De Novo Design of a Nanopore for DNA Detection that Incorporates a β-Hairpin Peptide

Keisuke Shimizu
Tokyo University of Agriculture and Technology
Batsaikhan Mijiddorj
National University of Mongolia
Masataka Usami
Tokyo University of Agriculture and Technology
Shuhei Yoshida
Konan University
Shiori Akayama
Konan University
Yoshio Hamada
Konan University
Akifumi Ohyama
Yokohama National University
Kenji Usui
Konan University https://orcid.org/0000-0003-4558-5762
Izuru Kawamura
Yokohama National University https://orcid.org/0000-0002-8163-9695
Ryuji Kawano (rjkawano@cc.tuat.ac.jp)
Tokyo University of Agriculture and Technology

Article

Keywords: nanopore, DNA detection, β-hairpin peptide

Posted Date: January 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-141682/v1

Version of Record: A version of this preprint was published at Nature Nanotechnology on November 22nd, 2021. See the published version at https://doi.org/10.1038/s41565-021-01008-w.

De Novo Design of a Nanopore for DNA Detection that Incorporates a β-Hairpin Peptide

K. Shimizu[1], B. Mijiddorj[2,3], M. Usami[1], S. Yoshida[4], S. Akayama[4], Y. Hamada[4], A. Ohyama[5], K. Usui[4], I. Kawamura[2,5,] and R. Kawano[1*]

[1]Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Japan
[2]Graduate School of Engineering, Yokohama National University, Japan
[3]School of Engineering & Applied Sciences, National University of Mongolia, Mongolia
[4]Faculty of Frontiers of Innovative Research in Science and Technology, Konan University, Japan
[5]Graduate School of Engineering Science, Yokohama National University, Japan

Abstract

The amino acid sequence of a protein encodes information on its three-dimensional structure and specific functionality. *De novo* protein design has emerged as a method to manipulate the primary structure for the development of artificial proteins and peptides with desired functionality. This paper describes the *de novo* design of a pore-forming peptide that has a β-hairpin structure and assembles to form a stable nanopore in a bilayer lipid membrane. This large synthetic nanopore is an entirely artificial device with practical applications. This peptide, named SV28, forms nanopore structures ranging from 1.6 to 6.2 nm in diameter assembled from 7 to 18 monomers. The nanopore formed with a diameter of 5 nm is able to detect long double-stranded DNA (dsDNA) with 1 kbp length. Moreover, the larger sized nanopore can discriminate and human telomeric DNA (G-quadruplex, G4). The blocking current signals allowed us to investigate the translocation behavior of dsDNA or G4 structure at the single molecule level. Such *de novo* design of peptide sequences has the potential to create novel nanopores, which would be applicable in molecular transporter between across lipid membrane.

Introduction

The folded structure of proteins is determined by their linear polypeptide sequence, as postulated in Anfinsen's dogma,[1] and gives rise to specific protein functionality. All proteins have a unique structure and size; the folded structure relies on the primary sequence of the amino acids, with this unique primary structure is a result of structural evolution such as the mutation and selection of amino acid residues over time. To reveal the relationship between this primary information and protein structure is one of the ultimate goals of science.

The *de novo* design of the primary sequence of artificial proteins has been studied in the last four decades,[2] recently also described as design "from scratch". [3-11] In early studies, the secondary structures of proteins - α-helix and β-sheet structures - were created synthetically by peptide

chemistry,[12,13] with these secondary structures subsequently connected through a loop sequence to construct the more complicated three-dimensional structure.[14] The design strategy at this time was manual and based on the physical model of proteins and peptides. Computational design guided by physicochemical principles has since been developed.[7] This method has recently been considered a powerful tool for the *de novo* design not only for constructing protein structure but also for expressing specific protein functions. D. Baker and co-workers proposed extensive artificial proteins designed using Rosetta algorithms such as a fluorescence-activating protein with a β-barrel,[15] and transmembrane proteins with α-helical[16] structures. *De novo* design has the potential not only to mimic the natural proteins but also to create artificial devices such as molecular machines. In the creation of manufactured devices for practical applications, pore-forming transmembrane structures are key targets because single molecular detection and DNA sequencing has been achieved using such pore-forming proteins.[17-20]

Nanopore sensing is a powerful tool for label-free single-molecule detection.[17,21-23] Once a nanopore-forming membrane protein has been reconstituted in a lipid bilayer to form a nano-sized pore, the target molecule is able to electrophoretically pass through the nanopore under an applied voltage. A partially blocked current is observed when the single molecule passes through the nanopore, which can be identified by analyzing the blocking amplitude and duration. The most promising application of nanopore technology is in DNA sequencing.[24] In principle, single-stranded DNA (ssDNA) shows four different blocking currents for the four different bases at the single DNA level. As a result of great research effort, the nanopore sequencer has been commercialized as an inexpensive and portable DNA sequencer device. Besides DNA sequencing, a wide variety of nanopore studies have been proposed, such as small molecule detection using an adapter[25] or DNA aptamer,[26] nanopore mass spectroscopy,[27] decoding of DNA computations,[28-33] and protein or peptide detection.[34-36] The choice of applicable target molecule is sometimes limited because the selectivity of nanopore sensing depends on the pore size, and the size variation of natural pore-forming proteins is insufficient for the detection of a range of molecules.[37] Bottom-up nanopore design has great potential to expand target variation, owing to the possibility to tailor size compatibility between the nanopore and target molecules. Moreover, this can offer improved accuracy of nanopore detection, with a potentially significant influence on peptide sequencing.

In designing an artificial protein nanopore, the process of pore insertion into the lipid membrane must also be considered. In the case of a natural system, membrane proteins are inserted into the cell membrane via chaperones or ER export. One way to facilitate membrane insertion is to use short peptides. For example, an α-helical barreled peptide (35 a.a.) based on the Wza protein, an outer-membrane auxiliary protein associated with exopolysaccharide assembly systems, was redesigned from the wild type and assembled to form monodisperse nanopores in lipid membranes.[38,39] However although the Wza-based nanopore was successfully assembled to form

the nanopore, tailoring its size and function remains challenging in terms of the design strategy. Consequently, *de novo* design would offer advantages in the design of pore-forming structures by utilizing the vast sequence space of amino acids.

We here focus on the β-barrel structure because the transmembrane region of the majority of biological nanopores, including αHL, has a β-barrel structure. However, the synthesis of β-sheet peptides is challenging due to their hydrophobicity that often leads to peptide aggregation during the synthesis process. In this study, a transmembrane peptide with a *de novo* design incorporating a β-hairpin structure was chemically synthesized using an isoacyl dipeptide method.[40] This method allowed facile synthesis of the hydrophobic β-sheet peptides. Our designed β-hairpin peptide assembled and formed a β-barreled structure with several different sizes of nanopore ranging from 1.6 to 6.2 nm in diameter. Using the peptide nanopore with 5 nm in diameter, we were successful in detecting long double-stranded DNA (dsDNA) with 1 kbp length and subsequent analysis of its translocation behavior at the single molecular level. Moreover, the peptide nanopore forms lager nanopore (>6.2 nm) and this large nanopore can discriminate to form human telomeric DNA (G-quadruplex, G4) comprising the repeat sequence 5'-TTAGGG-3'. The *de novo* designed β-barrel nanopore has the potential to adjust its size and shape to a target molecule, and to apply for detecting DNA or other extensive molecules.

Results

Design of β-hairpin peptide SV28

We designed the β-hairpin structure with three different regions of amino acid sequence: a β-strand backbone, a β-turn, and two terminal structures (Fig. 1a). We first considered the length of the peptide. The β-strands are necessary to have appropriate length in the transmembrane region. We decided to use ten amino acids to provide a length compatible with the thickness of lipid bilayer, because this is similar to the β-strand length of natural transmembrane β-barrel proteins.[41] The β-turn can form a bent structure with 4 amino acids.[42] The terminal structures require two amino acids with a random coil structure in the extra-membrane regions. Overall, a length of 28 amino acids was selected for use in this study (Fig. 1a). The following three strategies were then used to determine the type of amino acid residues in each region.

1) Construction of amphiphilic β-sheet structure as the β-strand backbone.

Alternation of hydrophilic and hydrophobic amino acids was used to promote formation of the β-sheet structure (Fig. 1b).[43] Separation of the hydrophilic and hydrophobic surfaces to be the inside and outside of the nanopore facilitates construction of the β-barrel structure in the lipid bilayer. Ser was selected as the hydrophilic residue, because Ser is the smallest hydrophilic amino acid, and may prevent steric hindrance in nanopore formation. Although Ala is the smallest hydrophobic

amino acid, it tends to form an α-helical structure, and so Val was selected instead as the hydrophobic residue. Furthermore, it has previously been reported that Val is the most common hydrophobic residue, except for aromatic residues, in natural β-barrel proteins.[44]

2) Stabilization of the membrane-spanning state using the snorkeling effect.

In natural α-helix and β-barrel membrane proteins, Trp and Tyr are sometimes localized at the interface between the aqueous and lipid phases, enhancing stability of the transmembrane structure.[45] This is known as the snorkeling effect (Fig. 1c), wherein the amphiphilic side chain of Tyr, common in β-barrels, works as an anchor at the interface between the hydrophilic and hydrophobic phase.[46] Based on this information, we selected Tyr, with the location of Tyr carefully decided at positions 4, 12, 18, and 26 from the N-terminus; such positions are located around the

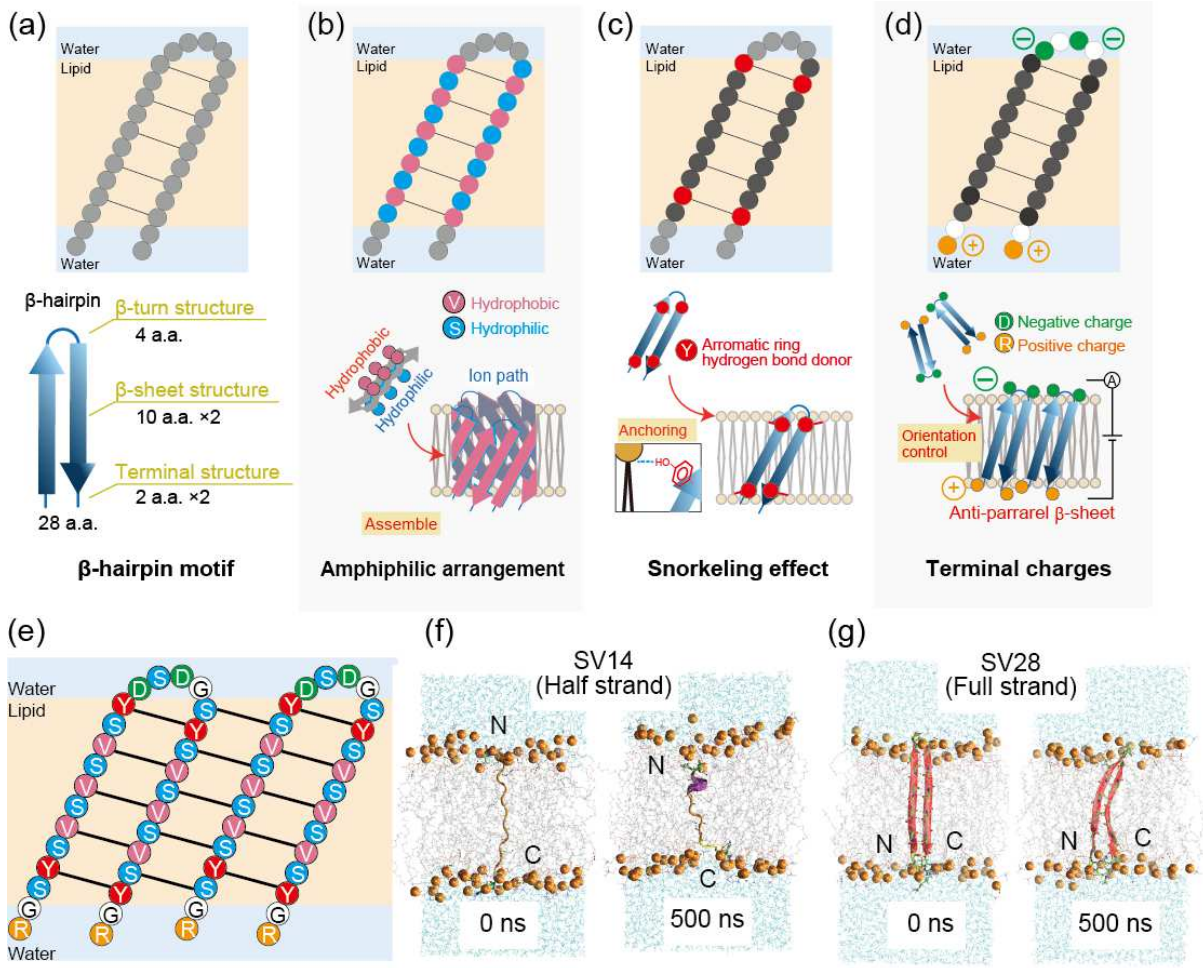

Fig. 1 Design strategies and structural confirmation of the β-hairpin peptide. (a) The design of β-hairpin peptide with 28 a.a. divided into three sections: β-turn, β-sheet transmembrane, and the terminals. (b) Hydrophilic and hydrophobic amino acids are arranged in an alternating fashion. (c) Interaction of aromatic rings stabilizes the β-barrel pore. (d) Designing specific charges at the terminus allows control of peptide orientation upon an applied voltage. (e) Amino acid sequence of designed structure, named SV28. The black lines indicate the hydrogen bonding. (f, g) MD simulations of the monomer structures. The 0 ns (left) and 500 ns (right) snapshots of the (f) half-length and (g) full length SV28 in a lipid bilayer membrane.

interface of the aqueous and lipid phases and direct towards the hydrophobic outside region (Fig. 1c).

3) Introduction of charged residues for controlling the orientation of SV28 using an applied voltage.

Anti-parallel β-sheets have more strong interactions than parallel β-sheets. To control the peptide orientation for forming the anti-parallel β-sheets structure, two negatively and positively charged residues were introduced at the β-turn and the terminal regions, respectively (Fig. 1d). The β-turn consists of four amino acids and the systematic study of the β-turn sequence has been previously reported.[42] Based on this report, we decided on the sequence of –DSDG– that has two negatively charged Asp residues. The N- and C-terminal regions were designed as RG– and –GR, respectively. Gly is the linker between charged Arg connecting to the β-sheet backbone. Combining all design strategies, the final sequence of SV28 was decided, as shown in Fig. 1e. The β-hairpin formation of the SV28 sequence was computationally confirmed using MINNOU simulator, software that predicts transmembrane domains of membrane proteins and peptides.[47]

We verified the formation of the nanopore structure from assembled SV28 peptides by all-atom MD simulations. Initially, we carried out 500 ns simulations of the monomer full- and half-length of SV28 to confirm the stability of the β-hairpin structure in the membrane (Fig. 1f and 1g). The full-length SV28 monomer kept the β-hairpin structure, whilst the half-length mostly showed random coil secondary structures during these simulations (Fig. S1). These results suggest that the hydrogen bond formations in the β-hairpin structure play an important role in maintaining a stable structure in the membrane.

To confirm the β-barrel formation of SV28 peptides, we performed 900 ns MD simulations of 5-mer and 11-mer pores in a DOPC membrane. Fig. 2a and S2 display the final snapshots of these simulations, with the 11-mer structure also shown in Fig. 2b. These simulations were started after 100 ns long equilibrations (see Methods for the explanation of the equilibrated system). Both simulations demonstrated stable nanopore structures. The 5-mer pore was slightly tilted in the membrane in Fig. 2a, which may indicate the adaptation of the structure to the movement of surrounding lipids. The central diameter of the pores was calculated at 1 ns intervals, as shown in Fig. 2c and S2. The average values are equal to 10.1±0.4 Å and 28.5±2.0 Å for 5-mer and 11-mer pores respectively.

Furthermore, the secondary structure profiles of two monomers of the pores as a function of time are shown in Fig. 2d. The 10^{th} (Val10) and 22^{nd} (Val22) positions form a β-sheet structure, and the 16^{th} position (Gly16) forms a turn structure. To compare to the experimental results, we calculated the distance between the nitrogen of Val10 and carbon of Val22 during the simulations (Fig. S3). The average distances of the 5-mer and 11-mer peptides are 4.2±0.1 Å and 4.1± 0.1 Å respectively. These distances are consistent with the usual value of the distance in the β-barrels of outer

membrane proteins. The range of the distance variation in the 5-mer was slightly larger than that in the 11-mer. This is likely due to the structural distortion of β-sheets in 5-mer pores.

Synthesis of hydrophobic SV28 using isoacyl dipeptide

Conventional solid-phase synthesis and subsequent purification of SV28 is challenging because the β-sheet structure is prone to aggregation. A method using isoacyl dipeptides[40] gives the aqueous-soluble SV28 preform; the final chemical structure can be derived from this preform after transition of the isoacyl group from the side to the main chain. The insertion of the isoacyl dipeptides into the β-sheet regions prevents the formation of β-sheet structures and hence prevents aggregation. The isoacyl peptides used in this study form an ester bond between the C-terminal of Val and the hydroxyl group of Ser. This ester bond was transferred to a peptide bond by incubation in a basic solution (pH=13) for 5 min before using the peptides in the experiments. Using this method, it was possible to synthesize and purify SV28 using conventional peptide synthesis.

The secondary structure of the transferred SV28 with DOPC liposomes was confirmed by circular dichroism (CD) spectroscopy; generally, the formation of a β-sheet structure can be seen from the positive and negative absorbances at 195 nm and 216 nm. The SV28 preform did not show a large negative absorbance at 216 nm (Fig. 2e). After the transference of isoacyl dipeptide, a large negative absorbance at 216 nm was observed (Fig. 2e), indicating that the transformation of isoacyl dipeptide facilitated formation of a β-sheet structure in SV28. Moreover, after 24 h incubation in a buffer solution, the SV28 was seen to maintain the β-sheet structure (Fig. S4).

The solid-state NMR measurements also supported the proposed conformation of SV28 in DOPC liposomes. ^{13}C and ^{15}N NMR signals of [^{13}C]Val10, [^{13}C]Gly16 (Fig. 2f and S5), and [^{15}N]Val22 were observed at 174.2, 45.3, and 124.1 ppm (Fig. 2g), corresponding to the β-sheet and the random coil structure at the transmembrane (Val10 and Val22) and the β-turn (Gly16) region.[48] Next, we confirmed the formation of hydrogen bonding between two β-strands using Val10 and Val22 (Fig. 2h), with the distance between them also checked by MD simulation as mentioned above. Rotational echo-double resonance (REDOR) was used to estimate the interatomic distance between the carbonyl carbon of Val10 and amide nitrogen of Val22 in SV28 to be around 4.4 Å (Fig. S6), implying formation of hydrogen bonds between the β-strands.

The NMR-deduced structure approximately corresponds to the structure predicted by the MD simulation. Overall, the information suggests the formation of the β-turn-β structure of SV28.

Confirmation of pore formation using channel current measurements

The pore-forming properties of SV28 were examined by the channel current recording in our lipid bilayer system (Fig. S7).[49-51] Several pore-opening states in which the current was raised, but did not plateau, were observed in the DOPC lipid bilayer under +100 mV (Fig. 3a). Sometimes step-

like signals were observed (Fig. 3b). Other different shapes of current signals were also observed. We have previously proposed current signal classification for several different signal shapes, and assigned these to various pore-forming models of α-helical peptides.[52-54] In this study, we also classified these signals into four types of current signals: step, square-top (Fig. 3c), multi-level, and erratic (Fig. 3d). The definition of signal classification is described in Fig. S8. To estimate the pore-

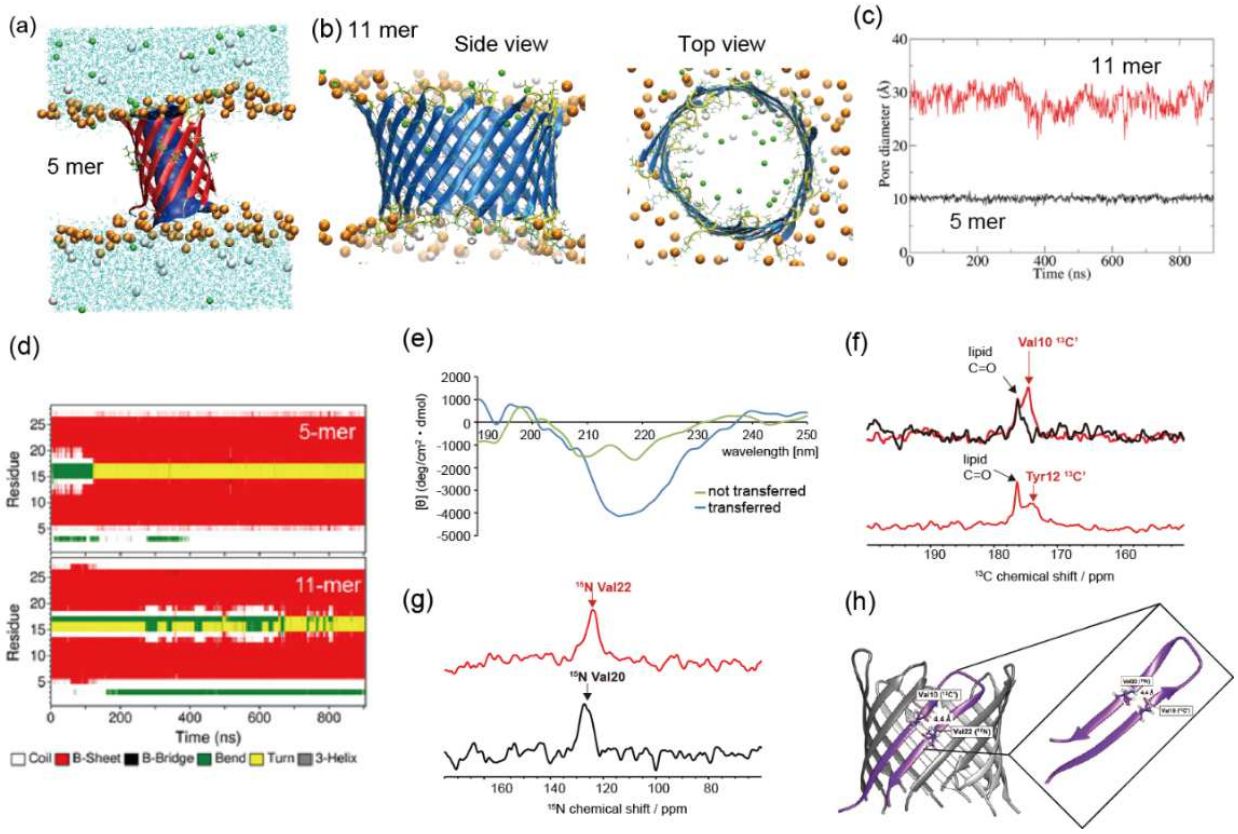

Fig. 2 (a) 5-mer structures of SV28 in a DOPC membrane in the MD simulation. Ribbons show the peptide structures, with the secondary structure indicated by the color of the ribbon (red: β-sheet, cyan: turn, white: random coil structure). Ribbon arrows indicate the direction of the backbone from N-terminal to C-terminal. The pore structures were analyzed by HOLE software and displayed as blue surfaces inside of the barrels. Val10 and Val22 amino acids showing central rim of the pores were displayed as the licorice models. Cyan lines indicate water molecules, and the lipid molecules were omitted for clarity (excluding phosphorus atoms as orange spheres). Green and white spheres indicate the potassium and chloride ions respectively. Structures were displayed by VMD software. (b) Molecular dynamics simulation of the SV28 nanopore formation in DOPC lipid membrane. The 11-mer-nanopore was simulated for 900 ns. Each color indicates: light blue: water, brown: lipid head, green: potassium ions, grey: chloride ion. (c) Central diameters of 5-mer (black) and 11mer (red) pores as a function of time were calculated using HOLE software at 1 ns interval. (d) Profile of the secondary structure of 5-mer and 11-mer pores as a function of time. (e) Circular dichroism spectra of non-transformed SV28 (green line) and transformed SV28 (blue line). (f) ^{13}C and ^{15}N CP-MAS NMR spectra of the triply isotope-labeled SV28 ([1-^{13}C]Val10, [2-^{13}C]Gly16, [^{15}N]Val22-labeled SV28) in DOPC liposomes. The black lines indicate the spectra from DOPC liposomes, and the red lines indicate the spectra of isotope-labeled SV28 with DOPC liposomes. (g) ^{13}C and ^{15}N CP-MAS NMR spectra of the triply isotope-labeled SV28 ([1-^{13}C]Tyr12, [2-^{13}C]Gly16, [^{15}N]Val20-labeled SV28). (h) The β-barrel structure of SV28 nanopore. The interatomic distance between the backbone amide of Val22 and the carbonyl carbon of Val10 in the SV28 is estimated to be around 4.4 Å as measured by solid-state NMR.

forming behavior of SV28, we here define that step and square-top signals reflect stable pore formations, and multi-level and erratic signals reflect unstable pore formations~~, as depicted in Fig. S2~~. In the initial measurement of SV28, the stable signals were observed only 3% of the time (Fig. 3e, left bar), which led us to investigate how to improve stable pore formation.

To optimize formation of stable pores, we investigated three different conditions as follows:

1) Optimization of applied voltage to align the SV28 monomers.

It is proposed that the orientation of SV28 can be controlled by applying a voltage, since we designed the positive and negative charged amino acids to be positioned at the turn and terminal regions. As predicted, the ratio of stable signals increased with increasing the applied voltage from +100 to +200 mV (Fig. 3e).

2) Incubation of SV28 monomers to form oligomer structures with the lipid monolayer surface.

To give an energetically stable structure of the assembled SV28, we attempted to incubate the SV28 monomers in a mixed aqueous-lipid/*n*-decane solution. We expected the SV28 monomers to assemble at the interface between aqueous and lipid/*n*-decane solution and form a stable pore during incubation. After incubating for 24 h at 37°C, the ratio of the stable signals dramatically increased from 27% to 61% (Fig. 3e). A similar phenomenon was previously reported wherein amyloid β(1-42), channel-forming β-sheet peptide, forms stable pores after incubation with lipid micelles.[55]

3) Adding cholesterol to the lipid bilayer.

Cholesterol addition enhanced formation of stable pores (Fig. 3e). This result may be explained in that cholesterol addition results in reduced fluidity of the DOPC bilayer, and subsequently there is reduced disturbance and pore dissociation. Combining the three optimized conditions led to a further increase in the ratio of stable signals to 67% (Fig. 3e, right). This ratio is close to the ratio we previously reported for pore-forming proteins.[37] The other typical current signals and the step signals are shown in Fig. S9 and S10.

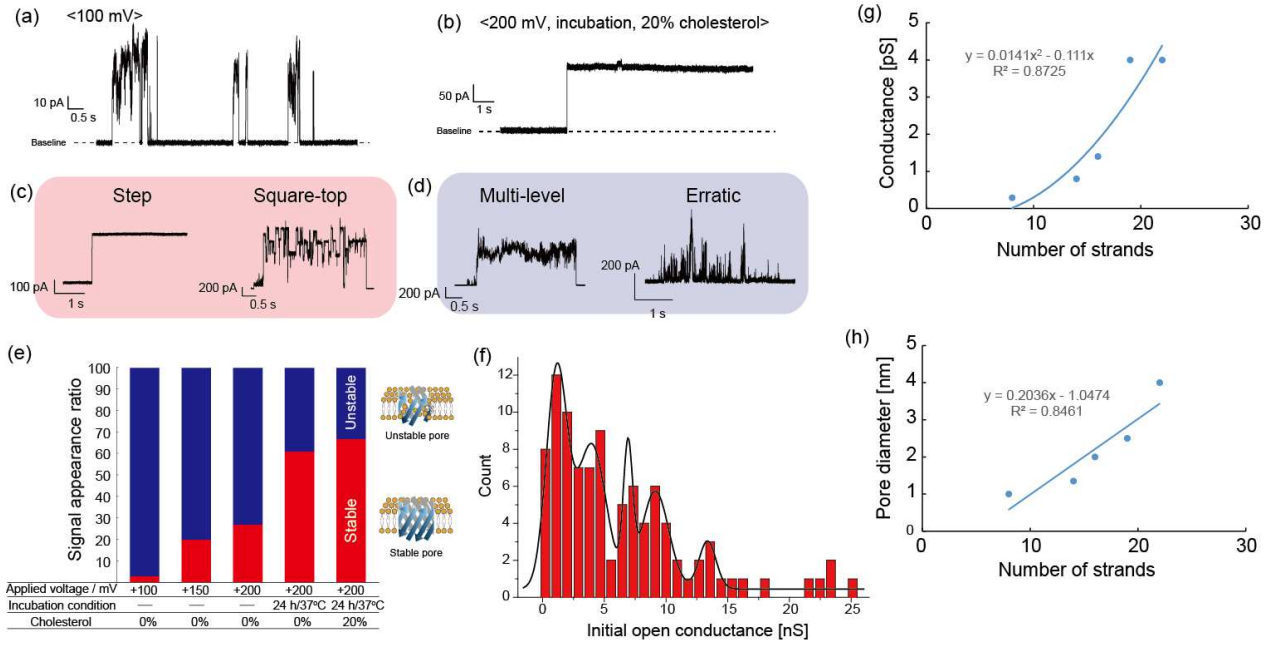

Fig. 3 (a) The typical current and time traces of SV28 at +100 mV with initial condition. (b) The step signal was occasionally observed under the initial conditions. The traces were atypical before optimization of the conditions. (c) Step and square-top signals were defined as stable pore formation. (d) Multi-level and erratic signals were defined as unstable pore formation. (e) The ratio of stable and unstable pore formation from the channel current analysis. (f) The histogram of the current conductance of the initial step from the 0 A increase of SV28 in the DOPC lipid bilayer. (g, h) The relationships in β-barrel membrane proteins OmpA, OmpG, OmpF, VDAC, and FhuA between (g) the channel conductance and the number of β-strands and (h) the pore diameter and number of β-strands.

Next, we considered how SV28 forms the nanopore with a β-barreled structure. There have been many reports on the pore/channel formation of α-helical peptides. These peptides initially bind to the surface of the lipid membrane and form α-helical structures, which subsequently assemble from the monomers to construct the transmembrane nanopore structure. Several pore-forming models have already been proposed for these structures, such as barrel-stave or toroidal models.[56] We have also proposed the assignment of current signals to these models in planar lipid bilayer experiments.[54] Although there are few studies on the pore-formation of β-sheet peptides, it has been reported that the β-sheet peptides also construct barrel-stave and toroidal pores.[57] Our electrophysiological measurements herein also presented the step and multi-level signals, analogous to those previously assigned to the barrel-stave and the toroidal models.

The pore diameter of the SV28 nanopore was calculated from the conductance of the open channel state and the Hille equation, which is a theoretical model that uses the resistance of a cylindrical pore to ion flow.[58] The open channel conductance was used as the initial step signals from the baseline (≈0 A). The histogram of the pore conductance of SV28 is shown in Fig. 3f. There are five identified peaks at conductances of 1, 4, 7, 9, and 14 nS, which give pore diameters of 0.8,

1.7, 2.3, 2.7, and 3.5 nm using the Hille model. The numbers of monomers used in nanopore assembly were mathematically calculated to be 4, 5, 7, 8, and 10 monomers using the diameters and the size of the β-hairpin molecule. Additionally, we estimated the pore size using a more precise method that is based on the experimental results instead of the theoretical Hille model. We assessed the relationship between the current conductance and the pore diameter of β-barrel proteins, measurements that were taken by electrophysiology, and by crystallography using microscopes. OmpA, OmpF, OmpG, VDAC, and FhuA were used in this estimation as the β-barrel transmembrane proteins (the detail is described in the Supplementary Information). Based on the relationship between the current conductance and the pore diameter of the natural proteins (Fig. 3g

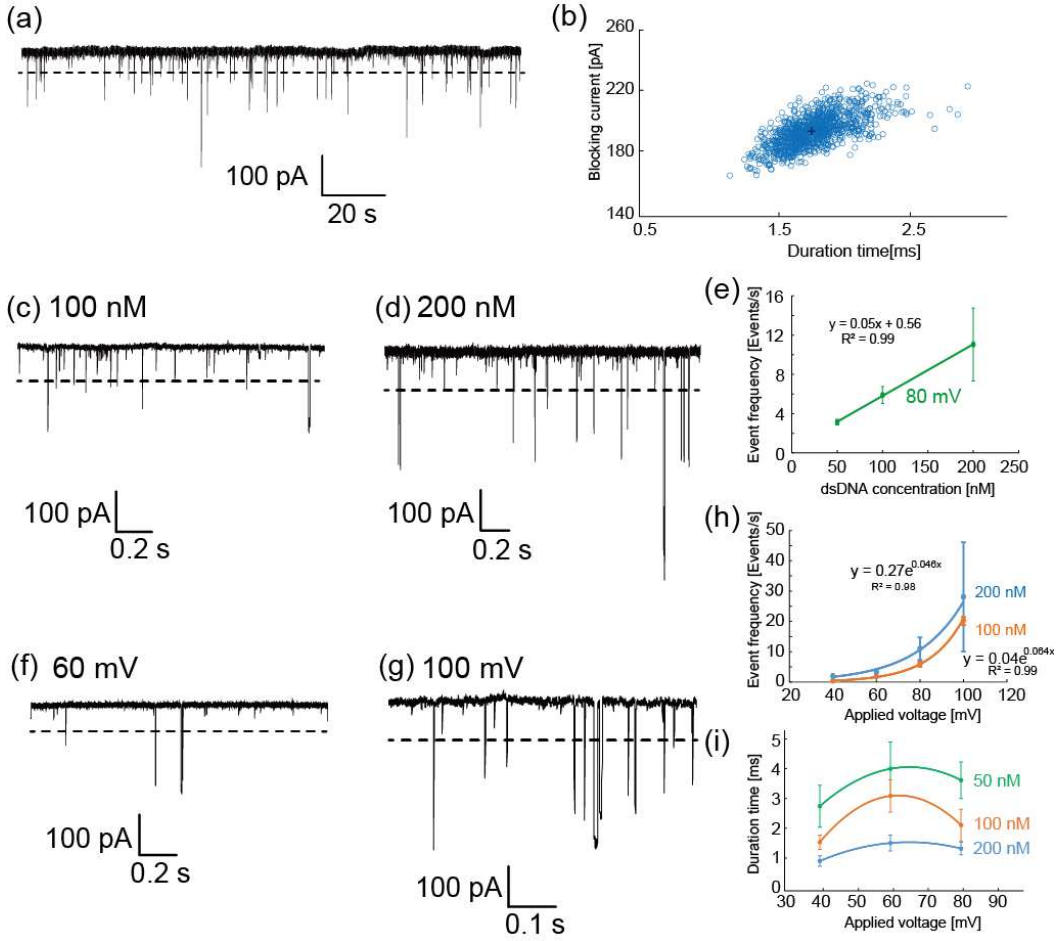

Fig. 4 The dsDNA (1 kbp) translocation through the SV28 nanopore with diameter of around 5 nm. (a) The current and time trace of the SV28 nanopore with 1 kbp dsDNA (100 nM) under application of 40 mV. The dashed lines indicate the threshold for the dsDNA translocation events as we defined. (b) The scatter plot of the translocation data of 100 nM 100 mV: the blocking rate and the duration after the bootstrapping. (c, d) The current and time traces of SV28 with dsDNA at (c) 100 nM and (d) 200 nM under 100 mV. The dashed lines indicate the threshold for the dsDNA translocation events as we defined. (e) The event frequency of the translocation as a function of the concentration of dsDNA. (f, g) The current and time traces of SV28 with dsDNA (100 nM) under (f) 60 mV and (g) 100 mV. The dashed lines indicate the threshold for the dsDNA translocation events as we defined. (h) The event frequency of the dsDNA translocation as a function of the applied voltage. Blue and orange lines indicate the dsDNA concentration at 100 nM and 200 nM to guide the eye. (i) Duration time of dsDNA with 50, 100, and 200 nM dependence on the applying voltages. Blue, orange, and green lines indicate each concentration to guide the eye.

and 3h), the five different pore sizes and the number of monomers in the case of SV28 are 1.6, 3.3, 4.4, 5.0, and 6.2 nm with 7, 11, 13, 15, and 18 monomers of SV28. The experimentally estimated pore sizes are larger than those from the theoretical Hille model.

Detection of double-stranded DNA using the SV28 nanopore

The SV28 forms nanopore structures with diameters ranging from 1.6 to 6.2 nm. In the conventional nanopore proteins for DNA detection such as αHL or MSPA,[59] ssDNA can translocate through the nanopore, whereas dsDNA cannot pass due to the size mismatch. The detection of dsDNA fragments using solid-state nanopores has recently been emerging in methylation analysis[60] without the need for complex procedures. Therefore, we used SV28 with a pore diameter of around 5.0 nm and attempted to detect dsDNA with lengths of 50 bp and 1 kbp. Although the translocation signal was not discriminated to the inherent blockage of the pore in the case of 50 bp dsDNA measurements at 1 and 10 μM concentrations (Fig. S11), the blocking currents were consistently observed using 1 kbp dsDNA ranging from 50 to 200 nM in concentration (Fig. 4a and S12). The scatter plot of the blocking rate and duration time of translocation at 100 nM under application of 100 mV is shown in Fig. 4b, with peak values of 195 pA and 1.8 ms respectively (scatter plots of other conditions without bootstrapping are presented in Fig. S13). Additionally, the translocation of multiple dsDNA through the pore was occasionally observed (Fig. S14).

We next examined the dependency of the translocation frequency of dsDNA on the concentration and applied voltage. The event frequency *versus* concentration showed a linear dependency ranging from 50 to 200 nM under the application of 80 mV (Fig. 4c-4e). Moreover, the event also depended exponentially on the voltage application from 40 to 100 mV at both 100 nM and 200 nM concentration (Fig. 4f-4h). These linear and exponential dependencies on the concentration and applied voltage are consistent with previously reported results of DNA translocation using the αHL nanopore. [61] The voltage dependency of the duration time did not show simple decreasing with voltage application, they had a peak voltage. Similar result has previously reported, and it can explain that the target molecule is rejected from the pore under low voltage application.[62] The scatter plots of each data are shown in Fig. S15.

Detection of G4 structure using the SV28 nanopore

Since the SV28 forms several different sizes of the nanopore, we here detect the G4 structure of DNA using a large SV28 nanopore (6.2 nm). The G4 DNA forms a hybrid hold with a 3.0 x 2.7 nm box-like structure and the G4 structure is strung out as five boxes in this salt condition (Fig. 5a and 5b).[63,64] In the case of G4 structure, the blocking signals are observed although the same length of DNA without G4 structure is not recognized as shown in Fig. 5c and 5d (the sequence are shown in supporting information). The location of the scatter plots in between with and without G4

structure obviously different. The peak top of the blockade of the G4 structure is 669 pA at 100 mV applied. The ratio of the blocking amplitudes between dsDNA and G4 is: I_{dsDNA}/I_{G4} = 0.3 (195 pA/669 pA). This value roughly consists of the ratio of their molecular volume: V_{dsDNA}/V_{G4} = 0.4 (12.5 nm^3/29.2 nm^3 with 4 nm length). The dwell time and translocation frequency of G4 depend on the applying voltages (Fig. 5f and 5g). The area of scatter plots with different voltages is positioned in a different area as shown in Fig. 5e. These scatter areas are also located at different positions of the inherent current blocking by SV28 nanopore (Fig. S16). The SV28 nanopore can discriminate the human telomeric DNA at the single-molecule level.

Discussion

We designed a pore-forming peptide with a β-sheet structure that forms 5 differently sized nanopores ranging in diameter from 1.6 to 6.2 nm in the lipid membrane. Using the SV28 assembled nanopore with a 5 nm diameter allowed us to detect a dsDNA with 1 kbp length. A particularly pleasing result was that we achieved success with our initially designed sequence of SV28; we did not need to adjust this initial sequence via trial-and-error experiments, unlike many proposed

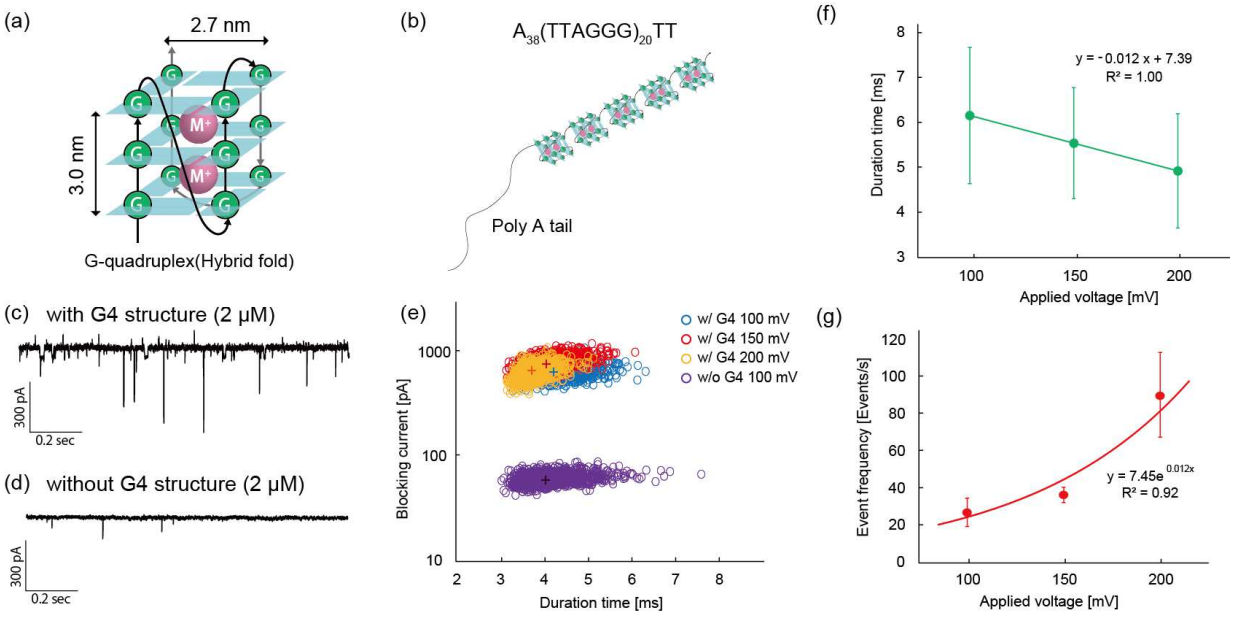

Fig. 5 Translocation of the DNA with G4 structure through the SV28 nanopore with diameter over 6.2 nm. (a, b) Schematic structure of G4 used in this study (a) and five G4s in the hybrid hold in a series (b). (c, d) The current and time traces of SV28 with G4 at 2 μM under 150 mV (c) and without G4 (d) structure at 2 μM under 100 mV. (e) Scatter plots of the blocking current and duration time after bootstrapping of DNA with G4 and without G4 structure in several voltage applications. The peak top of each conditions with G4 are: (duration, blocking current = 4.24 ms, 669 pA) at 100 mV, (4.07 ms, 797 pA) at 150 mV, (3.72 ms, 678 pA) at 200 mV, and without G4: (4.05 ms, 62 pA) at 100 mV. (f) The duration time of G4 (2 μM) as a function of the applying voltages at 100, 150, and 200 mV. (g) The event frequency of the translocation of G4 (2 μM) as a function of the concentration of DNA with G4 structure. The lines in (f, g) show the results of fitting by liner or exponential models.

protein structures that are first designed and expressed using computational methods. Through conducting this study, we found that the reduction of information hierarchy is the key to our design: in other words, the key is reduction of the sequence space. Our nanopore structure was designed with several limitations and requirements: a β-sheet structure, amphiphilic properties, and strong interaction between monomers. Moreover, the length of our peptide was decided as less than 30 amino acids so as to be suitable for chemical synthesis. These requirements limited the sequence and variety of amino acids incorporated into the transmembrane nanopore design. To that end, our designed peptide is an artificial device that can sense long dsDNA and human telomeric structure. These results provide insights into the application of the peptide nanopore as the potential technical hurdles DNA secondary structures may present to nanopore technology. Besides, our design strategy will also be a useful tool for creating molecular machines with utility as part of a molecular robot.[51]

Materials and Methods

MD simulation

The MD simulations of β-barrel structures of SV28 consisting of five and eleven peptides were performed in a DOPC membrane using GROMACS-5.1.4[65] and Charmm36 protein force field[65]. The structural modeling and simulation details are included in the Supporting Information. Briefly, the initial 3D structures of the SV28 peptide were modeled *via* a homology modeling technique, which allows prediction of the structure of proteins based on the sequence similarity.[66] HASR protein (PDB ID: 3CSL; chain A; sequence positions, 484-511)[67] was selected as a template structure for SV28 from the Protein Data Bank using BLAST search.[68] The sequence covering and charged amino acid positions were applied as the main criteria for the template selection (Fig. S17a). After modeling the 3D structure of the β-barrel SV28 containing 5-mer and 11-mer (Fig. S17b), the simulation systems were prepared by CHARMM-GUI membrane builder.[69] Then, the standard minimization and equilibration procedures of the builder[69,70] were completed, and an additional 100 ns equilibration carried out for each system (Fig. S17c). Finally, 900 ns MD simulations of 5-mer and 11-mer SV28 in DOPC membrane were performed under NPT conditions. Each system was simulated for a total of 1 μs. The final simulations of the systems were analyzed and discussed. The structures were represented by VMD software.[71] Analyses were performed using GROMACS packages, excluding the pore diameter, which was analyzed using HOLE software.[72]

Preparation of bilayer lipid membrane and SV28 pretreatments

Bilayer lipid membranes (BLMs) were prepared by the droplet contact method using a microdevice.[49] First, the DOPC (lipids/*n*-decane, 10 mg/mL) solution (2.3 μL) was poured into each chamber. Next, the buffer solution (4.7 μL) without any peptide was poured into the recording chamber. The buffer solution (4.7 μL) with peptide (final concentration 1 μM) was poured into the ground chamber. In this study, a buffer solution (1 M KCl, 10 mM MOPS, pH 7.0) was used. Before the measurement, the peptides were added to 100 mM KOH and incubated for 5 minutes in order to allow transfer of isoacyl dipeptide to the native dipeptide of **Val** and **Ser**. Then, HCl was added to make the buffer pH 7. A few minutes after adding the buffer solution, the two lipid monolayers connected to form BLMs. When the BLMs ruptured, they were reconstituted as BLMs by tracing with a hydrophobic stick between two droplets. The solutions were prepared comprising of 2 μM transformed SV28, 1 M KCl, 10 mM MOPS, and 10 mg/mL DOPC or DOPC: cholesterol = 4 : 1

(w/w) in *n*-decane at pH 7. The solution was agitated in a vortex for 30 s and incubated for 24 h at 37°C. The lipid and buffer solutions were added to the ground chamber.

Channel current measurements and data analysis

Channel current was monitored using a JET patch-clamp amplifier (Tecella, CA, USA) connected to the chambers. Ag/AgCl electrodes were already present in droplets when the solution was added to the chambers. A constant voltage ranging from 40 to 200 mV was applied to the recording chamber, and the other chamber was grounded. Pore formation in BLMs allowed ions to pass through the nanopore under the voltage gradient, giving the channel current signals. The signals were detected using a 4 kHz low-pass filter at a sampling frequency of 20 kHz. Analysis of channel current signals and duration time was performed using pCLAMP ver. 11.0.3 (Molecular Devices, CA, USA), Excel (Microsoft, Washington, USA) software, and in-house programs coded by a python. Channel current measurements were conducted at 22 ± 2 °C. The scatter plots of the blocking current and duration of the DNA translocation are described after the bootstrapping. The bootstrap method is based on the resampling of the original random sample drawn from a population with an unknown distribution. We used the exact bootstrap method, which availed the entire space of resamples. In the exact bootstrap method, the verification of accuracy will be made possible when the sample number is over 30. In this study, our bootstrap procedure took same amount of data randomly from the translocation data and the mean values for these samples were calculated with 65536 times.

Supplementary Materials

Acknowledgments

General: We thank Dr. Kazuyoshi Ueda at Yokohama National University for the technical comments of the calculations. The calculations were performed by the clusters of supercomputers of the Research Center for Computational Science, Okazaki, Japan. **Funding:** This work is partially supported by the KAKENHI (grant nos. 19H05382 and 19H00901) from MEXT. B.M. is thankful for the support of the Mongolian-Japan Engineering Education Development Program (J11B16). **Author contributions:** K.S. and R.K. conceived the original idea and K.S. designed SV28. K.S., S.Y., S.A, Y.H., and K.U. synthesized the peptides and performed the CD measurements. B.M. performed molecular dynamics simulations of SV28. A.O. and I.K. performed solid-state NMR measurements of SV28. K.S. and U.M. perform the channel current measurements. K.S. and R.K. wrote the entire manuscript and U.M., Y.H., B.M., K.U. and I.K. wrote each experimental details and results.

Corresponding author: rjkawano@cc.tuat.ac.jp

References and Notes

1 Anfinsen, C. B. Principles That Govern Folding Of Protein Chains. *Science* **181**, 223-230, (1973).

2 Gutte, B. Synthetic 70-Amino Acid Residue Analog of Ribonuclease S-Protein With Enzymic Activity. *J. Biol. Chem.* **250**, 889-904, (1975).

3 Baker, D. What has de novo protein design taught us about protein folding and biophysics? *Protein Sci.* **28**, 678-683, (2019).

4 Kortemme, T. & Baker, D. Computational design of protein-protein interactions. *Curr. Opin. Chem. Biol.* **8**, 91-97, (2004).

5 Korendovych, I. V. & DeGrado, W. F. De novo protein design, a retrospective. *Q. Rev. Biophys.* **53**, (2020).

6 Huang, P. S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320-327, (2016).

7 Dahiyat, B. I. & Mayo, S. L. De novo protein design: Fully automated sequence selection. *Science* **278**, 82-87, (1997).

8 Bolon, D. N., Voigt, C. A. & Mayo, S. L. De novo design of biocatalysts. *Curr. Opin. Chem. Biol.* **6**, 125-129, (2002).

9 Beesley, J. L. & Woolfson, D. N. The de novo design of alpha-helical peptides for supramolecular self-assembly. *Curr. Opin. Biotechnol.* **58**, 175-182, (2019).

10 Baltzer, L., Nilsson, H. & Nilsson, J. De novo design of proteins - What are the rules? *Chem. Rev.* **101**, 3153-3163, (2001).

11 Smith, B. A. & Hecht, M. H. Novel proteins: from fold to function. *Curr. Opin. Chem. Biol.* **15**, 421-426, (2011).

12 Kaiser, E. T. Design And Construction Of Biologically-Active Peptides And Proteins, Including Enzymes. *Bio. Chem. Hoppe-Seyler* **369**, 204-204, (1988).

13 Osterman, D. G. & Kaiser, E. T. Design And Characterization Of Peptides With Amphiphilic Beta-Strand Structures. *J. Cell. Biochem.* **29**, 57-72, (1985).

14 Mutter, M. & Vuilleumier, S. A Chemical Approach To Protein Design - Template-Assembled Synthetic Proteins (Tasp). *Angew. Chem.-Int. Edit.* **28**, 535-554, (1989).

15 Dou, J. Y., Vorobieva, A. A., Sheffler, W., Doyle, L. A., Park, H., Bick, M. J., Mao, B. C., Foight, G. W., Lee, M. Y., Gagnon, L. A., Carter, L., Sankaran, B., Ovchinnikov, S., Marcos, E., Huang, P. S., Vaughan, J. C., Stoddard, B. L. & Baker, D. De novo design of a fluorescence-activating beta-barrel. *Nature* **561**, 485-+, (2018).
16 Lu, P. L., Min, D. Y., DiMaio, F., Wei, K. Y., Vahey, M. D., Boyken, S. E., Chen, Z. B., Fallas, J. A., Ueda, G., Sheffler, W., Mulligan, V. K., Xu, W. Q., Bowie, J. U. & Baker, D. Accurate computational design of multipass transmembrane proteins. *Science* **359**, 1042-1046, (2018).
17 Ying, Y. L. & Long, Y. T. Nanopore-Based Single-Biomolecule Interfaces: From Information to Knowledge. *J. Am. Chem. Soc.* **141**, 15720-15729, (2019).
18 van Dijk, E. L., Jaszczyszyn, Y., Naquin, D. & Thermes, C. The Third Revolution in Sequencing Technology. *Trends Genet.* **34**, 666-681, (2018).
19 Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A. & Waterston, R. H. DNA sequencing at 40: past, present and future. *Nature* **550**, 345-353, (2017).
20 Ayub, M. & Bayley, H. Engineered transmembrane pores. *Curr. Opin. Chem. Biol.* **34**, 117-126, (2016).
21 Varongchayakul, N., Song, J. X., Meller, A. & Grinstaff, M. W. Single-molecule protein sensing in a nanopore: a tutorial. *Chem. Soc. Rev.* **47**, (2018).
22 Shi, W. Q., Friedman, A. K. & Baker, L. A. Nanopore Sensing. *Anal. Chem.* **89**, 157-188, (2017).
23 Misawa, N., Osaki, T. & Takeuchi, S. Membrane protein-based biosensors. *Journal of the Royal Society Interface* **15**, (2018).
24 Branton, D., Deamer, D. W., Marziali, A., Bayley, H., Benner, S. A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S. B., Krstic, P. S., Lindsay, S., Ling, X. S., Mastrangelo, C. H., Meller, A., Oliver, J. S., Pershin, Y. V., Ramsey, J. M., Riehn, R., Soni, G. V., Tabard-Cossa, V., Wanunu, M., Wiggin, M. & Schloss, J. A. The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **26**, 1146-1153, (2008).
25 Gu, L. Q., Braha, O., Conlan, S., Cheley, S. & Bayley, H. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature* **398**, 686-690, (1999).
26 Kawano, R., Osaki, T., Sasaki, H., Takinoue, M., Yoshizawa, S. & Takeuchi, S. Rapid Detection of a Cocaine-Binding Aptamer Using Biological Nanopores on a Chip. *J. Am. Chem. Soc.* **133**, 8474-8477, (2011).
27 Robertson, J. W., Rodrigues, C. G., Stanford, V. M., Rubinson, K. A., Krasilnikov, O. V. & Kasianowicz, J. J. Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8207-8211, (2007).
28 Hiratani, M., Ohara, M. & Kawano, R. Amplification and Quantification of an Antisense Oligonucleotide from Target microRNA Using Programmable DNA and a Biological Nanopore. *Anal. Chem.* **89**, 2312-2317, (2017).
29 Ohara, M., Takinoue, M. & Kawano, R. Nanopore Logic Operation with DNA to RNA Transcription in a Droplet System. *ACS Synth. Biol.* **6**, 1427-1432, (2017).
30 Zhang, H. L., Hiratani, M., Nagaoka, K. & Kawano, R. MicroRNA detection at femtomolar concentrations with isothermal amplification and a biological nanopore. *Nanoscale* **9**, 16124-16127, (2017).
31 Hiratani, M. & Kawano, R. DNA Logic Operation with Nanopore Decoding To Recognize MicroRNA Patterns in Small Cell Lung Cancer. *Anal. Chem.* **90**, 8531-8537, (2018).
32 Kawano, R. Nanopore Decoding of Oligonucleotides in DNA Computing. *Biotechnol. J.* **13**, 1800091, (2018).

33 Liu, P. & Kawano, R. Recognition of Single-Point Mutation Using a Biological Nanopore. *Small Methods*, (2020).

34 Sutherland, T. C., Long, Y. T., Stefureac, R. I., Bediako-Amoa, I., Kraatz, H. B. & Lee, J. S. Structure of peptides investigated by nanopore analysis. *Nano Lett.* **4**, 1273-1277, (2004).

35 Movileanu, L., Schmittschmitt, J. P., Scholtz, J. M. & Bayley, H. Interactions of peptides with a protein pore. *Biophys. J.* **89**, 1030-1045, (2005).

36 Restrepo-Perez, L., Joo, C. & Dekker, C. Paving the way to single-molecule protein sequencing. *Nat. Nanotechnol.* **13**, 786-796, (2018).

37 Watanabe, H., Gubbiotti, A., Chinappi, M., Takai, N., Tanaka, K., Tsumoto, K. & Kawano, R. Analysis of Pore Formation and Protein Translocation Using Large Biological Nanopores. *Anal. Chem.* **89**, 11269-11277, (2017).

38 Mahendran, K. R., Niitsu, A., Kong, L. B., Thomson, A. R., Sessions, R. B., Woolfson, D. N. & Bayley, H. A monodisperse transmembrane alpha-helical peptide barrel. *Nat. Chem.* **9**, 411-419, (2017).

39 Krishnan, R. S., Satheesan, R., Puthumadathil, N., Kumar, K. S., Jayasree, P. & Mahendran, K. R. Autonomously Assembled Synthetic Transmembrane Peptide Pore. *J. Am. Chem. Soc.* **141**, 2949-2959, (2019).

40 Sohma, Y., Sasaki, M., Hayashi, Y., Kimura, T. & Kiso, Y. Novel and efficient synthesis of difficult sequence-containing peptides through O-N intramolecular acyl migration reaction of O-acyl isopeptides. *Chem. Commun.*, 124-125, (2004).

41 Wimley, W. C. The versatile beta-barrel membrane protein. *Curr. Opin. Struct. Biol.* **13**, 404-411, (2003).

42 Chou, K. C. Prediction of beta-turns. *J. Pept. Res.* **49**, 120-144, (1997).

43 Mandel-Gutfreund, Y. & Gregoret, L. M. On the significance of alternating patterns of polar and non-polar residues in beta-strands. *J. Mol. Biol.* **323**, 453-461, (2002).

44 Cheley, S., Gu, L. Q. & Bayley, H. Stochastic sensing of nanomolar inositol 1,4,5-trisphosphate with an engineered pore. *Chem. Biol.* **9**, 829-838, (2002).

45 Killian, J. A. & von Heijne, G. How proteins adapt to a membrane-water interface. *Trends Biochem. Sci.* **25**, 429-434, (2000).

46 Hong, H. D., Park, S., Jimenez, R. H. F., Rinehart, D. & Tamm, L. K. Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. *J. Am. Chem. Soc.* **129**, 8320-8327, (2007).

47 Cao, B. Q., Porollo, A., Adamczak, R., Jarrell, M. & Meller, J. Enhanced recognition of protein transmembrane domains with prediction-based structural profiles. *Bioinformatics* **22**, 303-309, (2006).

48 Wang, Y. J. & Jardetzky, O. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci.* **11**, 852-861, (2002).

49 Kawano, R., Tsuji, Y., Sato, K., Osaki, T., Kamiya, K., Hirano, M., Ide, T., Miki, N. & Takeuchi, S. Automated Parallel Recordings of Topologically Identified Single Ion Channels. *Sci. Rep.* **3**, (2013).

50 Kawano, R., Horike, N., Hijikata, Y., Kondo, M., Carne-Sanchez, A., Larpent, P., Ikemura, S., Osaki, T., Kamiya, K., Kitagawa, S., Takeuchi, S. & Furukawa, S. Metal-Organic Cuboctahedra for Synthetic Ion Channels with Multiple Conductance States. *Chem* **2**, 393-403, (2017).

51 Kawano, R. Synthetic Ion Channels and DNA Logic Gates as Components of Molecular Robots. *ChemPhysChem* **19**, 359-366, (2018).

52 Sekiya, Y., Shimizu, K., Kitahashi, Y., Ohyama, A., Kawamura, I. & Kawano, R. Electrophysiological Analysis of Membrane Disruption by Bombinin and Its Isomer Using the Lipid Bilayer System. *ACS Applied Bio Materials* **2**, 1542-1548, (2019).

53 Saigo, N., Izumi, K. & Kawano, R. Electrophysiological Analysis of Antimicrobial Peptides in Diverse Species. *Acs Omega* **4**, 13124-13130, (2019).
54 Sekiya, Y., Sakashita, S., Shimizu, K., Usui, K. & Kawano, R. Channel current analysis estimates the pore-formation and the penetration of transmembrane peptides. *Analyst* **143**, 3540-3543 (2018).
55 Serra-Batiste, M., Ninot-Pedrosa, M., Bayoumi, M., Gairi, M., Maglia, G. & Carulla, N. Aβ42 assembles into specific beta-barrel pore-forming oligomers in membrane-mimicking environments. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 10866-10871, (2016).
56 Brogden, K. A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.* **3**, 238-250, (2005).
57 Mani, R., Cady, S. D., Tang, M., Waring, A. J., Lehrert, R. I. & Hong, M. Membrane-dependent oligomeric structure and pore formation of beta-hairpin antimicrobial peptide in lipid bilayers from solid-state NMR. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16242-16247, (2006).
58 Hille, B. *Ion channels of excitable membranes*. 3rd edn, (Sinauer, 2001).
59 Butler, T. Z., Pavlenok, M., Derrington, I. M., Niederweis, M. & Gundlach, J. H. Single-molecule DNA detection with an engineered MspA protein nanopore. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 20647-20652, (2008).
60 Shim, J., Banerjee, S., Qiu, H., Smithe, K. K. H., Estrada, D., Bello, J., Pop, E., Schulten, K. & Bashir, R. Detection of methylation on dsDNA using nanopores in a MoS_2 membrane. *Nanoscale* **9**, 14836-14845, (2017).
61 Henrickson, S. E., Misakian, M., Robertson, B. & Kasianowicz, J. J. Driven DNA transport into an asymmetric nanometer-scale pore. *Phys. Rev. Lett.* **85**, 3057-3060, (2000).
62 Huang, G., Voet, A. & Maglia, G. FraC nanopores with adjustable diameter identify the mass of opposite-charge peptides with 44 dalton resolution. *Nat. Commun.* **10**, 10, (2019).
63 An, N., Fleming, A. M., Middleton, E. G. & Burrows, C. J. Single-molecule investigation of G-quadruplex folds of the human telomere sequence in a protein nanocavity. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14325-14331, (2014).
64 An, N., Fleming, A. M., White, H. S. & Burrows, C. J. Nanopore detection of 8-oxoguanine in the human telomere repeat sequence. *ACS Nano* **9**, 4296-4307, (2015).
65 Van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. & Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701-1718, (2005).
66 Bradley, P., Misura, K. M. S. & Baker, D. Toward high-resolution de novo structure prediction for small proteins. *Science* **309**, 1868-1871, (2005).
67 Krieg, S., Huche, F., Diederichs, K., Izadi-Pruneyre, N., Lecroisey, A., Wandersman, C., Delepelaire, P. & Welte, W. Heme uptake across the outer membrane as revealed by crystal structures of the receptor-hemophore complex. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1045-1050, (2009).
68 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W. & Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402, (1997).
69 Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* **97**, 50-58, (2009).
70 Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B. & Lindahl, E. Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *J. Chem. Theory Comput.* **6**, 459-466, (2010).
71 Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics Modell.* **14**, 33-38, (1996).

72 Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graphics Modell.* **14**, 354-&, (1996).

Figures

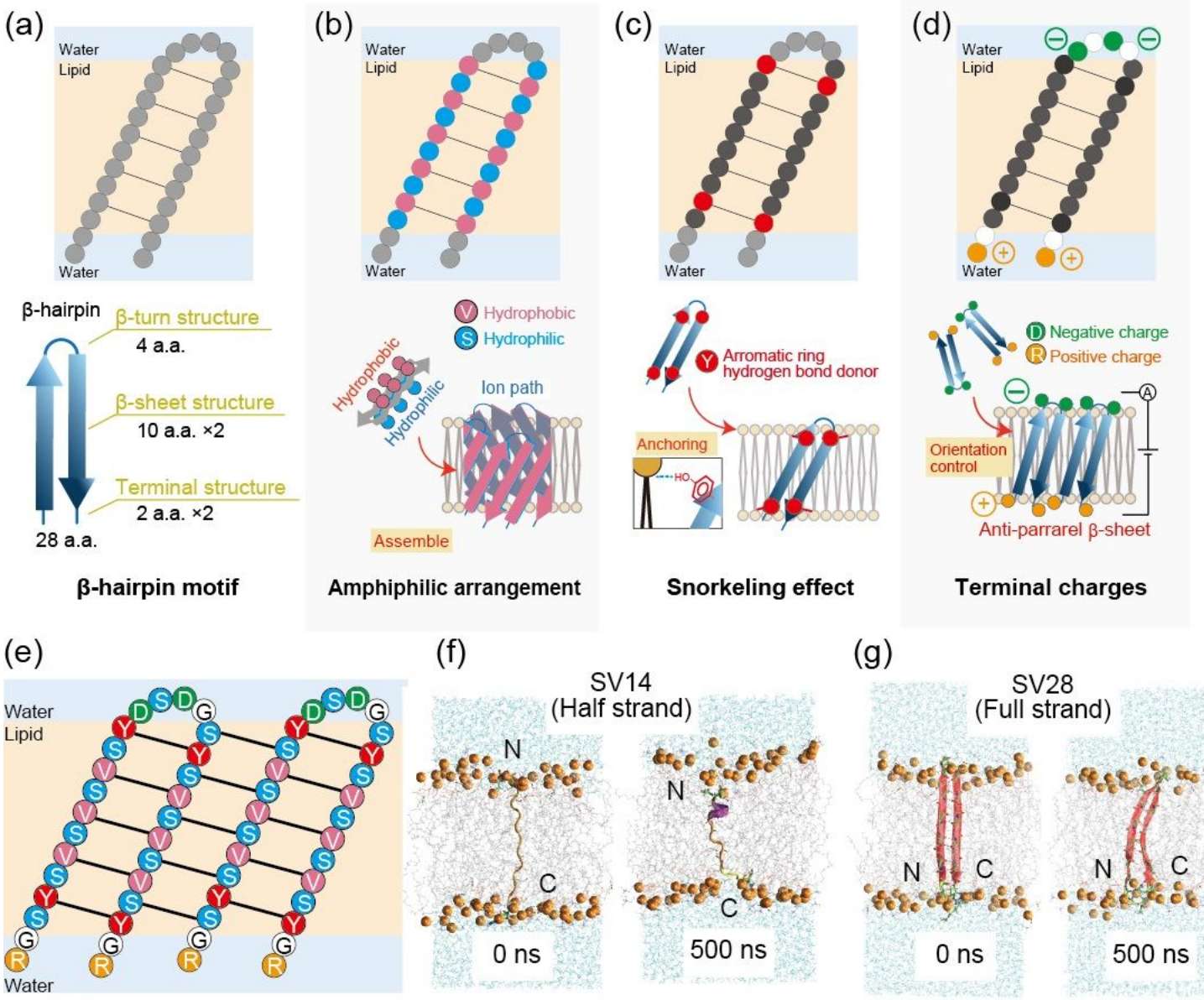

Figure 1

Design strategies and structural confirmation of the β-hairpin peptide. (a) The design of β-hairpin peptide with 28 a.a. divided into three sections: β-turn, β-sheet transmembrane, and the terminals. (b) Hydrophilic and hydrophobic amino acids are arranged in an alternating fashion. (c) Interaction of aromatic rings stabilizes the β-barrel pore. (d) Designing specific charges at the terminus allows control of peptide orientation upon an applied voltage. (e) Amino acid sequence of designed structure, named SV28. The black lines indicate the hydrogen bonding. (f, g) MD simulations of the monomer structures. The 0 ns (left) and 500 ns (right) snapshots of the (f) half-length and (g) full length SV28 in a lipid bilayer membrane.

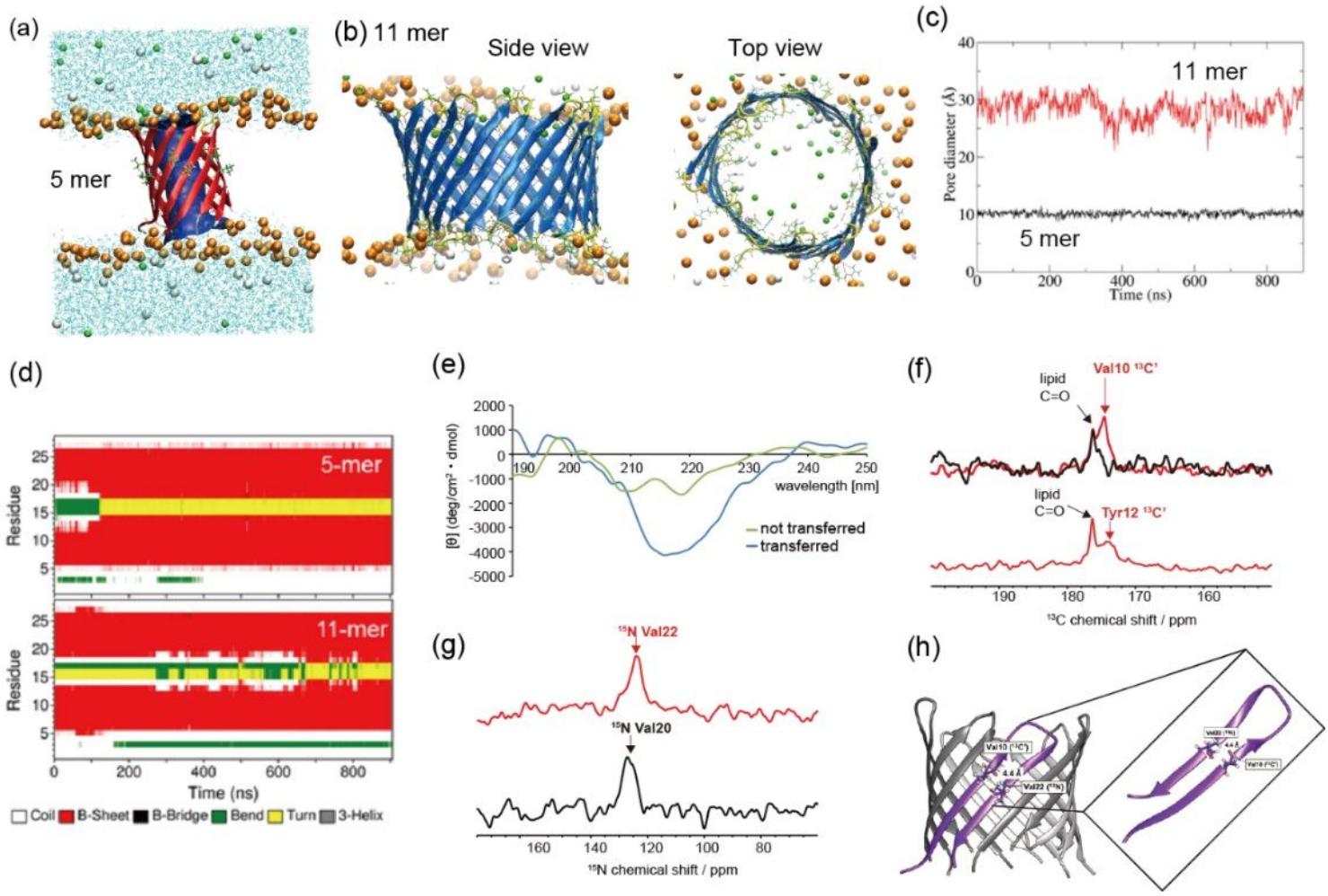

Figure 2

(a) 5-mer structures of SV28 in a DOPC membrane in the MD simulation. Ribbons show the peptide structures, with the secondary structure indicated by the color of the ribbon (red: β-sheet, cyan: turn, white: random coil structure). Ribbon arrows indicate the direction of the backbone from N-terminal to Cterminal. The pore structures were analyzed by HOLE software and displayed as blue surfaces inside of the barrels. Val10 and Val22 amino acids showing central rim of the pores were displayed as the licorice models. Cyan lines indicate water molecules, and the lipid molecules were omitted for clarity (excluding phosphorus atoms as orange spheres). Green and white spheres indicate the potassium and chloride ions respectively. Structures were displayed by VMD software. (b) Molecular dynamics simulation of the SV28 nanopore formation in DOPC lipid membrane. The 11-mer-nanopore was simulated for 900 ns. Each color indicates: light blue: water, brown: lipid head, green: potassium ions, grey: chloride ion. (c) Central diameters of 5-mer (black) and 11mer (red) pores as a function of time were calculated using HOLE software at 1 ns interval. (d) Profile of the secondary structure of 5-mer and 11-mer pores as a function of time. (e) Circular dichroism spectra of non-transformed SV28 (green line) and transformed SV28 (blue line). (f) 13C and 15N CP-MAS NMR spectra of the triply isotope-labeled SV28 ([1-13C]Val10, [2-13C]Gly16, [15N]Val22-labeled SV28) in DOPC liposomes. The black lines indicate the spectra from DOPC liposomes, and the red lines indicate the spectra of isotope-labeled SV28 with DOPC

liposomes. (g) 13C and 15N CP-MAS NMR spectra of the triply isotope-labeled SV28 ([1-13C]Tyr12, [2-13C]Gly16, [15N]Val20-labeled SV28). (h) The β-barrel structure of SV28 nanopore. The interatomic distance between the backbone amide of Val22 and the carbonyl carbon of Val10 in the SV28 is estimated to be around 4.4 Å as measured by solid-state NMR.

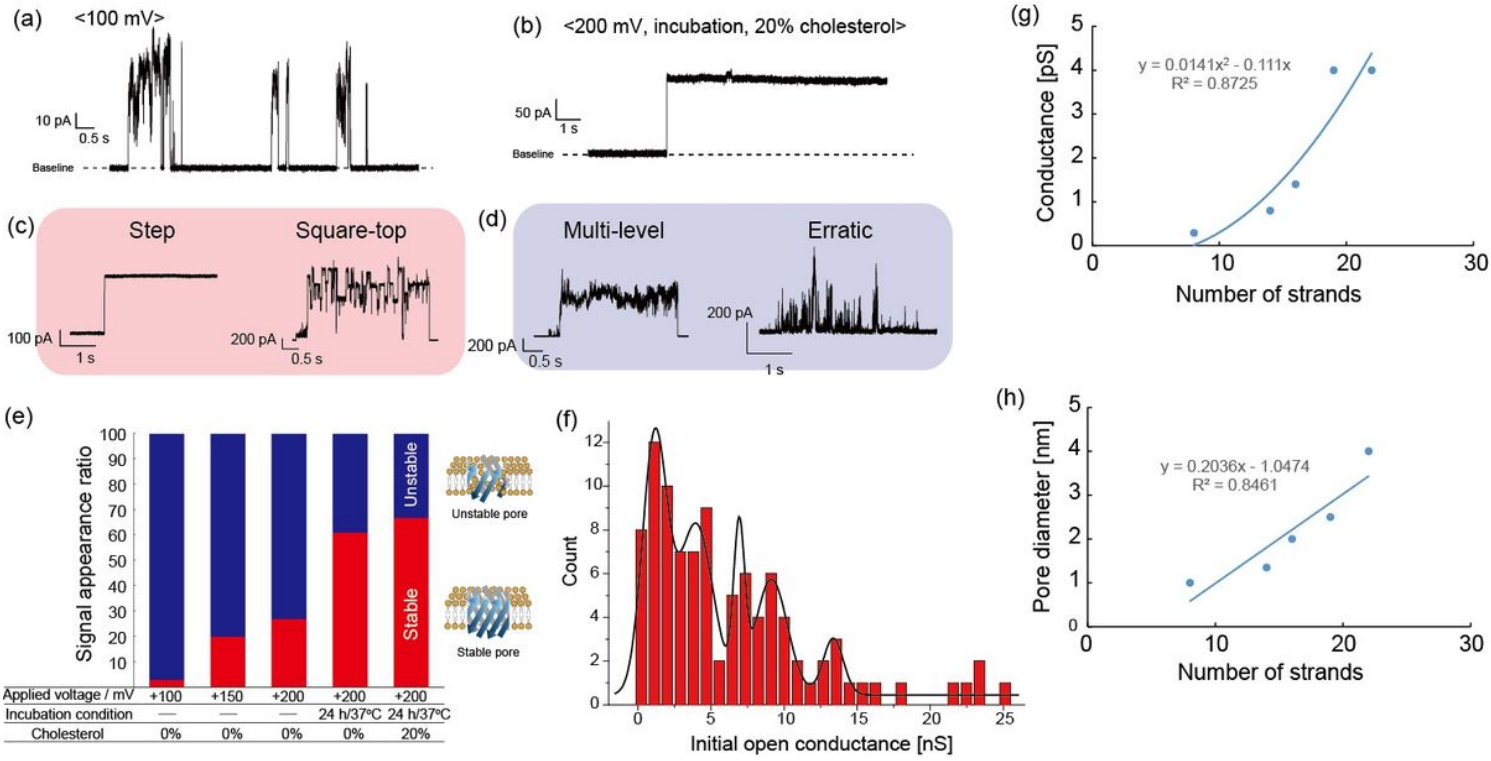

Figure 3

(a) The typical current and time traces of SV28 at +100 mV with initial condition. (b) The step signal was occasionally observed under the initial conditions. The traces were atypical before optimization of the conditions. (c) Step and square-top signals were defined as stable pore formation. (d) Multi-level and erratic signals were defined as unstable pore formation. (e) The ratio of stable and unstable pore formation from the channel current analysis. (f) The histogram of the current conductance of the initial step from the 0 A increase of SV28 in the DOPC lipid bilayer. (g, h) The relationships in β-barrel membrane proteins OmpA, OmpG, OmpF, VDAC, and FhuA between (g) the channel conductance and the number of β-strands and (h) the pore diameter and number of β-strands.

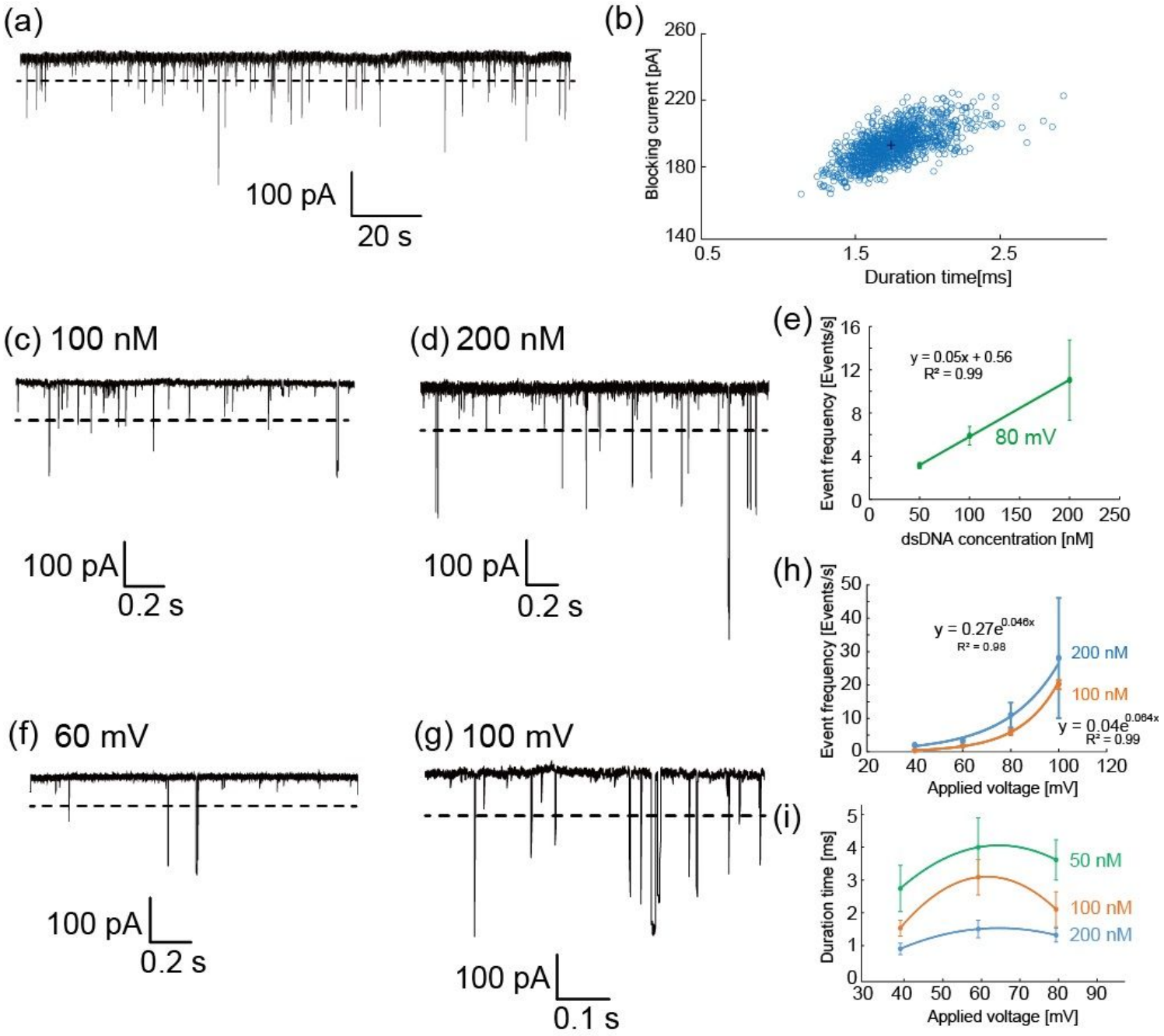

Figure 4

The dsDNA (1 kbp) translocation through the SV28 nanopore with diameter of around 5 nm. (a) The current and time trace of the SV28 nanopore with 1 kbp dsDNA (100 nM) under application of 40 mV. The dashed lines indicate the threshold for the dsDNA translocation events as we defined. (b) The scatter plot of the translocation data of 100 nM 100 mV: the blocking rate and the duration after the bootstrapping. (c, d) The current and time traces of SV28 with dsDNA at (c) 100 nM and (d) 200 nM under 100 mV. The dashed lines indicate the threshold for the dsDNA translocation events as we defined. (e) The event frequency of the translocation as a function of the concentration of dsDNA. (f, g) The current and time traces of SV28 with dsDNA (100 nM) under (f) 60 mV and (g) 100 mV. The dashed lines indicate the threshold for the dsDNA translocation events as we defined. (h) The event frequency of the

dsDNA translocation as a function of the applied voltage. Blue and orange lines indicate the dsDNA concentration at 100 nM and 200 nM to guide the eye. (i) Duration time of dsDNA with 50, 100, and 200 nM dependence on the applying voltages. Blue, orange, and green lines indicate each concentration to guide the eye.

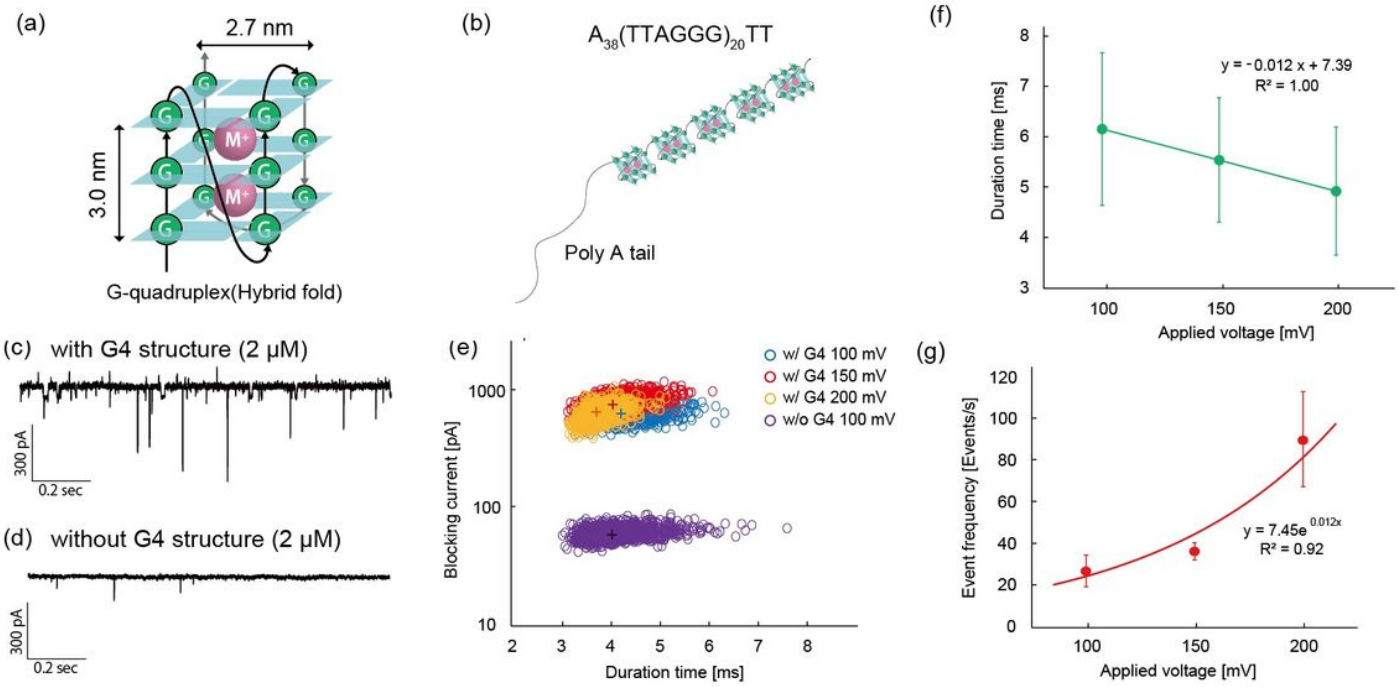

Figure 5

Translocation of the DNA with G4 structure through the SV28 nanopore with diameter over 6.2 nm. (a, b) Schematic structure of G4 used in this study (a) and five G4s in the hybrid hold in a series (b). (c, d) The current and time traces of SV28 with G4 at 2 μM under 150 mV (c) and without G4 (d) structure at 2 μM under 100 mV. (e) Scatter plots of the blocking current and duration time after bootstrapping of DNA with G4 and without G4 structure in several voltage applications. The peak top of each conditions with G4 are: (duration, blocking current = 4.24 ms, 669 pA) at 100 mV, (4.07 ms, 797 pA) at 150 mV, (3.72 ms, 678 pA) at 200 mV, and without G4: (4.05 ms, 62 pA) at 100 mV. (f) The duration time of G4 (2 μM) as a function of the applying voltages at 100, 150, and 200 mV. (g) The event frequency of the translocation of G4 (2 μM) as a function of the concentration of DNA with G4 structure. The lines in (f, g) show the results of fitting by liner or exponential models.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 20SIShimizuSV2809.docx