

# Resolving the Dynamic Non-Covalent Interaction inside Membrane Protein Channel by Single-Molecule Interaction Spectrum

Meng-Yin Li<sup>1†</sup>, Yi-Lun Ying<sup>1†</sup>, Xi-Xin Fu<sup>2</sup>, Jie Yu<sup>1</sup>, Shao-Chuang Liu<sup>1</sup>, Ya-Qian Wang<sup>1</sup>, Shuang Li<sup>1</sup>, Chan Cao<sup>1</sup>, Yong-Jing Wan<sup>2</sup> and Yi-Tao Long<sup>\*1</sup>

## Affiliations

<sup>1</sup> Key Laboratory for Advanced Materials & School of Chemistry and Molecular Engineering, East China University of Science and Technology, Shanghai, 200237, P. R. China.

<sup>2</sup> School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, P.R. China.

<sup>†</sup> These authors contributed equally to this work.

\* E-mail: ytlong@ecust.edu.cn (to Y.T.L.).

## Keywords

Non-covalent interaction • nanopore • methylcytosine • single-molecule interaction spectrum

## Abstract

Millions of years of evolution have produced membrane protein channels capable of efficiently moving ions across the cell membrane. The underlying fundamental mechanisms that facilitate these actions greatly contribute to the weak non-covalent interactions. However, uncovering these dynamic interactions and its synergic network effects still remains challenging in both experimental techniques and molecule dynamics (MD) simulations. Here, we present a rational strategy that combines MD simulations and frequency-energy spectroscopy to identify and quantify the role of non-covalent interactions in carrier transport through membrane protein channels, as encoded in traditional single channel recording or ionic current. We employed wild-type aerolysin transporting of methylcytosine and cytosine as a model to explore the dynamic ionic signatures with non-stationary and non-linear frequency analysis. Our data illuminate that methylcytosine experiences strong non-covalent interactions with the aerolysin nanopore at Region 1 around R220 than cytosine, which produces characteristic frequency-energy spectra. Furthermore, we experimentally validate the obtained hypothesis from frequency-energy spectra by designing single-site mutation of K238G which creates significantly enhanced non-covalent interactions for the recognition of methylcytosine. The frequency-energy spectrum of ions flowing inside membrane channels constitutes a single-molecule interaction spectrum, which bridges the gap between traditional

ionic current recording and the MD simulations, facilitating the qualitative and quantitative description of the non-covalent interactions inside membrane channels.

## Introduction

Ions in aqueous solution passing through membrane protein channels play crucial roles in cell-cell communications, shaping action potential and other key cellular functions<sup>1</sup>. The protein channels experience multiple reversible non-covalent interactions to execute the successive transport of target carriers<sup>2-5</sup>. Proper clarifying of the non-covalent interactions is essential to accurately describe the folding/unfolding of protein channel domains and specific recognition of target molecules<sup>6,7</sup>, which is highly relevant to understand the dynamic functions of the membrane protein channel. Although molecular dynamics (MD) simulations provide richly detailed information regarding the transient noncovalent interactions during the transportation of target carriers/ions through the membrane protein channel, there are few experimental approaches that can directly define the multiple but synergistic non-covalent interactions within the channel. The present techniques which include optimized optogenetic tools<sup>8,9</sup>, acoustic force spectroscopy<sup>10</sup>, single-molecule motions<sup>11</sup>, optical tweezers<sup>12</sup> and magnetic tweezers<sup>13</sup>, mostly lead to the interpretation of non-covalent interaction for the structurally stable proteins instead of dynamic membrane proteins. For example, nuclear magnetic resonance (NMR) spectroscopy combined with DFT calculations have been used to directly detect non-covalent interactions between side-chain methyl groups and aromatic amino acid residues in ubiquitin and GB3 protein<sup>14</sup>. As for conductive membrane proteins, the ionic current flow through the membrane proteins encodes the ion fluctuation from the dynamic binding state and conformational changes of target molecule inside this membrane protein. For example, reading the ionic amplitude through the nanopore could achieve biomolecule sensing<sup>15-24</sup> and further identify all four nucleobases<sup>25-29</sup>, while analyzing the duration could be used to measure single molecule force between protein and oligonucleotides<sup>30</sup>. In particular, nanopore techniques are enlightened by electrochemically confining single molecules into the membrane protein. The volume exclusion of a single molecule generates the distinguishable ionic flow through the membrane protein, which could be used to identify size<sup>31-35</sup>, charge<sup>36-40</sup>, conformation<sup>41-45</sup> and interaction between the membrane protein and target molecules<sup>46-48</sup>. More importantly, the MD simulations reveal that the transport of single molecules through the nanopore involves multiple non-covalent interactions between the sensing interface of the membrane protein and the analyte. Therefore, all of dynamic non-covalent interaction inside membrane protein has been recorded by the ionic flow.

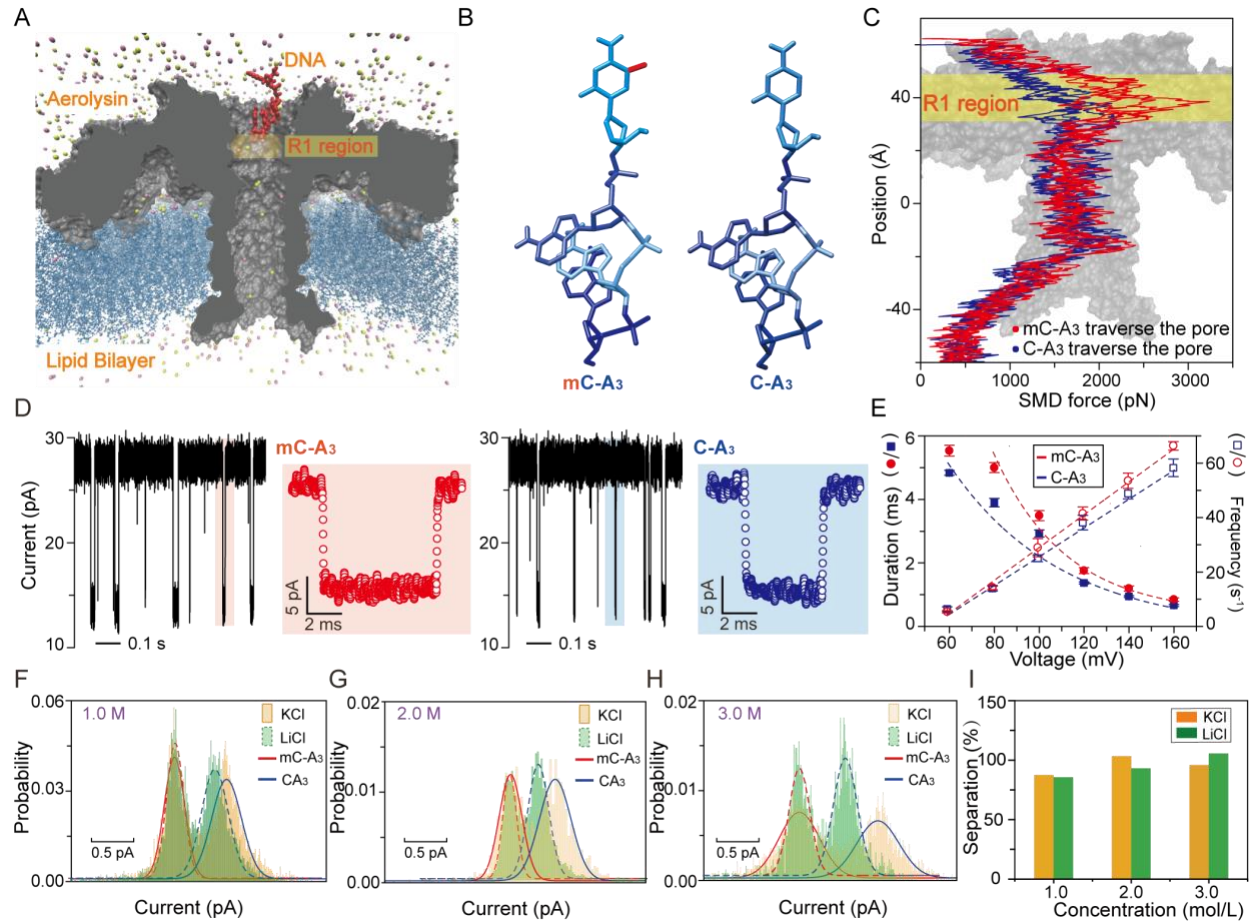
By reading the ionic flow, previous studies of the aerolysin nanopore have discriminated the single-

nucleotide length differences and single-nucleobase variations<sup>49-51</sup>. Furthermore, the reading of ionic current can achieve the recognition of methylcytosine and cytosine, which have a 1.12 pA difference in amplitude and 1.36 times difference in duration at an applied voltage of +80 mV<sup>52</sup>. All of these previous experiments on ionic measurements only extract the duration time and current amplitude from each ionic blockade, which is relative to the slow motions or changes of biomolecules inside the nanopore on a timescale ranging from milliseconds to seconds. Even so, the challenges still remain in disclosing the dynamic and weak non-covalent binding, much less the understanding of the synergistic neighboring effects.

To reveal the dynamic non-covalent interactions from the ionic flow, here, we present a rational strategy which employed aerolysin discrimination of methylcytosine and cytosine as a model to explore the dynamic ionic signatures with non-stationary and non-linear frequency analysis (Fig. 1A, 1B). By performing the traditional aerolysin nanopore experiments, we show that the enhanced events induced by methylcytosine nanopore are independent of applied voltage, counterions and ionic strength. Then all-atom MD simulations suggest that the amino acid residues in position 220 and 222 form a strong non-covalent interaction, including non-bonded interaction and hydrogen bonding with both oligonucleotides to restrict them in the sensing region, defined as region R1 by SMD simulation (Fig. 1C). Meanwhile, amino acid residues in position 218 and 274, also in region R1, greatly contribute to enhanced non-covalent interactions induced by methylcytosine rather than cytosine. Moreover, we extracted the frequency features hidden in the non-stationary current trace by using Hilbert–Huang Transform. The energy–frequency–time spectra reveal a stronger enhanced energy at 450 Hz for methylcytosine than cytosine, which could be determined as the single-molecule interaction spectrum for quantifying the non-covalent interaction between methylcytosine and region R1. With this knowledge in mind, a K238G mutant aerolysin nanopore was designed to enhance the non-covalent interaction in region R1, which produced a remarkable enhancement of frequency energy on current recordings. This rational method we reported here is analogous to that found for ionic channels, electrochemical potential-driven transporters, and primary active transporters where characterizing the non-covalent interactions inside the pore were shown to be the large challenges.

## Results

Aerolysin is one of the pore-forming toxins (PFTs) from *Aeromonas* spp that can self-assemble into the lipid bilayer and organize into circular heptameric pores<sup>53</sup>. Applied a transmembrane potential with a pair of Ag/AgCl electrodes, the negatively charged oligonucleotides are driven into the aerolysin nanopore one at a time, inducing a series of ionic current blockages. As shown in Fig. 1D, mC-A<sub>3</sub> exhibited a deeper current blockage with a longer translocation duration than that of the oligonucleotide of C-A<sub>3</sub>. Statistical results show that mC-A<sub>3</sub> and C-A<sub>3</sub> display a high separation of about 87.5% (details see in Supplementary Fig. S1 and S2). Previous studies suggest the blockage current of each analyte is mainly determined by the pore geometry, counterions, electrolyte solution and the dynamic interactions between the pore and every single analyte<sup>54</sup>. Unlike CsgG, MspA and  $\alpha$ -hemolysin which possesses a sensing site with a diameter of < 1 nm, the aerolysin are constituted by a long  $\beta$ -barrel about 10 nm with a relatively uniform diameter of 1 nm<sup>55</sup>. Therefore, the structure of aerolysin provides the confined space for accommodating single oligonucleotides but is not the key factor for the current separation of methylcytosine and cytosine. Instead, the dynamic interaction between oligonucleotides and aerolysin protein is suggested to be responsible for the high current separation. To support this suggestion, we carried out voltage-dependent experiments, changed counterions and varied ionic strength.



**Figure 1. Transporting C-A<sub>3</sub> and mC-A<sub>3</sub> through a wild-type aerolysin membrane channel. (A)** All-atom model of full-length aerolysin nanopore system. The aerolysin (gray) was inserted into lipid bilayer membrane (dark blue), while the nucleotide (red) placed at the entrance of the pore. The ions flow across the membrane through the aerolysin. **(B)** Structure of methylated and unmethylated oligonucleotide containing methylcytosine and cytosine, respectively. The only difference between them was the addition of a methyl group which was marked as red color. **(C)** SMD simulations for the translocation of mC-A<sub>3</sub> (red line) and C-A<sub>3</sub> (blue line) through the pore with a constant velocity of 0.25 Å/ps. The maximum difference of force on DNA is generated near R1 (yellow column). Results were performed in quintuplicate. **(D)** Raw single-channel recording traces of ionic current and corresponding typical events of mC-A<sub>3</sub> (red) and C-A<sub>3</sub> (blue) at +60 mV, respectively. **(E)** Voltage dependence of durations (left y axis) and event frequency (right y axis) for mC-A<sub>3</sub> (red line) and C-A<sub>3</sub> (blue line), respectively (details see in Supplementary Table S1 and S2). Event frequency of the pure C-A<sub>3</sub> and mC-A<sub>3</sub> increases linearly with the applied voltage. The experiments above were performed in 1 M KCl, 10 mM Tris, and 1 mM EDTA, pH 8.0, 20 ± 2 °C. **(F)-(H)** The current blockage histogram for the mixture of C-A<sub>3</sub> and mC-A<sub>3</sub> in KCl (orange) and LiCl (green) solution at

the salt concentration of 1.0 M (F), 2.0 M (G) and 3.0 M (H). Current histograms are all fitted to two Gaussian peaks drawn as a solid line and dash line in KCl and LiCl, respectively. The current was filtered at 5 kHz and sampled at 100 kHz. The I-V curve of aerolysin in 1M KCl is shown in Supplementary Fig. S3. **(I)** Separation of the two current peaks corresponding to mC-A<sub>3</sub> and C-A<sub>3</sub> in KCl (orange) and LiCl (green) at 1.0 to 3.0 M salt concentration. The separation is defined as  $R = 2(I_{C-A_3} - I_{mC-A_3}) / (W_{C-A_3} + W_{mC-A_3})$  based on our previous work<sup>56</sup>. All data were acquired in the presence of 2.0 μM at pH 8.0, 20 ± 2 °C.

As shown in Fig. 1E, the frequencies of both mC-A<sub>3</sub> and C-A<sub>3</sub> show a linear voltage dependence (also see in Supplementary Table S1). According to a previous study<sup>49</sup>, this relationship suggests that the capture of mC-A<sub>3</sub> and C-A<sub>3</sub> by the aerolysin protein are both limited by the biased diffusion rather than the energy barrier at the entry of the pore. Moreover, the durations of C-A<sub>3</sub> exhibit an exponential decay with potential from +60 mV to +160 mV (Fig. 1E and Supplementary Table S2), which could be scaled as  $\tau = a * \exp(-V/V_c)$ , where  $V_c$  represents the critical potential to overcome the free energy barrier to squeeze an oligonucleotide through a confined nanopore. However, mC-A<sub>3</sub> experiences the exponential dependence of duration on voltage is from +80 mV to +160 mV. Compared to C-A<sub>3</sub>, the anomalous duration of mC-A<sub>3</sub> at +60 mV indicates a higher translocation energy barrier according to previous study on α-hemolysin<sup>57</sup>. Based on the equation  $V_c = (k_B T) / (z_{inside} e)$ , we further estimated the effective charge ( $z_{inside}$ ) of mC-A<sub>3</sub> and C-A<sub>3</sub> inside the aerolysin. Surprisingly, in our experiments, the slower translocation of mC-A<sub>3</sub> exhibits a larger effective charge of 0.79 which is about 1.55 times greater than that of C-A<sub>3</sub> (Supplementary Table S3). Note that mC-A<sub>3</sub> and C-A<sub>3</sub> have the same number of nucleotides, therefore giving the same total number of negative charges. As suggested by the previous study, the effective charge is mainly affected by counterions that screen the elementary charges inside a nanopore<sup>58</sup>. Therefore, the difference of  $z_{inside}$  between mC-A<sub>3</sub> and C-A<sub>3</sub> may arise from the counterions binding. To further investigate the effects of counterions on the current separation, we performed the experiments in LiCl solution since Li<sup>+</sup> was reported to more effectively bind to DNA than K<sup>+</sup><sup>59</sup>. Contrary to the expectation, the effective charges of mC-A<sub>3</sub> and C-A<sub>3</sub> in LiCl solution increase to 1.03 and 0.91, respectively (Supplementary Table S4). These results do not show a stronger charge screening in LiCl. Moreover, the peak widths of the ionic current histogram in LiCl are narrower than those in KCl, but the peak current difference between the two populations of mC-A<sub>3</sub> and C-A<sub>3</sub> in LiCl are lower than those in KCl (Supplementary Fig. 4-5). As a result, the current separation between mC-A<sub>3</sub> and C-A<sub>3</sub> in LiCl is comparable to that in KCl solution (Fig. 1F and I), even at high concentration (Fig. 1G, H and I). Therefore, the

counterion is not the key factor for aerolysin to produce a distinct current drop as introducing the additional methyl group in cytosine. Considering the results above, the non-covalent interactions are the main contribution to the varied ionic current for mC-A<sub>3</sub> and C-A<sub>3</sub>. However, challenges still remain for the rational characterization of the non-covalent interactions for each carrier through the membrane proteins.

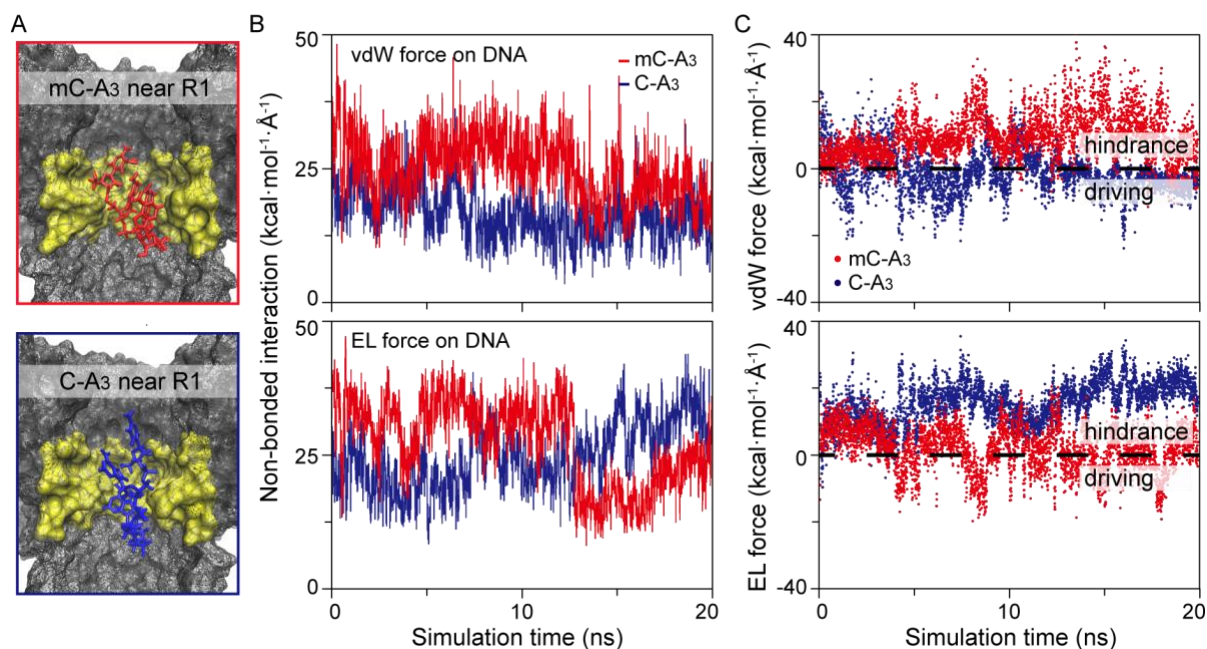
Thus, in the next step, we performed all-atom MD simulation of aerolysin nanopore system using program NAMD<sup>60</sup> to look insights into the non-covalent interaction between pore and C-A<sub>3</sub>/mC-A<sub>3</sub>. An all-atom model was constructed with a full-length aerolysin nanopore embedded into a lipid bilayer membrane and solvated in 1.0 M KCl electrolyte, which is consistent with our experiment condition. On the basis of our previous work<sup>61</sup>, oligonucleotides show a preferred translocation direction through aerolysin by its 3' end because of the additional steric hindrance needed to be counter balanced from 5' end. Thus, our simulation model assumed that the translocation of C-A<sub>3</sub> or mC-A<sub>3</sub> was initiated by 3' end. Accordingly, C-A<sub>3</sub> or mC-A<sub>3</sub> that has been also pre-equilibrated in 1.0 M KCl solution was placed in the system at the *cis* entrance of aerolysin with its 3' extreme orientated to the pore as depicted in Fig. 1A after aerolysin nanopore system equilibrated in NPT ensemble for 70 ns. Steered Molecular Dynamics (SMD) simulations were then performed to pull the oligonucleotides through the pore at a constant velocity of 0.25 Å/ps (Supplementary Fig. S6). The backbone of oligonucleotides, regarded as SMD atoms, was attached to a dummy atom via a virtual spring. During the whole translocating process, the steering force (SMD force) along with the traverse orientation would increase to overcome the dynamic interactions between the oligonucleotide and residues in the aerolysin lumen. As expected, SMD force increases rapidly as soon as oligonucleotides entered aerolysin pore (Fig. 1C), especially in the position from 29.8 Å to 47.4 Å and from -19.4 Å to -7.6 Å, which are denoted as R1 region and R2 region, respectively. Note that both C-A<sub>3</sub> and mC-A<sub>3</sub> needed a stronger SMD force to be threaded through these two regions, which is consistent with our previous finding that the aerolysin has two sensing region<sup>61,62</sup>. Furthermore, the SMD force applying on mC-A<sub>3</sub> becomes nearly 1.52-fold stronger than that in C-A<sub>3</sub> when the oligonucleotides traversed R1 region, while those on C-A<sub>3</sub> and mC-A<sub>3</sub> in R2 region were comparable. This result reveals a much stronger non-covalent interaction between mC-A<sub>3</sub> and R1 region of aerolysin protein than that between C-A<sub>3</sub> and R1 region. To verify these results, we slowed the constant velocity for ssDNA through aerolysin pore to 0.10 Å/ps and changed the orientation of ssDNA at the entrance of the pore. As shown in Supplementary Fig. S7-S8, mC-A<sub>3</sub> still reveals obviously stronger interaction than C-A<sub>3</sub> when traverses R1 region compared to other positions, which demonstrates that the R1 region is responsible for the distinguishable non-covalent interaction between the oligonucleotides (C-A<sub>3</sub> and mC-A<sub>3</sub>) and the aerolysin.

To further rationalized this mechanism, two frames shown in Fig. 2A with mC-A<sub>3</sub> and C-A<sub>3</sub> located in R1 region from SMD simulation were studied. In these two frames, the centers of mass for mC-A<sub>3</sub> and C-A<sub>3</sub> are both at position 35.2 Å in R1 region. Then, the MD simulations are carried out in the NPT ensemble for 20 ns with applying a harmonics constraint to restrain the backbone of C-A<sub>3</sub> or mC-A<sub>3</sub>. According to the trajectories calculated by above simulation (Fig. 2B), non-bonded pair interactions between the oligonucleotides and aerolysin protein were further estimated using NAMD-Energy plug-in. Non-bonded interactions are generally divided into two parts, van der Waals (vdW) interactions and electrostatic (EL) interactions. Modeling the unequal distribution of charge in a molecule, the force field framework places point charges at each of the atomic sites. Based on this model, the van der Waals interactions are taken consist of the all the interactions between atoms that are not covered by the electrostatic interaction, including dispersion, repulsion, and induction, among other interactions, while the electrostatic interactions between the point charges are generally modeled by a Coulomb potential. As shown in Fig. 2B, the value of vdW force and EL force exerted on mC-A<sub>3</sub> and C-A<sub>3</sub> exhibits a drastic fluctuation with the simulation time, demonstrating the adjustment of the conformation for both pore and analyte. Interestingly, the vdW force on mC-A<sub>3</sub> is significantly stronger than that on C-A<sub>3</sub> during about 70% of simulation time. The values of vdW force on the two oligonucleotides follows a Gaussian distribution (Supplementary Fig. S9), giving a mean vdW force of 23.5 kcal mol<sup>-1</sup> Å<sup>-1</sup> on mC-A<sub>3</sub> and 14.5 kcal mol<sup>-1</sup> Å<sup>-1</sup> on C-A<sub>3</sub>, respectively. Therefore, mC-A<sub>3</sub> experiences a stronger vdW force in its stable state. However, the EL forces on mC-A<sub>3</sub> is stronger than CA<sub>3</sub> in the first 60% of simulation time, while that on mC-A<sub>3</sub> is weaker than C-A<sub>3</sub> in the rest of simulation. Accordingly, the mean value of EL force on C-A<sub>3</sub> is 25.1 kcal mol<sup>-1</sup> Å<sup>-1</sup>, which is comparable to that on mC-A<sub>3</sub> (~ 27.6 kcal mol<sup>-1</sup> Å<sup>-1</sup>). The resultant of vdW and EL force which is the non-bonded force on mC-A<sub>3</sub> is indeed 1.35-fold stronger than C-A<sub>3</sub> (Supplementary Fig. S11). Note that the duration for mC-A<sub>3</sub> is also 1.36-fold longer than that for C-A<sub>3</sub>, suggesting a direct correlation with non-bonded force.

To further investigate the effects of non-bonded pair interaction, we take the direction of the forces into consideration. Fig. 3C plotted the decomposed vdW force (upper) and EL force (bottom), and the positive value of force represents the hindrance for the transfer of oligomers while the negative value means driving the translocation. The vdW force exerted on C-A<sub>3</sub> inclines to be more favorable for the transport of the oligomer compared to that on mC-A<sub>3</sub> during the whole simulation time. However, the EL force is apt to repel the translocation of C-A<sub>3</sub> rather than mC-A<sub>3</sub>, especially during last 15-ns simulation. The

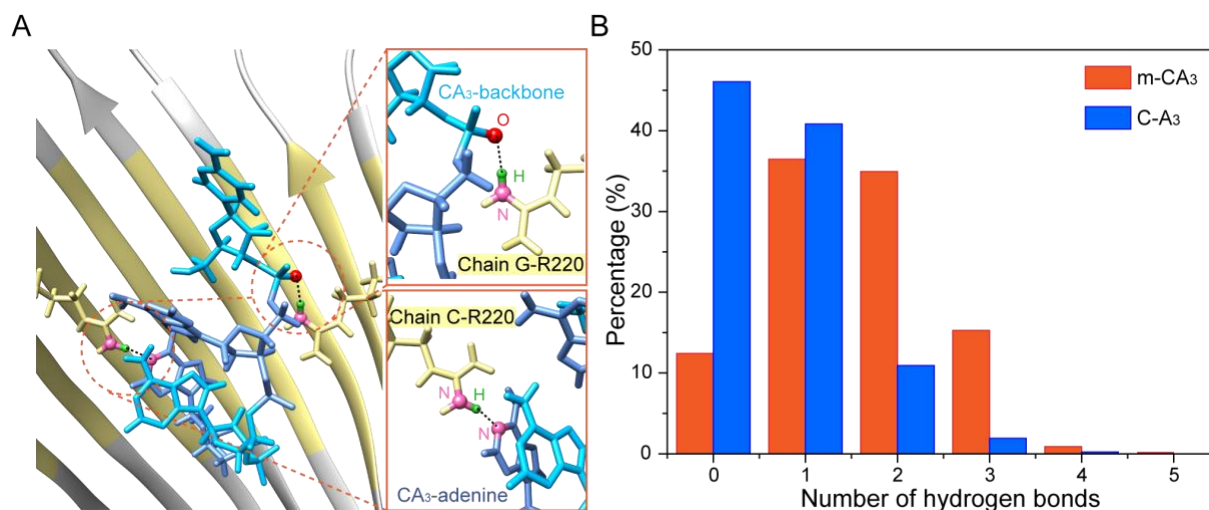


opposite effects of the vdW force and EL force could induce the conformational change of the oligonucleotides. Recent study on protein by MD simulation shows that the protein folding/unfolding inside the nanopore will change the ion mobility and further induce the substrate current blockage change<sup>41</sup>. Thus, the conformational change of mC-A<sub>3</sub> and C-A<sub>3</sub> caused by non-bonded interaction with R1 region would produce modulate the ion mobility inside aerolysin, leading to the distinguishable response of the ionic current. Moreover, two more frames with mC-A<sub>3</sub> and C-A<sub>3</sub> located in R1 were also computed (Supplementary Fig. S12). which further confirm the contribution of the vdW force and EL force between oligonucleotides and R1 region on the enhanced discrimination of methylcytosine.



**Figure 2. Non-bonded pair interaction between mC-A<sub>3</sub> (red line) or C-A<sub>3</sub> (blue line) with the aerolysin protein. (A)** The frames of mC-A<sub>3</sub> (red box) and C-A<sub>3</sub> (blue box) located near R1. The amino acid residues in R1 were marked as yellow color. **(B)** The vdW force (upper) and electrostatic force (lower) exerted on DNA by aerolysin protein. The distribution of vdW force and electrostatic force are shown in Supplementary Fig. S9-10. **(C)** Decomposed vdW force (upper) and electrostatic force (lower) along the transport orientation of oligomers. Both the force on mC-A<sub>3</sub> (red) and C-A<sub>3</sub> (blue) interacted by all amino acid residue in aerolysin lumen. The black dashed line corresponded to the value of zero.

Apart from the vdW force and EL force, hydrogen bonds are involved in the non-covalent interaction between the two oligonucleotides and aerolysin (Fig. 3A). Geometry criterion was used to define the hydrogen bonds based on the coordinates of all the atoms in the frames from above simulation. As shown in Fig. 3B, C-A<sub>3</sub> hardly form hydrogen bonds with R1 in around 46.0 % frames from the simulation, while mC-A<sub>3</sub> is favored for forming hydrogen bonds with R1 region. Besides, mC-A<sub>3</sub> is prone to form more than one hydrogen bond with R1 region at one time. These above results highly confirm that mC-A<sub>3</sub> exhibits a strong non-covalent interaction with R1 region, leading to the long duration and deep blockage amplitude.

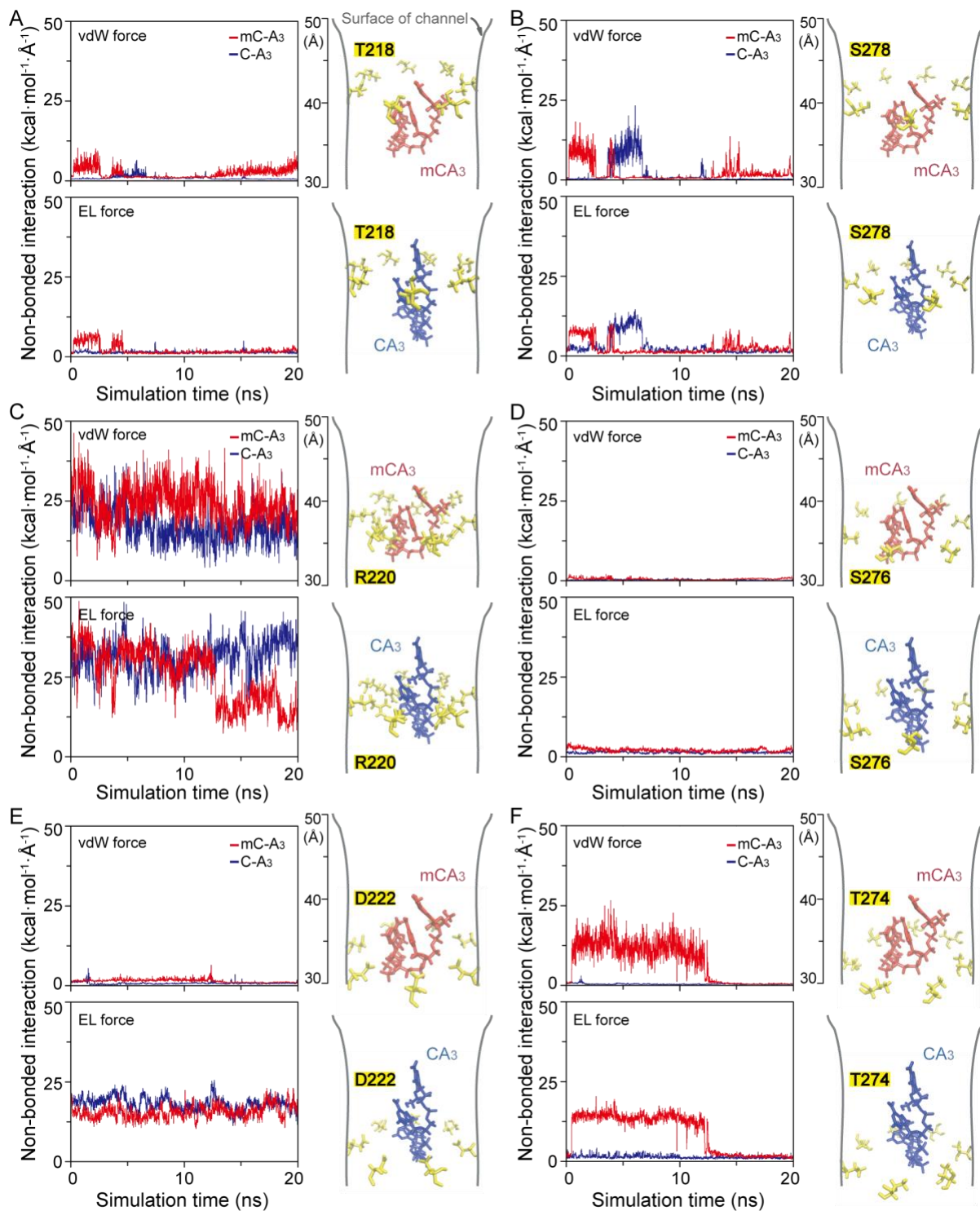


**Figure 3. Comparison of hydrogen bonds between mC-A<sub>3</sub> (red) or C-A<sub>3</sub> (blue) with R1 region of the aerolysin protein.** (A) Example of hydrogen bonds formed between CA<sub>3</sub> and amino acid residues in R1 region. The atoms that formed hydrogen bonds were drawn as ball module while the other part of oligonucleotide and residues were drawn as stick bonds. The black dash line represents the hydrogen bonds. (B) The percentage of numbers of hydrogen bonds between mC-A<sub>3</sub> (red) or C-A<sub>3</sub> (blue) and R1 region found in the frames from last 10-ns simulation. The zero value for the number of hydrogen bonds represents that no hydrogen bonds were founded between the oligonucleotides and R1 region. The hydrogen bonds were determined according to the geometry criterion, in which the H...A distance was shorter than 3.0 Å and the H-D...A angle was lower than 20 degrees. The hydrogen bonds forming between the two oligonucleotides and amino acid residues in R1 region are counted in each frame from the last 10-ns MD simulation.

During the whole dynamic interactions, every amino acid executes the unique interactions with oligonucleotide. Therefore, we further study the contribution of individual site of amino acid residues on

the non-covalent interactions. Due to the double  $\beta$ -barrel structure of aerolysin transmembrane region (Supplementary Fig. S13), two residues form a pair in each monomer. Therefore, R1 region contains 3 group of pairing residue positions: I) threonine at position 218 (T218) and serine at position 278 (S278) (Fig. 4A and 4B); II) arginine at position 220 (R220) and serine at position 276 (S276) (Fig. 4C and 4D); III) aspartic acid at position 222 (D222) and threonine at position 274 (T274) (Fig. 4E and 4F). Pair I that near the position of 43.8 Å are more inclined to interact with C-A<sub>3</sub> or mC-A<sub>3</sub> through vdW force than EL force as shown in Fig. 4A and 4B. Moreover, the mean vdW force and EL force between T218 and mC-A<sub>3</sub> are valued as high as 7.6 and 5.6 kcal mol<sup>-1</sup> Å<sup>-1</sup>, respectively; while those between T218 and C-A<sub>3</sub> remain lower than 2.0 kcal mol<sup>-1</sup> Å<sup>-1</sup> during the whole simulation time (Supplementary Fig. S14-S15). Therefore, mC-A<sub>3</sub> undergoes stronger interaction with T218 than C-A<sub>3</sub>. As for S278, its mean vdW forces and EL forces for both two oligonucleotides could reach above 9.0 kcal mol<sup>-1</sup> Å<sup>-1</sup> during 20% simulation time (Supplementary Fig. S16-S17). This result suggests that S278 undergoes relatively strong non-bonded interaction with two oligonucleotides but is not responsible for the difference in blockades. Surprisingly, the oligonucleotides show the strongest non-bonded interactions with R220 among all the sites in R1 (Fig. 4C). Compared to the entire R1 region, R220 possess the comparable values and fluctuations for both vdW and EL force (Supplementary Fig. S18-S19). Therefore, R220 dominates the non-bonded interaction between R1 and oligonucleotides. This finding is consistent with our previous works that R220 is one of the most sensitive spots of aerolysin for oligonucleotide analysis<sup>55,56</sup>. On the contrary, S276, also locating near position 37.0 Å, hardly interacts with C-A<sub>3</sub> and mC-A<sub>3</sub>, no matter via vdW or EL force (Fig. 4D and Supplementary Fig. S20-S21). At the bottom of R1 region, the D222 experiences stronger EL force of ~ 18 kcal mol<sup>-1</sup> Å<sup>-1</sup> than vdW force with the oligonucleotides (Fig. 4E and Supplementary Fig. S22-S23). However, T274, as another site in the pairing residue positions III provides a much stronger vdW force and EL force on mC-A<sub>3</sub> during 60% of simulation time compare to that on C-A<sub>3</sub> (Supplementary Fig. S24-S25). In addition, as presented in Supplementary Table S5, every individual amino acid residue in R1 region acts the similar role in the contribution of the hydrogen bonds compared to the non-bonded interactions. R220 devotes to the majority of the hydrogen bonds contribution towards the oligonucleotide with the occupancy of 115.26 % and 61.02 %, respectively. More interestingly, the hydrogen bonds formed by T274 only can be observed in mC-A<sub>3</sub> simulation with the occupancy of 37.13 %, which also agrees with the result that T274 forms much stronger non-bonded interaction with mC-A<sub>3</sub> than C-A<sub>3</sub>. However, hydrogen bonds formed by S276 or D222 on two oligonucleotides are not found during the whole simulation.

Based on the above results, here, we propose the mechanism of the R1 region interacting with mC-A<sub>3</sub> and C-A<sub>3</sub>. The residues of R220 and D222 form stronger non-covalent interaction with both two oligonucleotides, thus restricting the oligonucleotides in R1 region. In other words, R220 and D222 support the mC-A<sub>3</sub> and C-A<sub>3</sub> have the efficient long resident time to sufficiently interact with other residues in the confined R1 region. As a result, the residue of T218 and T274 in R1 region could assess to oligonucleotides. Both of these two sites are prone to provide a stronger non-covalent force on mC-A<sub>3</sub> compared to C-A<sub>3</sub>, leading to a prolonged duration and deep current blockades of mC-A<sub>3</sub>. Consequently, T218 and T274 are mainly responsible for the distinguishable difference in ionic blockages between mC-A<sub>3</sub> and C-A<sub>3</sub>. Note that the other residues inside aerolysin nanointerface also synergistic contribute to the non-covalent force difference in discriminating of mC-A<sub>3</sub> and C-A<sub>3</sub>.



**Figure 4. Non-bonded pair interaction between oligonucleotide and amino acid residues at R1 region of aerolysin lumen. (A)-(F)** The vdW force (upper of each panel) and electrostatic force (lower of each panel) on mC-A<sub>3</sub> (red line) or C-A<sub>3</sub> (blue line) interacted by amino acid residue **(A)** T218, **(B)** S278, **(C)** R220, **(D)**

S276, (E) D222 and (F) T274. Right halves of each panel: detailed DNA structure of mC-A<sub>3</sub> (red) and C-A<sub>3</sub> (blue) and corresponding amino acid residues (yellow).

Although the above simulations provide richly detailed information regarding to the difference of the non-covalent dynamics between mC-A<sub>3</sub> and C-A<sub>3</sub>, there are still short of experimental evidence to directly recognize the non-covalent force for quantitatively and qualitatively elucidating the characteristic interactions. The recorded blockage current in microseconds scale originates from the ensemble interactions between aerolysin sensing interface and the oligonucleotide. Due to the bandwidth limitation of the amplifier, the short-lived non-covalent interactions could not be directly read out by extracting the duration time and blockage current from the time-domain current traces. However, all these transient and dynamic interactions induce dynamic re-distribution of the surface charge along the oligonucleotides and the aerolysin sensing interface. As a result, each ion moving inside the nanopore interface endows a featured migration frequency as an oligonucleotide presence inside nanopore.<sup>63,64</sup> Note that the low-pass filter of the amplifier only attenuates the frequency energy from signals with frequencies higher than the cutoff frequency, but not restrain the recording of all frequency. Therefore, all these varied mobilities from dynamic moving ions have been recorded but buried in the noise of measured current signals with non-linear and non-stationary features. The traditional noise analysis methods such as Fourier transform, short-time Fourier transform (STFT) and Wavelets are used to acquire the frequency domain from time sequence data<sup>65-67</sup>, but are all subject to the Heisenberg uncertainty principle, which can not satisfy the high precise analysis of time and frequency at the same time. Moreover, the power spectral density<sup>68,69</sup> which are widely employed in single channel recording show the strength of the variations as a function of frequency. It only provides the ensemble features of the noise contribution from the fixed length of ionic current inside nanopore. Other noise analysis methods such as Hidden Markov Model<sup>70,71</sup> and Markov model<sup>72</sup> trace the hidden current state along the ionic current instead of characterizing ionic mobility inside nanopore. Therefore, here we used Hilbert–Huang Transform (HHT) to extract the frequency characters hidden in the time-sequence ionic current<sup>73</sup> which are nonlinear and nonstationary. Briefly, the recording current is decomposed into some individual monocomponent signals called intrinsic mode functions (IMFs) using ensemble empirical mode decomposition (EEMD) (details see in Supplementary Fig. S26). Next, the instantaneous frequencies in each IMFs are extracted by Hilbert Transform. Figure 5C displays the example of normalized energy at each frequency extracted from IMF6. In order to compare the ion mobility with and without the presence of oligonucleotide, both the baseline current and blockages current were performed HHT analysis as shown in Fig. 5D and 5E. The analysis

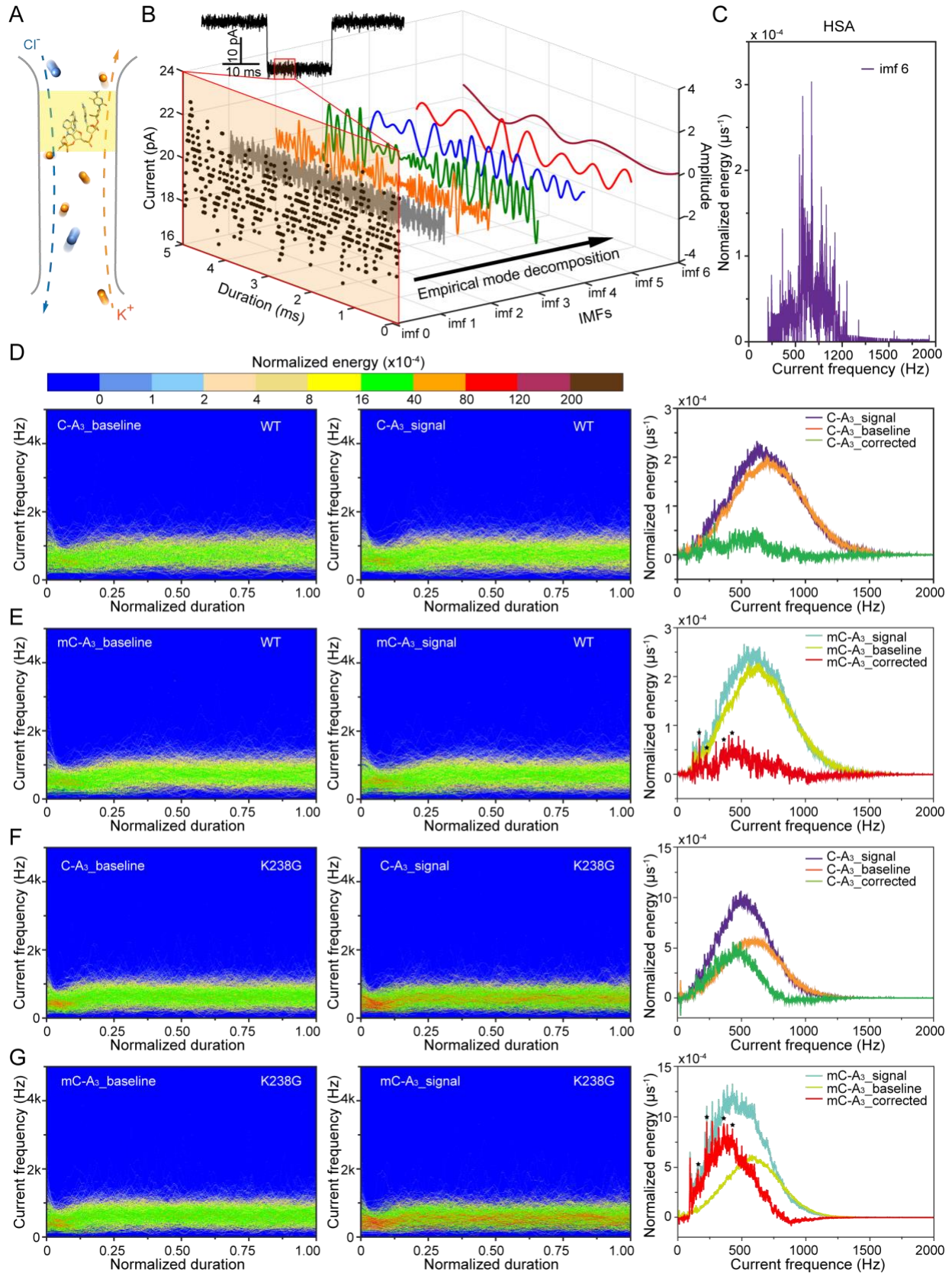
details are shown in Supplementary Methods. The average HHT spectra is a result from 600 segments of baseline current or 600 blockage events.

The HHT spectra from the baseline current show a relative stable distribution of frequency energy along the whole duration, illustrating the stability of ionic flow through the aerolysin channel in the absence of the analyte (left panel of Fig. 5D and 5E). Interestingly, the frequency distribution induced by C-A<sub>3</sub> or mC-A<sub>3</sub> inside the pore keep constant during most duration, proving that C-A<sub>3</sub> or mC-A<sub>3</sub> is mainly restrained in one sensing region during the entire translocation process. This is consistent with the MD simulations which suggest strong non-covalent interaction between C-A<sub>3</sub> or mC-A<sub>3</sub> and R1 region. Note that the energy in HHT spectra represents the amplitude of current at different frequency. Based on average HHT spectra, we counted the energy at each frequency and made the normalization of whole duration to obtain current frequency-normalized energy spectra. As shown in right panel of Fig. 5D, the C-A<sub>3</sub> inside the aerolysin pore enhances a 20% normalized energy, especially at frequency around 500 Hz. As introduced the mC-A<sub>3</sub> into the aerolysin, the current frequency shifts to a lower peak value of 450 Hz with a significant 40% increasement in normalized energy (right panel of Fig. 5E). These results can be attributed to the stronger interaction between mC-A<sub>3</sub> and R1 region, which largely enhances the vibration amplitude of the restricted ions moving through the slot between inner wall and the analyte. Therefore, the HHT analysis on ionic current traces further experimentally proves the non-covalent interaction between the oligonucleotides and aerolysin, which dominate the whole recognition and translocation process.

To further confirm this hypothesis, we further design a mutant aerolysin nanopore to strengthen the non-covalent interaction, which in turns produce a strong normalized energy in HHT spectra. Since there are two sensing regions shown in SMD simulation (Fig. 1C), we further replaced the positively charged lysine in position 238 in R2 region with glycine (denoted as K238G) to screen this positive charge and let the neighboring negative charge E258 unpair. As expected, the frequency energy from the baseline in K238G is approximately 2-fold larger than that from WT aerolysin, which shows peak frequency value at around 700 Hz. The positive shift of the peak frequency to 700 Hz indicates the violent interaction between ions and residues of K238G mutant pore. These phenomena also proved by the extra inhibition of electrostatic potential distribution in R1 region of K238G from MD simulations<sup>62</sup>. As driving the oligonucleotide into the K238G, it remarkably prolongs the duration of oligonucleotides inside the aerolysin channel. As shown in Supplementary Fig. S27, the durations of mC-A<sub>3</sub> and C-A<sub>3</sub> from K238G are 6.5 and 5.7 times larger than that from wild-type aerolysin, respectively. Therefore, compared to wild-type aerolysin, nearly 2-fold enhancement on normalized frequency energy is shown for K238G in the presence of oligonucleotide.

Still, the ions near R1 region brings the significant contribution on the totally ions' mobility inside aerolysin. Comparing to wild-type aerolysin, the frequency of both mC-A<sub>3</sub> and C-A<sub>3</sub> also exhibit a constant and similar distribution inside K238G during the most of translocation duration (Middle panel of Fig. 5F and 5G), proving the two oligonucleotides also spend most of time in R1 region. However, the K238G mutation does not affect the current separation in discriminating C-A<sub>3</sub> and mC-A<sub>3</sub>, which strongly supports that the R1 region is responsible for the sensitivity and selectivity of aerolysin nanopore. Similar to WT aerolysin, the frequency peak undergoes a negative shift while the oligonucleotides inside the K238G. For mC-A<sub>3</sub> inside K238G, it produces the larger corrected normalized energy of  $7.20 \times 10^{-4} \mu\text{s}^{-1}$  and lower peak frequency of 350 Hz compared to that of C-A<sub>3</sub> at K238G. These results further confirm that mC-A<sub>3</sub> owns a stronger non-covalent interaction with the aerolysin than C-A<sub>3</sub>. Similar to WT aerolysin, K238G interacting with mC-A<sub>3</sub> also exhibits the same sudden increases at a frequency of  $\sim 150$  Hz, 220 Hz, 360 Hz and 410 Hz (right panel of Fig. 5E and 5G), but with larger value in normalized energy. Note that the normalized energy was averaged by 600 events, the sudden increases are not caused by stochastic fluctuations. Thus, the same non-covalent interactions between mC-A<sub>3</sub> and a certain part of aerolysin induce the featured ion mobility, resulting in the above characteristic frequency spectra. Note that the sudden strong spike of 85 Hz and 255 Hz only present in K238G- mC-A<sub>3</sub> system, which might contribute to K238G induced characteristic changes in non-covalent interactions. We therefore validate that the accuracy of frequency spectra in classifying and identifying the non-covalent interaction for membrane ion channels.





**Figure 5. Hilbert–Huang Transform (HHT) analysis for the ionic current traces of mC-A<sub>3</sub> and C-A<sub>3</sub> through**

**wild type and K238G mutant aerolysin nanopore. (A)** Illustration of ions moving at different frequency with mCA3 inside the nanopore. The R1 region of the aerolysin was colored yellow. **(B)** A typical event of the ionic current for the mCA3 traversing through the wild type aerolysin nanopore at +80 mV and its corresponding intrinsic mode functions (IMFs) from the ensemble empirical mode decomposition (EEMD) results. **(C)** The normalized energy via current frequency extracted from IMF6 from single ionic blockages shown in Fig. 5B by Hilbert Transform. **(D)-(E)** Average HHT spectra of the energy-frequency-time distribution of IMF6 from current traces of baseline (*left panels*) and signals for (*middle panels*) C-A<sub>3</sub> **(D)** and mC-A<sub>3</sub> **(E)** through wild type aerolysin nanopore. The average normalized frequency-energy spectra for C-A<sub>3</sub> and mC-A<sub>3</sub> were shown in the *right panels* in **D** and **E**, respectively. **(F)-(G)** Average HHT spectra of the energy-frequency-time distribution of IMF6 from current traces of baseline (*left panels*) and signals for (*middle panels*) C-A<sub>3</sub> **(F)** and mC-A<sub>3</sub> **(G)** through K238G mutant aerolysin nanopore. The average normalized frequency-energy spectra for C-A<sub>3</sub> and mC-A<sub>3</sub> were shown in the *right panels* in **F** and **G**, respectively. The corrected normalized energy represented the differences between signal energy and baseline energy. The same sudden increases in right panel of **E** and **G** were labeled as star. The average HHT spectra and corresponding frequency-energy spectra were all acquired from 600 separate current traces.

## Discussion

Taken together, our data identify a role of non-covalent interaction for aerolysin membrane protein in transporting the carriers. Combined with a series of experiments, MD simulations and frequency-energy spectra of ionic flow, we show that (i) the non-covalent interaction dominate the recognition ability of aerolysin towards the methylcytosine and cytosine, yielding the distinguishable current events; (ii) methylcytosine experiences the strong non-covalent interaction with the aerolysin nanopore, which produces the characteristic frequency-energy spectra; (iii) the residue R220 and D222 around R1 constrain oligonucleotides in R1 region, which ensure T218 and T274 exert the discrepant non-covalent interaction on mC-A<sub>3</sub> and C-A<sub>3</sub>. Synergistic effects of amino acids in R1 region on the non-covalent interaction with oligonucleotide are credited with label-free sensing responsibility of aerolysin-based nanopore techniques on detection of epigenetic modifications, RNA modification and protein sequencing.

Since non-covalent interaction affects the ion mobility inside the nanopore, our data illuminate that the frequency-energy spectrum of ions flowing inside membrane channels bridges the gap between traditional ionic current recording and the MD simulations. We experimentally validate the obtained

hypothesis from frequency-energy spectrum by designing single-site mutation of K238G which behaves a significantly enhanced non-covalent interaction. The main frequency peak position in frequency-energy spectrum represents the main ion mobility inside membrane protein while its energy suggests the vibration amplitude of the related ions. The sudden spikes in the frequency-energy are considered as the fingerprint spectra for potential characterization of non-covalent interaction. We speculate that these spikes may originate from the specific interactions between amino acid and carriers. To date, most studies of membrane channels are performed by measuring the single channel conductance and optionally combining with genetic engineering. It is tempting to suggest that the frequency-energy spectrum of the single-channel ionic current, coupled with site-directed mutagenesis, will prove useful as a qualitative and quantify methods to describe the non-covalent interactions inside membrane channels. For example, designing the amino acids along the R1 region would allow complex patterns of non-covalent interaction to provide a characteristic frequency-energy spectrum for single molecule detection. The alternative negative and positive charges could be placed along R1 region to align the peptide with multiple charges, leaving the other side of R1 region available for peptide discriminations. Therefore, the specific spike in frequency-energy spectrum is enhanced for reading every amino acid along the peptide. Since the small organic molecule transverse through the aerolysin at a high speed ( $>$ bandwidth of the amplifier), the frequency analysis may recovery its “hidden” interaction with aerolysin at high frequency region which facilitates understanding of the ionic membrane proteins.

This presented work take the first step in developing the frequency-energy spectra to characterizing each type of non-covalent interaction between every single residue and the carriers inside membrane proteins. Because a model predicting the interaction will be complex, the frequency-energy spectrum to extract the underlying non-covalent interaction will likely involve measuring a library of characteristic spectra peak for all possible mutations. Ideally, transferring the frequency-energy spectrum from the ionic current into the voice frequency could possibly let us hear produce the rhythm from membrane channels.

## **Methods**

### **Materials and reagents.**

1,2-Diphytanoyl-*sn*-glycero-3-phosphocholine was obtained from Avanti Polar Lipids (Alabaster, AL, USA). DNA was synthesized and purified by Sangon Biotech (Shanghai, China). WT proaerolysin was purified and activated at the East China University of Science and Technology (Shanghai, China). K238G mutant

aerolysin was purified based on our previous work<sup>62</sup>. The Decane and trypsin-EDTA were purchased from Sigma-Aldrich (St. Louis, MO, USA).

### **Single channel assays.**

A Delrin bilayer cup (Warner Instruments, Hamden, CT, USA) was used to perform the experiments. DMPC solution was spread across a 50  $\mu\text{m}$  orifice in the Delrin partition. WT or K238G mutant aerolysin monomers were added into the grounded *cis* chamber to form a heptameric pore. A pair of Ag/AgCl electrodes were used to apply appropriate potentials across the membrane. Oligonucleotides (mC-A<sub>3</sub> or C-A<sub>3</sub>) were added into *cis* chamber for a final concentration of 5  $\mu\text{M}$  which for purpose of a high-speed collection for a large amount of data. All of the experiments were conducted by premixing DNA samples and electrolyte solution before the formation of a single aerolysin pore. Currents were recorded using an Axopatch 200B patch-clamp amplifier (Molecular Devices, Sunnyvale, CA, USA) coupled with a Digidata 1440A A/D converter (Molecular Devices, USA). The signals were filtered at a frequency of 5 kHz and acquired with Clampex 10.4 software (Molecular Devices, USA) at a sampling frequency of 100 kHz.

### **Data analysis.**

To remove the collision events, a reduction of events in durations < 0.2 ms was required. Data were analysed with MOSAIC<sup>74,75</sup> and OriginLab 8.0 (OriginLab Corporation, Northampton, MA, USA).

### **Molecular dynamics simulations.**

#### *Model Construction.*

The wild-type aerolysin cryo-EM structure coordinates were obtained from Protein Data Bank entry 5JZT<sup>55</sup>. The aerolysin was firstly solvated by SOLVATE<sup>76</sup> program to produces a 3-Å-thick shell of water around the aerolysin using 8 Gaussians to approximate the protein surface. The resulting structure was merged with a 200 Å x 200 Å path of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) lipid bilayer, aligned to the xy-plane plane with the hydrophobic  $\beta$ -barrel of aerolysin embedded in the bilayer while its hydrophilic cap protruded above the membrane. All the lipid and water molecules overlapping with the protein nanopore were removed. The nanopore-lipid system was solvated with pre-equilibrated TIP3P water molecules<sup>77</sup> using VMD<sup>78</sup>. K<sup>+</sup> and Cl<sup>-</sup> ions were added to the solution to neutralize the aerolysin system and achieve a concentration of 1.0 M. A ssDNA of 5'-CAAA-3' was derived from a model of double-stranded DNA, created with the 3D-DART web server<sup>79</sup>, while 5'-mCAAA-3' was generated by patching a methyl group to cytosine of 5'-CAAA-3' described by CHARMM36 force field<sup>80</sup>.

#### *System Equilibrium.*

All MD simulation were performed using program NAMD<sup>60</sup> with the visualization and analysis by VMD<sup>78</sup>. Aerolysin protein and oligonucleotide molecule were described by CHARMM36 force filed<sup>80</sup>. After a 5000-

step minimization, the aerolysin nanopore system was equilibrated in the NPT ensemble at 295 K and 1 atm for 2 ns with the heavy atoms (non-hydrogen) of protein restrained and allowing the relaxation of lipid, water, ions and protein hydration atoms. Afterwards, the alpha carbons of the protein were still restrained, and the equilibration simulation lasted 10 ns. Next, all the restraints were removed, and the system was relaxed for 60 ns. The 5'-CAAA-3' and 5'-mCAAA-3' were also equilibrated in NPT ensemble for 20 ns at 295 K in a 40 Å x 40 Å x 40 Å water box at 1.0 M KCl. The integration time step chosen was 2 fs. All NPT ensemble simulations were carried out with a Langevin piston for pressure control and Langevin dynamics for temperature control<sup>56</sup>. Periodic boundary conditions were employed in all directions. van der Waals energies were calculated using a 12 Å cutoff and the particle-mesh-Ewald (PME) method was used to treat long-range electrostatics<sup>81</sup>.

#### *Constant Velocity SMD and MD Simulation.*

Steered molecule dynamic simulation was used to pull ssDNA through the aerolysin nanopore. The oligonucleotide was put at the entrance of aerolysin with a certain position (-10.0, 0.0, 65.0) for the mass center of 5'-CAAA-3' and 5'-mCAAA-3'. The initial conformations of oligonucleotides were determined by the equilibrium of NPT ensemble and the translocation was initiated by its 3' end. The SMD atoms, backbone of oligomers, are attached to a dummy atom via a virtual spring with a spring constant of 7.0 kcal mol<sup>-1</sup> Å<sup>-2</sup>. The dummy atom is moved to traverse the pore from *cis* to *trans* with a constant velocity of 0.25 Å ps<sup>-1</sup>.

For the study of non-covalent interaction between oligonucleotides and aerolysin protein. Two frames, selected from the trajectories of C-A<sub>3</sub> and mC-A<sub>3</sub> in SMD simulation with the center of mass for the two oligonucleotides both at position 35.2 Å, were used to perform MD simulation for 20 ns. Harmonic constraints were imposed on the backbone of the oligomers to restrain the oligonucleotides around R1 region. Based on the trajectories from MD simulation, NAMD energy plugin in VMD was used to calculate the non-bonded interaction between oligonucleotides with the whole aerolysin protein or individual amino acid residues in R1 region using 12 Å cutoff for each snapshot of MD trajectory. Hydrogen bonds between the oligonucleotides and aerolysin protein were calculated with the donor-acceptor distance (H...A distance) shorter than 3.0 Å and the H-D...A angle lower than 20 degrees. The hydrogen bonds forming between the two oligonucleotides and amino acid residues in R1 region are counted in each frame from the last 10-ns MD simulation.

#### **Hilbert–Huang Transform**

The processes of Hilbert-Huang Transform (HHT) analysis was presented in Supplementary Fig. S26. The methods of Hilbert Transform and empirical mode decomposition were described in our previous work<sup>62</sup>.

Due to the exponential distribution of translocation duration for the two oligonucleotides, the time normalization was applied to avoid the interferences of events duration. For each sample, we used 600 events and 600 baselines from current recording. Due to the filter used in the experiments was 5 kHz, after empirical mode decomposition, imf6 was chosen to extract the instantaneous frequencies by Hilbert Transform. Average energies at each instantaneous frequency were calculated by 600 recording currents described above.

## References

1. Swartz, K. Sensing voltage across lipid membranes. *Nature* **456**, 891-897 (2008).
2. Ding, J. *et al.* Pore-forming activity and structural autoinhibition of the gasdermin family. *Nature* **535**, 111-116 (2016).
3. Mahadevi, A. S. & Sastry, G. N. Cooperativity in noncovalent interactions. *Chem. Rev.* **116**, 2775–825 (2016).
4. Guo, W., Tian, Y. & Jiang, L. Asymmetric ion transport through ion-channel-mimetic solid-state nanopores. *Acc. Chem. Res.* **46**, 2834–2846 (2013).
5. Ghai, I., Winterhalter, M. & Wagner, R. Probing transport of charged  $\beta$ -lactamase inhibitors through OmpC, a membrane channel from E. coli. *Biochem. Biophys. Res. Co.* **484**, 51–55 (2017)
6. Wzorek, J., Lee, J., Tomasek, D., Hagan, C. & Kahne, D. Membrane integration of an essential  $\beta$ -barrel protein prerequires burial of an extracellular loop. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2598–2603 (2017).
7. Patrick, J. *et al.* Allostery revealed within lipid binding events to membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 2976-2981 (2018).
8. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat. Neurosci.* **8**, 1263–1268 (2005).
9. Taslimi, A. *et al.* An optimized optogenetic clustering tool for probing protein interaction and function. *Nat. Commun.* **5**, 4925 (2014).
10. Sitters, G. *et al.* Acoustic force spectroscopy. *Nat. Methods* **12**, 47-50 (2015).
11. Wang, Q. & Moerner, W. Single-molecule motions enable direct visualization of biomolecular interactions in solution. *Nat. Methods* **11**, 555-558 (2014).
12. Neuman, K. C., Abbondanzieri, E. A., Landick, R., Gelles, J. & Block, S. M. Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* **115**, 437–47 (2003).
13. Neuman, K. & Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* **5**, 491–505 (2008).
14. Plevin, M., Bryce, D. & Boisbouvier, J. Direct detection of CH/ $\pi$  interactions in proteins. *Nat. Chem.* **2**, 466-471 (2010).
15. Bayley, H. & Cremer, P. Stochastic sensors inspired by biology. *Nature* **413**, 226–230 (2001).

16. Robertson, J., Kasianowicz, J. & Banerjee, S. Analytical approaches for studying transporters, channels and porins. *Chem. Rev.* **112**, 6227–6249 (2012).
17. Derrington, I. *et al.* Nanopore DNA sequencing with MspA. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 16060–16065 (2010).
18. Liu, N., Hou, R., Gao, P., Lou, X. & Xia, F. Sensitive Zn<sup>2+</sup> sensor based on biofunctionalized nanopores via combination of DNAzyme and DNA supersandwich structures. *Analyst* **141**, 3626–3629 (2016).
19. Fahie, M., Chisholm, C. & Chen, M. Resolved single-molecule detection of individual species within a mixture of anti-biotin antibodies using an engineered monomeric nanopore. *ACS Nano* **9**, 1089–1098 (2015).
20. Ghai, I. *et al.* A general method to determine the flux of charged molecules through nanopores applied to  $\beta$ -lactamase inhibitors and OmpF. *J. Phys. Chem. Lett.* **8**, 1295–1301 (2017).
21. Huang, S. *et al.* High-throughput optical sensing of nucleic acids in a nanopore array. *Nat. Nanotechnol.* **10**, 986–991 (2015).
22. Wang, Y., Zheng, D., Tan, Q., Wang, M. & Gu, L.-Q. Nanopore-based detection of circulating microRNAs in lung cancer patients. *Nat Nanotechnol* **6**, 668 (2011).
23. Gooding, J. & Gaus, K. Single-Molecule Sensors: Challenges and Opportunities for Quantitative Analysis. *Angew. Chem. Int. Ed.* **55**, 11354–11366 (2016).
24. Haywood, D., Saha-Shah, A., Baker, L. & Jacobson, S. Fundamental Studies of Nanofluidics: Nanopores, Nanochannels, and Nanopipets. *Anal. Chem.* **87**, 172–187 (2015).
25. Stoddart, D., Heron, A. J., Mikhailova, E., Maglia, G. & Bayley, H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7702–7707 (2009).
26. Clarke, J. *et al.* Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* **4**, 265–270 (2009).
27. Manrao, E., *et al.* Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat. Biotechnol.* **30**, 349–353 (2012).
28. Garalde, D. *et al.* Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* **15**, 201–206 (2018).
29. Feng, J. *et al.* Identification of single nucleotides in MoS<sub>2</sub> nanopores. *Nat. Nanotechnol.* **10**, 1070–1076 (2015).
30. Hornblower, B. *et al.* Single-molecule analysis of DNA-protein complexes using nanopores. *Nat. Methods* **4**, 315–317 (2007).
31. Reiner, J., Kasianowicz, J., Nablo, B. & Robertson, J. Theory for polymer analysis using nanopore-based single-molecule mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12080–12085 (2010).
32. Baaken, G. *et al.* High-resolution size-discrimination of single nonionic synthetic polymers with a highly charged biological nanopore. *ACS Nano* **9**, 6443–6449 (2015).
33. Piguet, F. *et al.* Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore. *Nat. Commun.* **9**, 966 (2018).
34. Uram, J., Ke, K., Hunt, A. & Mayer, M. Label-free affinity assays by rapid detection of immune complexes in submicrometer pores. *Angew. Chem. Int. Ed.* **118**, 2339–2343 (2006).
35. Soskine, M., Biesemans, A., Maeyer, M. & Maglia, G. Tuning the size and properties of ClyA nanopores assisted by directed evolution. *J. Am. Chem. Soc.* **135**, 13456–13463 (2013).

36. Wong, C. & Muthukumar, M. Polymer capture by electroosmotic flow of oppositely charged nanopores. *J. Chem. Phys.* **126**, 164903 (2007).
37. Deamer, D. & Branton, D. Characterization of nucleic acids by nanopore analysis. *Acc. Chem. Res.* **35**, 817–825 (2002).
38. Asandei, A. *et al.* Electroosmotic trap against the electrophoretic force near a protein nanopore reveals peptide dynamics during capture and translocation. *ACS Appl. Mater. Inter.* **8**, 13166–13179 (2016).
39. Li, S., Cao, C., Yang, J. & Long, Y.-T. Detection of peptides with different charges and lengths by using the aerolysin nanopore. *ChemElectroChem* (2018). DOI:10.1002/celec.201800288
40. Nivala, J., Marks, D. & Akeson, M. Unfoldase-mediated protein translocation through an  $\alpha$ -hemolysin nanopore. *Nat. Biotechnol.* **31**, 247–250 (2013).
41. Yusko, E. *et al.* Real-time shape approximation and fingerprinting of single proteins using a nanopore. *Nat. Nanotechnol.* **12**, 360–367 (2017).
42. Japrun, D. *et al.* Single-molecule studies of intrinsically disordered proteins using solid-state nanopores. *Anal. Chem.* **85**, 2449–2456 (2013).
43. Singer, A. *et al.* Nanopore based sequence specific detection of duplex DNA for genomic profiling. *Nano Lett.* **10**, 738–742 (2010).
44. Si, W & Aksimentiev, A. Nanopore sensing of protein folding. *ACS Nano* 2017, 11, 7091–7100 (2017).
45. Johnson, R. P., Fleming, A., Beuth, L., Burrows, C. & White, H. Base flipping within the  $\alpha$ -hemolysin latch allows single-molecule identification of mismatches in DNA. *J. Am. Chem. Soc.* **138**, 594–603 (2016).
46. Hornblower, B. *et al.* Single-molecule analysis of DNA-protein complexes using nanopores. *Nat. Methods* **4**, 315–317 (2007).
47. Wanunu, M. *et al.* Rapid electronic detection of probe-specific microRNAs using thin nanopore sensors. *Nat. Nanotechnol.* **5**, 807–814 (2010).
48. Soskine, M., Biesemans, A. & Maglia, G. Single-molecule analyte recognition with ClyA nanopores equipped with internal protein adaptors. *J. Am. Chem. Soc.* **137**, 5793–5797 (2015).
49. Cao, C. *et al.* Discrimination of oligonucleotides of different lengths with a wild-type aerolysin nanopore. *Nat. Nanotechnol.* **11**, 713–718 (2016).
50. Cao, C., Liao, D.-F., Yu, J., Tian, H. & Long, Y.-T. Construction of an aerolysin nanopore in a lipid bilayer for single-oligonucleotide analysis. *Nat. Protoc.* **12**, 1901–1911 (2017).
51. Cao, C., Yu, J., Li, M.-Y., Wang, Y.-Q., Tian, H. & Long, Y.-T., Direct readout of single nucleobase variations in an oligonucleotide. *Small* **13**, 1702011 (2017).
52. Yu, J., Cao, C. & Long, Y.-T. Selective and sensitive detection of methylcytosine by aerolysin nanopore under serum condition. *Anal. Chem.* **89**, 11685–11689 (2017).
53. Parker, M. *et al.* Structure of the *Aeromonas* toxin proaerolysin in its water-soluble and membrane-channel states. *Nature* **367**, 292–295 (1994).
54. Cao, C. & Long, Y.-T. Biological nanopores: confined spaces for electrochemical single-molecule analysis. *Acc. Chem. Res.* **51**, 331–341 (2018).
55. Iacovache, I. *et al.* Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process. *Nat. Commun.* **7**, 12062 (2016).
56. Hu, Z.-L. *et al.* Real-time and accurate identification of single oligonucleotide photoisomers via an aerolysin nanopore. *Anal. Chem.* **90**, 4268–4272 (2018).



57. Breton, M. *et al.* Exploration of neutral versus polyelectrolyte behavior of poly(ethylene glycol)s in alkali ion solutions using single-nanopore recording. *J. Phys. Chem. Lett.* **4**, 2202–2208 (2013).
58. 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).
59. Kowalczyk, S., Wells, D., Aksimentiev, A. & Dekker, C. Slowing down DNA translocation through a nanopore in lithium chloride. *Nano Lett.* **12**, 1038–1044 (2012).
60. Phillips, J. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
61. Cao, C. *et al.* Mapping the sensing spots of aerolysin for single oligonucleotides analysis. *Nat. Commun.* **9**, 2823 (2018).
62. Wang, Y.-Q. *et al.* Identification of essential sensitive regions of the aerolysin nanopore for single oligonucleotide analysis. *Anal. Chem.* **90**, 7790–7794 (2018).
63. Hoogerheide, D., Garaj, S. & Golovchenko, J. Probing surface charge fluctuations with solid-state nanopores. *Phys. Rev. Lett.* (2009) DOI: 10.1103/PhysRevLett.102.256804.
64. Kesselheim, S., Müller, W. & Holm, C. Origin of current blockades in nanopore translocation experiments. *Phys. Rev. Lett.* **112**, 018101 (2014).
65. Neild, S. A., McFadden, P. D. & Williams. A review of time-frequency methods for structural vibration analysis. *Engineering Structures* **25**, 713–728 (2003).
66. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *Ieee T Inform Theory* **36**, 961–1005 (1990).
67. Shao, X.-G., Leung, A. & Chau, F.-T. Wavelet: A New Trend in Chemistry. *Acc. Chem. Res.* **36**, 276–283 (2003).
68. Bezrukov, SM & Kasianowicz, JJ. Current noise reveals protonation kinetics and number of ionizable sites in an open protein ion channel. *Phys. Rev. Lett.* **70**, 2352–2355 (1993).
69. Bezrukov, S. M. & Vodyanoy, I. Noise-induced enhancement of signal transduction across voltage-dependent ion channels. *Nature* **378**, 362–4 (1995).
70. Robertson, J. *et al.* Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8207–8211 (2007).
71. Zhang, J. *et al.* High-bandwidth nanopore data analysis by using a modified hidden Markov model. *Nanoscale* **9**, 3458–3465 (2017).
72. Bodrenko, I., Wang, J., Salis, S., Winterhalter, M. & Ceccarelli, M. Sensing Single Molecule Penetration into Nanopores: Pushing the Time Resolution to the Diffusion Limit. *ACS Sens.* **2**, 1184–1190 (2017).
73. Liu, S.-C. *et al.* Measuring a frequency spectrum for single-molecule interactions with a confined nanopore. *Faraday Discuss.* (2018) DOI:10.1039/C8FD00023A.
74. Balijepalli, A., Etteedgui, J., Cornio, A. T., Robertson, J. W. F., Cheung, K. P., Kasianowicz, J. J. & Vaz, C., Quantifying short-lived events in multistate ionic current measurements. *ACS Nano* **8**, 1547–1553 (2014).
75. Forstater, J. H., Briggs, K., Robertson, J. W. F., Etteedgui, J., Marie-Rose, O., Vaz, C., Kasianowicz, J. J., Tabard-Cossa, V. & Balijepalli, A. MOSAIC: A modular single-molecule analysis interface for decoding multistate nanopore data. *Anal. Chem.* **88**, 11900–11907 (2016).
76. Grubmüller, H. & Groll, V. Solvate 1.0.1. <http://www.mpibpc.mpg.de/grubmueller/solvate>, 1996–2010.

77. Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, R. & Klein, M. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
78. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
79. Van Dijk, M. & Bonvin, A. 3D-DART: a DNA structure modelling server. *Nucleic Acids Res.*, **37**, W235-W239 (2009).
80. Jr, M. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
81. Batcho, P., Case, D. & Schlick, T. Optimized particle-mesh Ewald/multiple-time step integration for molecular dynamics simulations. *J. Chem. Phys.* **115**, 4003–4018 (2001).

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (21834001 and 61871183), Innovation Program of Shanghai Municipal Education Commission (2017-01-07-00-02-E00023) and the Fundamental Research Funds for the Central Universities (222201718001, 222201717003).

## Author contributions

M.Y.L., Y.L.Y. and Y.T.L. conceived the idea; J.Y. and C.C. performed the nanopore experiments on WT aerolysin, M.Y.L performed nanopore experiments on K238G mutant aerolysin; M.Y.L. and S.C.L. built the aerolysin system module and performed the molecular dynamics simulations; M.Y.L., X.X.F., Y.J.W and Y.L.Y. designed and carried out the HHT analysis; M.Y.L., J.Y., Y.Q.W. and S.L. performed image processing; M.Y.L., Y.L.Y., Y.T.L. and Y.J.W. interpreted the data; M.Y.L., Y.L.Y., J.Y. and Y.T.L. wrote the paper; Y.T.L. supervised the project.

## Competing financial interests

The authors declare no competing financial interests.