachment of the motor protein, eliminating the random positioning of motor and assembly with the MspA pore and thereby reducing the system variability. 4) It might be possible to create a specific path for single-stranded DNA inside the MspA pore, thereby decreasing the translocation velocity and eliminating the need for a motor protein altogether. While the possibilities to tailor the single-chain MspA pore for DNA sequencing by combining asymmetric mutations are almost endless, it would require an efficient assay to test these combinations. However, the currently used lipid bilayer systems are not high-throughput. By contrast, the single-molecule measurements are tedious and usually require the analysis of several pores before a pore suitable for sequencing is identified. Hence, the recent developments of computer modeling for structure predictions (51,52) and virtual drug screening (53) could provide a valuable alternative for testing and identifying beneficial combinations of mutations of the single-chain MspA pore in silico.

Heterogeneous channel conductances of MspA pores

Our study showed a wide variety of channel conductances for our single-chain MspA constructs, ranging from 0.5 to 4.5 nS for sc_8MspA M2 without tags (Fig. S4 *D*). This heterogeneity makes more difficult the identification of single-chain MspA pores suitable for sequencing. However, this is not a consequence of the covalent linkers between the subunits of single-chain MspA, as the channel conductances of the unmodified octameric MspA pore isolated from *M. smegmatis* also vary between 2 and 5 nS (Fig. S4 *A*). These measurements are consistent with early measurements (10,13) indicating that the conductance variability is an intrinsic property of the MspA pore. The mutations introduced in MspA M1 and MspA M2 to enable DNA translocation (22) appear to increase the intrinsic conductance variability but shift the major conductance peak from 4.5–5 nS to approximately 1.5 nS (Fig. S4, *B* and *C*), indicating a significant contribution of the constriction zone to the overall channel conductance of MspA. It is plausible that neutralizing the constriction zone by mutating the negatively charged aspartates D90 and D91 to asparagines leads to a smaller and more flexible constriction zone due to the lack of electrostatic repulsion, explaining both the shift to smaller channel conductances and their wider range. While we cannot exclude that the harsh procedures used to extract MspA from *M. smegmatis* such as organic solvents (10) or boiling in detergents (12) contribute to pore heterogeneity, it is striking that the recombinant single-chain MspA M2 pores, which were isolated from *E. coli*, showed a conductance profile similar to that of octameric MspA M2 isolated from *M. smegmatis*. Thus, it appears that pore conductance heterogeneity is largely an intrinsic property of MspA, driven mainly by the properties of the constriction zone. The hypothesis that the flexible constriction zone of the MspA variants M1 and M2 used in DNA-sequencing experiments and in the corresponding single-chain variant sc_8MspA M2 is major driver of the conductance heterogeneity of these pores is supported by an additional experimental observation. The MspA PN1 variant with a phenylalanine 97 at the tip of the loop 6, which connects the constriction zone with the subsequent β-strand (11,54), has a much smaller conductance distribution (Fig. S5 *A*), indicating that putative hydrophobic interactions of neighboring phenylalanines stabilize the constriction zone. Conductance heterogeneity is common in general diffusion pores from other bacteria. For example, OmpF, the main porin of *E. coli*, and porins of *Salmonella* and *Pseudomonas aeruginosa* show a broad range of channel conductances (55–58). By contrast, the channel conductances are much better defined for specific porins, which have specific substrate binding sites, and the much smaller ion channels, which are located in the cytoplasmic membrane. For example, LamB, a maltodextrin-specific porin of *E. coli* (59–61), OprO, a phosphate-specific porin of *P. aeruginosa* (62), and the bacterial amyloid secretion channel CsgG (63) have narrow conductance distributions.

Single-chain MspA pores with different subunit stoichiometries

The channel diameter is an important parameter determining the performance of a nanopore in sensing applications. Surprisingly, all four single-chain MspA constructs with less than eight subunits formed functional pore proteins, demonstrating a remarkable plasticity of the MspA structure, in particular of the β-barrel, for expansions and/or reductions of the number of subunits. Based on the previous observation that the octameric pore is the predominant form in a self-assembly process with the purified MspA monomer (11), we assume that the single-chain constructs with three and five

connected subunits assemble to larger oligomeric pores consisting of six, nine, and ten subunits, respectively, i.e., two and/or three sc_3MspA molecules and two sc_5MspA molecules form channels. The much broader channel conductance distributions of sc_8MspA PN1 and the other nonoctameric pores compared with octameric MspA PN1 indicate that the purification and/or refolding procedure of recombinant scMspA introduced channel heterogeneity, which is not observed for octameric MspA PN1 purified from *M. smegmatis*. This makes further interpretation of the single-channel conductances difficult. Structural analysis of the single-chain variants by electron microscopy will be necessary to obtain quantitative results about different scMspA assembly forms for each subunit composition and exact constriction zone diameters of the scMspA pores. Taken together, these results indicate a straightforward approach to optimize the pore diameter of MspA for different substrates, e.g., in nucleic acid and polypeptide sequencing, by altering the number of subunits using the single-chain MspA design. To our knowledge, this is first such experiment for any pore protein. However, more research is necessary to understand the kinetic folding pathways and thermodynamic constraints governing the formation of MspA with an altered number of subunits.

CONCLUSIONS

This study demonstrated the feasibility of producing single-chain MspA pores and their use in DNA sequencing. The single-chain MspA pore enables asymmetric mutations and/or altering the channel diameter and offers multiple pathways to improve the DNA-sequencing capability of MspA and for many other applications such as detection of RNA (64), proteins (65), and small molecules (66). We predict that the single-chain concept might be applicable to other oligomeric protein pores such as CsgG (67), ClyA (68), FraC (69), and α-HL (70), depending on the proximity of the C-terminus of one monomer to the N-terminus of the next monomer in the pore structure.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2022.01.022.

AUTHOR CONTRIBUTIONS

M.N. conceived the project; M.P. and M.N. designed the project and experiments; M.P., D.H., and L.Y. performed research; M.P., L.Y., M.W., and M.N. wrote and edited the manuscript; M.N., M.P., and M.W. directed research.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health grant R21 HG010543 to M.N. The funding agency had no involvement in preparation of the article, study design, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication.

CONFLICTS OF INTEREST

M.N. and M.P., along with the University of Alabama at Birmingham, hold patents describing the use of MspA for nanosequencing of DNA and single-chain MspA.

REFERENCES

1. Gu, L. Q., O. Braha, …, H. Bayley. 1999. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature.* 398:686–690.
2. Braha, O., L. Q. Gu, …, H. Bayley. 2000. Simultaneous stochastic sensing of divalent metal ions. *Nat. Biotechnol.* 18:1005–1007.
3. Bayley, H., and P. S. Cremer. 2001. Stochastic sensors inspired by biology. *Nature.* 413:226–230.
4. Bezrukov, S. M., and I. Vodyanoy. 1993. Probing alamethicin channels with water-soluble polymers. Effect on conductance of channel states. *Biophys. J.* 64:16–25.
5. Nekolla, S., C. Andersen, and R. Benz. 1994. Noise analysis of ion current through the open and the sugar-induced closed state of the LamB channel of Escherichia coli outer membrane: evaluation of the sugar binding kinetics to the channel interior. *Biophys. J.* 66:1388–1397.
6. Bayley, H., and C. R. Martin. 2000. Resistive-pulse sensing-from microbes to molecules. *Chem. Rev.* 100:2575–2594.
7. Deamer, D., M. Akeson, and D. Branton. 2016. Three decades of nanopore sequencing. *Nat. Biotechnol.* 34:518–524.
8. Branton, D., D. W. Deamer, …, J. A. Schloss. 2008. The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* 26:1146–1153.
9. Kasianowicz, J. J., E. Brandin, …, D. W. Deamer. 1996. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. U S A.* 93:13770–13773.
10. Niederweis, M., S. Ehrt, …, R. Benz. 1999. Cloning of the *mspA* gene encoding a porin from *Mycobacterium smegmatis*. *Mol. Microbiol.* 33:933–945.
11. Faller, M., M. Niederweis, and G. E. Schulz. 2004. The structure of a mycobacterial outer-membrane channel. *Science.* 303:1189–1192.
12. Heinz, C., and M. Niederweis. 2000. Selective extraction and purification of a mycobacterial outer membrane protein. *Anal. Biochem.* 285:113–120.
13. Stahl, C., S. Kubetzko, …, M. Niederweis. 2001. MspA provides the main hydrophilic pathway through the cell wall of *Mycobacterium smegmatis*. *Mol. Microbiol.* 40:451–464 (Authors' correction appeared in Mol. Microbiol. 457, 1509).
14. Stephan, J., J. Bender, …, M. Niederweis. 2005. The growth rate of *Mycobacterium smegmatis* depends on sufficient porin-mediated influx of nutrients. *Mol. Microbiol.* 58:714–730.
15. Wolschendorf, F., M. Mahfoud, and M. Niederweis. 2007. Porins are required for uptake of phosphates by *Mycobacterium smegmatis*. *J. Bacteriol.* 189:2435–2442.
16. Danilchanka, O., M. Pavlenok, and M. Niederweis. 2008. Role of porins for uptake of antibiotics by *Mycobacterium smegmatis*. *Antimicrob. Agents Chemother.* 52:3127–3134.
17. Jones, C. M., and M. Niederweis. 2010. Role of porins in iron uptake by *Mycobacterium smegmatis*. *J. Bacteriol.* 192:6411–6417.
18. Heinz, C., H. Engelhardt, and M. Niederweis. 2003. The core of the tetrameric mycobacterial porin MspA is an extremely stable beta-sheet domain. *J. Biol. Chem.* 278:8678–8685.
19. Niederweis, M., C. Heinz, …, S. H. Bossmann. 2001. Nanostructuring of carbon surfaces by deposition of a channel-forming protein and

subsequent polymerization of methyl-methacrylate-prepolymers. *Nano Lett.* 1:169–174.

20. Niederweis, M., C. Heinz, …, S. H. Bossmann. 2002. Nanostructuring by deposition of protein channels formed on carbon-surfaces. *Nano Lett.* 2:1263–1268.
21. Yu, L., X. Kang, …, M. Wanunu. 2021. Stable polymer bilayers for protein channel recordings at high guanidinium chloride concentrations. *Biophys. J.* 120:1–5.
22. Butler, T. Z., M. Pavlenok, …, J. H. Gundlach. 2008. Single-molecule DNA detection with an engineered MspA protein nanopore. *Proc. Natl. Acad. Sci. U S A.* 105:20647–20652.
23. Derrington, I. M., T. Z. Butler, …, J. H. Gundlach. 2010. Nanopore DNA sequencing with MspA. *Proc. Natl. Acad. Sci. U S A.* 107:16060–16065.
24. Fragasso, A., S. Schmid, and C. Dekker. 2020. Comparing current noise in biological and solid-state nanopores. *ACS Nano.* 14:1338–1349.
25. Manrao, E. A., I. M. Derrington, …, J. H. Gundlach. 2012. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat. Biotechnol.* 30:349–353.
26. Cherf, G. M., K. R. Lieberman, …, M. Akeson. 2012. Automated forward and reverse ratcheting of DNA in a nanopore at 5-A precision. *Nat. Biotechnol.* 30:344–348.
27. Brown, C. G., and J. Clarke. 2016. Nanopore development at Oxford nanopore. *Nat. Biotechnol.* 34:810–811.
28. Franceschini, L., T. Brouns, …, G. Maglia. 2016. DNA translocation through nanopores at physiological ionic strengths requires precise nanoscale engineering. *ACS Nano.* 10:8394–8402.
29. Wloka, C., N. L. Mutter, …, G. Maglia. 2016. Alpha-Helical fragaceatoxin C nanopore engineered for double-stranded and single-stranded nucleic acid analysis. *Angew. Chem. Int. Ed. Engl.* 55:12494–12498.
30. Cao, C., Y. L. Ying, …, Y. T. Long. 2016. Discrimination of oligonucleotides of different lengths with a wild-type aerolysin nanopore. *Nat. Nanotechnol.* 11:713–718.
31. Zhou, W., H. Qiu, …, W. Guo. 2020. Molecular insights into distinct detection properties of alpha-hemolysin, MspA, CsgG, and aerolysin nanopore sensors. *J. Phys. Chem. B.* 124:1611–1618.
32. Carter, J. M., and S. Hussain. 2017. Robust long-read native DNA sequencing using the ONT CsgG nanopore system. *Wellcome Open Res.* 2:23.
33. Laszlo, A. H., I. M. Derrington, …, J. H. Gundlach. 2014. Decoding long nanopore sequencing reads of natural DNA. *Nat. Biotechnol.* 32:829–833.
34. Nova, I. C., I. M. Derrington, …, J. H. Gundlach. 2017. Investigating asymmetric salt profiles for nanopore DNA sequencing with biological porin MspA. *PLoS One.* 12:e0181599.
35. Noakes, M. T., H. Brinkerhoff, …, J. H. Gundlach. 2019. Increasing the accuracy of nanopore DNA sequencing using a time-varying cross membrane voltage. *Nat. Biotechnol.* 37:651–656.
36. Ledbetter, M. P., J. M. Craig, …, F. E. Romesberg. 2020. Nanopore sequencing of an expanded genetic alphabet reveals high-fidelity replication of a predominantly hydrophobic unnatural base pair. *J. Am. Chem. Soc.* 142:2110–2114.
37. Agah, S., M. Zheng, …, A. B. Kolomeisky. 2016. DNA sequencing by nanopores: advances and challenges. *J. Phys. D Appl. Phys.* 49:413001.
38. Derrington, I. M., J. M. Craig, …, J. H. Gundlach. 2015. Subangstrom single-molecule measurements of motor proteins using a nanopore. *Nat. Biotechnol.* 33:1073–1075.
39. Schreiber, J., Z. L. Wescoe, …, M. Akeson. 2013. Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proc. Natl. Acad. Sci. U S A.* 110:18910–18915.
40. Pavlenok, M., I. M. Derrington, …, M. Niederweis. 2012. MspA nanopores from subunit dimers. *PLoS One.* 7:e38726.
41. Distefano, M. D., H. Kuang, …, A. Mazhary. 1998. The design of protein-based catalysts using semisynthetic methods. *Curr. Opin. Struct. Biol.* 8:459–465.
42. Pavlenok, M., and M. Niederweis. 2016. Hetero-oligomeric MspA pores in *Mycobacterium smegmatis*. *FEMS Microbiol. Lett.* 363:fnw046.
43. Prilipov, A., P. S. Phale, …, R. Koebnik. 1998. Coupling site-directed mutagenesis with high-level expression: large scale production of mutant porins from *E. coli*. *FEMS Microbiol. Lett.* 163:65–72.
44. Heinz, C., E. Roth, and M. Niederweis. 2003. Purification of porins from *Mycobacterium smegmatis*. *Methods Mol. Biol.* 228:139–150.
45. Sambrook, J., and D. W. Russell. 2006. Purification of expressed proteins from inclusion bodies. *CSH Protoc.* 2006:pdb.prot4089.
46. Kang, X., M. A. Alibakhshi, and M. Wanunu. 2019. One-pot species release and nanopore detection in a voltage-stable lipid bilayer platform. *Nano Lett.* 19:9145–9153.
47. Manrao, E. A., I. M. Derrington, …, J. H. Gundlach. 2011. Nucleotide discrimination with DNA immobilized in the MspA nanopore. *PLoS One.* 6:e25723.
48. Dohm, J. C., P. Peters, …, H. Himmelbauer. 2020. Benchmarking of long-read correction methods. *NAR Genom. Bioinform.* 2:lqaa037.
49. Jain, M., H. E. Olsen, …, M. Akeson. 2016. The Oxford nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol.* 17:239.
50. Lu, B., S. Fleming, …, J. Golovchenko. 2015. Thermal motion of DNA in an MspA pore. *Biophys. J.* 109:1439–1445.
51. Yang, J., I. Anishchenko, …, D. Baker. 2020. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. U S A.* 117:1496–1503.
52. Senior, A. W., R. Evans, …, D. Hassabis. 2020. Improved protein structure prediction using potentials from deep learning. *Nature.* 577:706–710.
53. Pereira, J. C., E. R. Caffarena, and C. N. Dos Santos. 2016. Boosting docking-based virtual screening with deep learning. *J. Chem. Inf. Model.* 56:2495–2506.
54. Huff, J., M. Pavlenok, …, M. Niederweis. 2009. Functions of the periplasmic loop of the porin MspA from *Mycobacterium smegmatis*. *J. Biol. Chem.* 284:10223–10231.
55. Benz, R., K. Janko, …, P. Lauger. 1978. Formation of large, ion-permeable membrane channels by the matrix protein (porin) of *Escherichia coli*. *Biochim. Biophys. Acta.* 511:305–319.
56. Benz, R., J. Ishii, and T. Nakae. 1980. Determination of ion permeability through the channels made of porins from the outer membrane of Salmonella typhimurium in lipid bilayer membranes. *J. Membr. Biol.* 56:19–29.
57. Benz, R., K. Janko, and P. Lauger. 1980. Pore formation by the matrix protein (porin) of Escherichia coli in planar bilayer membranes. *Ann. N Y Acad. Sci.* 358:13–24.
58. Benz, R., and R. E. Hancock. 1981. Properties of the large ion-permeable pores formed from protein F of Pseudomonas aeruginosa in lipid bilayer membranes. *Biochim. Biophys. Acta.* 646:298–308.
59. Nakae, T. 1979. A porin activity of purified lambda-receptor protein from *Escherichia coli* in reconstituted vesicle membranes. *Biochem. Biophys. Res. Commun.* 88:774–781.
60. Nakae, T., and J. Ishii. 1980. Permeability properties of *Escherichia coli* outer membrane containing, pore-forming proteins: comparison between lambda receptor protein and porin for saccharide permeation. *J. Bacteriol.* 142:735–740.
61. Klebba, P. E. 2002. Mechanism of maltodextrin transport through LamB. *Res. Microbiol.* 153:417–424.
62. Hancock, R. E., C. Egli, …, R. J. Siehnel. 1992. Overexpression in *Escherichia coli* and functional analysis of a novel PPi-selective porin, OprO, from *Pseudomonas aeruginosa*. *J. Bacteriol.* 174:471–476.

63. Goyal, P., P. V. Krasteva, …, H. Remaut. 2014. Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature.* 516:250–253.

64. Garalde, D. R., E. A. Snell, …, D. J. Turner. 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods.* 15:201–206.

65. Nivala, J., L. Mulroney, …, M. Akeson. 2014. Discrimination among protein variants using an unfoldase-coupled nanopore. *ACS Nano.* 8:12365–12375.

66. Guan, X., L. Q. Gu, …, H. Bayley. 2005. Stochastic sensing of TNT with a genetically engineered pore. *Chembiochem.* 6:1875–1881.

67. Cao, B., Y. Zhao, …, Y. Huang. 2014. Structure of the nonameric bacterial amyloid secretion channel. *Proc. Natl. Acad. Sci. U S A.* 111:E5439–E5444.

68. Mueller, M., U. Grauschopf, …, N. Ban. 2009. The structure of a cytolytic alpha-helical toxin pore reveals its assembly mechanism. *Nature.* 459:726–730.

69. Tanaka, K., J. M. Caaveiro, …, K. Tsumoto. 2015. Structural basis for self-assembly of a cytolytic pore lined by protein and lipid. *Nat. Commun.* 6:6337.

70. Song, L., M. R. Hobaugh, …, J. E. Gouaux. 1996. Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science.* 274:1859–1866.