(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 October 2014 (02.10.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/153625 A1

(51) International Patent Classification:
C07K 14/255 (2006.01)

(21) International Application Number:
PCT/BE2014/000013

(22) International Filing Date:
25 March 2014 (25.03.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/805,068 25 March 2013 (25.03.2013) US
1313477.0 29 July 2013 (29.07.2013) GB

(71) Applicant: **KATHOLIEKE UNIVERSITEIT LEUVEN** [BE/BE]; KU Leuven Research & Development, Waaistraat 6 - bux 5105, B-3000 Leuven (BE).

(72) Inventors: **MAGLIA, Giovanni**; E.Solvaysstraat 20, B-3010 Kessel-Lo (BE). **SOSKINE, Mikhael**; Celestijnenlaan 61/11B, B-3001 Heverlee (BE).

(74) Common Representative: **KATHOLIEKE UNIVERSITEIT LEUVEN**; KU Leuven Research & Development, Waaistraat 6 - bux 5105, B-3000 Leuven (BE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2014/153625 A1

(54) Title: NANOPORE BIOSENSORS FOR DETECTION OF PROTEINS AND NUCLEIC ACIDS

(57) Abstract: Described herein are nanopore biosensors based on a modified cytolysin protein. The nanopore biosensors accommodate macromoiecules including proteins and nucleic acids, and may additionally comprise ligands with selective binding properties.

NANOPORE BIOSENSORS FOR DETECTION OF PROTEINS AND NUCLEIC ACIDS

FIELD OF INVENTION

The present disclosure relates to nanopore biosensors based on a modified cytolysin protein. The nanopore biosensors accommodate large molecules including folded proteins and nucleic acids including double stranded (ds) DNA, and showed augmented activity, solubility and electrical properties, as compared with other nanopore biosensors.

BACKGROUND

The transport of ions or molecules across a biological membrane is a fundamental process in cellular life and is tightly regulated by ion channels, transporters and pores. Recently, researchers have adopted biological,[1] solid-state,[2] DNA origami[3] and hybrid[3a, b, 4] nanopores in single-molecule analysis.[5] Biological nanopores have advantages compared to their synthetic counterparts, mostly because they can be reproducibly fabricated and modified with an atomic level precision that cannot yet be matched by artificial nanopores. Biological nanopores, however, also have drawbacks. The mechanical stability of biological nanopores depends on individual cases. Alpha hemolysin from *Staphylococcus aureus* (αHL) and porin A from *Mycobacterium smegmatis* (MspA) nanopores remain open in lipid bilayers at high-applied potentials and can cope surprisingly well with extreme conditions of temperature,[6] pH[6b, 7] and denaturant concentrations.[6b, 8] However, most of other porines and channels are much less robust. Arguably, however, the biggest disadvantage of biological nanopores is their fixed size. For example, the dimensions of αHL, MspA or aerolysin nanopores allowed the analysis of single stranded nucleic acids, aptamers or small peptides,[9] but their small internal diameter (~1 nm) precludes the direct investigations of other important biological systems such as folded enzymes or ribozymes.

Recently a significant number of studies sampled the translocation of folded proteins through artificial nanopores.[10] However, the investigation of proteins with solid-state nanopores is difficult because proteins have a non-uniform charge distribution, they often adsorb to the nanopores surface and they translocate too quickly to be properly sampled.[10c] Further, because proteins have compact folded structure, the diameter of the nanopore should be similar to that of the protein.[10b] Recently, we have introduced Cytolysin A from *Salmonella typhi* (ClyA) as the first biological nanopore that allows the investigation of natively folded proteins.[7a] The ClyA structure is ideal for this task because proteins such as thrombin (37 kDa) or malate dehydrogenase (dimer, 35 kDa monomer) can be electrophoretically trapped between the wide cis entrance (5.5 nm, table 1) and the narrower trans exit (3.3 nm, table 1), and can therefore be sampled for several minutes. Ionic currents through ClyA are so sensitive to the vestibule environment that blockades of human

and bovine thrombin can be easily distinguished.[7a] Our work was based on a ClyA construct where the two native cysteine residues of ClyA-WT (C87 and C285) were replaced by serine (ClyA-SS).[7a] However, ClyA-SS monomers showed low water solubility and low activity when compared to ClyA-WT monomers (Figure S1), and in planar lipid bilayers spontaneously opened and closed (gated) at applied potentials higher than +60 mV or lower than -90 mV.

Thus, there remains a need in the art for nanopore biosensors with high sensitivity for target analytes as well as high water solubility and stability at a range of potentials. Nanopore biosensors should have favorable properties of oligomerization, voltage dependent gating, and electrical noise in single-channel current recordings. The present disclosure relates to engineered nanopores in which specific substitutions to the native cysteine residues and other residues confer additional properties as compared with ClyA-WT and ClyA-SS.

SUMMARY

One aspect of the present disclosure relates to a modified ClyA pore comprising a plurality of subunits, wherein each subunit comprises a polypeptide represented by an amino acid sequence at least 80% identical to SEQ ID NO:1 wherein exactly one Cys residue is substituted with Ser. In some embodiments, the Cys residue is C285. In certain embodiments, each subunit of the modified ClyA pore comprises at least one additional amino acid substitution selected from L99Q, E103G, F166Y, and K294R. For example, each subunit may comprise a polypeptide represented by an amino acid sequence of SEQ ID NO:2. In certain embodiments, exactly one Cys residue in each subunit is substituted with Ala. In some embodiments, each subunit of the modified ClyA pore comprises at least one additional amino acid substitution selected from L203V and H207Y. For example, each subunit may comprise a polypeptide represented by an amino acid sequence of SEQ ID NO:3.

In some embodiments, the modified ClyA pore comprises at least 12 subunits. For example, the modified ClyA pore may comprise 12 subunits or may comprise 13 subunits. In certain embodiments, the modified ClyA pore has a cis diameter of at least 3.5 nm. In certain embodiments, the modified ClyA pore has a trans diameter of at least 6 nm. In some embodiments, the modified ClyA pore remains open when the voltage across the modified ClyA pore ranges from -60 +90 to -150mV.

In some embodiments, a protein analyte binds within the lumen of the modified ClyA pore. A protein analyte may bind to more than one site within the lumen of the modified ClyA pore. In some embodiments, the protein analyte is a protein with a molecular weight in the range of 15-70 kDa.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 shows a ribbon representation of S. typhi ClyA nanopores constructed by homology modeling from the E. coli ClyA structure (PDB: 2WCD, 90% sequence identity).[17] In Fig. 1a, one protomer is highlighted, with the secondary structure elements shaded in dark grey from N to C terminii; other protomers are shown alternating in pale grey. The side chain of the amino acids changed by directed evolution experiments are displayed as spheres. The two native cysteine residues and Phenylalanine 166 are labeled. Fig. 1b indicates the sequence changes in ClyA-SS, ClyA-CS and ClyA-AS relative to ClyA-WT.

Figure 2 shows the oligomerization and nanopore formation of ClyA-WT, ClyA-SS and evolved ClyA variants. Fig. 2a shows the oligomerisation of ClyA nanopores examined by 4-20% BN-PAGE. Proteins (1 mg/ml) were pre-incubated with 0.5% DDM for 30 min at room temperature before loading into the gel (40 µg/lane). Lane 1: Protein ladder, lane 2: ClyA-WT, lane 3: ClyA-SS, lane 4: ClyA-AS and lane 5: ClyA-CS. Fig. 1b shows the unitary nanopore conductance distribution obtained from 100 nanopores reconstituted in planar lipid bilayers for ClyA-WT (top), ClyA-CS (middle) and ClyA-AS (bottom) nanopores after pre-oligomerization in 0.5% DDM. Recordings were carried out at -35 mV in 15mM Tris.HCl, pH 7.5, 150mM NaCl and the temperature was 28°C.

Figure 3 shows the unitary conductance of 62 ClyA-CS nanopores extracted from the lowest (Fig. 3a), second lowest (Fig. 3b) and third lowest (Fig. 3c) oligomeric band of ClyA-CS separated on 4-20% acrylamide BN-PAGE. ClyA-CS monomers were pre-incubated in 0.5% DDM and loaded on a BN-PAGE as described in Figure 2. The bands that were excised are boxed and marked with an arrow on the insets. Recordings were carried out at -35 mV, 28° C in 15mM Tris.HCl pH 7.5 containing 150mM NaCl.

Figure 4 shows the current blockades provoked by HT on the three types of ClyA-CS nanopores. In Fig. 4a, from left to right, cut through Type I, Type II and Type III ClyA nanopores (grey) containing HT (black) in the nanopore vestibule. Type II and Type III ClyA nanopores were modeled as 13mer (Type II) or 14mer (Type III) as described in supplementary information. Fig. 4b shows HT blockades to Type I (left), Type II (middle) or Type III (right) ClyA-CS nanopores at -35 mV. HT current blockades to Type I and Type II ClyA-CS switched between L1 (IRES% =56±1 and 68±1, respectively) and L2 (IRES% =23±3 and 31±1, respectively) current levels. The blockades lasted for several minutes, therefore only the first second of the current traces is shown. In Type I ClyA-CS, L1 was the most represented current blockade (70%), while in Type II ClyA-CS L2 was mostly populated (96%).

HT current blockades to Type III ClyA-CS nanopores only showed L2 current levels (IRES=32±9). Recordings were carried out at -35 mV in 15mM Tris.HCl pH 7.5 150mM NaCl. The traces in Fig. 4b were collected applying a Bessel low-pass filter with 2000 Hz cutoff and sampled at 10 kHz and the temperature was 28°C.

Figure 5 shows protein translocation through Type I and Type II ClyA-CS nanopores. Fig. 5a shows the voltage dependency of HT blockade dwell times determined for Type I (hollow circles) and Type II (filled rectangles) ClyA-CS nanopores. Lifetimes at each voltage were calculated from single exponential fits to cumulative distributions (n≤3) constructed from dwell times of at least 50 blockades. The lines indicate double exponential fits to the experimental points. Fig. 5b shows the HT current blockades to Type I (left) and Type II (right) ClyA-CS nanopores at -150 mV. The blockades showed only L2 current levels for both nanopores (IRES% =23±2 and 31±5, for Type I and Type II ClyA-CS respectively). Fig. 5c shows a typical HT current blockade on Type I ClyA-CS at -150 mV showing "shoulder" and "spike" current signatures. Recordings were carried in 15 mM Tris.HCl, pH 7.5, 150mM NaCl in presence of 20 nM HT. The traces in Fig. 5b were collected applying a Bessel low-pass filter with 2000 Hz cutoff and sampled at 10 kHz. The trace in c was collected applying a Bessel low-pass filter with10 kHz cutoff and sampled at 50 kHz. The temperature was 28°C. Errors are given as standard deviations.

Figure 6 shows the characterization of purified ClyA monomers. a) Solubility of purified ClyA monomers examined by 4-20% acrylamide BN-PAGE. Equal amounts (40 µg) of purified ClyA monomers (no detergent) were supplemented with ~10% glycerol and 1x of NativePAGETM Running Buffer and 1x Cathode Buffer Additive (InvitrogenTM) and loaded in each lane: Lane 1: markers, Lane 2: ClyA-WT, Lane 3: ClyA-SS, Lane 4: ClyA-CS, Lane 5: ClyA-AS. b) Overexpression of ClyA variants. Equal amounts of bacterial pellets derived from overnight cultures overexpressing ClyA variants were resuspended to ~100 mg/mL concentration and disrupted by sonication followed by centrifugation at 20'000g for 10 min (4ºC). 20 µL of the supernatant containing the soluble fraction of ClyA proteins were loaded on lanes 2, 4, 6 and 8 of a 12% acrylamide SDS-PAGE. The lysate pellets were brought to the original volume by adding a solution containing 15mM, Tris.HCl pH 7.5, 150 mM NaCl and 2% SDS w/v. 20 µL of such solution were loaded on lanes 3, 5, 7 and 9 of the same 12 % acrylamide SDS-PAGE. Therefore, Lane 1: protein marker, Lanes 2 and 3: ClyA-SS supernatant and pellet fractions, respectively; Lanes 4 and 5: ClyA-WT supernatant and pellet fractions, respectively; Lane 6 and 7: ClyA-CS supernatant and pellet fractions, respectively; and Lane 8 and 9: ClyA-AS supernatant and pellet fractions, respectively. c) Hemolytic assays. ClyA monomers (0.6 µM) were incubated with 100 µL of 1% horse erythrocytes suspension (110 µL final volume) and the decrease

of turbidity was measured at 650 nm (OD650nm). ClyA-WT is shown as a thick grey line, ClyA-SS as a thick black line, ClyA-CS as a thin black line and ClyA-AS as a dashed line. d) The rates of hemolysis (calculated as the inverse of the time to reach 50% of turbidity) plotted against protein concentration ClyA-WT (triangle, thick grey line), ClyA-SS (squares, thick black line), ClyA-AS (circles, dashed line) and ClyA-CS (diamonds, thin black line).

Figure 7 shows an example of the screening of the oligomerization of ClyA variants using 4-20% acrylamide BN-PAGE. a) Round 4. Lane 1, 2: ClyA-CS, lane 3 and 4: 4ClyA5, lane 5 and 6: 4ClyA3, lane 7 and 8: 4ClyA1, lane 9 and 10: 4ClyA2, lane 11 and 12: 4ClyA6. Samples were prepared as explained in Figure 6a supplemented with 0.05% (even lane number) or 0.1% (odd lane numbers) SDS. SDS was used to counter the "smearing" effect of large quantity of DDM in the samples. b) Round 5. Lane 1, 2: 5ClyA2, lane 3: 5ClyA1, lane 4: ClyA-AS. Samples were supplemented with 0.05% SDS. Oligomerization was triggered by the addition of 1% DDM and ClyA variants were partly purified by Ni-NTA affinity chromatography as described in methods.

Figure 8 shows noise characteristics of three types of ClyA nanopores under -35 mV potential in 150 mM NaCl, 15 mM Tris.HCl. Current power spectral densities of the Type I (dashed line), Type II (dotted line) and Type III (light gray solid line) ClyA-CS nanopores at -35 mV obtained from 0.5 s traces. The current power spectral density at 0 mV is shown in black. Each line corresponds to the average of power spectra calculated from 3 recordings carried out on different single channels.

Figure 9 shows the duration of HT (a) and FP (b) blockades on Type I ClyA-SS pores. Traces were recorded at a sampling rate of 10 kHz with an internal low-pass Bessel filter set at 2 kHz. Each plotted value corresponds to the average determined using at least 3 different single channels. Errors are given as standard deviations.

Figure 10 shows the typical HT current blockades on Type I ClyA at -150 mV showing "shoulder" and "spike" current signatures. Recordings were carried in 15 mM Tris.HCl, pH 7.5, 150mM NaCl in presence of 20 nM HT. The traces were filtered with a Gaussian low-pass filter with 10,000 Hz cutoff filter and sampled at 50,000 Hz.

Figure 11 shows dsDNA translocation through ClyA nanopores. On the right of the current traces, ClyA (pore shaped structure) and neutravidin (dark gray) are depicted , and the dsDNA is shown as a black line. a) Sections through of ClyA from S. typhi constructed by homology modeling from the E. coli ClyA structure (PDB: 2WCD, 90% sequence identity).[22] ClyA is shown with pore measurements including the Van der Waals radii of the atoms, and the approximate location of

residue 103 (serine in the WT sequence) on the pore indicated. b) at +100 mV (level I_{O+100}=1.74±0.05 nA), the addition of 0.12 μM of 290 bp dsDNA 1 to the cis side of a ClyA nanopore provokes short current blockades of I_{RES}=0.63±0.01 (level 1^*_{+100} = 1.10±0.03 nA, n=3) that are due to the translocation of dsDNA through the pore. Addition of neutravidin to the cis chamber converted the short current blockades to higher conductive and long-lasting current blockades with I_{RES}=0.68±0.01 (level 1_{+100}=1.19±0.01) due to the partial translocation of DNA through the pore. The open pore current was restored by reversing the potential to -100 mV (red asterisk). The inset shows a typical transient current blockade. c) Details for a current blockade due to the formation of a cis pseudorotaxanes at +100 mV d) Formation of a trans pseudorotaxanes at -100 mV by threading the dsDNA:neutravidin complex (see above) from the trans side (level 1_{-100} = 1.02±0.03 nA, I_{RES}=0.62±0.01, n=4). The electrical recordings were carried out in 2.5 M NaCl, 15 mM Tris.HCl pH 7.5 at 22°C. Data were recorded by applying a 10 kHz low-pass Bessel filter and using a 20 μs (50 kHz) sampling rate.

Figure 12 shows formation of a nanopore-DNA rotaxane. a) representation of the hybridisation of the DNA molecules used to form the rotaxane. Arrowheads mark the 3′ ends of strands. b) rotaxane formation. At -100 mV following the addition of the DNA hybrid 3 (1 μM) complexed with neutravidin (0.3 μM) and oligo 4 (1 μM) to the trans compartment, the open pore current of ClyA-2 (I_{O-100} = -1.71±0.07 nA, n=4) is reduced to level 1_{-100} = 1.1±0.04 nA (I_{RES} value of 0.64±0.02, n=4), indicating that dsDNA threads the pore from the trans side. Stepping to +100 mV (red asterisk) produced a current block with I_{RES} = 0.77±0.04 (level 2_{+100} = 1.31±0.09 nA, n=4), indicating that the DNA is still occupying the pore at positive applied potentials. Level 2 most likely corresponds to ssDNA occupying the vestibule of the pore. Successive switching to positive applied potentials did not restore the open pore current (Figure 16), confirming that a rotaxane is permanently formed. c-e) Rotaxane removal. c) at +100 mV the ionic current through ClyA-2 nanopores showed a multitude of fast current blockades (Figure 15), suggesting that the ssDNA molecules attached at the cis entrance of the pore transiently occupy the lumen of the pore. d) after the rotaxane is formed (Figure 12b), at +100 mV the nanopore shows a steady ionic current (level 2_{+100}) suggesting that a single DNA molecule is occupying the pore. e) 20 minutes after the addition of 20 mM DTT to the cis compartment the DNA molecules atop the ClyA pore are removed and the open pore current at +100 mV is restored (I_{O+100} = 1.78±0.07, n=4). Recordings were performed as described in Figure 11.

Figure 13 shows the ssDNA blockades to ClyA-CS. At +100 mV, the addition of 2 μM of a biotinylated ssDNA (5a) to the cis side of ClyA-CS nanopores in the presence of 0.6 μM neutravidin

provoked transient current blocks (1.24±0.02 nA, I_{RES} = 0.69±0.04, n=3), indicating that ssDNA can enter the lumen of the pore but only transiently. The subsequent addition of the complementary ssDNA strand (5b) converted the current blockades into level 1_{+100} blocks (1.22±0.13 nA, I_{RES} = 0.67±0.01, n=3), indicating that the dsDNA can now translocate the entire length of the pore.

Figure 14 shows the transport of DNA through ClyA nanopores. a) schematic representation of the strand displacement reaction that promotes the release of DNA from the pore. b) at +50 mV and in the presence of 3 (0.3 μM) ClyA-2 showed a steady open pore current (I_{O+50}= 0.85±0.01 nA, n=3), showing that the ssDNA strands attached to the pore do not thread through the lumen of the pore and prevent the translocation of dsDNA form solution. c) the addition of the ssDNA strands 6 (40 nM) to the cis chamber produced long lasting current blockades with I_{RES} = 0.70±0.02 , (level 2_{+50} = 0.59±0.02 nA, n=5) indicating that the dsDNA hybrid is threaded the pore. d) the subsequent addition of 1 μM of 7 to the trans chamber (+50 mV), which also contains 0.3 μM neutravidin, promoted the release of the DNA thread and restored the open pore current. Subsequently, dsDNA molecules are sequentially captured and released as shown by multiple blocked and open pore currents. For the sake of clarity, neutravidin is not included in the cartoon representation. The electrical recordings were carried out in 2.5 M NaCl, 15 mM Tris.HCl pH 7.5 at 22°C. Data were recorded by applying a 10 kHz low-pass Bessel filter and using a 20 μs (50 kHz) sampling rate. The current signal in panel d was digitally filtered at 2 kHz with a post- acquisition low-pass Gaussian filter. The applied potential of this experiment was set to +50 mV to facilitate the observation of the multiple block and release (Figure 14d), as at higher applied potentials the capture of the DNA in *cis* is very fast.

Figure 15 shows the results of control experiments showing that all components of Figure 12 are necessary to form a DNA rotaxane. a) the absence of the bridging sequence 4 does not allow the linkage between ClyA-2 and 3. After the complex is captured at -100 mV it is readily expelled from the pore at +100 mV (asterisk) as shown by the typical current signature of an open pore current for ClyA-2 at +100 mV. b) the absence of 2 from the pore top (e.g. after cleavage with DTT) does not allow 3•4 DNA hybrid to bind to the pore when captured at -100 mV. Upon reversing the potential to +100 mV (red asterisk) the open pore current is restored. c) Typical current recording for a ClyA-2 nanopore at +100 mV. d) all points histogram (5 pA bin size) for 20 seconds of the current trace shown in a. Level I_{O+100} = 1.71±0.07 nA, level A = 1.62±0.12 nA, level B = 1.43±0.05 nA and level C = 1.28±0.06 nA, corresponding to the I_{RES} values of 0.94±0.04, 0.84±0.03 and 0.75±0.03, respectively. The values, calculated from 12 experiments, might represent the DNA strands lodging within the lumen of ClyA.

Figure 16 shows the Current versus voltage (IV) relationships for ClyA-2 nanopores a) a typical curve for ClyA-2 before (lightest grey line) and after (black line) rotaxane formation. The medium-grey line indicates the same nanopore after the DNA molecules attached to the nanopore are removed by the addition of DTT. The current recordings were measured by applying an automated protocol that ramped the voltage from -100mV to +100 mV in 4 seconds. b) I-V curves calculated from the average of four experiments showing the steady-state (1 s) ClyA-2 open pore current levels (white spheres) and ClyA-2 open pore current levels in a rotaxane configuration (black spheres). The unitary conductance values of the nanopores as calculated from the slopes of the I-V curves were 17.1 nS for ClyA-2 at both positive and negative bias, 10.8 nS for the rotaxane at negative bias and 13.0 nS at positive bias. The rotaxanes were prepared as described in Figure 12.

Figure 17 shows dsDNA current blockades to ClyA-CS. On the right of each current trace the cartoon represents the physical interpretation of the current recordings. a-d) current recordings for ClyA-2 after hybridisation with a 6•7 (indicated with asterik) at different applied potentials. e) at +50 mV, upon hybridisation with 6 (40 nM) and 2,the DNA duplex 3 (0.3 μM) is transported through the pore as shown by the drop in the ionic current from I_{O+50}= 0.85±0.01 nA, to a level 2_{+50} block (I_{RES}=0.70±0.02). Reversal of the applied potential to -50 mV restores the open pore current (I_{O-50} = 0.83 ± 0.00). f) the subsequent addition of 1 μM neutravidin (black) to the trans chamber locked the DNA thread within the pore as revealed after the reversal of the potential to -50 mV when a blocked pore level (I_{RES} = 0.67±0.02) was observed.

Figure 18 shows selective DNA translocation through ClyA-2 pores. a) left, at +50 mV the ionic current through ClyA-2 nanopores showed fast and shallow current blockades, suggesting that the ssDNA molecules attached at the cis entrance might transiently occupy the lumen of the pore. Middle, after dsDNA strand 1 (50 nM) is added to the cis chamber the current signals did not change, indicating that dsDNA does not translocate ClyA-2. Right, 20 minutes after the addition of 20 mM DTT to the cis compartment the DNA molecules atop the ClyA pore are removed and the DNA can translocate through the pore. b) same experiment as described in panel a but in the presence of 1 μM of neutravidin. c) same as in panel b, but at +100 mV.

Figure 19 shows Voltage dependence of the interaction of a ssDNA-dsDNA hybrid construct with ClyA-CS pores. a) depiction of the DNA molecules used in the experiments described in panels b (left) and c (right). b) current blockades of the DNA hybrid 3 (panel a, left) in complex with neutravidin at +50 mV (left), +70 mV (middle) and +100 mV (right). c) current blockades of the DNA hybrid 3:6 (panel a, right) in complex with neutravidin at +50 mV (left), +70 mV (middle) and +100 mV (right).

Figure 20 shows strand release with Boolean logic. The OR gate is represented at the top and the AND gate at the bottom. DNA is represented as directional lines, with the arrow head denoting the 3' end. The sections within the same DNA strand represent DNA domains that act as a unit in hybridization, branch migration or dissociation. Domains are represented by the letter D followed by a number. T denotes a toehold domain. A starred domain represents a domain complementary in sequence to the domain without a star. Figure 14 shows the implementation of the first reaction of the OR gate.

Figure 21 shows alternative transport of DNA through ClyA nanopores. a) schematic of the release of the DNA thread by strand displacement showing the implementation of the second reaction of the OR gate described in Figure 18. b) open pore current of ClyA-2 at +35 mV. c) addition of DNA strands 3 (1 μM) and 8 (0.5 μM) to the cis chamber produce long lasting current blockades, indicating that the dsDNA hybrid is threading the pore. d) at +35 mV the subsequent addition of 1 μM of 9 to the trans chamber, which also might contain 1 μM neutravidin, promoted the release of the DNA thread and restored the open pore current. Subsequently, other dsDNA molecules are captured and then released as shown by the cycles of blocked and open pore currents. Data were recorded by applying a 10 kHz low-pass Bessel filter and using a 20 μs (50 kHz) sampling rate.

DETAILED DESCRIPTION

The present disclosure relates to nanopore biosensors which may be used for a variety of applications, including but not limited to detection and quantification of proteins and translocation of DNA. The nanopores are based on pore-forming bacterial cytotoxins, which act as channels for large macromolecules such as proteins and nucleic acids. Nucleic acids may include DNA, for example, single stranded DNA (ssDNA) or double stranded DNA (dsDNA), RNA, peptide nucleic acids, and/or nucleic acid analogs.

Modified Nanopore Biosensors

One aspect of the present disclosure relates to nanopore biosensors made from modified pore proteins. Exemplary pore proteins include, but are not limited to cytolysins, hemolysins, porins, DNA packaging protein, and motor proteins. Pore proteins may be bacterial or viral proteins.

In certain embodiments, the modified pore protein is a pore-forming cytotoxic protein, for example, a bacterial cytolysin. The cytolysin may be a Cytolysin A (ClyA) from a gram-negative bacteria such as *Salmonella* or *Escherichia coli* (*E. coli*). In some embodiments, the modified cytolysin

is a modified ClyA from *Salmonella typhi* (*S. typhi*) or *Salmonella paratyphi* (*S. paratyphi*). In some embodiments, the modified ClyA pore comprises a plurality of subunits, wherein each subunit comprises a polypeptide represented by an amino acid sequence at least 80% identical to SEQ ID NO: 1. In certain embodiments, the subunits are represented by an amino acid sequence at least 85% identical, 90% identical, 95% identical, or 100% identical to SEQ ID NO:1. Identical may refer to amino acid identity, or may refer to structural and/or functional identity. Accordingly, one or more amino acids may be substituted, deleted, and/or added, as compared with SEQ ID NO:1. Modifications may alter the pore lumen in order to alter the size, binding properties, and/or structure of the pore. Modifications may also alter the ClyA pore outside of the lumen.

In certain embodiments, each subunit comprises a polypeptide represented by an amino acid sequence at least 80% identical to SEQ ID NO:1, wherein exactly one Cys residue is substituted with Ser. Each subunit may be represented by an amino acid sequence that is at least 85%, 90%, 95%, 96%, 96%, 98%, or 99% identical to SEQ ID NO:1, and additionally exactly one Cys residue may be substituted with Ser. The Cys residue may be Cys 87 and/or Cys 285 in SEQ ID NO:1. In some embodiments, the Cys residue is Cys285. Other amino acid residues may be substituted, for example, with amino acids that share similar properties such as structure, charge, hydrophobicity, or hydrophilicity. In certain embodiments, substituted residues are one or more of L99, E103, F166, and K294. For example, the substituted residues may be one or more of L99Q, E103G, F166Y, and K294R. An exemplary subunit may comprise substitutions L99Q, E103G, F166Y, K294R, and C285S. Thus, each subunit may comprise a polypeptide represented by an amino acid sequence of SEQ ID NO:2. An exemplary modified ClyA pore comprising subunits in which exactly one Cys residue is substituted with Ser may be called ClyA-CS.

The modified ClyA pore may comprise a plurality of subunits, wherein each subunit comprises a polypeptide represented by an amino acid sequence that is at least 80% identical to SEQ ID NO:1, wherein exactly one Cys residue is substituted with Ala. The cysteine residue may be Cys 87 or Cys 285. Each subunit may be represented by an amino acid sequence that is at least 85%, 90%, 95%, 96%, 96%, 98%, or 99% identical to SEQ ID NO:1, and exactly one Cys residue may be substituted with Ser and/or exactly one Cys residue may be subsituted with Ala. In some embodiments, each subunit is represented by an amino acid sequence that is at least 85%, 90%, 95%, 96%, 96%, 98%, or 99% identical to SEQ ID NO:1, and additionally exactly one Cys residue may be substituted with Ser and exactly one Cys residue may be subsituted with Ala. Other amino acid residues may be substituted, for example, with amino acids that share similar properties such as structure, charge, hydrophobicity, or hydrophilicity. In certain embodiments, substituted residues

are one or more of L99, E103, F166, K294, L203 and H207. For example, the substituted residues may be L99Q, E103G, F166Y, K294R, L203V, and H207Y. An exemplary subunit may comprise L99Q, E103G, F166Y, K294R, L203V, and H207Y, and C285S. Accordingly, each subunit may comprise a polypeptide represented by an amino acid sequence of SEQ ID NO:3. An exemplary modified ClyA pore comprising subunits in which exactly one Cys residue is substituted with Ser and exactly one Cys residue is substituted with Ala may be called ClyA-AS.

The present disclosure further relates to nucleic acids encoding the modified ClyA pores. In some embodiments, a nucleic acid encoding a modified ClyA pore is represented by a nucleotide sequence that is at least 80%, 90%, 95%, 96%, 96%, 98%, or 99% identical to SEQ ID NO:4. A nucleic acid may be represented by SEQ ID NO: 5 or SEQ ID NO:6. Nucleotide sequences may be codon optimized for expression in suitable hosts, for example, *E. coli*.

The modified ClyA pore may have a pore lumen of at least 3 nm in diameter, for example, the diameter may measure 3 nm, 3.5 nm, 4 nm, 4.5 nm, 5 nm, 5.5 nm, 6 nm, 6.5 nm, 7 nm, or greater. The size of the pore lumen may depend on the analyte to be detected by the modified ClyA pore.The *cis* diameter of the pore lumen may be at least 3.5 nm and/or the *trans* diameter of the ClyA pore may be at least 6 nm. In general, *cis* refers to the end of the modified ClyA pore to which an analyte is added, while *trans* refers to the end of the modified ClyA pore through which the analyte exits after translocating the length of the pore lumen. In artificial lipid bilayers, for example, the *trans* end of a pore may be inserted in the lipid bilayer, while the *cis* end of the pore remains on the same side of the lipid bilayer. Accordingly, the *cis* diameter of the pore is the diameter of the opening at the *cis* end of the pore, and the opening to which an analyte is added, while the *trans* diameter of the pore is the diameter at the opening of the *trans* end of the pore, from which an analyte exits.

The size of the pore lumen may also depend on the number of subunits in the modified ClyA pore. For example, larger pores which are made up of 13 or 14 subunits may have larger lumens than pores made up of 7 subunits. In some embodiments, the modified ClyA pores comprise 12 or more subunits. In certain embodiments, the modified ClyA pores comprise 12 subunits. In certain embodiments, the modified ClyA pores comprise 13 subunits, or comprise 14 subunits. The subunits may preferentially assemble in 12mers and/or 13mers, depending on the amino acid sequence of the subunits. In some embodiments, each subunit comprises a polypeptide as disclosed herein. Within a single modified ClyA pore, each of the subunits may be identical, or the subunits may be different, so that subunits in a modified ClyA pore may comprise sequences that differ from sequences of other polypeptide subunits in the same modified ClyA pore. In certain embodiments,

modified ClyA pores as disclosed herein, such as ClyA-CS pores, may form more than one subtype depending on subunit composition. For example, there may be at least 2 or 3 different subtypes (i.e., Type I, Type II, Type III) of modified ClyA-CS pore, depending on subunits. Each subtype may have different conductance measurements, as compared with other subtypes. Subtypes may be preferentially formed by subunits of a particular polypeptide sequence.

The substitutions in specific residues may confer new properties on the modified ClyA pores, as compared with wild-type ClyA pores found in nature. Voltage dependent opening and closing (gating) of the pore at specific voltages is one property. In planar lipid bilayers, for example, ClyA-SS spontaneously opens and closes at applied potentials that are greater than +60mV or lower than -90 mV. In some embodiments, the modified ClyA pores as described herein remain open when the voltage across the pore (i.e., the voltage across the membrane which the modified ClyA pore is in) ranges from +90 mV to -150 mV. Accordingly, the modified ClyA pore may remain open when the voltage across the pore is held at +90, +85, +80, +75, +70, +65, +60, +55, +50, +45, +40, +35, +30, +25, +20, +15, +10, +5, 0, -5, -10, -15, -20, -25, -30, -35, -40, -45, -50, -60, -65, -70, -75, -80, -85, -90, -95, -100, -110, -115, -120, -125, -130, -135, -140, -145, -150 mV, and/or the voltage across the pore is adjusted between +90 mV and -150 mV (inclusive), or any subrange of voltages in between. In certain embodiments, the modified ClyA pores show low electrical noise as compared with the signal (i.e., the current block measured). Thus, the noise inherent in a modified ClyA pore is reduced when pores as described herein are used. An exemplary modified ClyA pore shows noise measurements of _1.5 pA rms__ to 3 pA rms__under __-35mV in 150mM NaCl 15mM Tris·HCl pH 7.5___ conditions. Notably, it is possible to reduce noise by increasing the salt concentration and/or altering the length of time during which a current block is measured.

In some embodiments, the modified ClyA pores show solubility properties that differ from wild-type ClyA pores. For examples, monomers of the modified ClyA pores may be soluble in water, and/or in other solutions where surfactants such as SDS or DDM are not present. Stable oligomers are modified ClyA pores that are capable of withstanding applied potentials of +150 mV to -150 mV across membranes or lipid bilayers into which the modified ClyA pores are inserted.

Nanopores with Ligands

A further aspect of the present disclosure relates to nanopore biosensors in which modified ClyA pore proteins are combined with ligands that have selective binding properties. In some embodiments, these modified pores and ligands are used to identify protein analytes in complex

biological samples, for example, in a tissue and/or a bodily fluid. The target protein analyte may be present in a low concentration as compared to other components of the sample. In some embodiments, ligands may also be used to target subpopulations of macromolecular analytes based on conformation or on functional properties of the analytes. The presence of a ligand may increase the association of the target protein analyte with the modified pore. For example, the ligands may act as a selectivity filter at the entrance of the pore, increasing capture of the target protein while repulsing other non-target proteins in the sample.

Exemplary ligands include but are not limited to aptamers, antibodies, receptors, and/or peptides that bind to the target protein. In some embodiments, ligands may be inhibitors of the target protein, which suppress the binding of the target protein to the modified ClyA pore.

In certain embodiments, a ligand that binds to a target protein analyte is added to a sample prior to the detection steps described above. This step may provide additional confirmation that a target protein analyte is present. Thus, a method for detecting at least one target protein in a sample may comprise comprises (a) contacting the sample with a ligand that binds to a target protein; (b) contacting the sample with a modified ClyA pore as disclosed herein; (b) applying an electrical potential across the modified ClyA pore; (c) measuring electrical current passing through the modified ClyA pore at one or more time intervals; and (d) comparing the electrical current measured at one or more time intervals with a reference electrical current, wherein a change in electrical current relative to the reference current indicates that the presence of the target protein in the sample. In addition, the change in electrical current may be compared with a sample that was not contacted with a ligand prior to measuring the electrical current through the modified ClyA pore. If a target protein analyte is indeed present, the addition of a ligand will suppress the binding of the target protein to the modified ClyA pore, and a current block would not be detected. In contrast, a current block would be detected when the ligand was not added. With both results together, the presence and the concentration of the target protein could be determined. For example, in a given sample containing many different proteins including a target protein analyte, the sample may initially give X blockades per second. After addition of an excess of a specific ligand, the sample may give (X-n) blockades per second. Thus, n may reflect the blockades per second produced by the target protein analyte in the original sample, which, in turn, may provide information about the concentration of the target protein analyte in the original sample.

In certain embodiments, varying electrical potentials are applied across the modified ClyA pore. For example, the electrical potential applied across the modified ClyA pore may range from -90 mV to + 90 mV. The electrical potential may be -90 mV, -85 mV, -80 mV, -75 mV, -70 mV, -65 mV, -60

mV, -55 mV, -50 mV, -45 mV, -40 mV, -35 mV, -30 mV, -25 mV, -20 mV, -15 mV, -10 mV, -5 mV, 0 mV, +5 mV, +10 mV, +15 mV, +20 mV, +25 mV,+30 mV, +35 mV, +40 mV, +45 mV, +50mV, +55 mV, +60 mV, +65 mV, +70 mV , +80 mV, +85 mV, and/or +90 mV . For each potential, the electrical current may be measured and compared with one or more reference electrical currents. In some embodiments, a reference electrical current is measured in an open, unblocked pore. In some embodiments, a reference electrical current is measured in a modified ClyA pore that is bound to a known protein, for example, a protein whose presence or absence will be determined in solution.

In some embodiments, a modified ClyA pore as described herein may be conjugated to one or more aptamers. When more than one aptamer is conjugated, the aptamers may be the same aptamer or may be different aptamers. The one or more aptamers may be conjugated to a cysteine residue in the modified ClyA pore. In some embodiments, a modified ClyA pore comprises a cysteine residue in place of another amino acid residue in the pore. This cysteine residue substitution may be combined with other amino acid substitutions, deletions, and/or additions made relative to the wild-type pore protein. For example, modifications within the pore lumen may be engineered to alter the size, binding properties, and/or structure of the pore. In some embodiments, cysteine residues are substitued with other amino acids such as serine residues. The modified pore protein may be a pore comprising multiple subunits, for example, 12 subunits, in which at least one subunit comprises a modified amino acid. In certain embodiments, the modified ClyA is from a gram-negative bacteria such as *Salmonella* or *Escherichia coli* (*E. coli*). In some embodiments, the modified cytolysin is a modified ClyA from *Salmonella typhi* (*S. typhi*) or *Salmonella paratyphi* (*S. paratyphi*). In some embodiments, the modified A modified ClyA pore comprises 12 subunits, each subunit comprising a sequence shown in SEQ ID NO: 2.

An aptamer may be a nucleic acid aptamer comprising DNA, RNA, and/or nucleic acid analogs. An aptamer may be a peptide aptamer, such as a peptide aptamer that comprises a variable peptide loop attached at both ends to a scaffold. Aptamers may be selected to bind to a specific target protein analyte. In certain embodiments, two or more aptamers are conjugated to the same modified ClyA pore. For example, 5, 6, 7, 8, or 9 aptamers may be conjugated to a modified ClyA pore. Alternatively, 10, 11, or 12 aptamers may be conjugated to a modified ClyA pore. If more than one aptamer is conjugated to the modified ClyA pore, the aptamers may be positioned at least 2 nm apart.

In certain embodiments, a modified ClyA pore as described herein is combined with one or more peptide ligands. Peptide ligands may be attached to the modified ClyA pores via disulfide linkages, cross-linking, and/or chemical ligation. The modified ClyA pore may also be engineered as

a fusion protein in which one or more peptide ligands is fused to at least one subunit of the modified ClyA pore. In some embodiments, the modified ClyA pore is combined more than one peptide ligand. The peptide ligands may be the same ligand or may be different ligands. Exemplary peptide ligands include, but are not limited to, receptors, antibodies, inhibitors, activators, and/or other peptide ligands that bind to target proteins.

Target Analytes

1. Protein analytes

In some embodiments, the modified ClyA pore protein is engineered to allow protein analytes to bind within the lumen of the pore. This binding mediates a robust, reproducible current block, which is readily distinguished from the unblocked ionic currents measured in unbound pores. The protein analytes may range from 15-70 kDa in molecular weight, for example, exemplary protein analytes may have a molecular weight of 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, or 70 kDa. The analyte may be a dimer or other multimer of a smaller protein.

In some embodiments, the modified ClyA pore proteins comprise more than one site of binding and/or residence for protein analytes. For example, a modified ClyA pore protein may comprise two sites: level 2 may be associated with residence of the protein analyte at a deep, more sterically constrained site, while the level 1 may be associated with residence of the protein analyte at a position closer to the wider cis entrance of the pore. Protein analyes may move between the two sites, thus eliciting two current levels seen within the same current blockade event. Thus, a modified ClyA pore proteins may provide more than one (i.e., two) current level measurement upon binding to a protein analyte.

In some embodiments, the modified ClyA-CS pores as described herein are capable of detecting and quantifying protein analytes. The modified ClyA pores may distinguish between homologs of the same protein, for example, bovine thrombin and human thrombin.

a. Detection and Identification of Proteins

Another aspect of the present disclosure relates to detection of specific proteins in a sample. In some embodiments, a method for detecting the presence of at least one protein analyte in a sample comprises (a) contacting the sample with a modified ClyA pore as disclosed herein; (b) applying one or more electrical potentials across the modified ClyA pore; (c) measuring current passing through the modified ClyA pore at each of the one or more electrical potentials; and (d) comparing measured currents with reference currents, wherein a change in currents relative to the reference currents indicates the presence of the protein analyte in the sample. In some

embodiments, the change in currents is a decrease in current. In some embodiments, a target protein has a molecular weight in the range of 15-50 kDa, for example, a molecular weight of 15, 20, 25, 30, 35, 40, 45, or 50 kDa. The reference current may be a current measured through the modified ClyA pore in the absence of a ligand, and/or a current measured through the modified ClyA pore in the presence of a reference ligand. In some embodiments, the reference ligand and the protein analyte are identical. In certain embodiments, the reference ligand and the protein analyte are at least 75%, 80%, 85%, 90%, or 95% identical. Thus, in some embodiments, a method for identifying a protein analyte in a sample comprises (a) contacting the sample with a modified ClyA pore as described herein; (b) applying one or more electrical potentials across the modified ClyA pore; (c) measuring currents passing through the modified ClyA pore at each of the one or more electrical potentials; and (d) comparing measured currents with one or more reference currents from a known ligand, wherein a match between the measured currents and the reference currents indicates the protein analyte and the known ligand are identical. Similarly, a non-match between the measured currents and the reference currents may indicate that the protein analyte and the known ligand are not identical.

In some embodiments, the modified ClyA pore comprises amino acid substitutions, deletions, and/or additions, as compared with the wild-type pore protein. For example, modifications within the pore lumen may be engineered to alter the size, binding properties, and/or structure of the pore. In some embodiments, cysteine residues are substitued with other amino acids such as serine residues. The modified pore protein may be a pore comprising multiple subunits, for example, between 7-11 subunits, 12 subunits, 13 subunits, or 14 subunits, in which at least one subunit comprises a modified amino acid. In certain embodiments, the modified ClyA is from a gram-negative bacteria such as *Salmonella* or *Escherichia coli* (*E. coli*). In some embodiments, the modified cytolysin is a modified ClyA from *Salmonella typhi* (*S. typhi*) or *Salmonella paratyphi* (*S. paratyphi*). In some embodiments, the modified A modified ClyA pore comprises 12 subunits, each subunit comprising a sequence shown in SEQ ID NO: 1.

In certain embodiments, varying electrical potentials are applied across the modified ClyA pore. For example, the electrical potential applied across the modified ClyA pore may range from -90 mV to + 90 mV. The electrical potential may be -90 mV, -85 mV, -80 mV, -75 mV, -70 mV, -65 mV, -60 mV, -55 mV, -50 mV, -45 mV, -40 mV, -35 mV, -30 mV, -25 mV, -20 mV, -15 mV, -10 mV, -5 mV, 0 mV, +5 mV, +10 mV, +15 mV, +20 mV, +25 mV,+30 mV, +35 mV, +40 mV, +45 mV, +50mV, +55 mV, +60 mV, +65 mV, +70 mV , +80 mV, +85 mV, and/or +90 mV. At each voltage, the electrical current may be measured and compared with one or more reference electrical currents. In some embodiments, a

reference electrical current is measured in an open, unblocked pore. In some embodiments, a reference electrical current is measured in a modified ClyA pore that is bound to a known protein, for example, a protein whose presence or absence will be determined in solution.

In certain embodiments, the interaction of a protein analyte with a modified ClyA pore provokes current blocks as compared to open, unbound pores. The amplitude of the current block may be measured as the residual current percentage (IRES) of the open pore current. In some embodiments, characteristic current blocks are used to identify a specific protein analyte. For example, current blocks may be short, quick current spikes. Current blocks may be transient, or may last for more than 1 minute. Current blocks may be shallow or may be deep. A shallow current level may be indicated by an IRES value of about 41-100%, indicating that the target protein analyte leaves a residual current that is about 41%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% of the open pore current. Conversely, a deep current level may be indicated by an IRES value of about 0-40%, indicating that the target analyte leaves a residual current that is about 0%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, or 40% of the open pore current. In some embodiments, a current block comprises more than one current level, for example, a current block may comprise a shallow current level and a deep current level. In one exemplary modified ClyA pore, a target analyte may provoke a first current block with an IRES value of about 70% as well as a second current block with an IRES value of about 15%. In certain embodiments, the presence and/or identity of a protein analyte is determined by comparing the shallow and deep current levels from the protein analyte with shallow and deep current levels from a reference ligand. The inter-conversion between shallow and deep current levels may be compared, for example, IRES values may be compared, the rate of conversion between shallow and deep current levels may be compared, the relative durations of the shallow and deep currents levels may be compared, and/or the number of shallow and deep current levels may be compared.

For any target protein analyte, when the voltage across the modified ClyA pore is varied, the percentage of shallow and/or deep current levels may also vary. The distribution of shallow and/or deep current levels may differ from one target protein analyte to another target protein analyte. Thus, in some embodiments, two or more target analytes are distinguished from one another on the basis of their current level measurements. Two or more target protein analytes may be distinguished by the intensity and duration of their current blocks and/or by their distributions of current levels. For example, a first protein may show a large decrease in the percentage of shallow levels as the voltage across the pore is increased, while a second protein may show a more gradual decrease.

In certain embodiments, the two or more target proteins about 95%, 90%, 85%, 80%, 75%, or 70% identical. In certain embodiments, the two or more target proteins or their subunits are about 65%, 60%, 55%, 50%, or 45% identical. In some embodiments, the two or more target proteins are species-specific forms of the same protein.

2. Translocation of nucleic acids

One aspect of the present disclosure is a modified ClyA pore that is capable of binding, detecting, identifying, translocating nucleic acids and/or modulating transport of nucleic acids. Nucleic acids which translocate through the modified ClyA pores include but are not limited to DNA, which may or may not carry posttranslational modifications such as methylation; RNA such as transfer RNA (tRNA), messenger RNA (mRNA), ribosomal RNA (rRNA), small interfering RNA (siRNA), micro RNA (miRNA), small nucleolar RNA (snoRNA), small nuclear RNA (snRNA), extracellular RNA (exRNA), Piwi-interacting RNA (piRNA), and any non-coding RNA (ncRNA). Nucleic acid analogs such as 2'-O-methyl substituted RNA, locked nucleic acid (LNA), morpholino, and peptide nucleic acid (PNA), dideoxynucleotides, etc.

In certain embodiments, the modified ClyA pores are engineered to translocate DNA such as ssDNA or dsDNA. While ssDNA in particular may not enter a modified ClyA pore at physiological salt conditions due to replusive charges from negatively-charged residues in the lumen, DNA translocates through modified ClyA pores as described herein under conditions of high ionic strength. In some embodiments, conditions of high ionic strength are (or have the equivalent to the ionic strength of) a 2.5 M NaCl solution. In certain embodiments, conditions of high ionic strength are higher than the ionic strength of a 2.5 M NaCl, for example, conditions of high ionic strength may be be (or may be equivalent to the ionic strength of) a 2.75 M, 3 M, 3.25 M, 3.5 M, 3.75 M, 4 M, 4.25 M, 4.5 M, 4.75 M, or 5 M NaCl solution and/or may be (or may be equivalent to) 2.75 M, 3 M, 3.25 M, 3.5 M, 3.75 M, 4 M, 4.25 M, 4.5 M, 4.75 M, or 5 M KCl solution. The negatively charged residues lining the internal lumen of the ClyA pore may be screened under these conditions. For example, in 2.5 M NaCl, 15 mM Tris-HCl at pH 7.5, addition of ds DNA to a modified ClyA pore as described herein at +100 mV may produce a short current blockade due to translocation of dsDNA through the pore. As demonstrated herein, dsDNA translocates the modified ClyA pores under these conditions. In addition, ssDNA may translocate the modified ClyA pores, for example under conditions of high ionic strength and/or in a folded structure. ssDNA may translocate modified ClyA pores at a different rate than dsDNA.

Accordingly, one aspect of the present disclosure is a method for translocating DNA, for example dsDNA through a modified ClyA pore that is capable of translocating DNA, comprising the

steps of obtaining a modified ClyA pore as described herein, applying a voltage of at least +50 mV across the modified ClyA pore, adding a sample containing the DNA to the *cis* opening of the modified ClyA pore, and measuring the current flowing through the pore. A current blockade indicates translocation of the DNA. Current may be restored by reversing the potential to a negative potential, such as -100 mV or -50 mV. In some embodiments, the modified ClyA pore is used under conditions of high ionic strength.

In some embodiments, the modified ClyA pore for translocating DNA comprises a modified ClyA pore as described herein. The modified ClyA pores may be engineered to translocate nucleic acids. For example, the modified ClyA pore may comprise at least 12 subunits, wherein each subunit comprises a polypeptide represented by an amino acid sequence at least 80% identical to SEQ ID NO:1, wherein exactly one Cys residue is substituted with Ser. Each subunit may be represented by an amino acid sequence that is at least 85%, 90%, 95%, 96%, 96%, 98%, or 99% identical to SEQ ID NO:1, and additionally exactly one Cys residue may be substituted with Ser. The Cys residue may be Cys 87 and/or Cys 285 in SEQ ID NO:1. In some embodiments, the Cys residue is Cys285. Other amino acid residues may be substituted, for example, with amino acids that share similar properties such as structure, charge, hydrophobicity, or hydrophilicity. In certain embodiments, substituted residues are one or more of L99, E103, F166, and K294. For example, the substituted residues may be one or more of L99Q, E103G, F166Y, and K294R. Thus, each subunit may comprise a polypeptide represented by an amino acid sequence of SEQ ID NO:2. An exemplary modified ClyA pore comprising subunits in which exactly one Cys residue is substituted with Ser may be called ClyA-CS.

In some embodiments, the modified ClyA pore recognizes and chaperones a specific DNA molecule across a biological membrane under a fixed transmembrane potential. The reaction mechanism may be based on DNA strand displacement. For example, a DNA moiety may be conjugated to a nanopore in order to allow the transport of selected DNA molecules across the nanopore via the DNA strand displacement reaction. In certain embodiments, the modified ClyA pore comprises at least 12 subunits, wherein each subunit comprises C87S, C285S, and D103C substitutions, and each subunit is conjugated to an oligonucleotide. Accordingly, in some embodiments, a method for translocating dsDNA comprises (a) obtaining a modified ClyA pore comprising at least 12 subunits, wherein each subunit comprises C87S, C285S, and D103C substitutions and is conjugated to an oligonucleotide; (b) applying a voltage of +50 mV across a modified ClyA pore as described herein, (c) adding a sample containing a dsDNA to the cis opening of the modified ClyA pore, (d) adding to the *cis* opening of the modified ClyA pore a first single stranded nucleic acid comprising (i) a sequence that is complementary to at least 15 nucleobases of

the oligonucleotide that is conjugated to the ClyA pore, and (ii) a sequence is that is complementary to at least 12 nucleobases of the double stranded DNA; and measuring the current across the modified ClyA pore,wherein a decrease in current after step (d) indicates translocation of the double stranded DNA through the modified ClyA pore.

In some embodiments, the method optionally comprises adding to a *trans* opening of the modified ClyA nanopore a second single stranded nucleic acid comprising a sequence that is complementary to the first single stranded nucleic acid. Here, an increase in current across the modified ClyA pore after adding the second single stranded nucleic acid to a *trans* opening of the modified ClyA pore indicates that the double stranded DNA has translocated completely through the pore.

A further aspect of the present disclosure relates to a device for translocating DNA, comprising: a fluid-filled compartment separated by a membrane into a first chamber and a second chamber; electrodes capable of applying potential across the membrane; one or more nanopores inserted in the membrane; a solution of high ionic strength in one chamber of the membrane, wherein DNA translocates through the nanopore from the first chamber to the second chamber. In certain embodiments, the DNA is double stranded. The nanopores may be ClyA pores, for example, the modified ClyA pores described herein. The pores, such as modified ClyA pores, may have an inner diameter of at least 2.2 nm. In some embodiments, the membrane is an artificial lipid bilayer. In some embodiments, the potential across the membrane ranges from -100 mV to +100 mV. The solution of high ionic strength may comprise 2.5M NaCl.

EXAMPLES

Having provided a general disclosure, the following examples help to illustrate the general disclosure. These specific examples are included merely to illustrate certain aspects and embodiments of the disclosure, and they are not intended to be limiting in any respect. Certain general principles described in the examples, however, may be generally applicable to other aspects or embodiments of the disclosure.

Example 1. Tuning the property of ClyA by directed evolution

Directed evolution approaches for tailoring enzymes with desired properties were used to improve the activity of ClyA-SS, reasoning that mutations that compensate for the deleterious effects of C87S and C285S substitutions would also increase stability of the nanopore in lipid bilayers. Random libraries were generated on the background of ClyA-SS by error prone PCR (approximately 1-4 mutations per gene per round) and screened for hemolytic activity (Figure 6).

The most active variants were then purified by Ni-NTA affinity chromatography and tested for oligomer formation by BN-PAGE (Figure 7). Selected nanopore variants were finally screened in planar lipid bilayers for the desired behavior (low electrical noise and ability to remain open at high applied potentials), which served as final and critical criteria for selection. After just four rounds of screening, ClyA-CS variants were isolated (Table 1) that showed low electrical noise (Figure 8) and remained open in planar lipid bilayers from +90 to -150. Remarkably, the serine at position 87 converted back to cysteine, the original residue in the wild-type gene. In order to obtain a cysteine-less ClyA variant amenable to site-specific chemical modification, ClyA-CS was subjected to saturation mutagenesis at position 87, the resulting library was screened, and cysteine-less ClyA-AS with desired electrical properties (SI) was selected. In contrast to ClyA-SS, evolved ClyA nanopores expressed in *E.coli* cells in the soluble fraction (Figure 6) and the monomers could be purified in one-step by affinity chromatography, which allowed a ten fold increase in the production yield (~0.6 mg per 10 ml culture).

Example 2. Isolation of ClyA nanopores with different size

ClyA oligomers, formed by incubation of ClyA monomers with 0.5% w/v β-dodecyl maltoside (DDM), revealed multiple bands on a BN-PAGE (Figure 2a), suggesting that ClyA might assemble into several oligomeric states. This is particularly intriguing, since the exact stoichiometry of *E. coli* ClyA oligomerisation is controversial. The ClyA crystal structure (PDB_ID: 2WCD) revealed a dodecamer with a 5.5 nm opening on the cis side and a 3.3 nm opening at the trans entrance (including Van der Waals radii of the amino acid side chains), while earlier cryo-EM structures revealed nanopores with 8[11] or 13[12] subunits.

In planar lipid bilayers, ClyA-WT pre-incubated with DDM showed a wide distribution of open nanopore conductances spanning approximately 2 nS (Figure 2b, top), suggesting that also in lipid bilayers ClyA-WT might assemble into nanopores of different size and/or geometry. The major peak in the distribution of ClyA-WT unitary conductance (Type I ClyA) represented only 24% of the reconstituted nanopores and showed an average conductance of 1.83±0.06 nS in the conductance range from 1.7 to 1.9 nS (-35 mV, 15 mM Tris.HCl pH 7.5 and 150mM NaCl). The distribution of ClyA-CS open pore conductance showed two major peaks: the first included 37% of the reconstituted nanopores and corresponded to Type I ClyA-CS (1.79±0.05 nS); while the second (Type II ClyA-CS) included 23% of the nanopores and showed an average conductance of 2.19±0.09 nS (conductance range 2.1 - 2.4 nS, Figure 2b, middle). ClyA-WT and ClyA-AS also showed small percentage of Type II ClyA (18% and 16%, respectively). The unitary conductance of ClyA-AS was especially uniform with 52% of the reconstituted nanopores corresponding to Type I ClyA (Figure 2b, bottom).

To establish whether the different bands of ClyA oligomers corresponded to nanopores with different size, ClyA-CS were extracted from the three major oligomeric bands in the BN-PAGE, and measurements were made of the unitary open nanopore conductance of 62 nanopores derived from each band within two days from gel extraction. 62% of ClyA-CS oligomers from the lowest band formed Type I ClyA-CS nanopores (1.78±0.04 nS, Figure 3a), while 68% of nanopores extracted from the second lowest band (Figure 3b) reconstituted as Type II ClyA-CS nanopores (2.19 ± 0.09 nS). Interestingly, 42% of the nanopores extracted from the third band reconstituted in lipid bilayers as a third nanopore type (Type III ClyA) that showed an average conductance of 2.81±0.11 nS in the conductance range 2.5 - 3.0 nS (Figure 3c).

Taken together, these results show that the three major bands of ClyA oligomers observed on the BN-PAGE correspond to three distinct nanopore types with different size and different unitary conductance. This finding is consistent with reports that high order symmetrical oligomeric structures are often permissive with respect to subunit stoichiometry.[13] Therefore it can be hypothesized that Type I ClyA most likely represents the 12mer of the crystal structure, while Type II ClyA might be the 13mer observed in earlier cryo-EM studies. Both nanopores showed low electrical noise (Figure 8) and remained open over a wide range of applied potentials (from +90 mV to -150 mV). Type III ClyA-CS nanopores, which showed higher noise than Type I and Type II nanopores (Figure 8) and frequently gated especially at applied potentials lower than -40 mV and higher than +50 mV, may correspond to a 14mer version of ClyA not observed before.

Example 3. HT as molecular caliper to test ClyA nanopores of different size

The ability to employ nanopores with identical amino acid composition but different size is a new feature in the biological nanopore field and is important because the size of a nanopore defines its ability to capture and study a particular molecule.[10b, 14] It has been previously shown[7a] that at -35 mV HT (human thrombin, 37kDa) inflicted well-defined current blockades to Type I ClyA-SS nanopores that lasted for several minutes. The blockade signal switched rapidly between two current levels, level 1 [percentage of the open nanopore current ($I_{RES\%}$) = 56 ± 1%] and level 2, ($I_{RES\%}$ =23 ± 1%, Table 1, and Figure 4a), reflecting two residence sites for HT within the lumen of the ClyA nanopore. Level 2 is most likely associated to HT residence at a deep site, while level 1 is associated to the residence of HT closer to the cis entrance of the nanopore.[7a] Because thrombin provoked such a well-defined pattern of current blockades HT was used here as a molecular caliper to compare the geometries of the different ClyA nanopores.

At -35 mV HT current blockades to Type I ClyA-CS nanopores were identical to that of Type I ClyA-SS nanopore (Table 1), confirming that mutations accumulated in the variants disclosed herein

most likely did not change the size and geometry of the ClyA nanopore. HT current blockades to Type II ClyA also switched between the two current levels, but their relative distribution was different. In Type I ClyA-CS HT mostly lodged at the more superficial binding site (70% occupancy), while in Type II ClyA-CS HT preferred the binding site deeper within the nanopore (96% occupancy). These results suggest that both ClyA Types most likely retain similar overall nanopore architecture but provide different steric hindrance to HT. HT blockades to Type III ClyA were fast (55±48 ms) and showed only a level 2 current block ($I_{RES\%}$ of 32±9%), suggesting that Type III ClyA is large enough to allow unhindered translocation of HT through the nanopore (see below).

Example 4. Protein translocation through ClyA nanopores.

Thrombin is a globular protein with a molecular volume that can be approximated to a sphere with diameter of 4.2 nm (SI). Therefore, assuming that Type I, II and III ClyA correspond to nanopores with different oligomeric state (see above), HT should not easily translocate through Type I and Type II ClyA nanopores (trans diameter, including the Van der Waals radii of the atoms, of 3.3 and 3.7 nm, respectively, Table 1). Contrary, HT has the same diameter as Type III ClyA (Table 1), suggesting that the protein might be capable of translocating through this nanopore. A powerful method to assess whether molecules pass through nanopores is to investigate the voltage dependence of the duration of the proteins current blockades. The decrease of the duration of the current blockades with increasing potential is strong evidence that the molecules translocate through the nanopore. By contrast, an increase in the duration of the current blockades with the voltage suggests that the proteins are driven into the nanopore but do not translocate through it.[10f, 15] From -5 mV to -25 mV the dwell time of HT blockades to Type I ClyA-SS nanopores increased with the applied potential (Figure 9a), suggesting that HT does not translocate through the nanopores in this voltage interval. Taking advantage of the fact that Type I and Type II evolved ClyA nanopores remained open at applied potentials up to -150 mV, HT blockades were characterized on Type I and II ClyA-CS nanopores from -60 to -150 mV. At applied potentials higher than -100 mV for Type I ClyA-CS nanopores and -60 mV for Type II ClyA-CS nanopores the dwell times of HT current blockades strongly decreased with increasing potential (Figure 5a), suggesting that in this potential range HT molecules translocate through the nanopores. Similarly, Dendra2_M159A (FP, a GFP like protein, 30 kDa protein) displayed an initial increase (from -25 mV to -40 mV) followed by a decrease of the duration of the current blockades (from -50 mV to -70 mV), suggesting that FP translocates through Type I ClyA-SS nanopores at potentials higher than -50 mV (Figure 9b). Comparing the duration of HT blockades to Type I and Type II ClyA-CS nanopores at the same potential revealed that the translocation of HT through Type II ClyA-CS was about two orders of magnitude faster than for Type I

ClyA-CS (Figure 5b, Table 1), which is in line with the interpretation that Type I and Type II ClyA describe nanopores with different size.

Example 5. Folded versus unfolded translocation

When a molecule is lodged within the lumen of a nanopore, the ionic current block is proportional to the atomic volume of the electrolytes being excluded by the molecule.[10d, 10f, 16] Therefore, if the molecule translocates through the nanopore with a folded structure, the $I_{RES\%}$ should remain constant at different applied voltages. In contrast, if a protein unfolds upon translocation, the $I_{RES\%}$ is expected to change, giving that the volume and shape of the unfolded polypeptide chain is different to that of the globular protein. The $I_{RES\%}$ values during the translocation of HT through Type I and Type II nanopores were identical at -35 mV and -150 mV (level 2, Table 1), suggesting that in this potential range HT does not unfold while in the nanopore. Interestingly, solely with Type I ClyA-CS nanopores, the current blockades of HT at potentials below -90mV often terminated with a current block of higher $I_{RES\%}$ (shoulder) followed by a current increase (spike) with respect to the open nanopore current (Figure 5c and 10). Although the shoulder in the $I_{RES\%}$ values might indicate that HT unfolds upon translocation, the current spikes that follows the protein translocation suggest otherwise that ClyA nanopores may need to deform in order to allow the translocation of folded HT through Type I nanopores.

Example 6: Translocation of dsDNA through a ClyA pore

In this work ClyA-CS was selected for its enhanced activity, solubility and favourable behaviour in planar lipid bilayers when compared to Wild Type ClyA The internal diameter of the ClyA dodecamers (3.8 nm at its narrower entrance,[17] Figure 11a) is larger than the diameter of dsDNA (2.2 nm for the B form), indicating that dsDNA should readily electrophoretically translocate through the pore. However, most likely because of the negatively charged residues lining the lumen of ClyA (pI = 5.1), at physiological salt concentrations ssDNA does not enter the nanopore.[7a] In view of previous work with alpha hemolysin (αHL) nanopores at high alkaline pH,[7b, 7c] the ability of DNA to translocate through ClyA-CS nanopores was tested at high ionic strength, where the internal charges of the pore are screened. In 2.5 M NaCl, 15 mM Tris.HCl pH 7.5 and under +100 mV applied potential, the addition of 0.12 μM of biotinylated dsDNA **1** (290 bp, Table 2) to the *cis* compartment produced transient current blockades (I_B) to the open pore current (I_O) showing a residual current ($I_{RES} = I_B / I_O$) of 0.63±0.01 (level 1^*_{+100} = 1.10±0.03 nA, n=3 experiments), with 2.0±0.6 ms dwell time, due to the entrance of the DNA into the lumen of the pore (Figure 11b). The subsequent addition to the *cis* compartment of 0.3 μM of neutravidin, which forms a tight complex with biotin, converted the transient blockades into long lasting current blockades (level 1_{+100} = 1.19±0.01 nA, n=4) with a

higher residual current value (I_{RES} = 0.68±0.01). The open pore current could be restored by reversal of the applied potential to -100 mV (Figure 11b). These results suggest that neutravidin prevents the full translocation of DNA through ClyA nanopores by forming a *cis* protein:DNA complex where the DNA occupies the full length of the pore (Figure 11c). *Trans* complexes could also be formed at -100 mV (level 1_{-100} = 1.02±0.03 nA, I_{RES}=0.62±0.01, n=4) when the dsDNA:neutravidin complex is threaded through the *trans* side (Figure 11d).

Example 7. A rotaxane system traps a dsDNA within a ClyA nanopore

Rotaxanes are supramolecular interlocked systems in which a linear unit (thread) is translocated through a microcyclic ring and is tapped by two bulky substituents (stoppers). Such mechanically joined molecules have applications for example as switches in molecular electronics or as components in molecular machineries. Rotaxanes have been made from a variety of molecules including dsDNA[22] or by locking a biotinylated ssDNA molecules threaded through αHL nanopore with streptavidin on one side and with a DNA hairpin on the other side (dsDNA can not translocate through αHL).[23] Here to prove the translocation of dsDNA through ClyA nanopores a rotaxane system was built in which a dsDNA molecule added to the *trans* side of a ClyA nanopore hybridize with a second DNA strand on the *cis* side after threading through the lumen of the nanopore. A ClyA nanopore ClyA-**2** was used, which contains 12 ssDNA molecules **2** (51 bases, Table 2, Figure 12a) covalently attached at their 5' ends to cysteine residues introduced at the *cis* entrance of the pore (at position 103, Figure 11a) via disulphide linkages. **2** is designed to act as a rotaxane stopper. The thread **3** is a dsDNA molecule (59 bp, Table 2) with an additional 31 nucleobases stretch of ssDNA at the 5' end that is designed to hybrize with the stopper at the *cis* side through hybridisation with oligo **6**; and a 3' biotinylated linker that is used for complexation with neutravidin at the *trans* side. The linkage between the thread and stopper on the *cis* side is mediated by a bridging ssDNA molecule **4** (Table 2, Figure 12a) that is complementary to the first 16 nucleobases of **2** and to the last 25 nucleobases of **3.** When **3** and **4** are added to the *trans* compartment, at -100 mV the DNA thread is captured by the pore and not released from the pore upon reversing of the potential to +100 mV (level 2_{+100}, I_{RES} = 0.77±0.04, n=4), indicating that a DNA rotaxane is formed (Figure 12b). Interestingly, at +100 mV the residual current of the rotaxane was higher than the I_{RES} values of the *cis* or *trans* pseudorotaxane threads (0.68±0.01 and 0.62±0.01, respecively), suggesting that an unhybridized ssDNA stretch of **2** is likely to span the pore at this potential (Figure 12b). The rotaxane could be disassembled by addition of 20 mM DTT to the *cis* chamber, which reduced the disulphide bond between **2** and ClyA and restored the open pore current at +100 mV (Figure 12c-12e).

Selective translocation through ClyA-2 pores.

The data indicate that ClyA-**2** excludes non-specific DNA from the pore lumen of the pore, suggesting that the mesh ssDNA molecules attached to the pore might produce a steric and or electrostatic impediment for non-tethered DNA translocation. In addition, the DNA attached to the pore often occupies the pore lumen (Figure 15c) thus preventing the entrance of other DNA molecules. When a specific DNA molecule is hybrised to ClyA-2, rapid DNA capture is observed at positive applied potentials. The concentration of ssDNA attached to the pore was previously estimated to be ~ 20 mM.[21] Therefore, augmented concentration of the dsDNA in the proximity of the pore mouth may enhance the capture of the specific DNA strands. In this case, unhybridized strands **2** might still be tethered to the pore, in which case the dsDNA construct might have to compete with the ssDNA **2** to enter the pore lumen.

Example 8. A nanopore:DNA device utilizing a strand displacement reaction

At high positive applied potentials, the ionic current of ClyA-**2** nanopores fluctuated between the open pore level and several blocked pore levels (Figure 12c, Figure 15 and 16a,), suggesting the ssDNA molecules tethered to the top of the pore enter the lumen of ClyA but do not permanently thread to the *trans* side of the pore.[1c] Further suggesting this interpretation; at +100 mV the addition of a 90mer ssDNA **5a** (Table 2) in complex with neutravidin to the *cis* side of ClyA-CS provoked transient current blockades (Figure 13) that converted into long lasting DNA translocation events upon the subsequent addition of equimolar concentrations of the complimentary ssDNA **5b** (Table 2, Figure 13). The translocation of DNA through nanopores is often observed above a threshold potential[24-27] that can be tuned by modulating the charge distribution of the lumen of the pore,[26,27] or by changing the ionic strength of the solution.[7c,28] Therefore, most likely because of its lower charge density and/or higher flexibility, these findings indicate that ssDNA has higher threshold for DNA translocation than dsDNA (Figure 13).

These results suggest that despite the applied potentials a ssDNA molecule attached to the *cis* entrance of ClyA is likely to sample the *cis* solution. On the other hand if the DNA molecule becomes double stranded (e.g. by strand hybridisation) at positive applied potential the dsDNA strand will translocate through the pore and sample the *trans* solution. Therefore, a nanopore:DNA device was designed in which the hybridisation of a specific DNA strand to the *cis* side of the nanopore promotes the translocation of the DNA hybrid through the pore. The DNA:nanopore complex is then disassembled by a strand displacement reaction (Figure 14a), which will promote the transport of DNA across the bilayer and the return of **2** to sample the *cis* chamber. Conveniently, at positive applied potentials the addition of dsDNA molecules to the *cis* side of a ClyA-2 do not produce current blockades (Figure 18), indicating that the ssDNA molecules attached to ClyA-**2**

prevent or drastically reduce the translocation of DNA from solution. Therefore, the DNA unit atop of the nanopore might infer specificity to the system by promoting the translocation of specific DNA molecules while creating a barrier for the translocation of non-specific DNA. At +50mV, the addition of **3** to the *cis* side of ClyA-**2** did not produce current blockades, further confirming that the ssDNA molecules attached to ClyA-**2** prevent the DNA in solution from entering the pore (Figure 14b). Nonetheless, after the addition to the *cis* chamber of **6**, which is complementary to the first 15 bases of **2** and to the last 12 nucleobases of **3** (Table 2), the dsDNA hybrid nanopore showed permanent current blockades with I_{RES} = 0.70±0.02 (level 2_{+50} = 0.59±0.02 nA, Figure 14c, n=5), the hallmark of DNA capture. The translocation of **3** to the *trans* side was confirmed by the formation of a rotaxane upon addition of neutravidin to the *trans* chamber (Figure 14c-d, Figure 17e-f). Crucially, **6** was designed to include a 10 nucleobases 5'-single-strand extension to serve as *toehold* for the dissociation of the rotaxane (Figure 14a). Notably, the addition to the *trans* chamber of **7**, a ssDNA molecule complementary to **6** (Table 2), released **3** from the nanopore by first hybridising to the *toehold* and then promoting strand displacement (Figure 14a). This was observed by the restoration of the open pore current, because the ssDNA molecule tethered to ClyA returned to the *cis* side after **3** and **6** are released from the pore (Figure 14d). Remarkably, the nanopore showed a succession of open and blocked current levels, reflecting a cycle of capture, translocation and release, as the DNA cargos are captured from the *cis* chamber, transported through the pore and released to the *trans* chamber (Figure 14d).

The DNA strands at the top of ClyA-**2** nanopores drastically reduced the capture of non-specific DNA at both +50 and +100 mV (Figure 18), suggesting that the DNA at the top of the pore prevented the translocation of dsDNA from solution. Notably, the applied potential was set to +50 mV and not at +100 mV because at +100 mV ClyA-**2** occasionally produced long current blockades that were similar to typical events provoked by the capture of non-specific DNA (Figure 18). Such current blockades were less frequent at lower applied potentials; hence the experiment was performed at +50 mV (Figure 14) and +35 mV (Figure 20). An additional reason to work at lower applied potential, as explained in the legend of Figure 14, is that the frequency of DNA capture is reduced with the potential, thus at lower potentials the cycles of capture and release are more easily observed.

ssDNA vs dsDNA translocation

In the rotaxane configuration ssDNA molecules may be spanning the entire length of the pore at positive applied potentials. This is likely because at +100mV the I_{RES} value of the rotaxane (0.77) is higher than the I_{RES} values of the cis- and trans- pseudorotaxanes (0.68 and 0.62,

respectively, Figures 11c and 11d), which have a dsDNA immobilised within the pore. In addition, the unitary conductance values of the rotaxane as calculated from the slopes of the I-V curves were lower at positive applied potentials (10.8 nS) than at positive bias (13.0 nS, Figure 16). Since ssDNA has a diameter (d = 1nm) is smaller than dsDNA (d = 2.2 for the B form), these results further suggesting that ssDNA occupies the pore at positive bias. To further investigate the ability of ssDNA to span the pore, the ability of DNA hybrid **3**, which is formed by a 3' biotinylated dsDNA section of 59 nucleobasepairs followed by a ssDNA stretch of 31 nucleobases at the 5' end, was tested for ability to translocate through ClyA-CS pores. At +100 mV, the addition of **3** in complex with neutravidin to the *cis* side of a ClyA-CS pore provoked long lasting current blockades with I_{RES}=0.67 (Figure 19), the same I_{RES} of a *cis*-pseudorotaxane (0.68), suggesting that at this potential the DNA hybrid translocate through ClyA and dsDNA spans the lumen of the pore. Since the translocation of **3** can only be initiated from the ssDNA end these results suggest that at +100 mV the ssDNA leading sequence is capable of translocating the lumen of ClyA. Interestingly, at +50 mV and +70 mV the current blockades were only transient (Figure 19), suggesting that at this voltage ssDNA cannot pass the ClyA pore. The addition of strand 6, which is complimentary to the last 12 nucleobases of **3** produced long lasting current blockades +70 mV (Figure 19), suggesting that threshold for DNA translocation is reduced.

MATERIALS & METHODS

Screening of ClyA nanopores.

ClyA was expressed in *E. cloni*® EXPRESS BL21 (DE3) cells (Lucigen) by using a pT7 plasmid. Transformants were prescreened on Brucella Agar with 5% Horse Blood (BBL™, Becton, Dickinson and Company), and individually grown and overexpressed in 96-deep-wells plates. Monomers from cell lysates were first screened for hemolytic activity on horse erythrocytes (bioMérieux) and then purified by using Ni-NTA affinity chromatography. Purified monomers were oligomerized in the presence of 0.5% β-dodecyl maltoside (GLYCON Biochemicals GmbH)[12] and loaded on native gel electrophoresis gels to check for oligomerisation. The electrical properties of ClyA oligomers were then screened in planar lipid bilayers.

Purification of evolved ClyA nanopore

ClyA was expressed in *E. cloni*® EXPRESS BL21 (DE3) cells by using a pT7 plasmid. Monomers were purified by using Ni-NTA affinity chromatography and oligomerized in the presence of 0.2% β-dodecyl maltoside (GLYCON Biochemicals GmbH).

Electrical recordings.

Ionic currents were measured by recording from planar bilayers formed from diphytanoyl-*sn*-glycero-3-phosphocholine (Avanti Polar Lipids, Alabaster, AL). Currents were measured with Ag/AgCl electrodes by using a patch-clamp amplifier (Axopatch 200B, Axon Instruments, Foster City, CA).[18]

Construction of ClyA libraries by error-prone PCR.

Libraries were constructed by amplifying the ClyA genes from plasmid DNA using T7 promoter and T7 terminator primers (Table 3). In the first mutagenesis round we used as a template a plasmid containing the synthetic gene encoding for ClyA-SS from *Salmonella Typhi*. From the second mutagenesis round we used the DNA plasmids that were derived from the previous round of selection. In ClyA-SS, the WT sequence was modified by the substitution of the two Cys residues (positions 87 and 285) with Ser and by the attachment of DNA encoding a Gly-Ser-Ser linker followed by a C-terminal hexahistidine tag.[7a]

DNA amplification was performed by error prone PCR: 400 µL of the PCR mix (200 µl of REDTaq ReadyMix, 8 µM final concentration of forward and reverse primers, ~400 ng of plasmid template and ddH_2O up to 400 µl) was split into 8 reaction mixtures containing 0-0.2 mM of $MnCl_2$ and cycled for 27 times (pre-incubation at 95ºC for 3 min, then cycling: denaturation at 95ºC for 15 sec, annealing at 55ºC for 15 sec, extension at 72ºC for 3 min). These conditions typically yielded 1-4 mutations per gene in the final library. The PCR products were pooled together, gel purified (QIAquick Gel Extraction Kit, Qiagen) and cloned into a pT7 expression plasmid (pT7-SC1) by MEGAWHOP procedure:[19] ~500 ng of the purified PCR product was mixed with ~300 ng of ClyA-SS circular DNA template and the amplification was carried out with Phire Hot Start II DNA polymerase (Finnzymes) in 50 µL final volume (pre-incubation at 98ºC for 30s, then cycling: denaturation at 98ºC for 5 sec, extension at 72ºC for 1.5 min for 30 cycles). The circular template was eliminated by incubation with Dpn I (1 FDU) for 2 hr at 37ºC. The resulted mixture was desalted by dialysis against agarose gel (2.5 % agarose in Milli-Q water) and transformed into *E. cloni*® 10G cells (Lucigen) by electroporation. The transformed bacteria were grown overnight at 37ºC on ampicillin (100 µg/ml) LB agar plates typically resulting in >10^5 colonies, and were harvested for library plasmid DNA preparation.

Construction of saturation mutagenesis library at position 87 of evolved ClyA

In order to have a library containing cysteine-free ClyA variants, the gene encoding for 4ClyA4 (Table 2) was amplified using the 87NNS primer (containing a degenerate codon at position 87 encoding for the complete set of amino acids, Table 3) and T7 terminator primers. PCR

conditions: 0.3 mL final volume of PCR mix (ReadyMix™), containing ~400 ng of template plasmid, cycled for 30 times (pre-incubation at 95ºC for 3 min, then cycling: denaturation at 95ºC for 15 sec, annealing at 55ºC for 15 sec, extension at 72ºC for 3 min). The resulting PCR product was cloned into pT7 expression plasmid by MEGAWHOP procedure using 4ClyA4 circular template (see above).

Construction of the ClyA-WT

ClyA-SS gene was amplified using 87C and 285C primers (Table 3). PCR conditions: 0.3 mL final volume of PCR mix (150 µl of REDTaq ReadyMix, 6µM of forward and reverse primers, ~400 ng of template plasmid), cycled for 27 times (pre-incubation at 95ºC for 3 min, then cycling: denaturation at 95ºC for 15 sec, annealing at 55ºC for 15 sec, extension at 72ºC for 3 min). The resulting PCR product was cloned into pT7 expression plasmid by the MEGAWHOP procedure described above, using ClyA-SS circular template.

Screening of ClyA libraries and hemolytic assay

Since ClyA-SS displays "border of detection" hemolytic activity, during the first two rounds of the mutagenesis libraries were only screened for activity on Brucella Agar with 5% Horse Blood. From the third selection round, colonies displaying hemolytic activity on Brucella Agar were further screened for hemolytic activity on horse erythrocytes from the crude lysate after overexpression. The goal of this work was to obtain ClyA variants that oligomerise well in β-dodyl maltoside (DDM) and form nanopores with low electrical noise and uniform unitary conductance in lipid bilayers. Therefore, screening for hemolytic activity alone could not serve as the sole criteria for selection of such variants (for example WT-ClyA is very hemolytically active, but shows non-uniform unitary current distribution). Thus, from the 4th round, proteins from the crude lysate were purified by Ni-NTA affinity chromatography and the oligomerisation of ClyA was tested on BN-PAGE after incubation in DDM. ClyA nanopores that oligomerised well were then tested in planar lipid bilayers (with particular stress on uniformity of formed channels, low electrical noise and stability at high applied potentials).

Screening for Hemolytic activity on Brucella horse blood agar plates

After the plasmid DNA was electroporated into *E. cloni*® EXPRESS BL21 (DE3) cells (Lucigen), transformants were prescreened on Brucella Agar with 5% Horse Blood (BBL™, Becton, Dickinson and Company) supplemented with 100 µg/ml ampicillin. Clones displaying hemolytic activity, which was observed by a lytic aura around the colonies after overnight growth at 37 ºC, were individually grown in 96-deep-wells plates overnight by shaking at 37ºC (0.5 mL 2xYT medium containing

100µg/mL ampicillin). The obtained cultures were either pooled for preparation of plasmid DNA (QIAprep, Qiagen) that served as a template for the next round (rounds 1 and 2), or used as starters for protein overexpression (rounds 3 to 5).

Screening for hemolytic activity of crude lysates after ClyA overexpression

Rounds 3 to 5: 50 µl of the starter cultures from 400-600 clones (see above) were inoculated into 450 µl of fresh medium in new 96-deep-wells plates and the cultures were grown at 37ºC until OD_{600}~0.8. Then IPTG (0.5 mM) was added to induce overexpression, and the temperature was reduced to 25ºC for an overnight incubation. Next day, bacteria were harvested by centrifugation at 3000 x g for 15 min at 4ºC, the supernatant was discarded and pellets were incubated at -70ºC for few a hours to facilitate cell disruption. Then pellets were resuspended in 0.3 mL of lysis buffer (15mM Tris.HCl pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$, 10 µg/ml lysozyme and 0.2 units/mL DNAse I) and lysed by shaking at 1300 RPM for 30 min at 37 ºC. 5-30 µL of lysates were then added to 100 uL of ~1% horse erythrocytes suspension. The latter was prepared by centrifuging horse blood (bioMérieux) at 6000 x g for 5 mins at 4 ºC, the pellet was resuspended in 15 mM Tris.HCl pH 7.5, 150mM NaCl (If the supernatant showed a red color, the solution was centrifuged again and the pellet resuspended in the same buffer). The hemolytic activity was monitored by the decrease in OD at 650nm over time (~3-10 min intervals, measured using Multiskan GO Microplate Spectrophotometer, Thermo Scientific). The hemolytic activity of the clones was compared with the activity of 2-4 parent clones from the previous round that were grown in the same plates as a reference.

Screening for oligomerization and nanopore formation of evolved variants.

Rounds 4: 6-12 of the most hemolytically active variants were partially purified from the same lysates that were used for the screening of hemolytic activity by using Ni-NTA affinity chromatography: 0.2 mL of crude lysate containing monomeric ClyA was brought to 1 mL with 15mM Tris.HCl pH 7.5, 150mM NaCl supplemented with 1% DDM (to trigger ClyA oligomerization) and 10 mM imidazole, incubated at ambient temperature for 20 min and centrifuged at 20'000 x g for 10 min at 4ºC. The clarified lysates were allowed to bind to 20 µL (bead volume) of Ni-NTA agarose beads (Qiagen) for 1 hr by gentle mixing at 4ºC. The unbound fraction were removed by centrifugation at 20'000 x g for 10 min at 4ºC (the supernatant was discarded). Finally, oligomerized ClyA proteins were eluted with 50 µL of 600 mM imidazole in 15mM Tris.HCl pH 7.5 150mM NaCl 0.2% DDM. Typically 40 µg of ClyA were supplemented with ~10% glycerol and 1x of NativePAGE™ Running Buffer and 1x Cathode Buffer Additive (Invitrogen™) and then loaded on BN-PAGE (Figure

7). Variants that formed oligomers on BN-PAGE were then tested in planar lipid bilayers (with particular stress on uniformity of formed channels).

In round 5 the ClyA-CS clones that were subjected to saturation mutagenesis at position 87 were screened in order to obtain a cysteine-less ClyA variant amenable to site-specific chemical modification. Therefore, since that clones containing cysteine at position 87 were expected to show the highest activity, in this round the clones that showed medium to high hemolytic activity in the crude lysates were isolated. ClyA variants were then partially purified and selected as explained above.

Proteins overexpression and purification

E. cloni® EXPRESS BL21 (DE3) cells were transformed with the pT7 plasmid containing the ClyA gene. Transformants were grown overnight at 37ºC on LB agar plates supplemented with 100 mg/L ampicillin. The resulting colonies were pooled together and innoculated into 50 mL of LB medium containing 100 mg/L of ampicillin. The culture was grown at 37ºC, with shaking at 200 rpm, until it reached an OD_{600} of 0.6 to 0.8. The expression of ClyA was then induced by the addition of 0.5 mM IPTG and the growth was continued at 20ºC. The next day the bacteria were harvested by centrifugation at 6000 x g for 10 min and the pellets were stored at -70ºC.

The pellets containing monomeric ClyA were thawed and resuspended in 20 mL of wash buffer (10 mM imidazole, 150 mM NaCl, 15 mM Tris.HCl, pH 7.5), supplemented with 1 mM $MgCl_2$ and 0.05 units/mL of DNAse I and the bacteria were lysed by sonication. The crude lysates were clarified by centrifugation at 6000 x g for 20 min and the supernatant was mixed with 200 µL of Ni-NTA resin (Qiagen) pre-equilibrated with wash buffer. After 1 h, the resin was loaded into a column (Micro Bio Spin, Bio-Rad) and washed with ~5 ml of the wash buffer. ClyA was eluted with approximately ~0.5 mL of wash buffer containing 300 mM imidazole. Protein concentration was determined by Bradford assay and the purified proteins were stored at 4ºC.

Construct of ClyA models

The 13mer and 14mer ClyA nanopores were modeled from the 12mer of the crystal structure (PDB code: 2WCD) as follow: A central axis was constructed through the length of the 12mer (ax12), and the distance between the C-alpha atom of residue 114 in the monomer A (114Ca-A) and the ax12 was measured giving the approximate radius of the pore (r12). The distance (d) between the 114Ca-A and the equivalent atom in monomer B (114Ca-B) was used to calculate the approximate circumference of the 12mer (c12 = d x 12), 13mer (c13 = d x 13) and 14mer (c14 = d x

14). The radius of the three oligomers (r12, r13 and r14) was then calculated from the circumference using simple trigonometry. The 12mer, 13mer and 14mer were then built by placing the monomers at distances r12, r13 and r14, respectively, from the central ax and rotated over an angle of 360°/12, 360°/13 and 360°/14, respectively. The 12mer that was built using this method reproduced perfectly the 12mer of the X-ray crystal structure (RMS=0.29 Å), showing the high degree of symmetry and feasibility to construct higher order pores.[20]

The size of the nanopore opening was obtained by increasing the Van der Waals radii of the atoms of ClyA until the pore closed. Then the increased value of the Van der Waals radii was taken as the radius of the pore. The diameter of thrombin was calculated from a sphere corresponding to the measured molecular volume of the protein.[20]

Electrical recordings in planar lipid bilayers.

The applied potential refers to the potential of the trans electrode. ClyA nanopores were inserted into lipid bilayers from the cis compartment, which was connected to the ground. The two compartments were separated by a 25-μm thick polytetrafluoroethylene film (Goodfellow Cambridge Limited) containing an orifice ~100 μm in diameter. The aperture was pretreated with ~5 μL of 10% hexadecane in pentane and a bilayer was formed by the addition of ~10 μL of 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC) in pentane (10 mg/mL) to both electrophysiology chambers. Typically, the addition of 0.01-0.1 ng of oligomeric ClyA to the cis compartment (0.5 mL) was sufficient to obtain a single channel. Electrical recordings were carried out in 150 mM NaCl, 15 mM Tris.HCl pH 7.5. The temperature of the recording chamber was maintained at 28°C by water circulating through a metal case in direct contact with the bottom and sides of the chamber.

Data recording and analysis

Electrical signals from planar bilayer recordings were amplified by using an Axopatch 200B patch clamp amplifier (Axon Instruments) and digitized with a Digidata 1440 A/D converter (Axon Instruments). Data were recorded by using Clampex 10.2 software (Molecular Devices) and the subsequent analysis was carried out with Clampfit software (Molecular Devices). Open pore currents and HT blockades were recorded by applying a 10 kHz low-pass Bessel filter and sampling at 50 kHz if not otherwise stated. For unitary channel conductance distributions data collection, traces were further filtered digitally with Gaussian low-pass filter with 200 Hz cutoff. The open pore currents were determined for all inserted channels at both +35 and -35 mV to ensure that the pore inserted with the correct orientation (the unitary conductance of ClyA is higher at positive applied potentials when reconstituted from the cis side) and values corresponding to -35 mV were used to construct

the distributions. Open pore current values (I_O) for ClyA and blocked pore current values (I_B) for HT were calculated from Gaussian fits to all points histograms (0.3 pA bin size).[3] Histograms for HT and BT blockades were prepared from at least 10 current blockade at least 0.5 s long. The residual current values (I_{RES}) were calculated as: $I_{RES\%} = I_B / I_O$ %. When HT produced two current levels within the same blockade, their relative contributions (Figure 4b, Table 1) were deduced from the area of the peaks obtained from Gaussian fits to the all points histogram.[3] HT blockade lifetimes were calculated by fitting the cumulative distribution of the block dwell times for at least 50 events to a single exponential.[3] From -5 to -20 mV HT blockade lifetimes were measured by applying a cyclic sweep voltage protocol consisting of 3 steps. In the first "capture" step, the applied potential was set to -60 mV for 2 seconds. In the second "release" step the applied potential was decreased to the voltage of interest (-5 to -20 mV) for 2-10 sec where HT released from the pore. Finally in the "regeneration" step the potential was briefly reversed to + 35 mV for 0.2 seconds to regenerate a new unblocked open pore state. At least 50 sweeps were averaged and the part of the trace corresponding to release step was fit to single exponential. The duration of the FP blockades (dwell times), which occasionally showed both level 1 and level 2 currents, were distributed over two orders of magnitude and were not fit well with exponential functions. Therefore, median dwell times are quoted for FP. The traces were recorded at a sampling rate of 20 kHz with an internal low-pass Bessel filter set at 5 kHz. The measurements were performed in 150 mM NaCl, 15 mM Tris.HCl, pH 7.5. Graphs were made with Origin (OriginLab Corporation) and the temperature set at 28°C. All values quoted in this work are based on the average of at least three separate recordings, unless otherwise specified.

ClyA pores for DNA translocation

DNA preparation

ssDNA molecules were purchased from Integrated DNA Technologies (IDT). **1** was made by PCR where one of the two primers was 5' biotinylated. **3** was formed by incubating two complementary ssDNA molecules, one of which contained a biotin moiety at the 3' end. The DNA hybrid was then purified from the excess ssDNA by affinity chromatography. **5** and **6** were HPLC purified by IDT.

Preparation of ClyA pores

ClyA was expressed in *E. coli* (DE3) pLysS cells by using a pT7 plasmid. Monomers containing a C-terminal oligo-histidine tag were expressed in *E. coli* cells and the soluble fraction purified by Ni-NTA affinity chromatography. ClyA dodecamers were formed by the addition of 0.2% β-dodecyl

maltoside (DDM), and were separated from monomers by blue native poly-acrylamide gel electrophoresis. The lowest band of oligomeric ClyA-CS was extracted and stored at 4ºC.

ClyA-**2** nanopores were prepared by covalently attaching **2** to a ClyA protein where the two WT cysteine residues (positions 87 and 285) were substituted with serine (ClyA-SS), and a cysteine was introduced at position 103 (aspartate in the WT gene; ClyA-SSC_{103}). ClyA-SSC_{103} was constructed from ClyA-SS, which also encoded a Gly-Ser-Ser linker followed by a C-terminal hexahistidine tag, by using the megaprimer method[7a] using Phire® Hot Start DNA Polymerase (Finnzymes). The DNA (5'-TTTTTTTTTATCTACGAATTCATCAGGGCTAAAGAGTGCAGAGTTACTTAG-3'), containing a protected thiol group attached to the 5' hydroxyl of the DNA strand via a C6 linker (5ThioMC6-D, IDT), was then conjugated to ClyA-SSC_{103} monomers, purified and oligomerised as described. Purified oligomers were stored at -80 °C in 20% glycerol.

DNA preparation

1 was made by PCR amplification of a pT7-ClyA-WT DNA template using a 5' biotinylated forward primer (bio-5' TAATACGACTCACTATAGGG-3') and a non-biotinylated reverse primer (5'-CATCAGCAGCACTTTGATATCGCCCACC-3') using Taq DNA Polymerase from REDTaq® ReadyMix™ PCR Reaction Mix (Sigma). After a maximum number of 35 cycles the PCR product of 24 reaction tubes (50 µL each tube) was purified by using a PCR quick purification kit (QIAGEN) and the size of the construct checked by using a 2% agarose gel (TAE buffer). The typical sample concentration was ~200 ng/µL.

3 was formed by incubating a 3' biotinylated ssDNA molecule (5'-GGATGACCTGATCCAGATATTTATTATACAGGTCCAGCGCACCGTCAGCCCAATCGCACTTTTCACAAAAAGA GAGAGAGATCGATTACC-3'-bio, **3a**) with a 20% excess of a partially complimentary ssDNA (5'-GGTAATCGATCTCTCTCTCTTTTTGTGAAAAGTGCGATTGGGCTGACGGTGCGCTGGAC-3', **3b**, Table 4). The temperature was brought to 95°C for one minute and then decreased stepwise to room temperature. At around the estimated annealing temperature, the temperature was decreased in 2°C steps, each held for one minute. The hybrid DNA was then purified from the excess of ssDNA by affinity chromatography, using a biotin-binding column containing monomeric avidin immobilised on agarose beads (Thermo Scientific Pierce). **3** was eluted in Biotin Blocking/Elution Buffer according to the protocol. Typically a DNA concentration of ~400 ng/µL was obtained. The size of the dsDNA was checked by using a 2% agarose gel in TAE buffer. The purified dsDNA was stored at -20 °C in the presence of 1 µM EDTA.

Electrical recordings

If not otherwise specified, the signal was collected at sampling rate of 50 KHz using a 10 kHz Bessel filter. The lipid bilayer was formed by the addition of 1 to 2 μL of a 10% solution of 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC) in pentane (w/v). The electrical potential was applied by using Ag/AgCl electrodes submerged in agar bridges (3% w/v low melt agarose in 2.5 M NaCl buffer). The applied potential refers to the potential of the working electrode connected to the *trans* compartment of the chamber. ClyA nanopore solutions (0.01-0.1 ng) were added to the *cis* compartment, which was connected to the ground electrode. After the insertion of a single channel, excess protein was removed by several cycles of perfusion. Electrical recordings were carried out in 2.5 M NaCl, 15 mM Tris.HCl pH 8.0, at 22°C. The errors indicate the standard deviation from the average for at least three independent repeats, which are indicated with the letter *n*.

INCORPORATION BY REFERENCE

All publications and patents mentioned herein are hereby incorporated by reference in their entirety as if each individual publication or patent was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

EQUIVALENTS

While specific embodiments of the subject invention have been discussed, the above specification is illustrative and not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of this specification and the claims below. The full scope of the invention should be determined by reference to the claims, along with their full scope of equivalents, and the specification, along with such variations.

REFERENCES

1. (a) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W., Characterization of individual polynucleotide molecules using a membrane channel. Proc Natl Acad Sci U S A 1996, 93 (24), 13770-3; (b) Vercoutere, W.; Winters-Hilt, S.; Olsen, H.; Deamer, D.; Haussler, D.; Akeson, M., Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. Nat Biotechnol 2001, 19 (3), 248-52; (c) Howorka, S.; Cheley, S.; Bayley, H., Sequence-specific detection of individual DNA strands using engineered nanopores. Nat Biotechnol 2001, 19 (7), 636-9.

2. (a) Dekker, C., Solid-state nanopores. Nat Nanotechnol 2007, 2 (4), 209-15; (b) Li, J.; Stein, D.; McMullan, C.; Branton, D.; Aziz, M. J.; Golovchenko, J. A., Ion-beam sculpting at nanometre length scales. Nature 2001, 412 (6843), 166-9.

3. (a) Wei, R.; Martin, T. G.; Rant, U.; Dietz, H., DNA origami gatekeepers for solid-state nanopores. Angew Chem Int Ed Engl 2012, 51 (20), 4864-7; (b) Bell, N. A.; Engst, C. R.; Ablay, M.; Divitini, G.; Ducati, C.; Liedl, T.; Keyser, U. F., DNA origami nanopores. Nano Lett 2012, 12 (1), 512-7; (c) Langecker, M.; Arnaut, V.; Martin, T. G.; List, J.; Renner, S.; Mayer, M.; Dietz, H.; Simmel, F. C., Synthetic lipid membrane channels formed by designed DNA nanostructures. Science 2012, 338 (6109), 932-6.

4. Hall; Scott, A.; Rotem, D.; Mehta, K.; Bayley, H.; Dekker, C., Hybrid pore formation by directed insertion of alpha hemolysin into solid-state nanopores. Nature Nanotechnology 2011, In press.

5. Venkatesan, B. M.; Bashir, R., Nanopore sensors for nucleic acid analysis. Nat Nanotechnol 2011, 6 (10), 615-24.

6. (a) Jung, Y.; Bayley, H.; Movileanu, L., Temperature-responsive protein pores. J Am Chem Soc 2006, 128 (47), 15332-40; (b) Heinz, C.; Engelhardt, H.; Niederweis, M., The core of the tetrameric mycobacterial porin MspA is an extremely stable beta-sheet domain. J Biol Chem 2003, 278 (10), 8678-85.

7. (a) Soskine, M.; Biesemans, A.; Moeyaert, B.; Cheley, S.; Bayley, H.; Maglia, G., An Engineered ClyA Nanopore Detects Folded Target Proteins by Selective External Association and Pore Entry. Nano Lett 2012, 12 (9), 4895-900; (b) Franceschini, L.; Mikhailova, E.; Bayley, H.; Maglia, G., Nucleobase recognition at alkaline pH and apparent pKa of single DNA bases immobilised within a biological nanopore. Chem Commun (Camb) 2012, 48 (10), 1520-2; (c) Maglia, M.; Henricus, M.; Wyss, R.; Li, Q.; Cheley, S.; Bayley, H., DNA strands from denatured duplexes are translocated through engineered protein nanopores at alkaline pH. Nano Letters 2009, 9, 3831-3836.

8. (a) Pastoriza-Gallego, M.; Oukhaled, G.; Mathe, J.; Thiebot, B.; Betton, J. M.; Auvray, L. C.; Pelta, J., Urea denaturation of alpha-hemolysin pore inserted in planar lipid bilayer detected by single nanopore recording: Loss of structural asymmetry. FEBS Lett 2007, 581 (18), 3371-3376; (b) Japrung, D.; Henricus, M.; Li, Q. H.; Maglia, G.; Bayley, H., Urea Facilitates the Translocation of Single-Stranded DNA and RNA Through the alpha-Hemolysin Nanopore. Biophysical Journal 2010, 98 (9), 1856-1863.

9. (a) Mohammad, M. M.; Iyer, R.; Howard, K. R.; McPike, M. P.; Borer, P. N.; Movileanu, L., Engineering a Rigid Protein Tunnel for Biomolecular Detection. J Am Chem Soc 2012; (b) Bikwemu, R.; Wolfe, A. J.; Xing, X.; Movileanu, L., Facilitated translocation of polypeptides through a single nanopore. J Phys Condens Matter 2010, 22 (45), 454117; (c) Wolfe, A. J.; Mohammad, M. M.; Cheley, S.; Bayley, H.; Movileanu, L., Catalyzing the translocation of polypeptides through attractive interactions. Journal of the American Chemical Society 2007, 129 (45), 14034-14041; (d) Movileanu, L.; Schmittschmitt, J. P.; Scholtz, J. M.; Bayley, H., Interactions of peptides with a protein pore. Biophys J 2005, 89 (2), 1030-45; (e) Payet, L.; Martinho, M.; Pastoriza-Gallego, M.; Betton, J. M.; Auvray, L.; Pelta, J.; Mathe, J., Thermal unfolding of proteins probed at the single molecule level using nanopores. Anal Chem 2012, 84 (9), 4071-6; (f) Pastoriza-Gallego, M.; Rabah, L.; Gibrat, G.; Thiebot, B.; van der Goot, F. G.; Auvray, L.; Betton, J. M.; Pelta, J., Dynamics of unfolded protein transport through an aerolysin pore. J Am Chem Soc 2011, 133 (9), 2923-31; (g) Oukhaled, G.; Mathe´, J.; Biance, A.-L.; Bacri, L.; Betton, J.-M.; Lairez, D.; Pelta, J.; Auvray, L., Unfolding of Proteins and Long Transient Conformations Detected by Single Nanopore Recording. Phys. Rev. Lett. 2007, 98, 158101; (h) Stefureac, R. I.; Kachayev, A.; Lee, J. S., Modulation of the translocation of peptides through nanopores by the application of an AC electric field. Chem Commun (Camb) 2012, 48 (13), 1928-30; (i) Stefureac, R. I.; Lee, J. S., Nanopore analysis of the folding of zinc fingers. Small 2008, 4 (10), 1646-50; (j) Stefureac, R.; Waldner, L.; Howard, P.; Lee, J. S., Nanopore analysis of a small 86-residue protein. Small 2008, 4 (1), 59-63; (k) Stefureac, R.; Long, Y. T.; Kraatz, H. B.; Howard, P.; Lee, J. S., Transport of alpha-helical peptides through alpha-hemolysin and aerolysin pores. Biochemistry 2006, 45 (30), 9172-9.

10. (a) Firnkes, M.; Pedone, D.; Knezevic, J.; Doblinger, M.; Rant, U., Electrically Facilitated Translocations of Proteins through Silicon Nitride Nanopores: Conjoint and Competitive Action of Diffusion, Electrophoresis, and Electroosmosis. Nano Letters 2010, 10 (6), 2162-2167; (b) Plesa, C.; Kowalczyk, S. W.; Zinsmeester, R.; Grosberg, A. Y.; Rabin, Y.; Dekker, C., Fast Translocation of Proteins through Solid State Nanopores. Nano Lett 2013; (c) Niedzwiecki, D. J.; Grazul, J.; Movileanu, L., Single-Molecule Observation of Protein Adsorption onto an Inorganic Surface. Journal of the American Chemical Society 2010, 132 (31), 10816-10822; (d) Fologea, D.; Ledden, B.; McNabb, D. S.; Li, J. L., Electrical characterization of protein molecules by a solid-state nanopore. Applied Physics Letters 2007, 91 (5); (e) Han, A.; Creus, M.; Schurmann, G.; Linder, V.; Ward, T. R.; de Rooij, N. F.; Staufer, U., Label-free detection of single protein molecules and protein-protein interactions using synthetic nanopores. Anal Chem 2008, 80 (12), 4651-8; (f) Stefureac, R. I.; Trivedi, D.;

Marziali, A.; Lee, J. S., Evidence that small proteins translocate through silicon nitride pores in a folded conformation. J Phys Condens Matter 2010, 22 (45), 454133.

12. Eifler, N.; Vetsch, M.; Gregorini, M.; Ringler, P.; Chami, M.; Philippsen, A.; Fritz, A.; Muller, S. A.; Glockshuber, R.; Engel, A.; Grauschopf, U., Cytotoxin ClyA from Escherichia coli assembles to a 13-meric pore independent of its redox-state. EMBO J 2006, 25 (11), 2652-61.

13. (a) Pogoryelov, D.; Klyszejko, A. L.; Krasnoselska, G. O.; Heller, E. M.; Leone, V.; Langer, J. D.; Vonck, J.; Muller, D. J.; Faraldo-Gomez, J. D.; Meier, T., Engineering rotor ring stoichiometries in the ATP synthase. Proc Natl Acad Sci U S A 2012, 109 (25), E1599-608; (b) Bayfield, O. W.; Chen, C. S.; Patterson, A. R.; Luan, W.; Smits, C.; Gollnick, P.; Antson, A. A., Trp RNA-binding attenuation protein: modifying symmetry and stability of a circular oligomer. PLoS One 2012, 7 (9), e44309.

14. Niedzwiecki, D. J.; Iyer, R.; Borer, P. N.; Movileanu, L., Sampling a Biomarker of the Human Immunodeficiency Virus across a Synthetic Nanopore. ACS Nano 2013.

15. (a) Clarke, J.; Wu, H.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H., Continuous base identification for single-molecule nanopore DNA sequencing. Nature Nanotechnology 2009, 4, 265-270; (b) Rincon-Restrepo, M.; Mikhailova, E.; Bayley, H.; Maglia, G., Controlled Translocation of Individual DNA Molecules through Protein Nanopores with Engineered Molecular Brakes. Nano Lett 2011, 11 (2), 746-50.

16. (a) Freedman, K. J.; Jurgens, M.; Prabhu, A.; Ahn, C. W.; Jemth, P.; Edel, J. B.; Kim, M. J., Chemical, thermal, and electric field induced unfolding of single protein molecules studied using nanopores. Anal Chem 2011, 83 (13), 5137-44; (b) Niedzwiecki, D. J.; Movileanu, L., Monitoring protein adsorption with solid-state nanopores. J Vis Exp 2011, (58); (c) Talaga, D. S.; Li, J., Single-molecule protein unfolding in solid state nanopores. J Am Chem Soc 2009, 131 (26), 9287-97.

17. Mueller, M.; Grauschopf, U.; Maier, T.; Glockshuber, R.; Ban, N., The structure of a cytolytic alpha-helical toxin pore reveals its assembly mechanism. Nature 2009, 459 (7247), 726-U135.

18. Maglia, G.; Heron, A. J.; Stoddart, D.; Japrung, D.; Bayley, H., Analysis of Single Nucleic Acid Molecules with Protein Nanopores. Methods in Enzymology, Vol 475: Single Molecule Tools, Pt B 2010, 474, 591-623.

19. Miyazaki, K., MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. Methods Enzymol 2011, 498, 399-406.

20. Delhaise, P.; Bardiaux, M.; Demaeyer, M.; Prevost, M.; Vanbelle, D.; Donneux, J.; Lasters, I.; Vancustem, E.; Alard, P.; Wodak, S. J., The Brugel Package - toward Computer-Aided-Design of Macromolecules. J Mol Graphics 1988, 6 (4), 219-219.

21. King, N.P. et al. Computational design of self-assembling protein nanomaterials with atomic level accuracy. Science 336, 1171-1174 (2012).

22. Ackermann, D. et al. A double-stranded DNA rotaxane. Nat Nanotechnol 5, 436-442 (2010).

23. Sanchez-Quesada, J., Saghatelian, A., Cheley, S., Bayley, H. & Ghadiri, M.R. Single DNA rotaxanes of a transmembrane pore protein. Angew Chem Int Ed Engl 43, 3063-3067 (2004).

24. Meller, A., Nivon, L. & Branton, D. Voltage-driven DNA translocations through a nanopore. Phys Rev Lett 86, 3435-3438 (2001).

25. 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).

26. Maglia, G., Rincon Restrepo, M., Mikhailova, E. & Bayley, H. Enhanced translocation of single DNA molecules through α-hemolysin nanopores by manipulation of internal charge. Proc Natl Acad Sci U S A 105, 19720-19725 (2008).

27. 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).

28. Wanunu, M., Morrison, W., Rabin, Y., Grosberg, A.Y. & Meller, A. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. Nat Nanotechnol 5, 160-165 (2010).

TABLES

Table 1. Parameters for ClyA-SS and the three types of ClyA-CS nanopores. The diameter of Type I ClyA-SS and ClyA-CS was taken from the crystal structure of *E. coli* ClyA. The diameter of Type II and Type III ClyA-CS was measured from models that were created by adding one (Type II) or two (Type III) subunits to the structure of the 12mer ClyA (crystal structure) as described in the supplementary information. The diameters of the nanopores were determined including the Van der Waals radii of

the atoms (supplementary information). Errors are given as standard deviations. (*Data taken from[7a])

Parameters	Type I* ClyA-SS	Type I ClyA-CS	Type II ClyA-CS	Type III ClyA-CS
Trans diameter, nm	3.3	3.3	3.7	4.2
Cis diameter, nm	5.5	5.5	5.9	6.5
Nanopore conductance at -35 mV, nS	1.8 ± 0.1	1.78 ± 0.04	2.19 ± 0.09	2.81 ± 0.11
Nanopore conductance at -150mV, nS	NA	1.50 ± 0.03	1.85 ± 0.06	NA
HT Occupancy of L2 at -35 mV, %	22±5	30 ± 10	96 ± 2	100
HT Level 1 at -35 mV, I_{RES}%	56 ± 1	56 ± 1	68 ± 1	NA
HT Level 2 at -35 mV, I_{RES}%	23±1	23 ± 3	31 ± 1	32±9
HT Occupancy of L2 at -150 mV, %	NA	100	100	NA
HT Level 2 (-150 mV), I_{RES}, %	NA	23 ± 2	31 ± 5	NA
HT Dwell time at -150 mV, ms	NA	1.0 ± 0.4	235 ± 186	NA

Table 2. Mutations accumulated during the directed evolution rounds of the ClyA-SS gene.

Round	Name	Clone ID	Sequence changes relative to WT-ClyA
0	**ClyA-SS**	dSClyA	**C87S, C285S**
3		3ClyA1	**C87S**, F166Y, K230R, **C285S**
3		3ClyA2	Q73R, F166Y, **C285S**
3		3ClyA3	Q33R,Q56H, **C87S**,D122G, **C285S**
4		4ClyA1	I4T,N128S,S145I,**C285S**
4		4ClyA2	S110I, **C285S,** F166Y, T223A
4		4ClyA3	T39I,**C285S**, F166Y, K230R
4	**ClyA-CS**	*4ClyA4*	L99Q,E103G, F166Y, **C285S**, K294R
4		4ClyA5	I4T, Q73R, **C285S**
4		4ClyA6	Q73R, F166Y, **C285S**
5		5ClyA1	**C87A,** L99Q,E103G, **C285S**, F166Y,N220S, Q289R, K294R, H307Y
5		5ClyA2	**C87A,** L99Q,E103G, **C285S**, F166Y, Q289R, K294R, H307Y
5	**ClyA-AS**	*5ClyA3*	**C87A,** L99Q,E103G, **C285S**, F166Y, I203V, K294R, H307Y

Table 3. Primers: N stands for A, G, C, or T; S is G or C, thus NNS codon encodes for the full set of amino acids.

Name	Sequence	SEQ ID NO
87NNS	GAAGCTACCCAAACGGTTTACGAATGGNNSGGTGTGGTTACCCAGCTGCT G	17
T7 promoter	TAATACGACTCACTATAGGG	18
T7 terminator	GCTAGTTATTGCTCAGCGG	19
87C	GTTTACGAATGGTGTGGTGTGGTTACCCAG	20
285C	CGCTGCTGATATTCATTACAGGTATTAATCATTTTC	21

Table 4. Summary of DNA molecules used in this work. **1** was prepared by PCR as described in methods. **3** was formed by incubating **3a** with a 20% excess of **3b** and purified by affinity chromatogtraphy as described in methods. The complimentary sequences in the two DNA strands are shown in italics.

Name		DNA sequence	SEQ ID NO
1	1a	Bio-5'TAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGACGGGTATCTTTGCGGAACAGACGGTGGAAGTTGTGAAAAGTGCGATTGAAACGGCTGACGGTGCGCTGGACCTGTATAATAAATATCTGGATCAGGTCATCCCGTGGAAAACCTTTGACGAAACGATTAAAGAACTGAGCCGTTTCAAACAGGAATACAGTCAAGAAGCGTCCGTCCTGGTGGGCGATATCAAAGTGCTGCTGATG3'	7
	1b	5'CATCAGCAGCACTTTGATATCGCCCACCAGGACGGACGCTTCTTGACTGTATTCCTGTTTGAAACGGCTCAGTTCTTTAATCGTTTCGTCAAAGGTTTTCCACGGGATGACCTGATCCAGATATTTATTATACAGGTCCAGCGCACCGTCAGCCGTTTCAATCGCACTTTTCACAACTTCCACCGTCTGTTCCGCAAAGATACCCGTCATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGAAACCGTTGTGGTCTCCCTATAGTGAGTCGTATTA3'	8
2		5'TTTTTTTTTATCTACGAATTCATCAGGGCTAAAGAGTGCAGAGTTACTTAG3'	9
3	3a	5'GGATGACCTGATCCAGATATTTATTATACA*GGTCCAGCGCACCGTCAGCCCAATCGCACTTTTCACAAAAAGAGAGAGAGATCGATTACC*3'-bio	10
	3b	5'*GGTAATCGATCTCTCTCTCTTTTTGTGAAAAGTGCGATTGGGCTGACGGTGCGCTGGAC*-3'	11
4		5'AATAAATATCTGGATCAGGTCATCCCTAAGTAACTCTGCAC3'	12
5a		5'GGATGACCTGATCCAGATATTTATTATACAGGTCCAGCGCACCGTCAGCCCAATCGCACTTTTCACAAAAAGAGAGAGAGATCGATTACC3'bio	13
5b		5'GGTAATCGATCTCTCTCTCTTTTTGTGAAAAGTGCGATTGGGCTGACGGTGCGCTGGACCTGTATAATAAATATCTGGATCAGGTCATCC3'	14
6		5'GCCCTATATTATCAGGTCATCCCTAAGTAACTCTGCA3'	15

Name	DNA sequence	SEQ ID NO
7	5'TGCAGAGTTACTTAGGGATGACCTGATAATATAGGGC3'	16

ClyA SEQUENCES

Description	Sequence	SEQ ID NO
Protein sequence for *S. typhi* ClyA (ClyA-WT)	MIMTGIFAEQTVEVVKSAIETADGALDLYNKYLDQVIPWKTFDETIKEL SRFKQEYSQEASVLVGDIKVLLMDSQDKYFEATQTVYEWCGVVTQLLS AYILLFDEYNEKKASAQKDILIRILDDGVKKLNEAQKSLLTSSQSFNNASG KLLALDSQLTNDFSEKSSYFQSQVDRIRKEAYAGAAAGIVAGPFGLIISYS IAAGVIEGKLIPELNNRLKTVQNFFTSLSATVKQANKDIDAAKLKLATEIA AIGEIKTETETTRFYVDYDDLMLSLLKGAAKKMINTCNEYQQRHGKKTL FEVPDV	SEQ ID NO:1
Protein sequence for ClyA with C285S substitution (ClyA-CS)	MTGIFAEQTVEVVKSAIETADGALDLYNKYLDQVIPWKTFDETIKELSRF KQEYSQEASVLVGDIKVLLMDSQDKYFEATQTVYEWCGVVTQLLSAYI QLFDGYNEKKASAQKDILIRILDDGVKKLNEAQKSLLTSSQSFNNASGKL LALDSQLTNDFSEKSSYYQSQVDRIRKEAYAGAAAGIVAGPFGLIISYSIA AGVIEGKLIPELNNRLKTVQNFFTSLSATVKQANKDIDAAKLKLATEIAAI GEIKTETETTRFYVDYDDLMLSLLKGAAKKMINTSNEYQQRHGRKTLFE VPDVGSSHHHHHH*	SEQ ID NO:2
Protein sequence for ClyA-AS	MTGIFAEQTVEVVKSAIETADGALDLYNKYLDQVIPWKTFDETIKELSRF KQEYSQEASVLVGDIKVLLMDSQDKYFEATQTVYEWAGVVTQLLSAYI QLFDGYNEKKASAQKDILIRILDDGVKKLNEAQKSLLTSSQSFNNASGKL LALDSQLTNDFSEKSSYYQSQVDRIRKEAYAGAAAGIVAGPFGLIISYSIA AGVVEGKLIPELNNRLKTVQNFFTSLSATVKQANKDIDAAKLKLATEIAA IGEIKTETETTRFYVDYDDLMLSLLKGAAKKMINTSNEYQQRHGRKTLF EVPDVGSSYHHHHH*	SEQ ID NO:3
Nucleotide sequence for *S. typhi* ClyA	CCTGCGTAGATAAGCAGGAAGCAGGCAGTATTTCCAGCTTCTGGAA TGTTAAAGCTACAAAAGTTGTCTGGAGGTAATAGGTAAGAATACTT TATAAAACAGGTACTTAATTGCAATTTATATATTTAAAGAGGCAAAT GATTATGACCGGAATATTTGCAGAACAAACTGTAGAGGTAGTTAAA AGCGCGATCGAAACCGCAGATGGGGCATTAGATCTTTATAACAAAT ACCTCGACCAGGTCATCCCCTGGAAGACCTTTGATGAAACCATAAA AGAGTTAAGCCGTTTTAAACAGGAGTACTCGCAGGAAGCTTCTGTT TTAGTTGGTGATATTAAAGTTTTGCTTATGGACAGCCAGGACAAGT ATTTTGAAGCGACACAAACTGTTTATGAATGGTGTGGTGTCGTGAC GCAATTACTCTCAGCGTATATTTTACTATTTGATGAATATAATGAGA AAAAAGCATCAGCCCAGAAAGACATTCTCATTAGGATATTAGATGA TGGTGTCAAGAAACTGAATGAAGCGCAAAAATCTCTCCTGACAAGT TCACAAAGTTTCAACAACGCTTCCGGAAAACTGCTGGCATTAGATA GCCAGTTAACTAATGATTTTTCGGAAAAAAGTAGTTATTTCCAGTCA	SEQ ID NO: 4
	CAGGTGGATAGAATTCGTAAGGAAGCTTATGCCGGTGCTGCAGCC GGCATAGTCGCCGGTCCGTTTGGATTAATTATTTCCTATTCTATTGCT GCGGGCGTGATTGAAGGGAAATTGATTCCAGAATTGAATAACAGG CTAAAAACAGTGCAAAATTTCTTTACTAGCTTATCAGCTACAGTGAA ACAAGCGAATAAAGATATCGATGCGGCAAAATTGAAATTAGCCACT GAAATAGCAGCAATTGGGGAGATAAAAACGGAAACCGAAACAACC AGATTCTACGTTGATTATGATGATTTAATGCTTTCTTTATTAAAAGG AGCTGCAAAGAAAATGATTAACACCTGTAATGAATACCAACAAAGA CACGGTAAGAAGACGCTTTTCGAGGTTCCTGACGTCTGATACATTTT CATTCGATCTGTGTACTTTTAACGCCCGATAGCGTAAAGAAAATGA GAGACGGAGAAAAAGCGATATTCAACAGCCCGATAAACAAGAGTC	
	GTTACCGGGCTGACGAGGTTATCAGGCGTTAAGCTGGTAG	

Description	Sequence	SEQ ID NO
Nucleotide sequence for ClyA with C285S substitution (ClyA-CS)	ATGACGGGTATCTTTGCGGAACAGACGGTGGAAGTTGTGAAAAGT GCGATTGAAACGGCTGACGGTGCGCTGGACCTGTATAATAAATATC TGGATCAGGTCATCCCGTGGAAAACCTTTGACGAAACGATTAAAGA ACTGAGCCGTTTCAAACAGGAATACAGTCAAGAAGCGTCCGTCCTG GTGGGCGATATCAAAGTGCTGCTGATGGATTCTCAGGACAAATATT TTGAAGCTACCCAAACGGTTTACGAATGGTGTGGTGTGGTTACCCA GCTGCTGTCCGCATATATTCAGCTGTTCGATGGATACAACGAGAAA AAAGCGAGCGCGCAGAAAGACATTCTGATCCGCATTCTGGATGACG GCGTGAAAAAACTGAATGAAGCCCAGAAATCGCTGCTGACCAGCTC TCAATCATTTAACAATGCCTCGGGTAAACTGCTGGCACTGGATAGCC AGCTGACGAACGACTTTTCTGAAAAAAGTTCCTATTACCAGAGCCA AGTCGATCGTATTCGTAAAGAAGCCTACGCAGGTGCCGCAGCAGGT ATTGTGGCCGGTCCGTTCGGTCTGATTATCTCATATTCGATTGCTGC GGGCGTTATCGAAGGTAAACTGATTCCGGAACTGAACAATCGTCTG AAAACCGTTCAGAACTTTTTCACCAGTCTGTCTGCTACGGTCAAACA AGCGAATAAAGATATCGACGCCGCAAAACTGAAACTGGCCACGGA AATCGCTGCGATTGGCGAAATCAAAACCGAAACGGAAACCACGCG CTTTTATGTTGATTACGATGACCTGATGCTGAGCCTGCTGAAAGGTG CCGCGAAGAAAATGATTAATACCTCTAATGAATATCAGCAGCGTCA CGGTAGAAAAACCCTGTTTGAAGTCCCGGATGTGGGCAGCAGCCAC CACCATCATCACCACTAAAAGCTTGGATCCGGCTGCTAACAAAGCCC GAA	SEQ ID NO: 5
Nucleotide sequence for ClyA-AS	ATGACGGGTATCTTTGCGGAACAGACGGTGGAAGTTGTGAAAAGT GCGATTGAAACGGCTGACGGTGCGCTGGACCTGTATAATAAATATC TGGATCAGGTCATCCCGTGGAAAACCTTTGACGAAACGATTAAAGA ACTGAGCCGTTTCAAACAGGAATACAGTCAAGAAGCGTCCGTCCTA GTGGGCGATATCAAAGTGCTGCTGATGGATTCTCAGGACAAATATT TTGAAGCTACCCAAACGGTTTACGAATGGGCGGGTGTGGTTACCCA GCTGCTGTCCGCATATATTCAGCTGTTCGATGGATACAATGAGAAA AAAGCGAGCGCGCAGAAAGACATTCTGATCCGCATTCTGGATGACG GCGTGAAAAAACTGAATGAAGCCCAGAAATCGCTGCTGACCAGCTC TCAATCATTTAACAATGCCTCGGGTAAACTGCTGGCACTGGATAGCC AGCTGACGAACGACTTTTCTGAAAAAAGTTCCTATTACCAGAGCCA AGTCGATCGTATTCGTAAAGAAGCCTACGCAGGTGCCGCAGCAGGT ATTGTGGCCGGTCCGTTCGGTCTGATTATCTCATATTCAATTGCTGC GGGCGTTGTCGAAGGTAAACTGATTCCGGAACTGAACAATCGTCTG AAAACCGTTCAGAACTTTTTCACCAGTCTGTCTGCTACGGTCAAACA AGCGAATAAAGATATCGACGCCGCAAAACTGAAACTGGCCACGGA AATCGCTGCGATTGGCGAAATCAAAACCGAAACGGAAACCACGCG CTTTTATGTTGATTACGATGACCTGATGCTGAGCCTGCTGAAAGGTG CCGCGAAGAAAATGATTAATACCTCTAATGAATATCAGCAGCGTCA CGGTAGAAAAACCCTGTTTGAAGTCCCGGATGTGGGCAGCAGCTAC CACCATCATCACCACTAAAAGCTT	SEQ ID NO:6

CLAIMS

1. A modified ClyA pore comprising a plurality of subunits, wherein each subunit comprises a polypeptide represented by an amino acid sequence at least 80% identical to SEQ ID NO:1 wherein exactly one Cys residue is substituted with Ser.

2. The modified ClyA pore of claim 1, wherein the Cys residue is C285.

3. The modified ClyA pore of claim 1 or claim 2, wherein each subunit comprises at least one additional amino acid substitution selected from L99Q, E103G, F166Y, and K294R.

4. The modified ClyA pore of any one of claims 1-3, wherein each subunit comprises a polypeptide represented by an amino acid sequence of SEQ ID NO:2 .

5. The modified ClyA pore of any one of claims 1-3, wherein exactly one Cys residue is substituted with Ala.

6. The modified ClyA pore of claim 5, wherein each subunit comprises at least one additional amino acid substitution selected from L203V and H207Y.

7. The modified ClyA pore of claim 5 or 6, wherein each subunit comprises a polypeptide represented by an amino acid sequence of SEQ ID NO:3 .

8. The modified ClyA pore of any one of claims 1-7, wherein the modified ClyA pore comprises at least 12 subunits.

9. The modified ClyA pore of claim 8, wherein the modified ClyA pore comprises 12 subunits.

10. The modified ClyA pore of claim 8, wherein the modified ClyA pore comprises 13 subunits.

11. The modified ClyA pore of any one of claims 1-10, wherein the modified ClyA pore has a cis diameter of at least 3.5 nm.

12. The modified ClyA pore of any one of claims 1-11, wherein the modified ClyA pore has a trans diameter of at least 6 nm.

13. The modified ClyA pore of any one of claims 1-12, wherein the modified ClyA pore remains open when the voltage across the modified ClyA pore ranges from -60 +90 to -150mV .

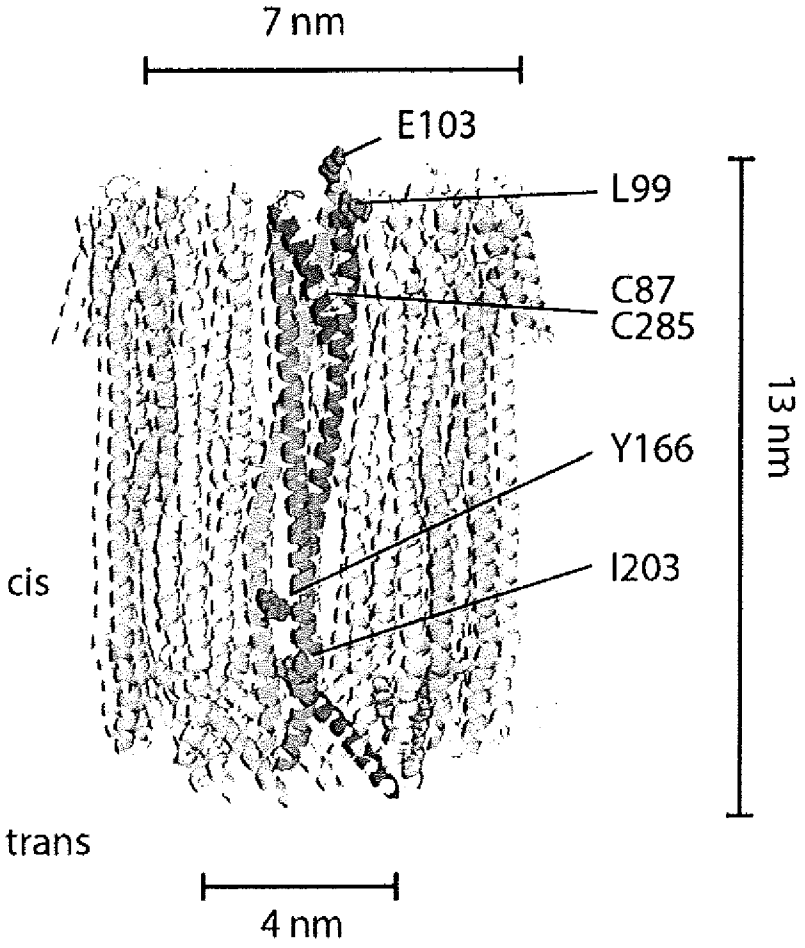
14. The modified ClyA pore of any one of claims 1-13, wherein a protein analyte binds within the lumen of the modified ClyA pore.

15. The modified ClyA pore of any one of claims 1-14, wherein the protein analyte binds to more than one site within the lumen of the modified ClyA pore.

16. The modified Cly A pore of claim 15, wherein the protein analyte is a protein with a molecular weight in the range of 15-70 kDa.

Figure 1

a)

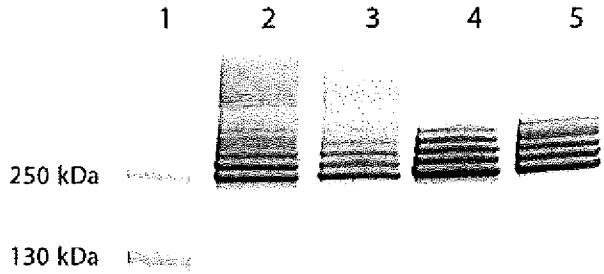


b)

Round	Name	Sequence changes
0	ClyA-SS	C87S, C285S
4	ClyA-CS	L99Q, E103G, F166Y, C285S, K294R
5	ClyA-AS	C87A, L99Q, E103G, F166Y, I203V, C285S, K294R, H307Y

Figure 2

a)


1
2
3
4
5
250 kDa
130 kDa

b)

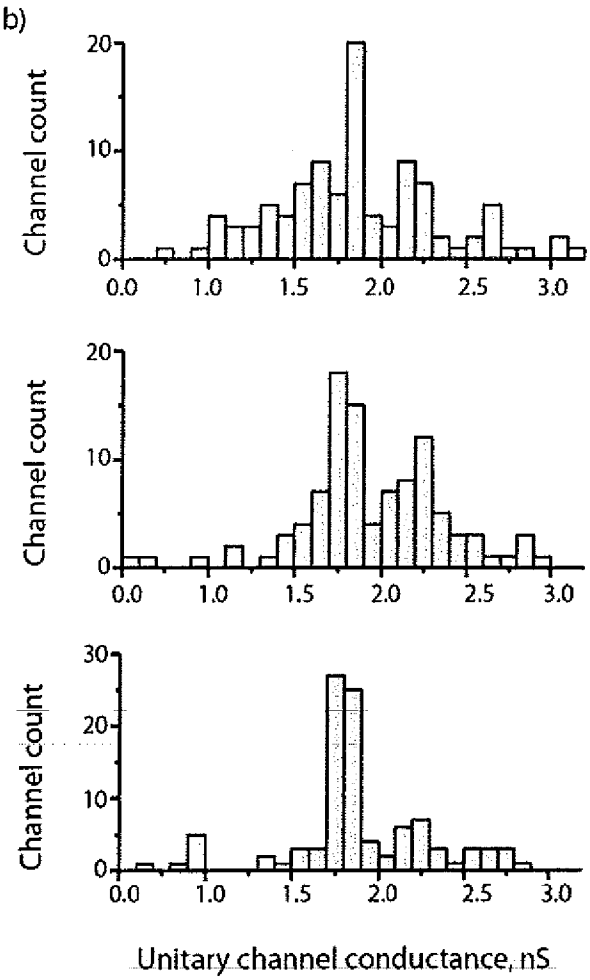
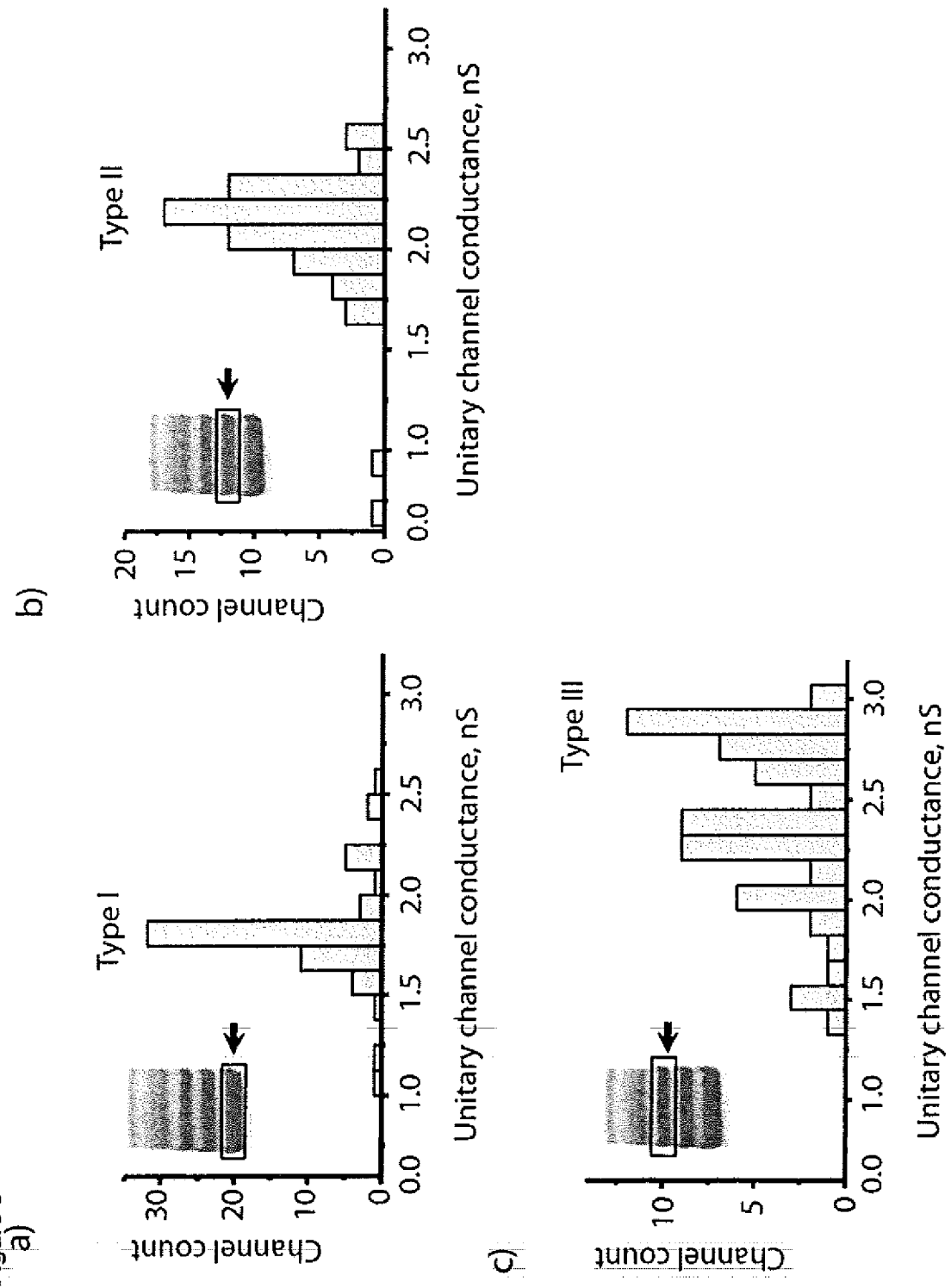

Channel count
20
10
0
0.0
1.0
1.5
2.0
2.5
3.0
Channel count
20
10
0
0.0
1.0
1.5
2.0
2.5
3.0
Channel count
30
20
10
0
0.0
1.0
1.5
2.0
2.5
3.0
Unitary channel conductance, nS

Figure 3


a)
Type I
Channel count
Unitary channel conductance, nS
b)
Type II
Channel count
Unitary channel conductance, nS
c)
Type III
Channel count
Unitary channel conductance, nS

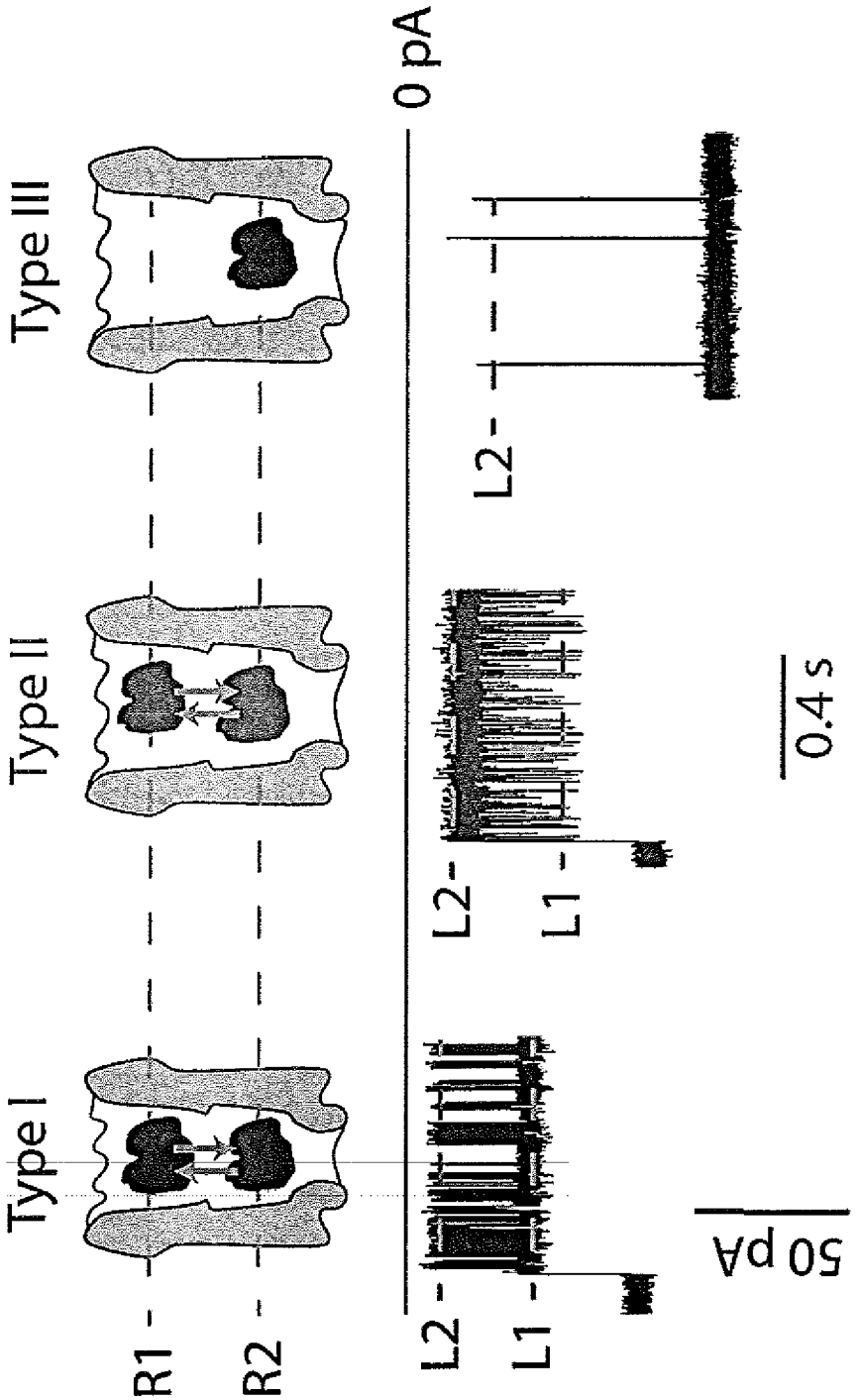


Figure 4

Figure 5

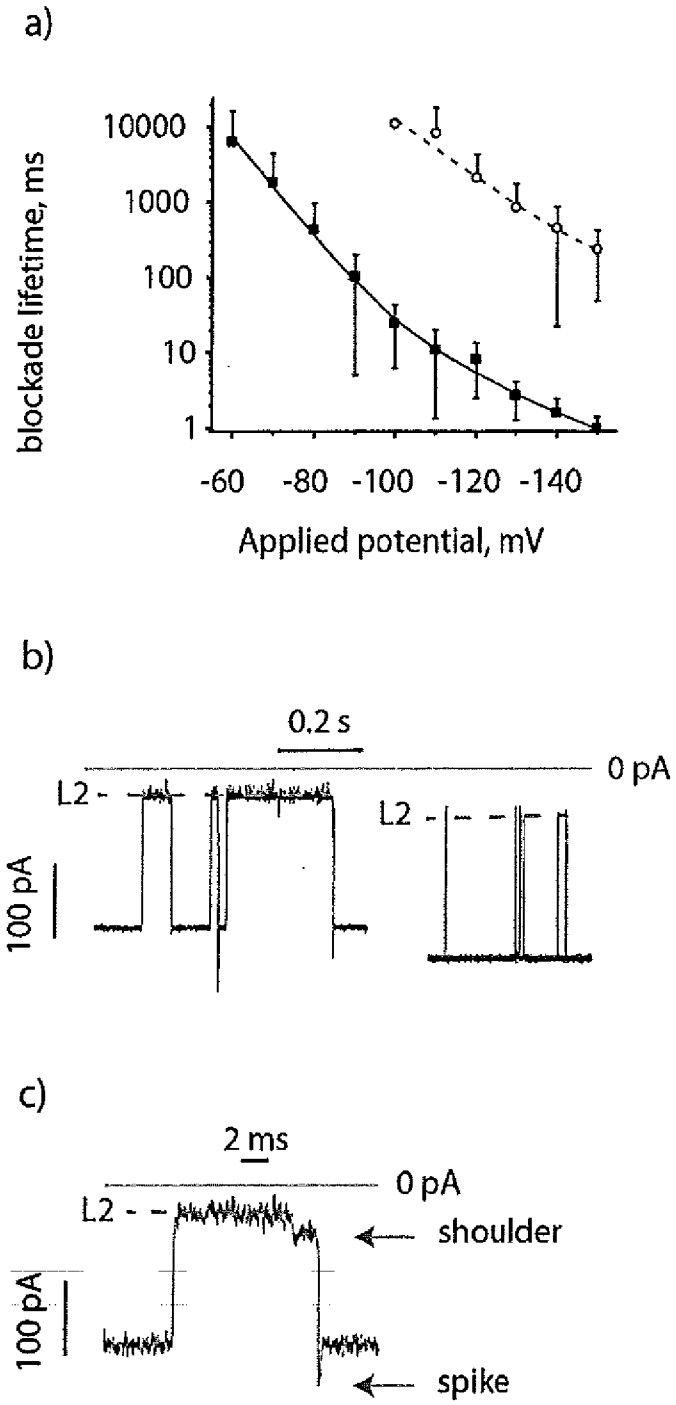

a)
blockade lifetime, ms
10000
1000
100
10
1
-60 -80 -100 -120 -140
Applied potential, mV
b)
0.2 s
0 pA
L2
L2
100 pA
c)
2 ms
0 pA
L2
shoulder
100 pA
spike

Figure 6

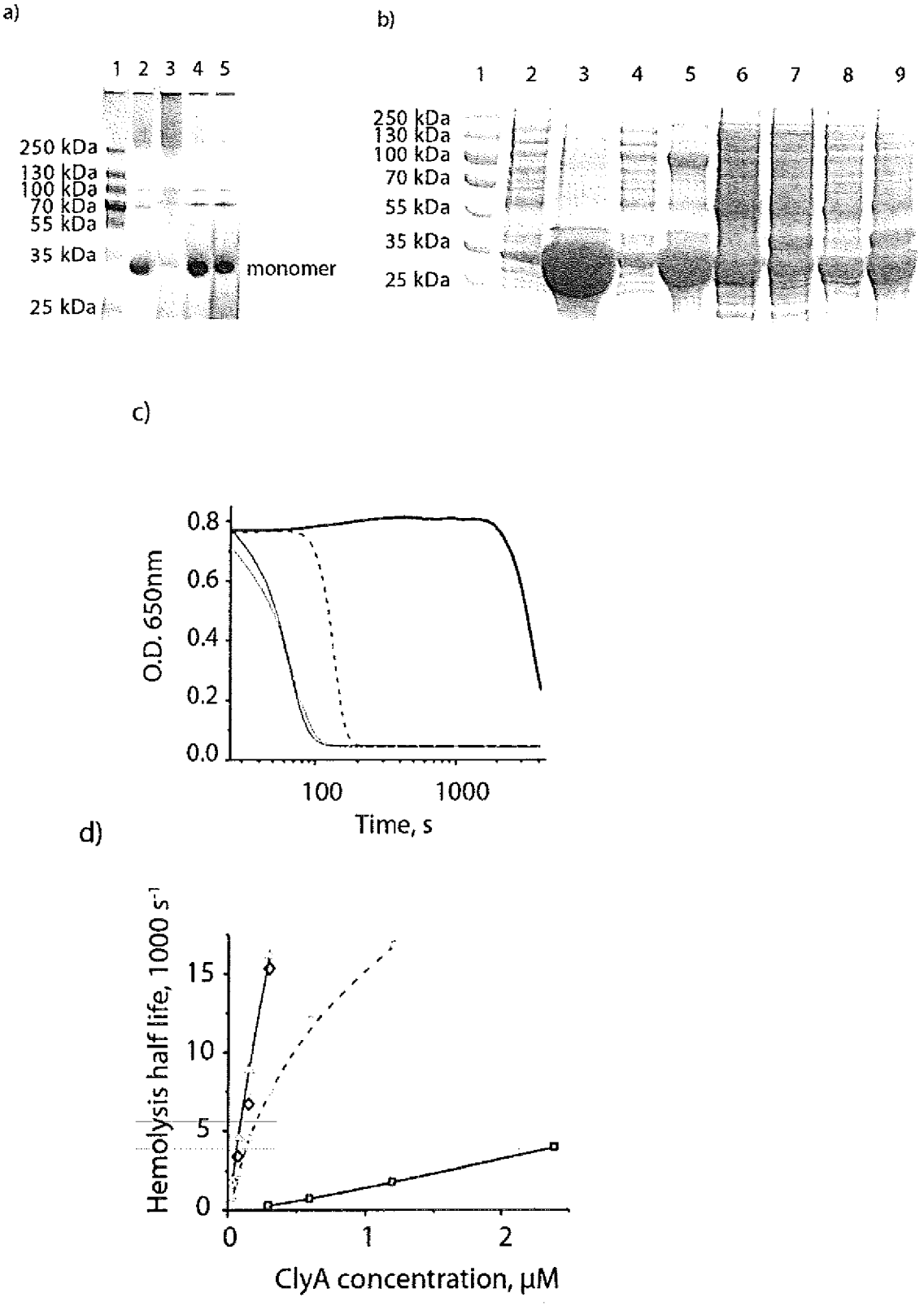

a)
1 2 3 4 5
250 kDa
130 kDa
100 kDa
70 kDa
55 kDa
35 kDa
25 kDa
monomer
b)
1 2 3 4 5 6 7 8 9
250 kDa
130 kDa
100 kDa
70 kDa
55 kDa
35 kDa
25 kDa
c)
O.D. 650nm
0.8
0.6
0.4
0.2
0.0
100
1000
Time, s
d)
Hemolysis half life, 1000 s-1
15
10
5
0
0
1
2
ClyA concentration, µM

Figure 7

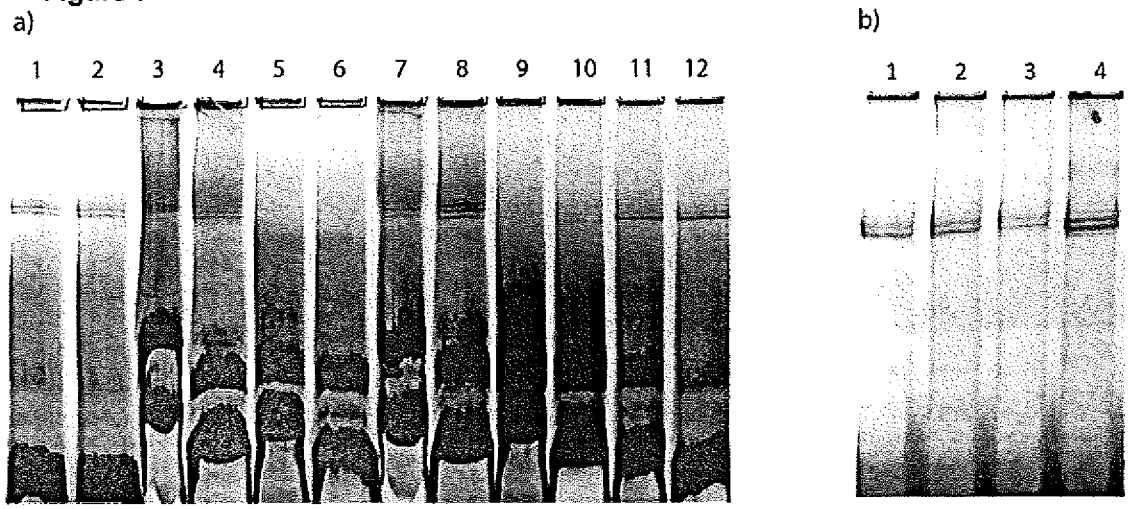
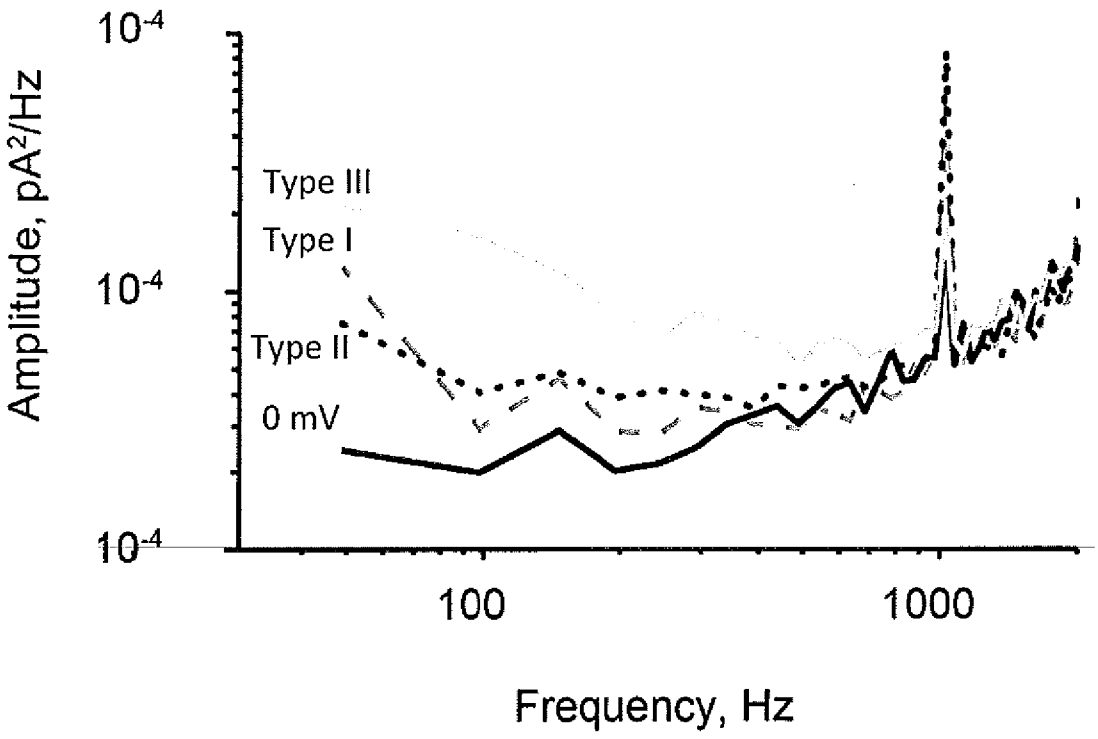


Figure 8



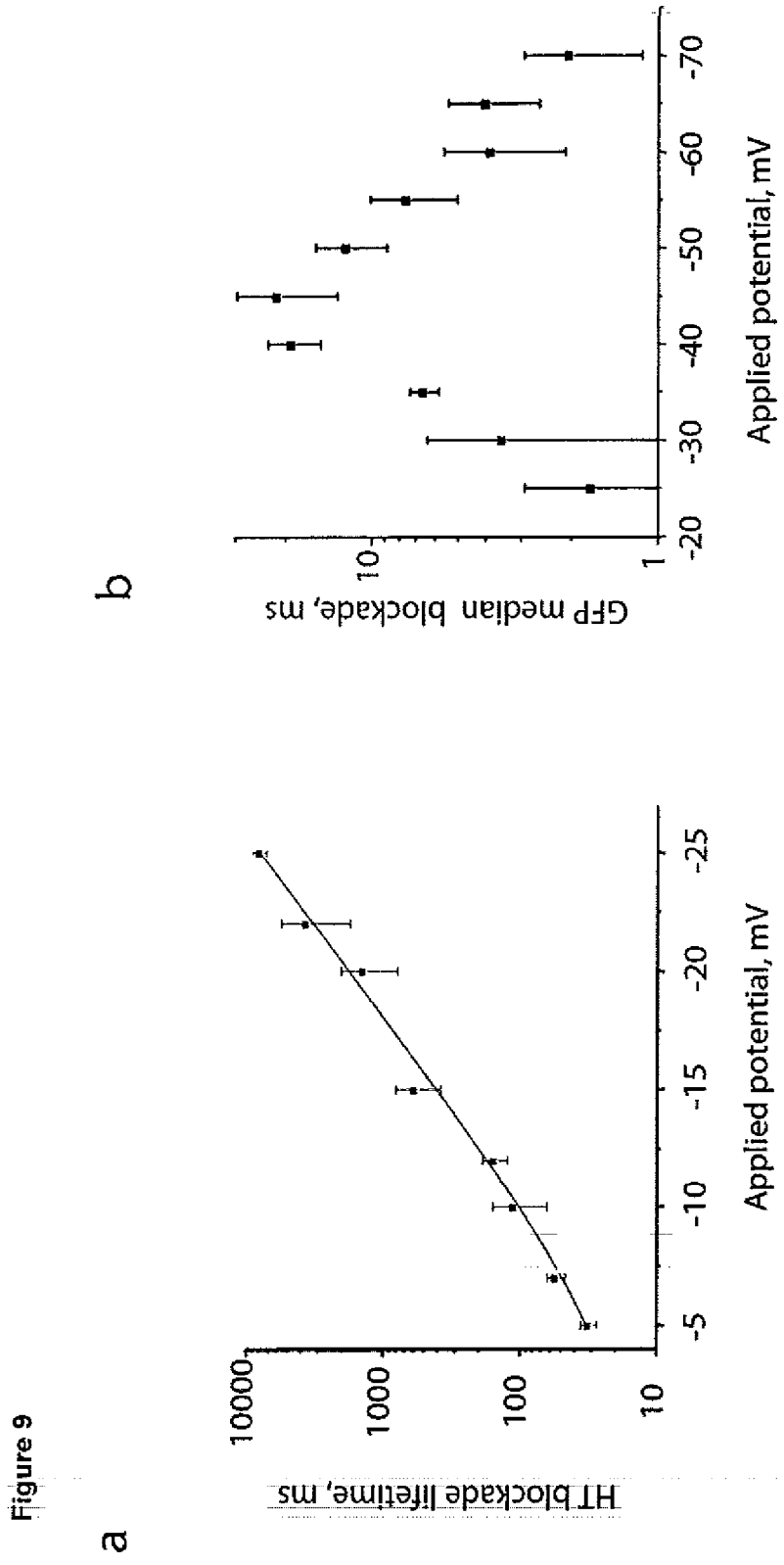
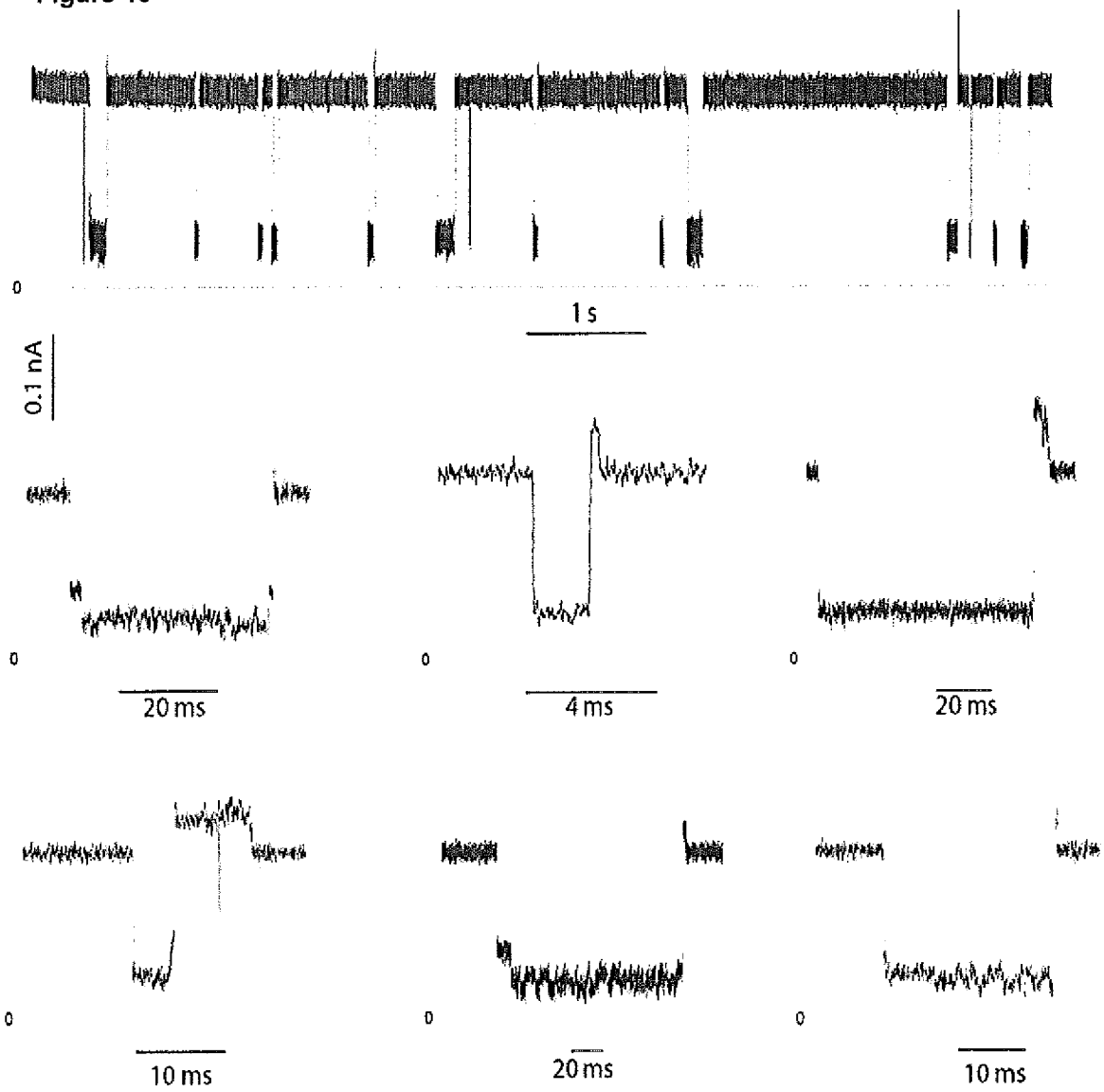

a
HT blockade lifetime, ms
10000
1000
100
10
-5
-10
-15
-20
-25
Applied potential, mV
b
GFP median blockade, ms
10
1
-20
-30
-40
-50
-60
-70
Applied potential, mV

Figure 9

Figure 10



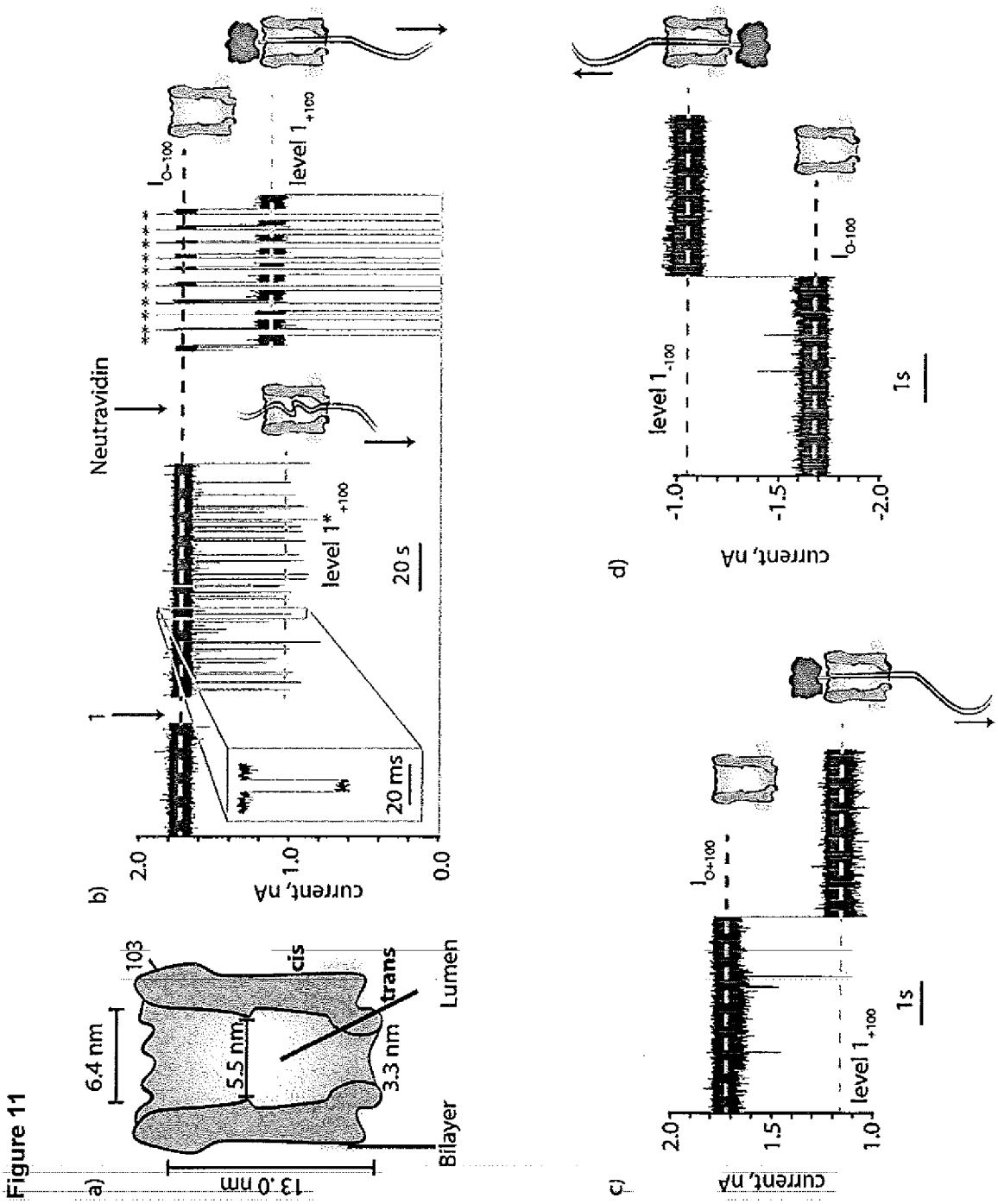
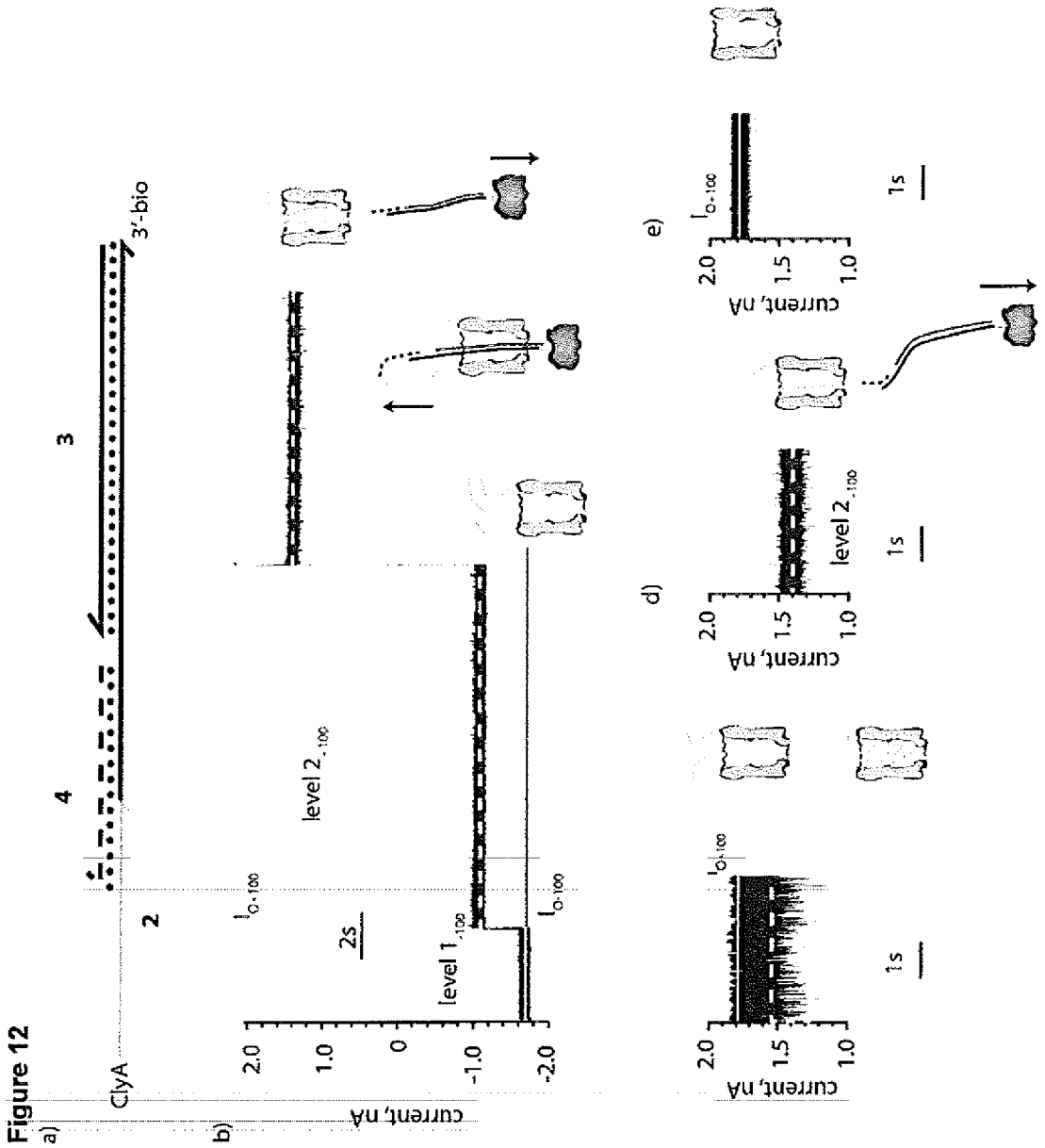

a)
6.4 nm
5.5 nm
3.3 nm
13.0 nm
cis
trans
Lumen
Bilayer
b)
Neutravidin
1
level 1*+100
level 1+100
I0-100
20 ms
20 s
current, nA
2.0
1.0
0.0
c)
I0+100
level 1+100
1s
current, nA
2.0
1.5
1.0
d)
level 1-100
I0-100
1s
current, nA
-1.0
-1.5
-2.0

Figure 11

Figure 12


a)
ClyA
2
4
3
3'-bio
b)
current, nA
2.0
1.0
0
-1.0
-2.0
level 2 -100
level 1 -120
I O-100
2s
d)
e)
1s

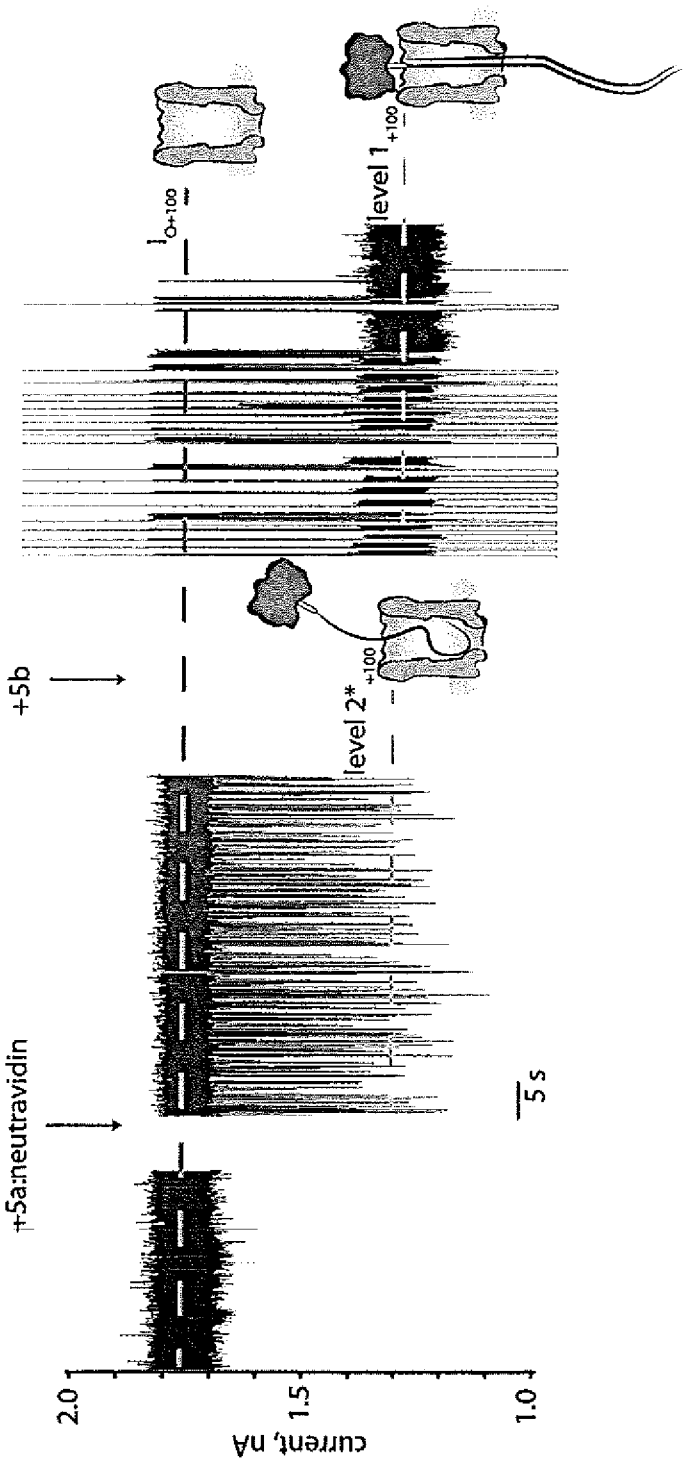
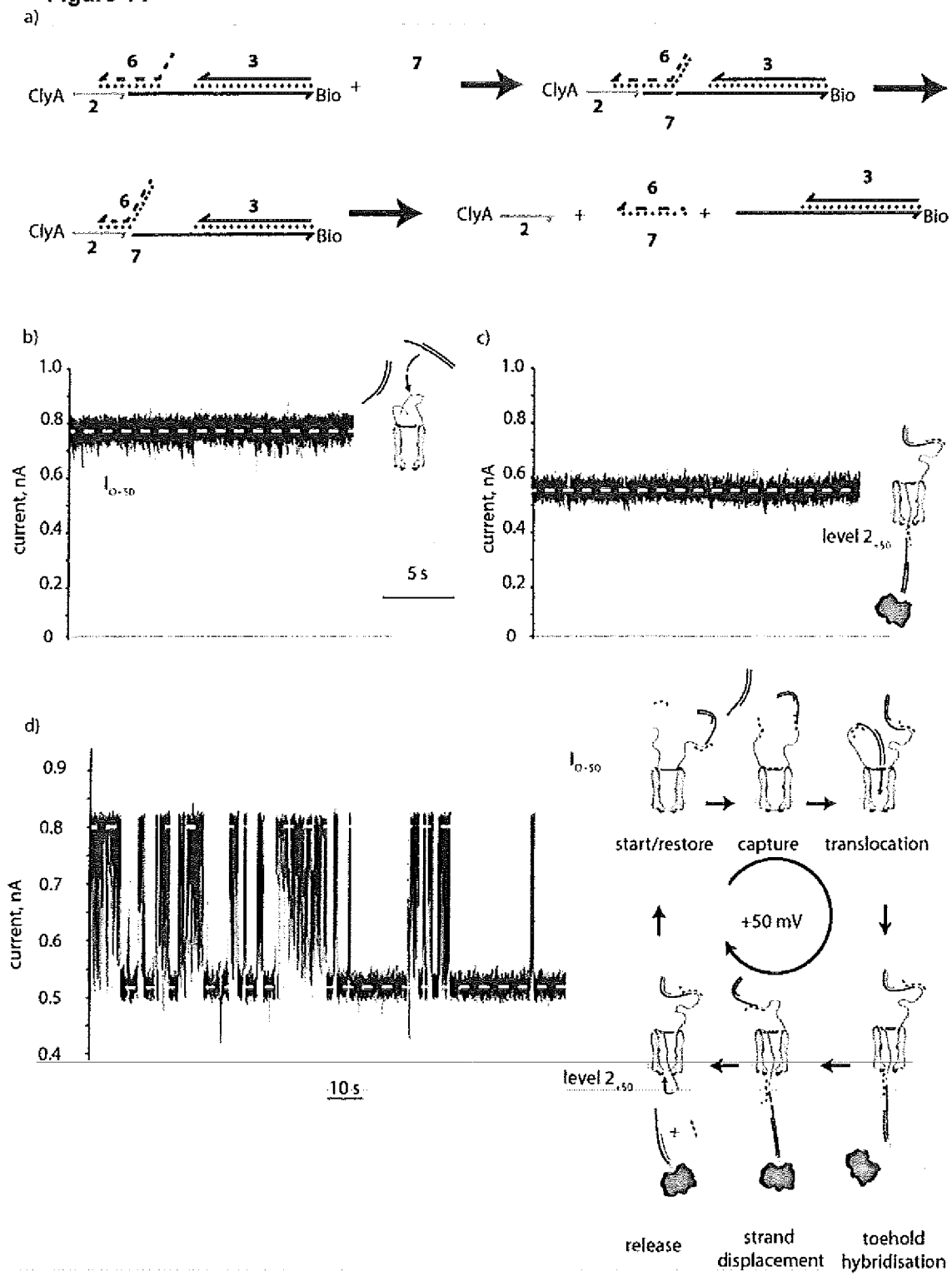

level 1 +100
I0+100
level 2* +100
+5b
+5a:neutravidin
5 s
current, nA
2.0
1.5
1.0

Figure 13

Figure 14


a)
ClyA
Bio
b)
current, nA
5 s
c)
level 2-50
d)
10 s
start/restore
capture
translocation
+50 mV
release
strand displacement
toehold hybridisation

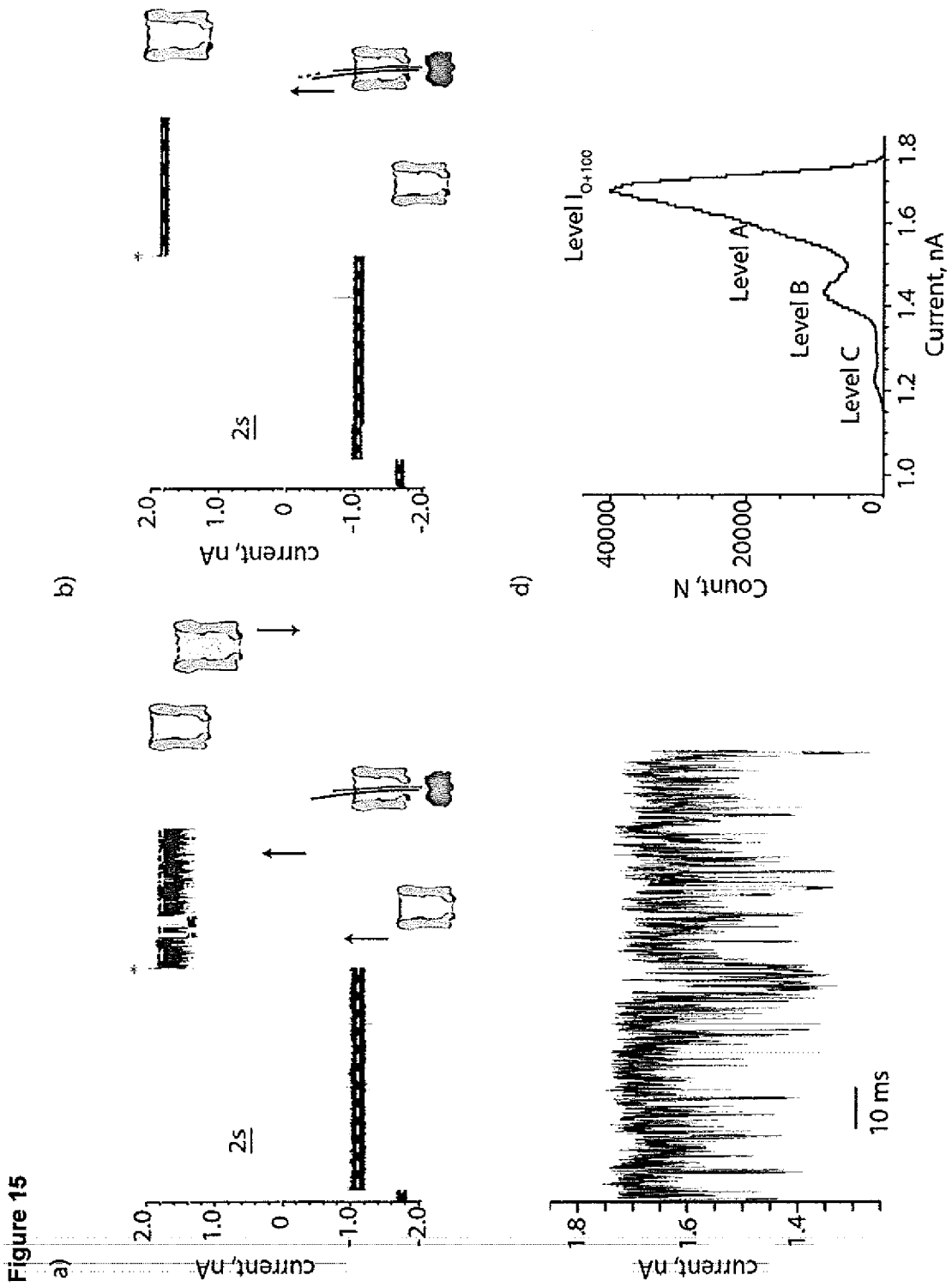

a)
current, nA
2.0
1.0
0
-1.0
-2.0
2s
b)
current, nA
2.0
1.0
0
-1.0
-2.0
2s
d)
Level I_{O+100}
Level A
Level B
Level C
Current, nA
1.0
1.2
1.4
1.6
1.8
Count, N
40000
20000
0
current, nA
1.8
1.6
1.4
10 ms

Figure 15

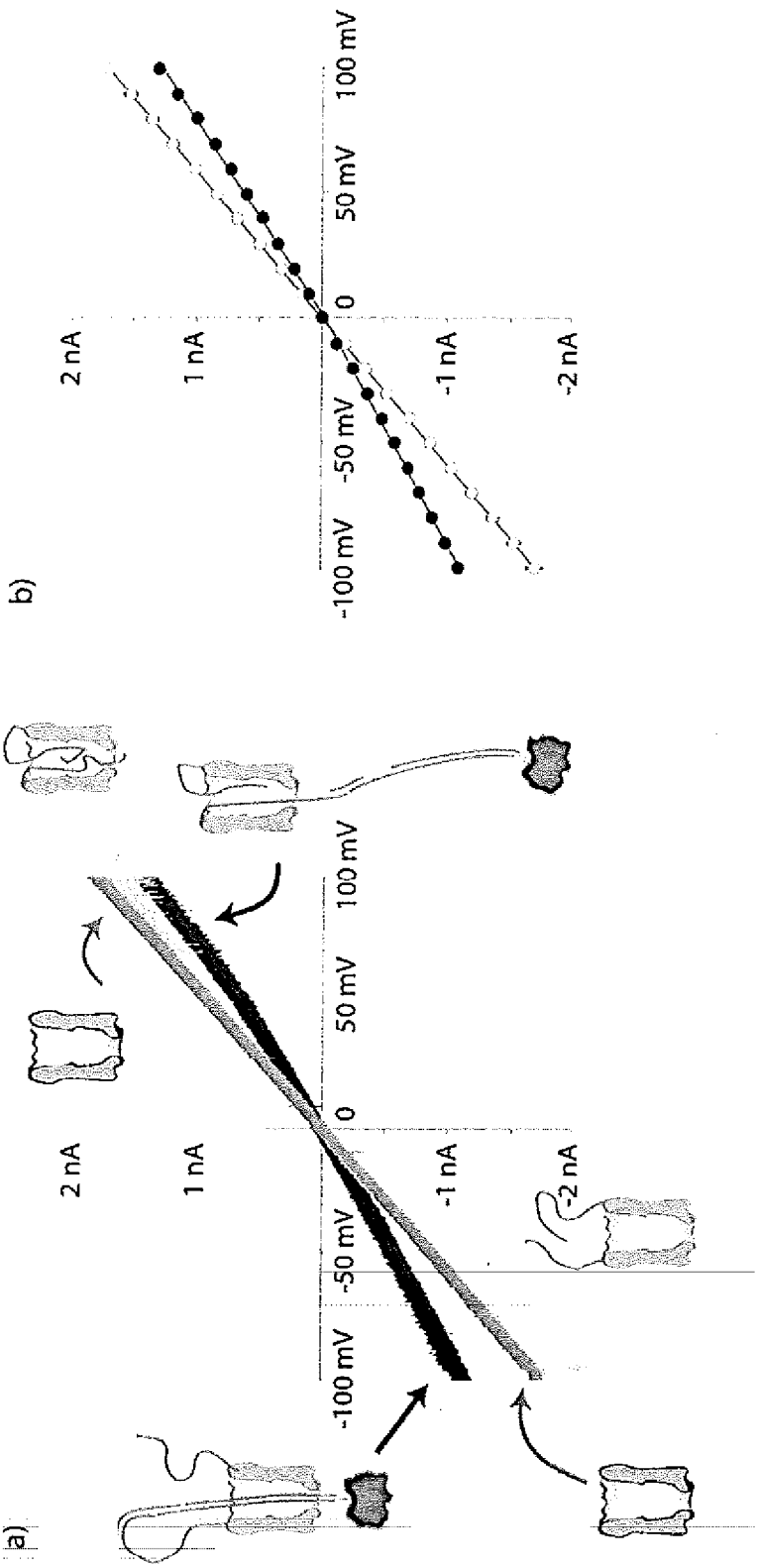


Figure 16

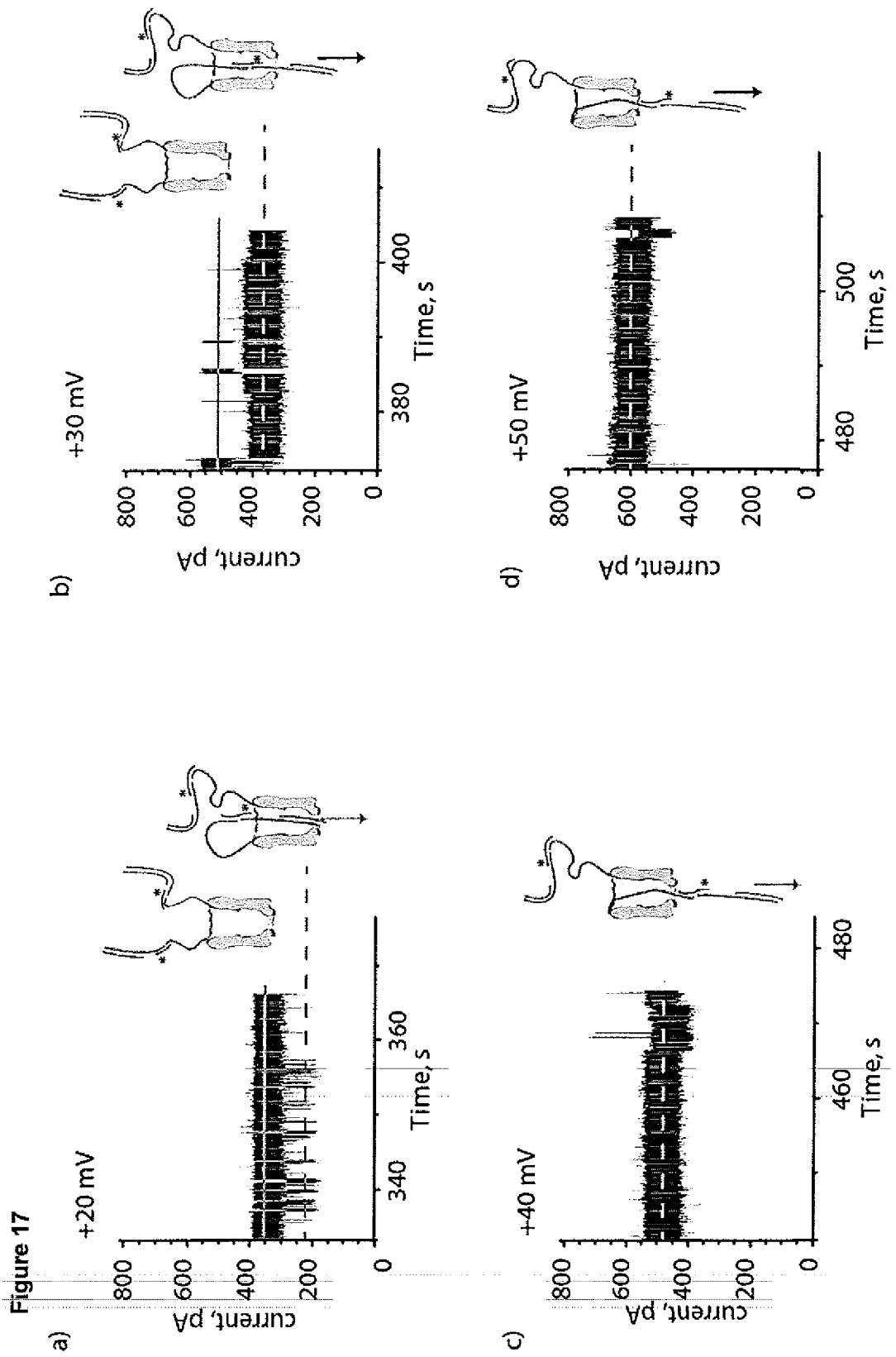
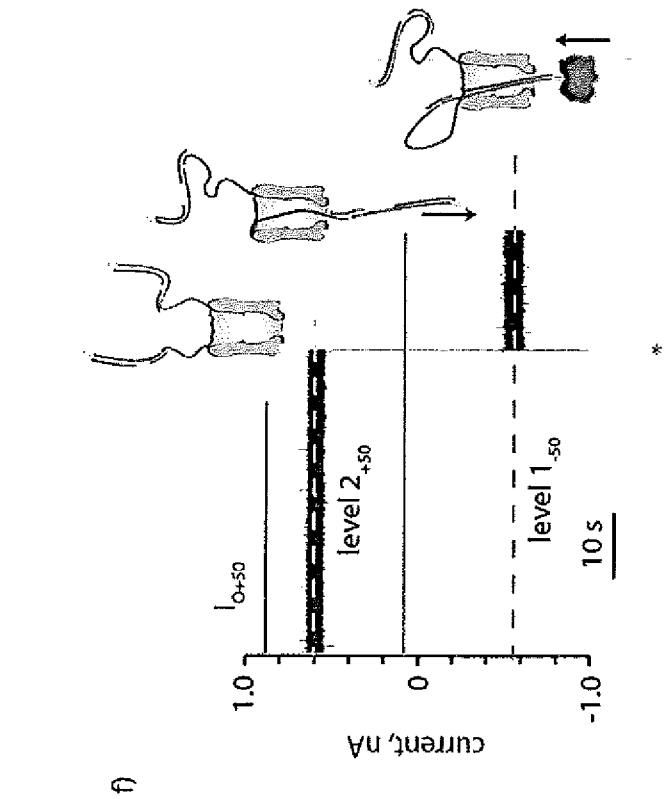

a)
+20 mV
current, pA
800
600
400
200
0
340
360
Time, s
b)
+30 mV
current, pA
800
600
400
200
0
380
400
Time, s
c)
+40 mV
current, pA
800
600
400
200
0
460
480
Time, s
d)
+50 mV
current, pA
800
600
400
200
0
480
500
Time, s

Figure 17



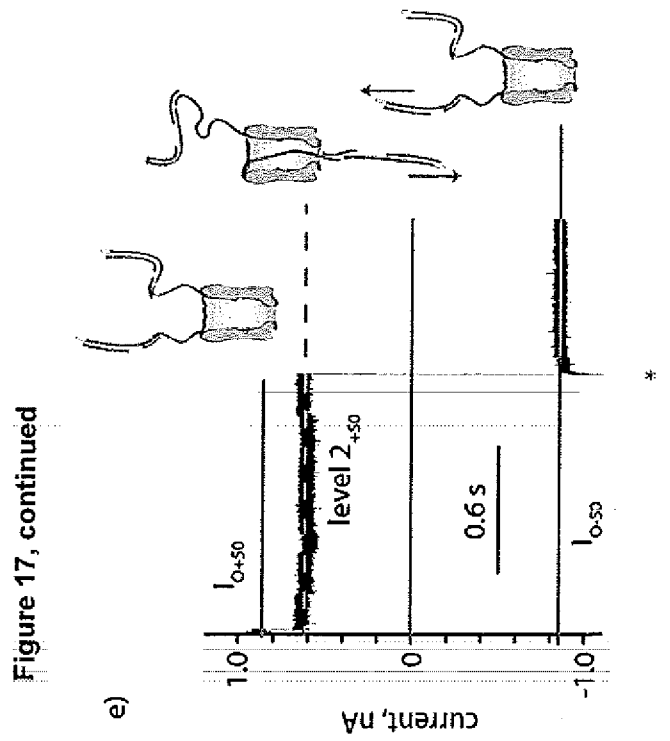


Figure 17, continued

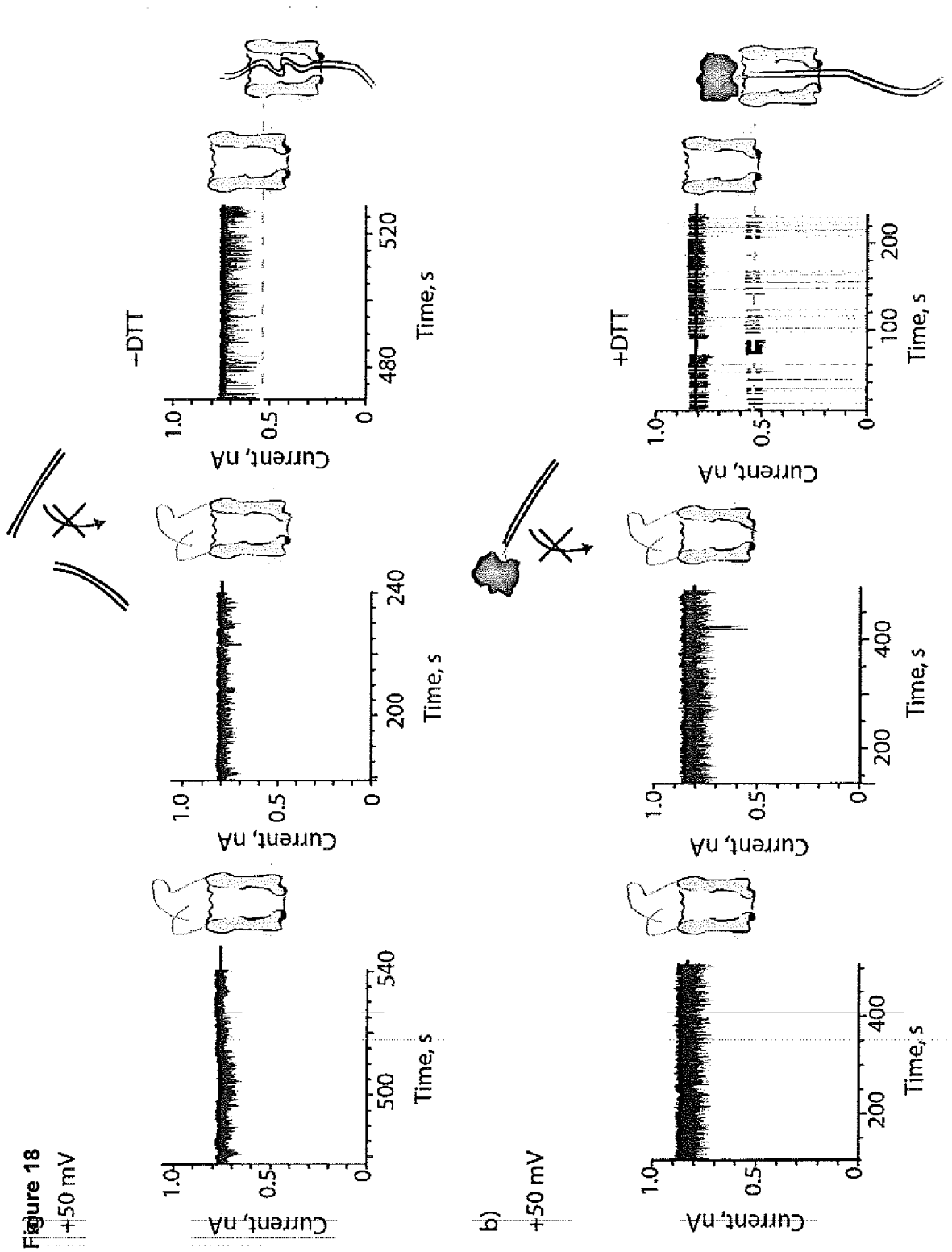


Figure 18

Figure 18, continued

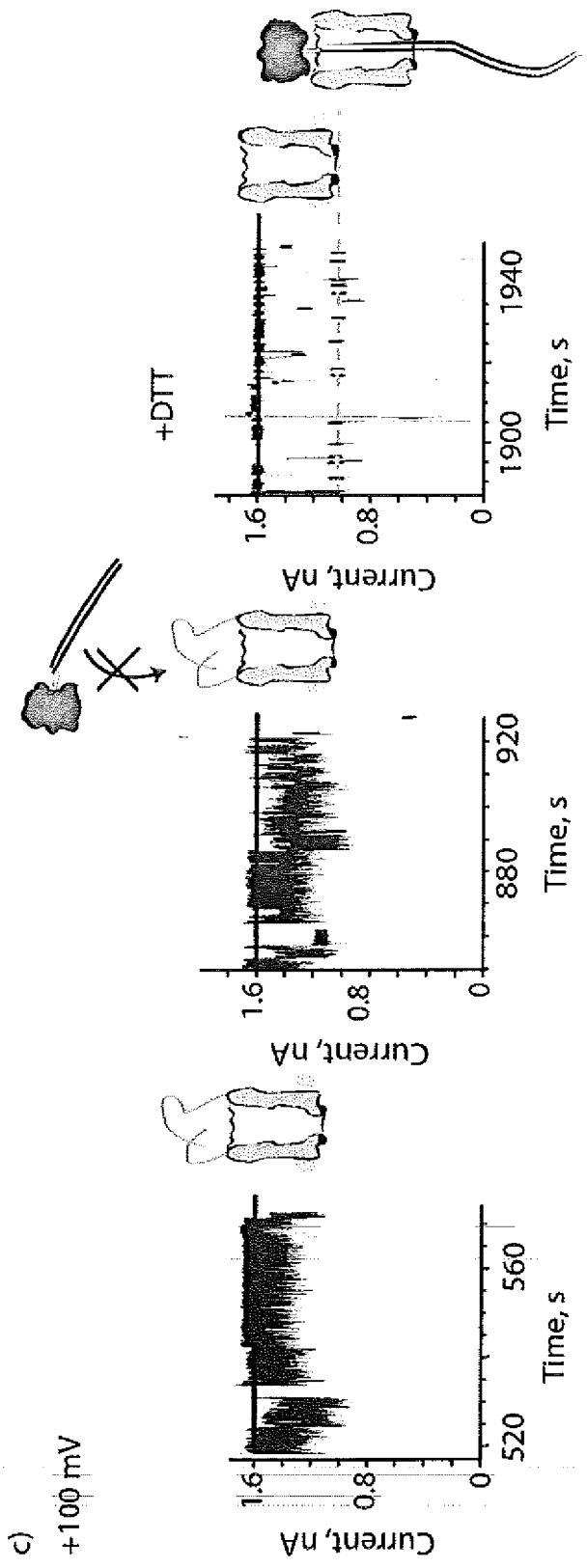
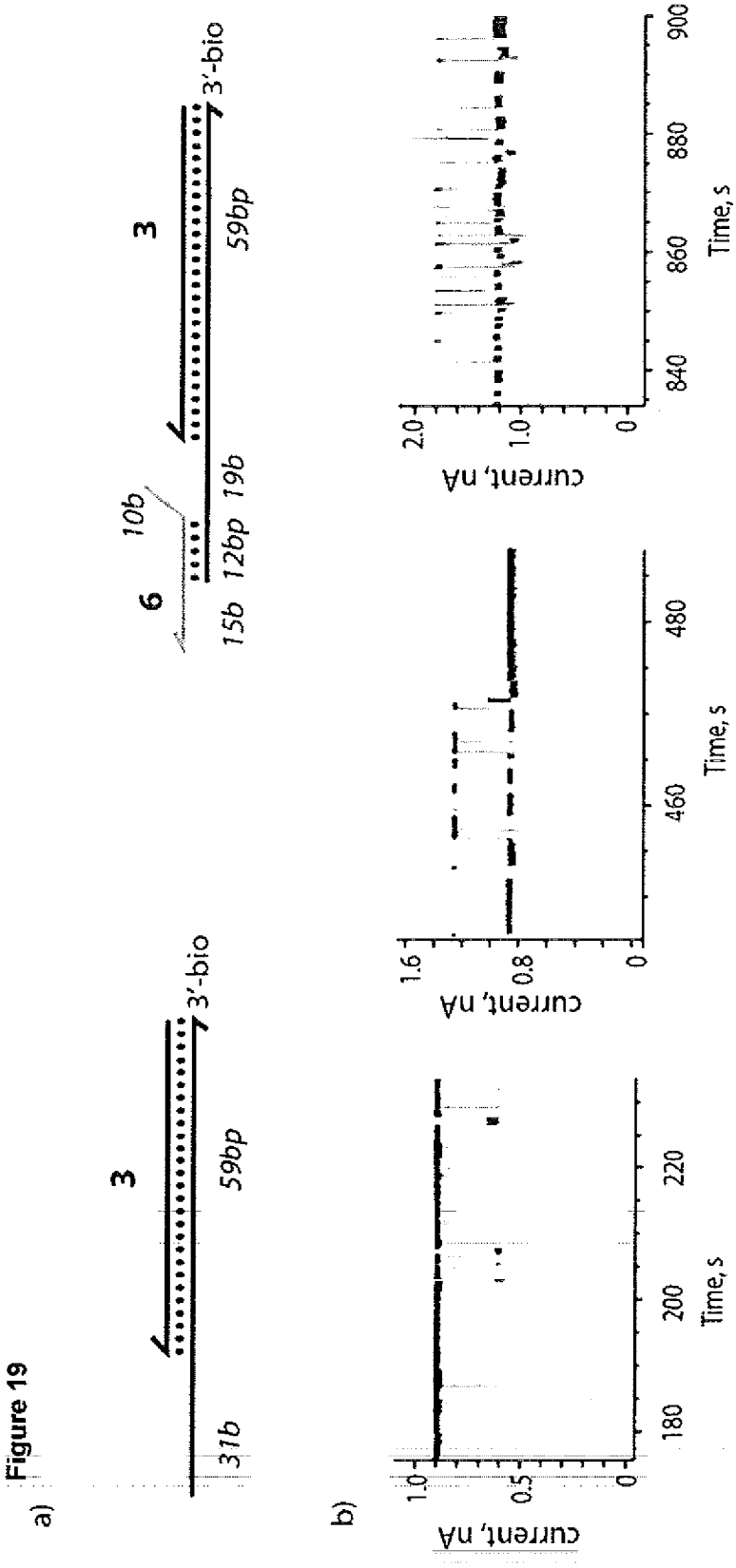


Figure 19


a)
3
3'-bio
59bp
31b
6
10b
15b 12bp 19b
3
3'-bio
59bp
b)
current, nA
Time, s
1.0
0.5
0
180
200
220
1.6
0.8
0
460
480
2.0
1.0
0
840
860
880
900

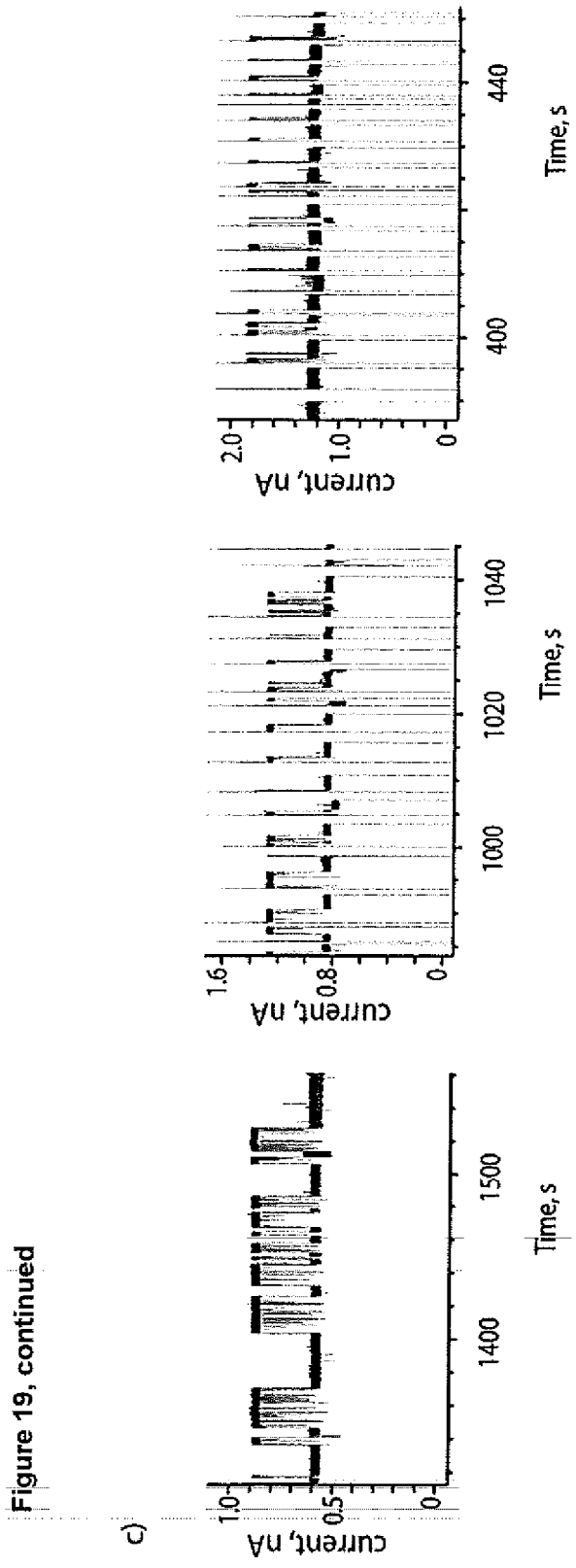


Figure 19, continued

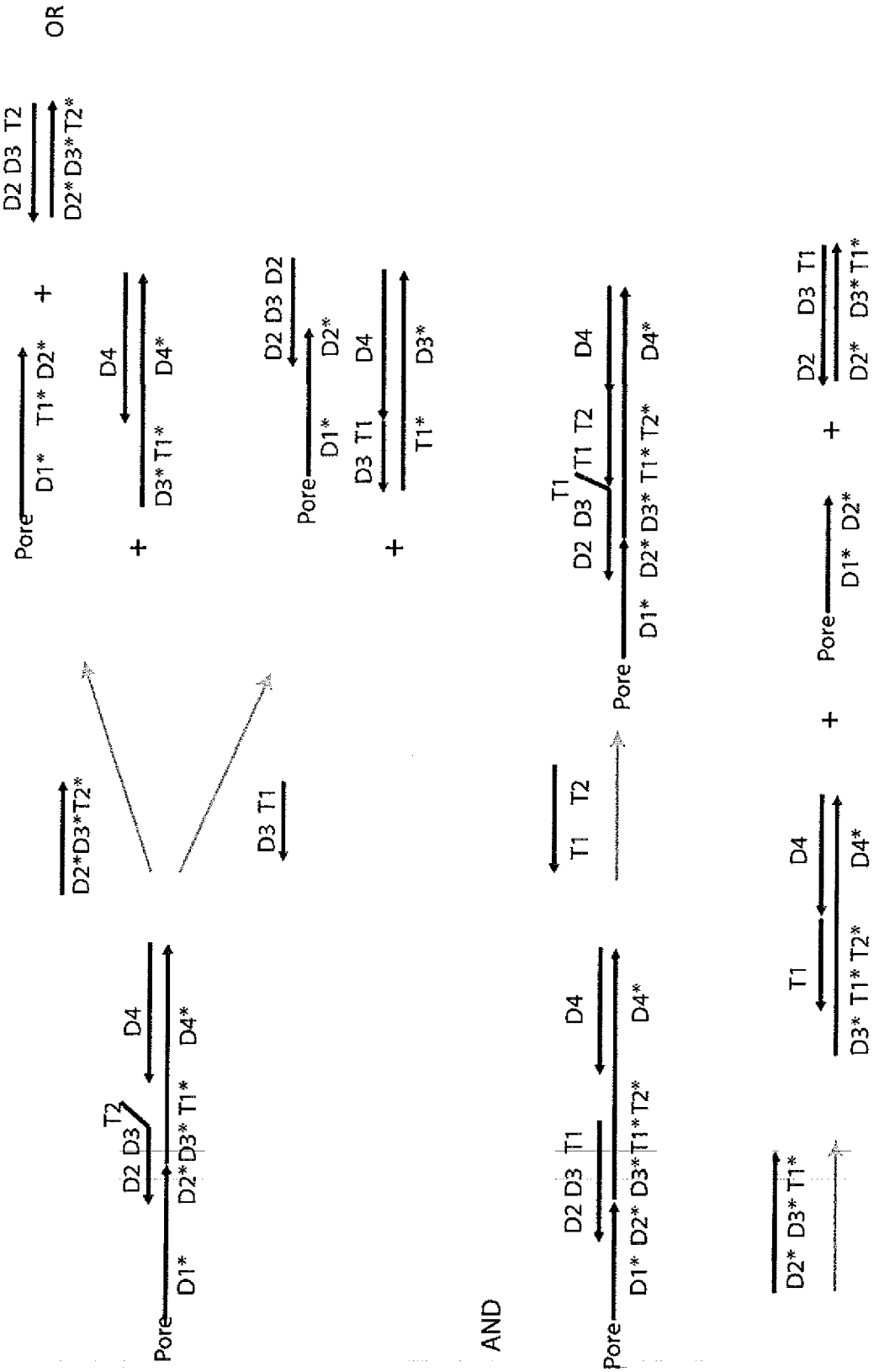

OR
D2 D3 T2
D2* D3* T2*
Pore
D1* T1* D2*
D4
D3* T1* D4*
D2 D3 D2
D1* D2*
D3 T1 D4
T1* D3*
D2*D3*T2*
D3 T1
T2
D2 D3
D2* D3* T1* D4*
D1*
T1
D2 D3 T1
D2* D3* T1*
AND
T1 T2
D1* D2* D3* T1* T2* D4*
D1* D2*
D3* T1* T2* D4*

Figure 20

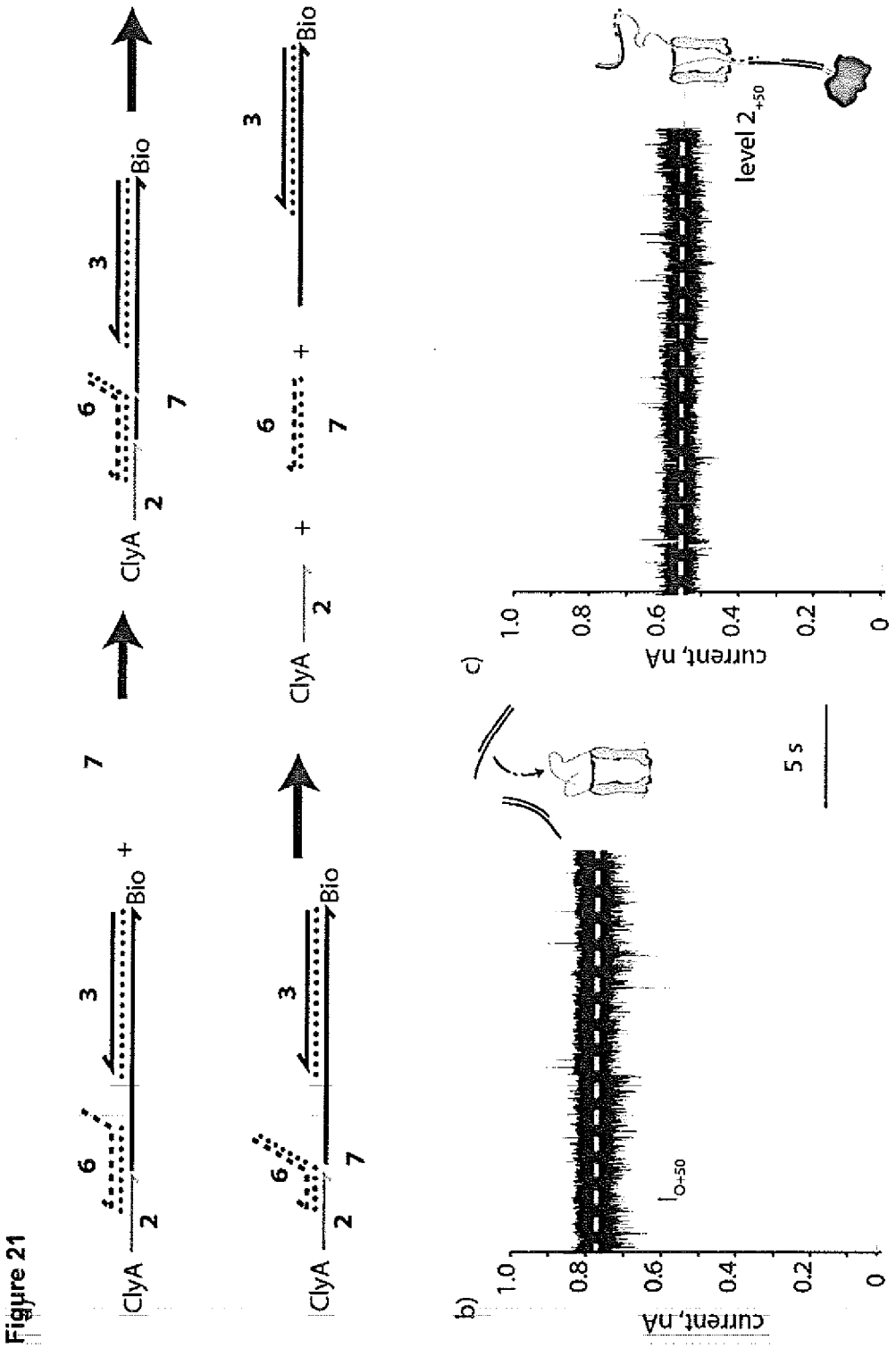


Figure 21

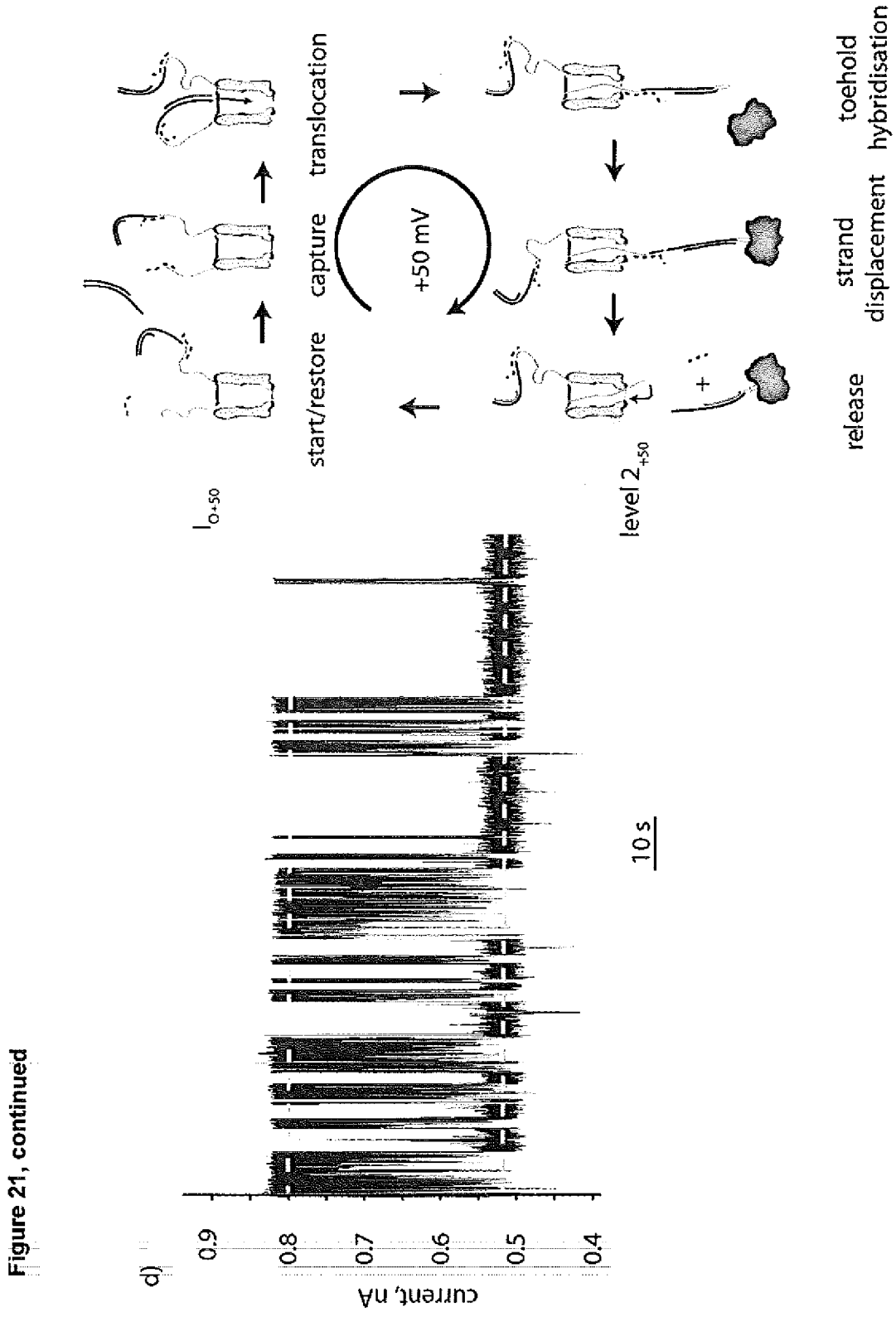


Figure 21, continued

INTERNATIONAL SEARCH REPORT

International application No
PCT/BE2014/000013

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/255
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	A. ATKINS: "Structure-Function Relationships of a Novel Bacterial Toxin, Hemolysin E. THE ROLE OF alpha G", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 275, no. 52, 22 December 2000 (2000-12-22), pages 41150-41155, XP55006624, ISSN: 0021-9258, DOI: 10.1074/jbc.M005420200 the whole document ----- -/--	1-6,8-16

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 July 2014	05/08/2014

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Chavanne, Franz

3

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	SOSKINE MISHA ET AL: "Tuning the Size and Properties of ClyA Nanopores Assisted by Directed Evolution", JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 135, no. 36, September 2013 (2013-09) , pages 13456-13463, XP002727574, ISSN: 0002-7863 the whole document	1-16
A	SOSKINE MISHA ET AL: "An engineered ClyA nanopore detects folded target proteins by selective external association and pore entry.", NANO LETTERS 12 SEP 2012, vol. 12, no. 9, 12 September 2012 (2012-09-12), pages 4895-4900, XP002727575, ISSN: 1530-6992 cited in the application the whole document	1-16
A	MAGLIA GIOVANNI ET AL: "Engineering a Biomimetic Biological Nanopore to Selectively Capture Folded Target Proteins", BIOPHYSICAL JOURNAL, vol. 104, no. 2, Suppl. 1, January 2013 (2013-01), page 518A, XP55130841, & 57TH ANNUAL MEETING OF THE BIOPHYSICAL-SOCIETY; PHILADELPHIA, PA, USA; FEBRUARY 02 -06, 2013 the whole document	1-16