Identification of single amino acid chiral and positional isomers using an electrostatically asymmetric nanopore

Jiajun Wang[a]§, Jigneshkumar Dahyabhai Prajapati[b]§†, Fan Gao[a]§, Yi-Lun Ying*[a,c], Ulrich Kleinekathöfer[b], Mathias Winterhalter*[d], Yi-Tao Long[a]

[a] State Key Laboratory of Analytical Chemistry for Life Science, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, China

[b] Department of Physics and Earth Sciences, Jacobs University Bremen, 28759, Germany

[c] Chemistry and Biomedicine Innovation Center, Nanjing University, Nanjing 210023, China

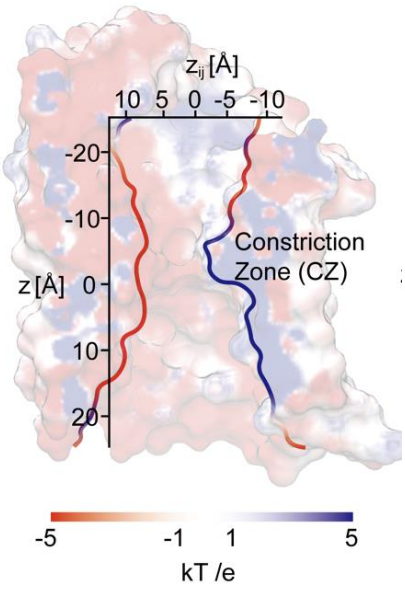
[d] Department of Life Sciences and Chemistry, Jacobs University Bremen, 28759 Bremen, Germany

***Corresponding authors: yilunying@nju.edu.cn; m.winterhalter@jacobs-university.de**

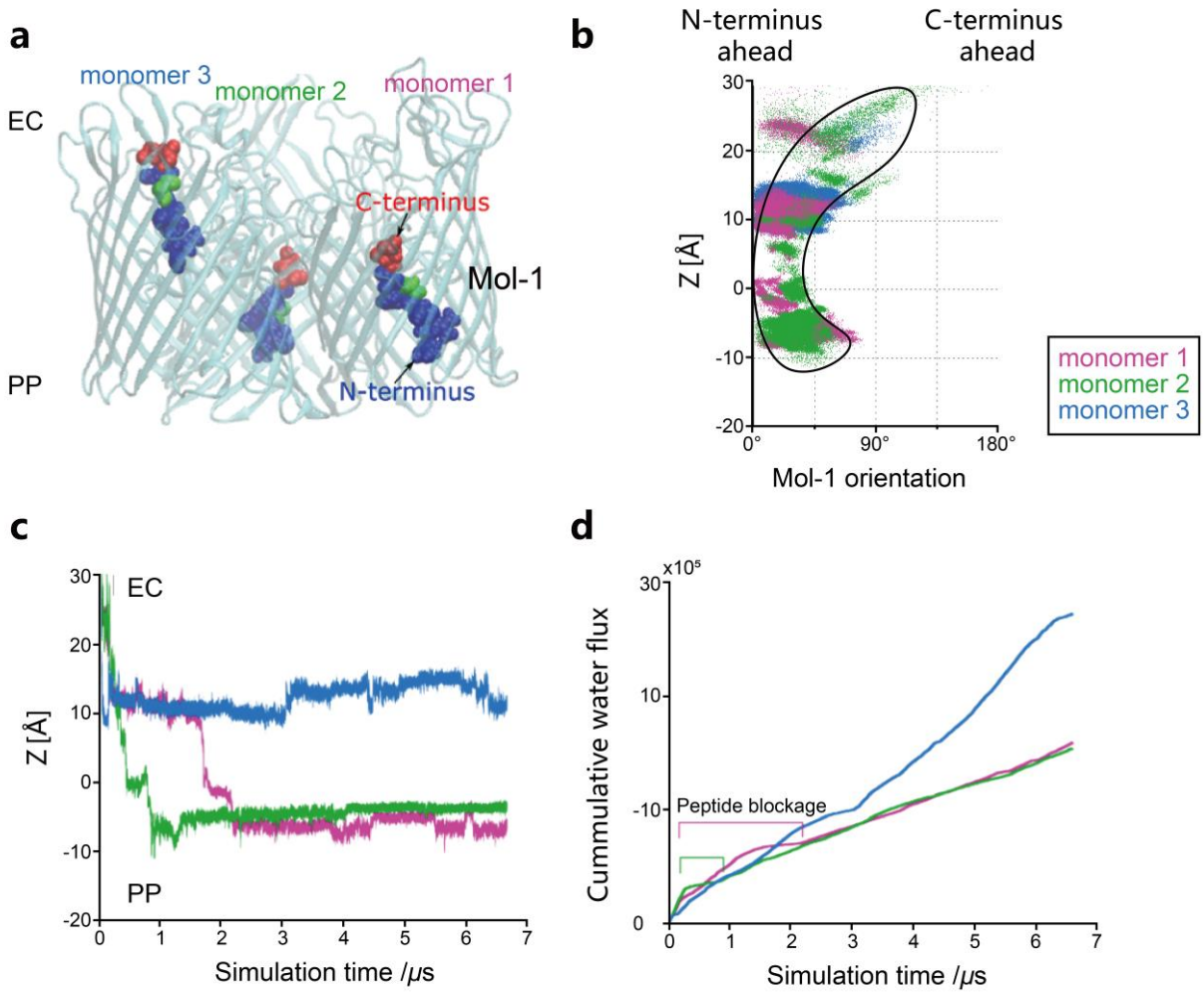
§J.W., J.D.P. and F.G. contributed equally.

Supporting Information

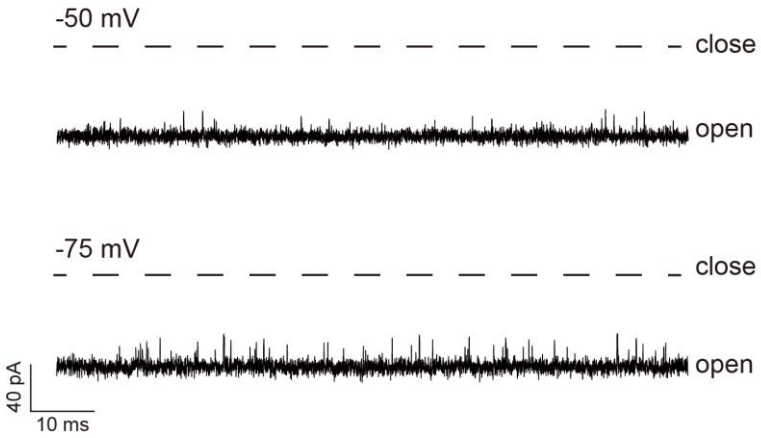
Table of Contents



Supplementary Figure S1 The electrostatic potential surface of the OmpF monomer illustrating the two distinctly charged regions within the constriction zone (CZ) of the pore.



Supplementary Figure S2 Results from a 6.8 μs-long applied-field MD simulation performed at +1 V applied voltage. (a) Confirmations of the peptide inside the OmpF monomers at 6.8 μs, where the N-terminus is always seen to orient towards the periplasmic side (PP) during the permeation. Mol-1 has initially been placed at the extracellular (EC) side. The peptides in monomer 1 and monomer 2 have been seen to translocate at this time point, while the peptide in monomer 3 has not yet reached the periplasmic (PP) side. (b) Distribution of the peptide positions and orientations of Mol-1 inside the different OmpF monomers. The Mol-1 interacts with the OmpF with its N-terminus while there is almost no interaction if C-terminus is placed ahead. (c) Position of the N-terminus of the Mol-1 peptides along the pore axis as a function of simulation time. The three lines represent the dynamics of the peptides inside the three monomers of OmpF. Peptides in monomer 1 and monomer 2 reached -10 Å suggesting that translocation has taken place. (d) Cumulative water flux through the three monomers as a function of simulation time. The water flux was seen to rise slowly but never stopped, indicating that the peptides never block the channel completely.



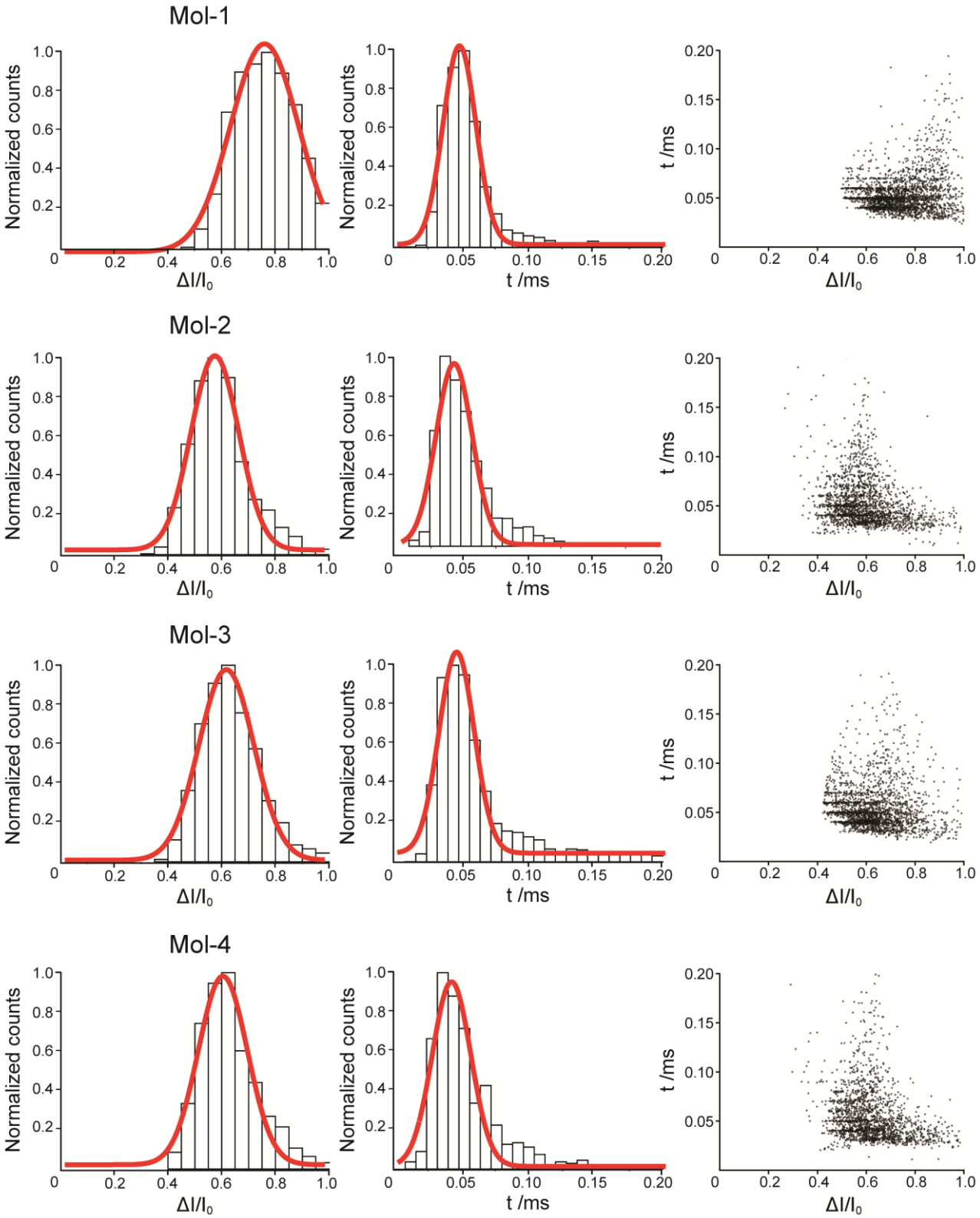
Supplementary Figure S3 Baseline recording of the OmpF in 1.0 M KCl, pH 8.0 under -50 mV and -75 mV. The OmpF produces small gating spikes (<10% of open channel current) from the baseline which is distinct from the interaction events with peptides (>70% of open channel current).

Supplementary Table S1 The blockade $\triangle I/I_0$ and residence time of peptide sensing with OmpF

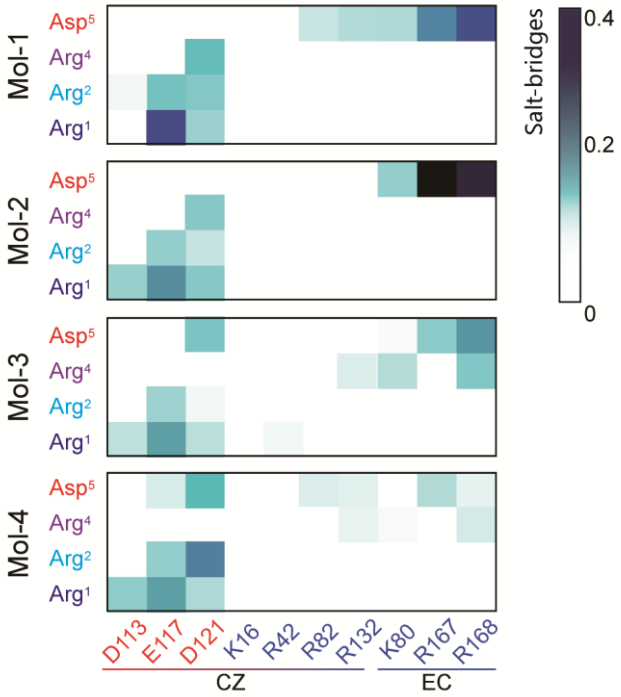
Peptide	$\triangle I/I_0$ *	t /μs *	Electrolyte Concentration#
Mol-1	65.9 ± 8.7%	47.8 ± 1.2	1.0 M KCl
Mol-2	61.1 ± 0.9%	46.8 ± 0.7	
Mol-3	59.4 ± 1.6%	47.2 ± 0.3	
Mol-4	60.6 ± 0.3%	47.6 ± 0.2	
Mol-1	68.0 ± 9.8%	66.1 ± 2.0	0.5 M KCl
Mol-5	68.5 ± 7.2%	112.3 ± 33.0	
Mol-6	74.4 ± 3.9%	137.9 ± 37.0	
Mol-7	76.9 ± 3.1%	64.9 ± 0.3	
Mol-8	73.7 ± 9.1%	104.9 ± 23.0	
Mol-9	73.7 ± 3.6%	61.3 ± 0.1	

*Errors are from three independent experiments.

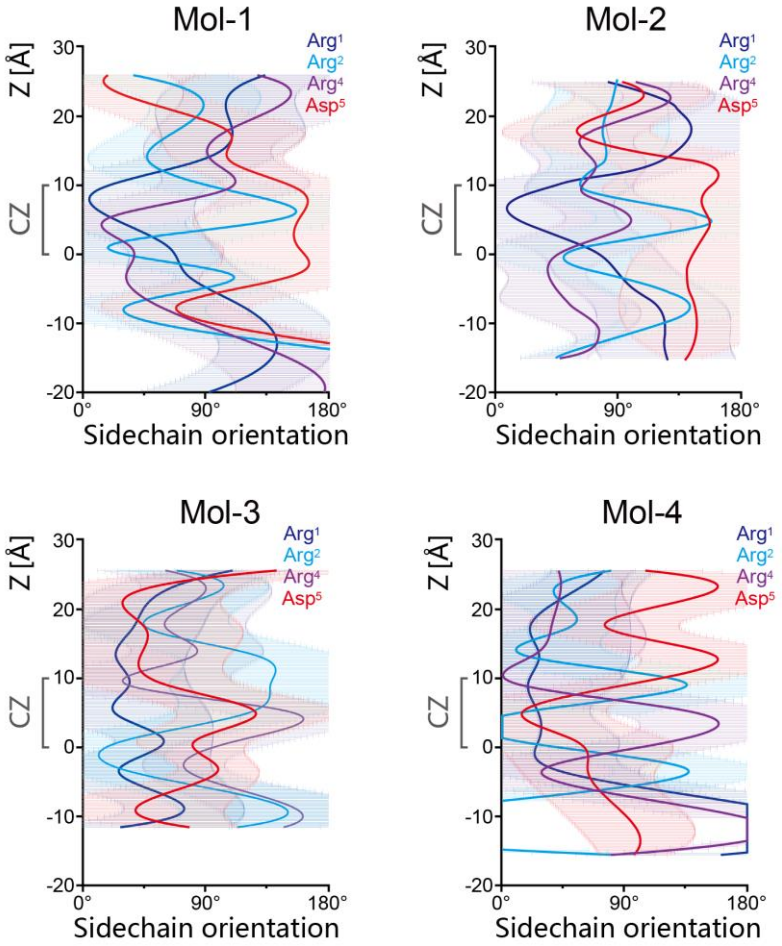
#There are two measure conditions differing in the electrolyte concentration. 1.0 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0 corresponds to Mol-1, 2, 3, 4, and 0.5 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0 corresponds to Mol-1, 5, 6, 7, 8, 9.



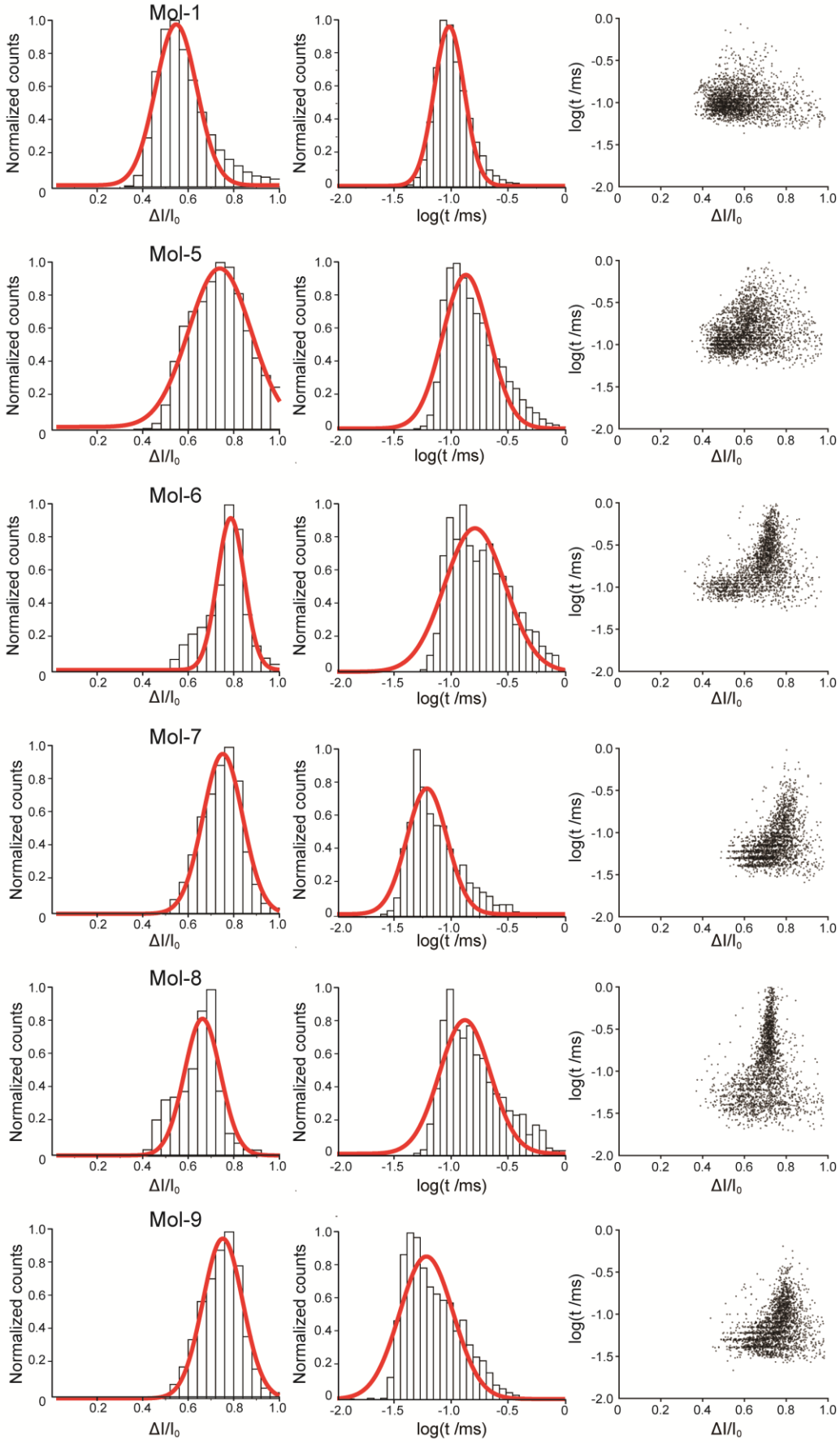
Supplementary Figure S4 Histogram and scatter analysis of the blockade $\triangle I/I_0$ and residence time *t* plotted for Mol-1, 2, 3, and 4 detailing the color maps in Figure 1e. The single peptide events were measured at a bias voltage of -50 mV in 1.0 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0. At least 2000 single-molecule events have been analyzed per peptide. The data processing method has been obtained as described in Supplementary Note 3.



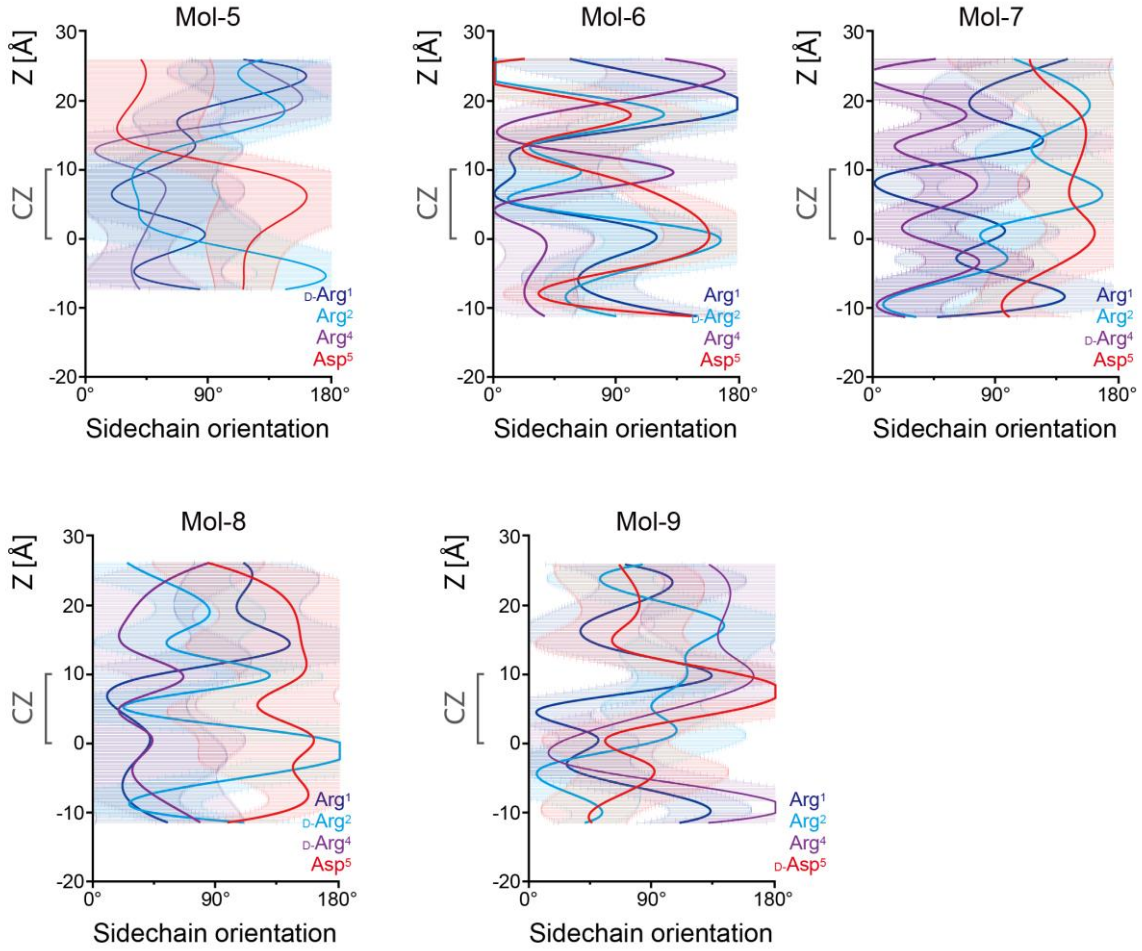
Supplementary Figure S5 Calculated the average number of salt bridges between the charged residues of the peptides and the pore from the steered MD trajectories. The number of salt bridges has been averaged over the total simulation time. A cut-off distance of 3.2 Å was used to identify salt bridges.[33,34]



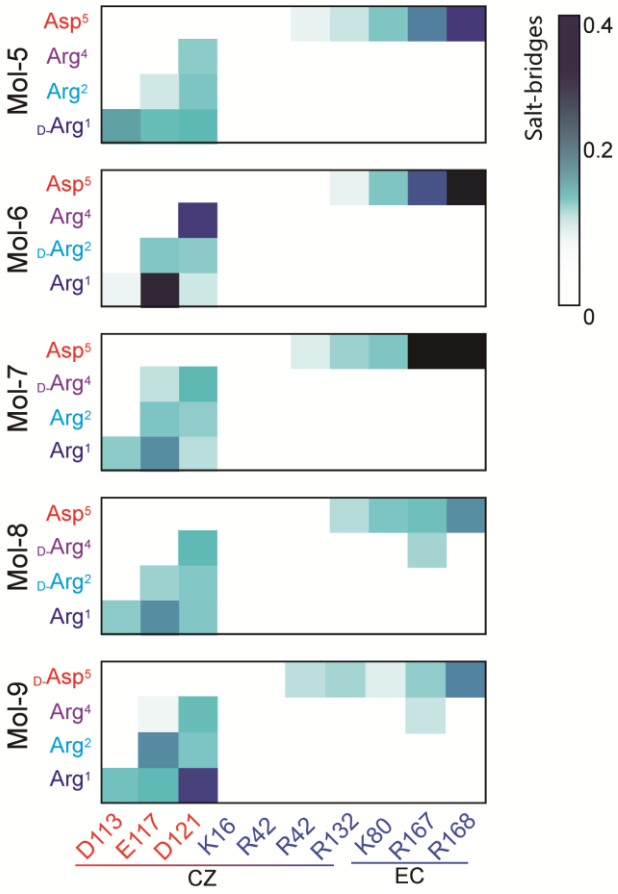
Supplementary Figure S6 Dynamics of the sidechain orientations of the residues in Mol-1, 2, 3, and 4 along the pore axis. An angle of 0° indicates that the side chain points towards the acidic group in the negatively charged pocket, while it is oriented towards the basic ladder at 180°. Results from average of 5 times simulation experiments were given. CZ indicated the narrow constriction zone in the OmpF.



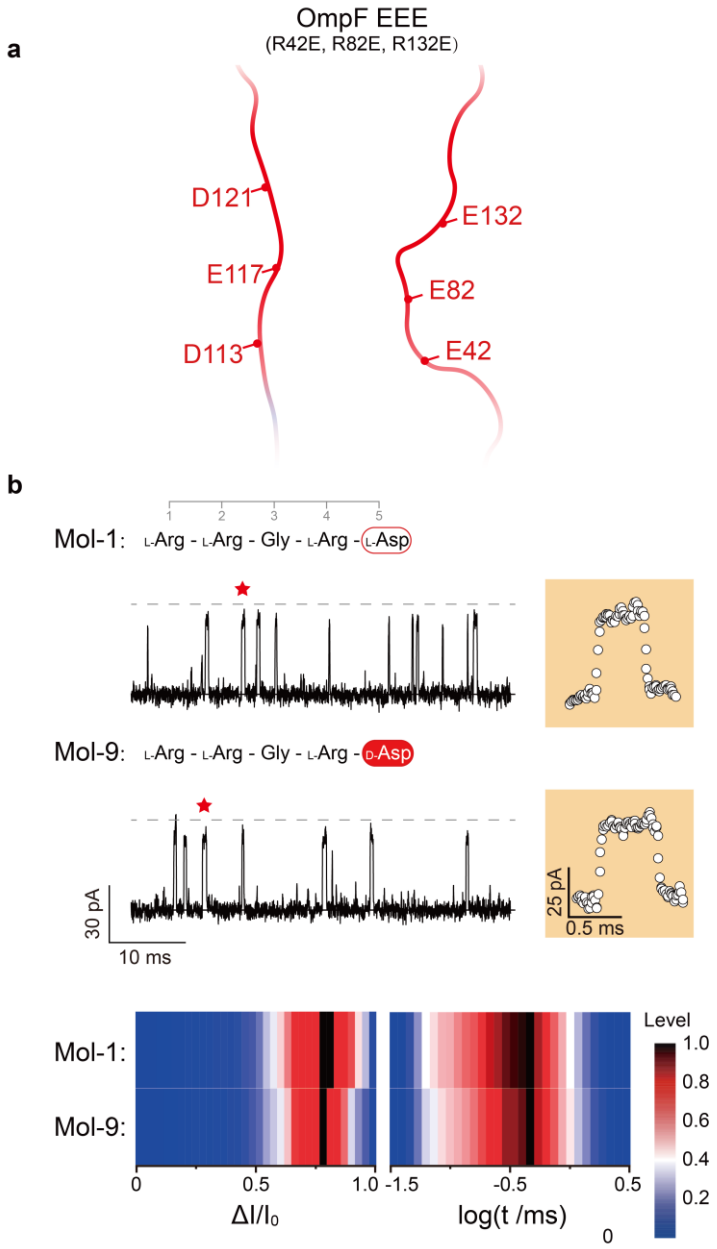
Supplementary Figure S7 Histogram and scatter analysis of the blockade $\triangle I/I_0$ and the logarithm of the residence time *t* plotted for Mol-1, 5, 6, 7, 8, and 9 detailing the color maps in Figure 2c. The single peptide events were measured at a bias voltage of -50 mV in 0.5 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0. At least 2000 single-molecule events have been analyzed per peptide. The data processing method has been obtained as described in Supplementary Note 3.



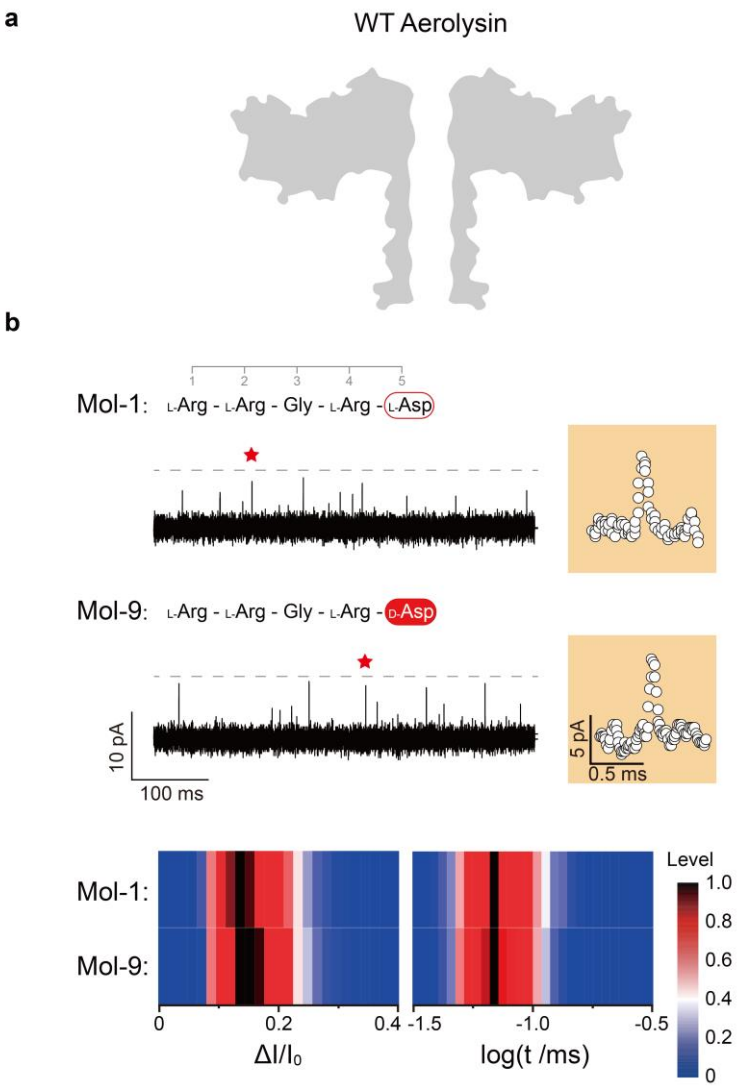
Supplementary Figure S8 Dynamics of the sidechain orientations of the residues in Mol-5, 6, 7, 8, and 9 along the pore axis. An angle of 0° indicates that the side chain points towards the acidic group in the negatively charged pocket, while it is oriented towards the basic ladder at 180°. Results from average of 5 times experiments were given. CZ indicated the narrow constriction zone in the OmpF.



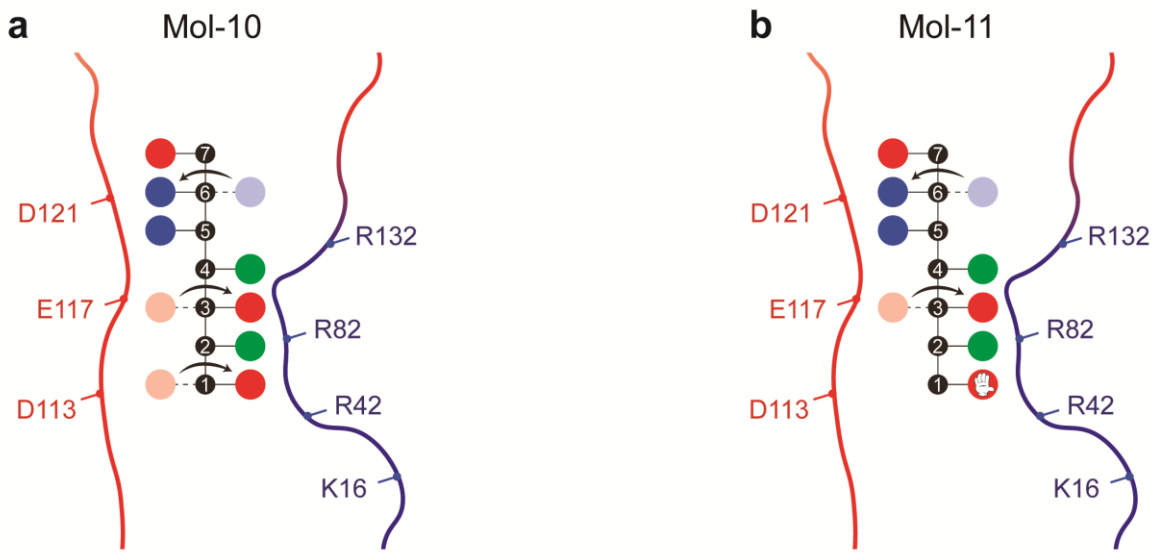
Supplementary Figure S9 Same as in Figure S5 but for peptides Mol-5, 6, 7, 8, and 9.



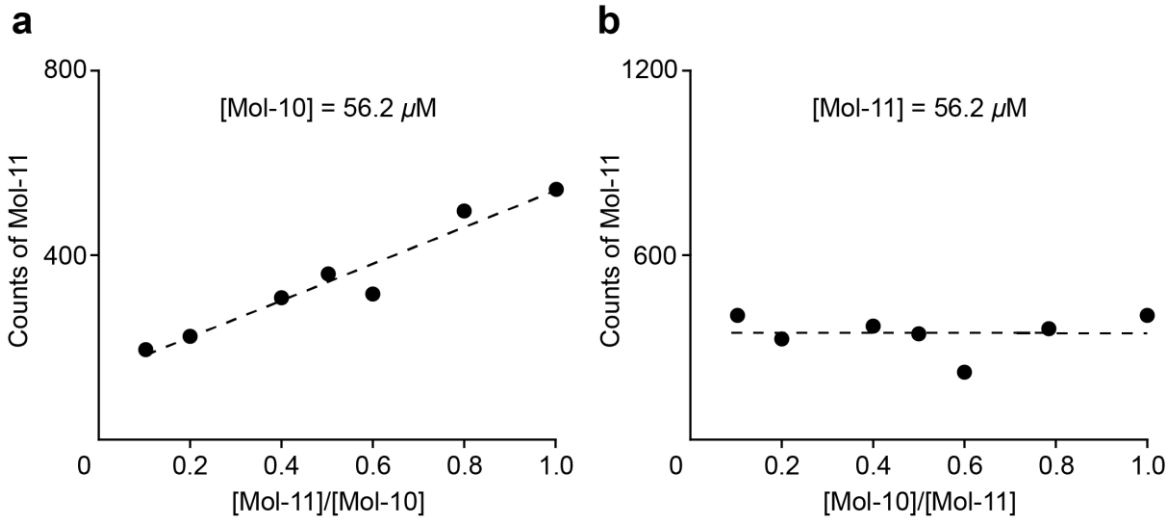
Supplementary Figure S10 Demonstration of OmpF triple mutant (R42E, R82E, R132E) examining chiral amino acid contained peptide. (a) Scheme showing the charged residue distribution in WT OmpF and OmpF triple mutant EEE. (b) upper: Current traces of Mol-1 and Mol-9 interaction with OmpF EEE. Zoomed-in color panels highlighted single molecule event with red star in left panel. Lower: Normalized peaks of $\triangle I/I_0$ and log(t) for both analytes interaction with OmpF EEE. The elimination of asymmetric charge distribution gives the similar distribution of $\triangle I/I_0$ and log(t) for Mol-1 and Mol-9. Therefore, OmpF EEE could not differentiate the chiral amino acid contained peptides, which is distinct from WT OmpF as shown in Figure 2c. A final concentration of 0.4 μM has been reached for Mol-1 and Mol-9. Measure condition: 0.5 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0. At least 10000 single-molecule events have been analyzed.



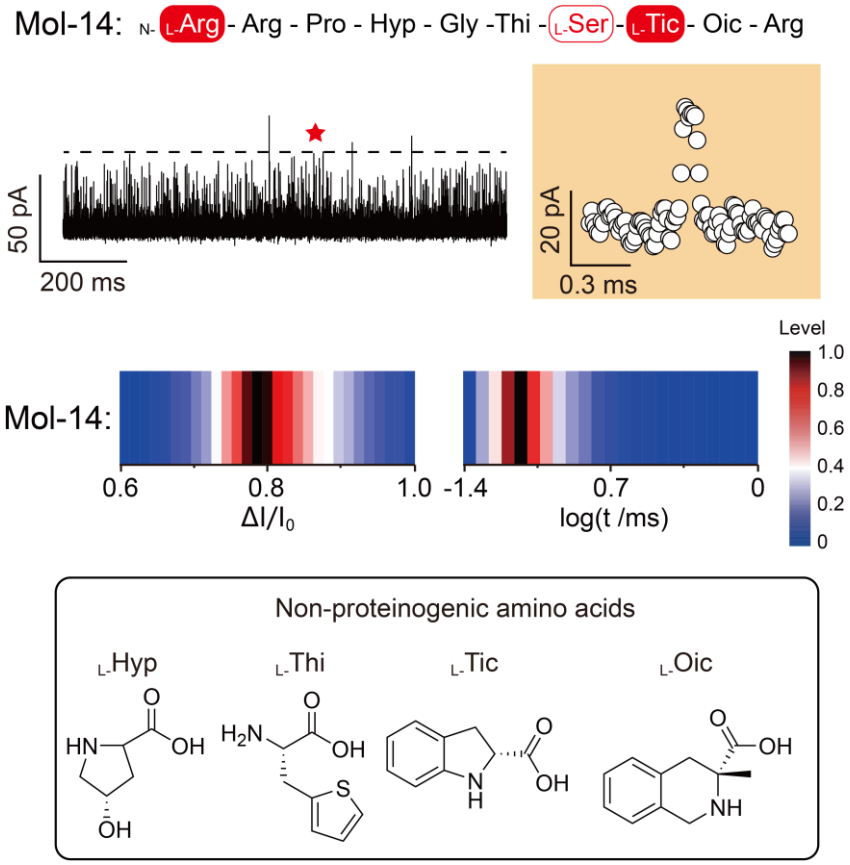
Supplementary Figure S11 Results for using wild type aerolysin to detect Mol-1 and Mol-9. (a) Scheme showing symmetric structure of aerolysin. (b) upper: current trace of Mol-1 and Mol-9 interaction with aerolysin. Zoomed-in color panels give star highlighted single molecule event. Lower: Histogram analysis of the level of blockade $\triangle I/I_0$ and logarithm of the residence time log(t). The normalized main peaks of both analytes are overlapped suggesting insufficient identifying capability and the frequency of interacting events is quite low. The WT aerolysin is not able to identify these two molecules. A final concentration of 227.7 μM for Mol-1 and Mol-9 has been used during the experiments. Measure condition: 1.0 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0. At least 5000 single molecule events have been analyzed.



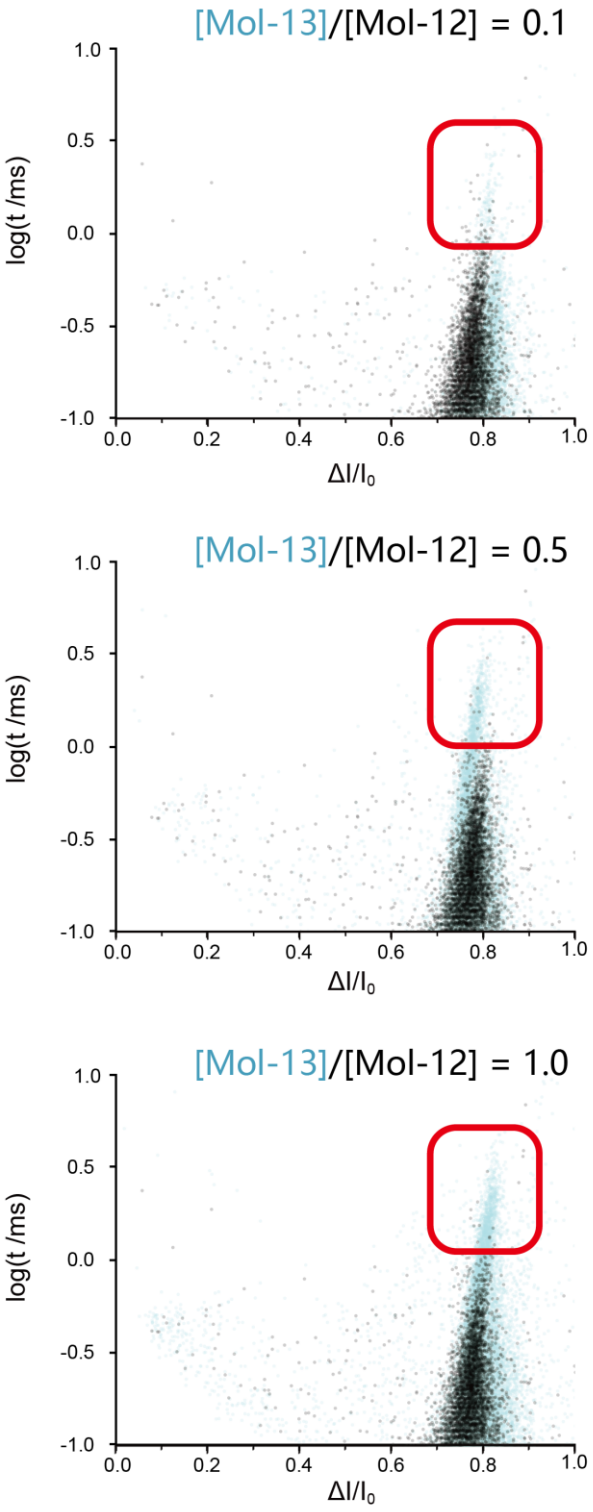
Supplementary Figure S12 Schematic representation of the sidechain orientation of (a) the *β*-amyloid peptide key segment (Mol-10) and (b) its $_{D}$-Asp^{1} analogue (Mol-11) within OmpF (orientations not to scale). The light colors represent the zig-zag alignment of each amino acid side chain of the peptide in bulk solution while the arrows indicate the proposed reorientations when entering OmpF during translocation (dark colors).



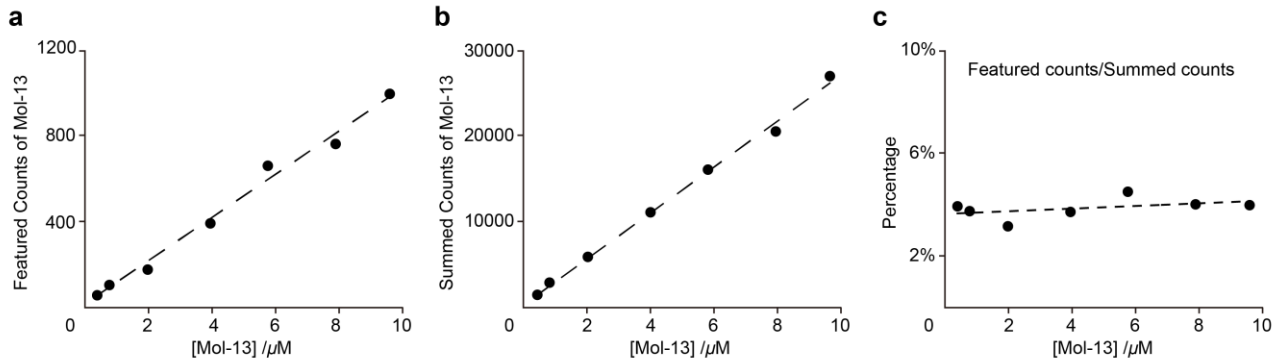
Supplementary Figure S13 (a) Detection of Mol-11 from mixture of Mol-10 and Mol-11 with fixed concentration of Mol-10 at 56.2 μM. (b) The relationship between the counts of Mol-11 and the concentration ratio of Mol-10 and Mol-11 at a fixed concentration of Mol-11 (56.2 μM). The counts of Mol-11 based from a 10-min ionic current trace. Measure condition: 0.5 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0.



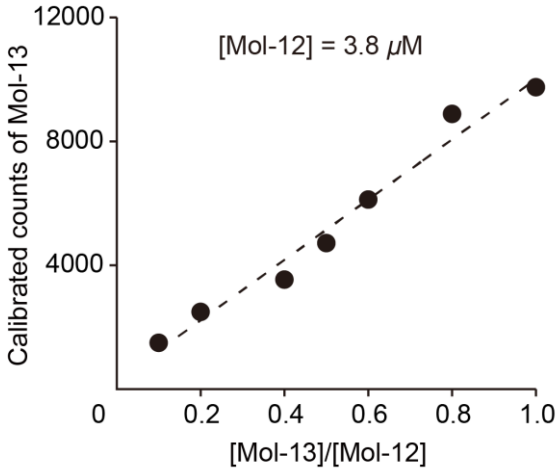
Supplementary Figure S14 (a) Current trace of the all L-type amino acid HOE140 (Mol-14) measured using a single OmpF protein at a bias voltage of -50 mV. Red star points denote enlarged events (orange panel) in the original trace. Bottom section: Electrochemical characterization of Mol-14 using the OmpF nanopore assay. A final concentration of 19.2 μM for Mol-14 has been used during the experiments. Measure condition: 0.5 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0. At least 10000 single-molecule events have been analyzed. (b) Chemical structures of non-proteomic amino acid synthesized in the HOE140 peptide drug.



Supplementary Figure S15 Scatter plots of single-molecule events from Mol-12 and sequentially addition of Mol-13 at molar ratios of 0.1, 0.5 and 1.0. The data has been processed as described in Supplementary Note 3. With an increasing ratio of Mol-13, the density of the impurities detected becomes more obvious as indicated in the signature region (t > 1 ms). Measure condition: 0.5 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0.



Supplementary Figure S16 Characterizing the Mol-13 featured events from signature region (t > 1 ms) and summed events. (a) Featured counts of Mol-13 against the concentration. (b) Similar plot to Figure S16a but summed counts of Mol-13. (c) Percentage of featured events over summed events against Mol-13 concentration. The counts of Mol-13 based from a 10-min ionic current trace. Measure condition: 0.5 M KCl, 10 mM Tris, and 1 mM EDTA at pH 8.0.



Supplementary Figure S17 Characterization of Mol-13 events from mixture of Mol-12 and Mol-13. Similar figure as Figure 3d but decrease the concentration of Mol-12 to 3.8 μM. The counts of Mol-13 based from a 10-min ionic current trace. Measure condition: 0.5 M KCl, 10 mM Tris, and 1 mM EDTA.

Supplementary Note 1: All-atom molecular dynamics simulations

The three-dimensional structures of the peptides were built using the ProBuilder webserver. Moreover, the pre-equilibrated OmpF setup used in a recent study[34] has been adapted. The system consisted of the OmpF trimeric protein inserted in the 1,2-Di-phytanoyl-sn-glycero-3-phosphocholine lipid bilayer, TIP3P water molecules, K^+ ions for neutralization, and peptide molecules placed towards the EC side of each monomer. Placing three copies of the respective peptide (Mol-1 to Mol-9) in front of the trimeric OmpF and adding 1 M KCl, results in a system size of roughly 131,000 atoms for all nine model peptides. The hydrogen atoms of the channel, lipids, and peptides were treated as virtual sites to achieve an integration time step of 5 fs. Subsequently, all systems were equilibrated by maintaining the temperature at 300 K and the pressure at 1 bar for a timescale of approximately 15 ns in several steps using the protocol from a previous study.[35] For the system containing the Mol-1 peptide, we have carried out a 6.8 μs long applied-field simulation with an external voltage of +1 V defined as the potential difference between the EC and the PP sides of the pore. Steered molecular dynamics (SMD) simulations for all systems were performed by pulling the N-terminus of the peptides towards the PP side of the channel.[35] Note that all three peptides were parallelly pulled through three monomers, using the reaction coordinate z in each case which was defined as the center of the mass difference between the N-terminus residue of the peptide and the Cα atoms of the respective barrel. These pulling simulations were carried out at a constant velocity of 1 $Å \cdot ns^{-1}$ using a spring constant of 100 $kJ \cdot mol^{-1} \cdot nm^{-2}$. During the simulations, additional wall constraints were applied on the peptides to prevent their entrance into the neighboring monomers using the same parameters as the previous study.[32] For all systems, the simulations were repeated five-time for a duration of roughly 100 ns, resulting in 15 pulling trajectories because of the parallel runs in each monomer. All simulations were performed using the CHARMM36 force field with the GROMACS 2019.4 package. The other simulation parameters were the same as in previous studies.[34-36]

Firstly, steered MD simulations were performed for Mol-1 to Mol-9. Moreover, the salt-bridge networks (Figure S5) and the sidechain orientations of the individual amino acids have been calculated for the translocation throughout OmpF (Figure S6). Both properties suggest that the interactions between basic peptide residues and the acidic pore residues on L3 at the extracellular side mainly change due to variations in the peptides. The basic ladder (K^{16}, R^{42}, R^{82}, R^{132}) inside the CZ plays an indirect role, which ensures that the acidic and basic residues of the peptides face the opposite side. To this end, we also analyzed how the sidechains of each peptide residue reorient themselves in response to the heterogeneous pore environment during translocation (Figure S6). Note that we also estimated the percentage of pore blockages by measuring the water flux from these simulations, which confirmed that the pore is blocked by roughly 91% (Figure S2d).

Next, we discuss the peptides containing chiral amino acids. Regarding Mol-5 and Mol-6, the salt-bridge interactions in both cases were seen to get stronger, suggesting a rise in affinity (Figure S9). As an underlying reason, we find that the sidechains of the arginine in either $_{D\text{-}}Arg^1Arg^2$ or $Arg^1{}_{D\text{-}}Arg^2$ orient towards the acidic residues much better compared to Mol-1, which favors an energetically more stable conformation inside the CZ (Figure S8). As a result, both Mol-5 and Mol-6 permeate slower and spend more time inside the channel. The sidechain reorientation profiles revealed that Arg^2 fails to reorient toward the acidic residues. We believe that this is the primary reason why the molecules lose affinity and reside shorter inside the pore in experiments.

Supplementary Note 2: Nanopore experiments and apparatus

The employed Ag/AgCl electrodes were homemade. An Axon 200B device has been used to acquire the nanopore signals which were digitized by Digidata 1440B at a 10 kHz cut-off frequency and a sampling frequency under 250 kHz. In a typical experiment, a single trimeric OmpF is reconstituted into an artificial lipid bilayer that separates the measuring chamber into cis and trans sides. Specifically, two Teflon home-made chambers have been clamped with a Teflon sheet (FP301200, Goodfellow Cambridge Ltd., UK) in middle. Each Teflon sheet contains a hole (50 to 100 μm in diameter) fabricated using an electric spark generator in a homemade setup and visually checked under a microscope. 1,2-Di-phytanoyl-sn-glycero-3-phosphocholine (Avanti Polar Lipids, USA) was dissolved in a Decane solution to a final concentration of 2.0 mg/70 μL and then added to both cis and trans chamber to form a lipid bilayer membrane with an appropriate thickness.[37] The WT OmpF and the triple mutant OmpF EEE protein has been purified as detailed in a previous study.[38,39] To perform single protein reconstitution, the stock solution (1 mg/mL) has been diluted in 1% Genapol X-080 (Sigma Aldrich, USA). Then the protein solution has been added directly into the cis chamber to reach a final concentration of 8-10 mg/mL for single protein insertion. Note that the extracellular (EC) side of the OmpF after reconstitution is commonly identical to the addition side.[40] OmpF protein is added on the *cis* side and an external bias voltage, *e.g.*, -50 mV, is applied from the *trans* side which connects to the headstage of the current acquisition equipment. A steady open trimeric pore current ($3 \times I_0$) is observed after a single OmpF reconstitution. Additional target peptides on the *cis* side induce transient current blockades (Figure 1d), indicating one molecule crosses one of the monomers at a time. In all reported experiments, the temperature was 24 ± 2 °C. The analytes were all synthesized by GL Biochem (Shanghai) Ltd. (CHN). All peptides were 98% purified using HPLC. Subsequently, the samples were freshly dissolved in the measuring buffer and then added into the measurement cuvette at the cis side until reaching the final concentration.

Supplementary Note 3: Data analysis and histogram color mapping

A nanopore-based electrochemical analysis has been performed. The MOSAIC[32] approach has been applied to recover the single-molecule events that were beyond the acquisition rate[31] along with a supporting vector machine to eliminate inadequate points. Subsequently, a Gaussian fit was employed to extract the level of blockade ($\triangle I/I_0$) and the residence time ($\log(t)$). The ratio of blockade current $\triangle I$ and the open pore current I_0 results in the (relative) current drop $\triangle I/I_0$. The residence time t is determined using a Gaussian fit and indicates how long the peptide is residing in OmpF. The MOSAIC approach[32] has been used to analyze the single-molecule events. To functionalize it, a current drop threshold has been set for the single-channel search of OmpF interactions with target peptides. The slightly gating events from protein channels are cumulated at the left-up corner of the scatter plots with long durations and low current amplitude. To excluded these interference events, we used support vector machine to classify and extract the peptide events. Once single-molecule peptide events have been counted, a scatter plot of current drop versus residence time was plotted. Subsequently, the histograms of both current drop and residence time for each target peptide were plotted assuming a certain bin width. To differentiate the amperometric fluctuation, a logarithmic scale has been used to the residence time. Making use of OriginPro (OriginLab Corporation, USA) to fit the histograms using Gaussian distributions, the peak values of the distributions have been obtained directly from the fitting results. All histograms have been normalized to one concerning the respective peak value. To make the stochastical analysis more comparable, the histogram plots have then been converted into 1D color maps as shown in Figure 1, Figure 2, and Figure 3. Concerning the color code, the corresponding density increases from blue and white to red and finally, the maximum is shown in black. In the text, we mainly compare and discuss the maxima shown in black.

Supplementary Note 4: Mass spectrometry analysis

MS determination of Mol-10, 11, 12, 13 has been performed. Mol-10 and Mol-11 were dissolved in ultrapure water to 6.25 μg/mL, Mol-12 and Mol-13 were dissolved in ultrapure water containing 0.1% formic acid (FA) to 100 μg/mL. Samples were identified then quantified using an Orbitrap Q Exactive high resolution mass spectrometer (Thermo Scientific, MA, USA) equipped with a heated electrospray ionization (HESD) source. Following generic MS tune parameters: spray voltage = 3.5 kV (positive mode); capillary temperature = 320ºC, sheath gas (nitrogen) flow = 35 arbitrary units; auxiliary gas (nitrogen) flow = 10 arbitrary units; probe heater temperature = 400ºC. Full-MS scan mode was performed for each run in the range of m/z 200-3000 with a resolving power of 140,000 FWHM. The automatic gain control (AGC) target value and maximum injection time used for the full-MS scan were 3e6 and 50 ms respectively. All the data in raw format obtained from ESI-Quadrupole-Orbitrap MS were processed using Thermo Xcalibur software version 4.1.

Supplementary Note 5: Qualitatively and quantitatively determination of $_{D}$-Ser^{7} HOE140 (Mol-13) from HOE140 (Mol-12)

As shown in Figure S15, the featured events of $_{D}$-Ser^{7} HOE140 (Mol-13) in the signature region (t > 1 ms) are distinct from Mol-12. This region could qualitatively report the Mol-13 from the Mol-12. In order to demonstrate the quantitative ability of OmpF discrimination, we performed sensing of Mol-13 at varied concentration, and compared the percentage of featured events (t > 1 ms) in summed events (Figure S16c). Here, the events counts were based on 10-min ionic current traces. The calculated results showed that the featured events take 3.8 ± 0.4% of the summed events of Mol-13, which is independent on the concentration of Mol-13. Therefore, the number of featured events within 10 min is proportional to the concentration of Mol-13, which could be used for the quantitation. In the following data processing, we take 3.8 % as a coefficient to calibrate the summed events of Mol-13 from featured events in the mixture.

For the mixture assay, the concentration of Mol-12 was fixed to either 9.6 μM or 3.8 μM, and concentration Mol-13 was gradually increased in the mixture until the same molar ratio. Then, the featured events within 5-min recording in the signature region (t > 1 ms) were counted. To reduce the experimental deviation, the linear relationship between the concentration and featured event counts (t > 1 ms) of Mol-13 (Figure S16) were used as a calibration curve. To estimate the calibrated summed events number of Mol-13, the calibrated featured event counts were divided by the coefficient of 3.8 %. The resulted linear relationship between calibrated counts of Mol-13 and increased molar ratio of Mol-13 in Mol-12 are shown in Figure 3d and Figure S16, which demonstrates that quantitatively determine the concentration of Mol-13 from Mol-12 is feasible.

References

(32) J. H. Forstater et al., MOSAIC: a modular single-molecule analysis interface for decoding multistate nanopore data. *Anal. Chem.* **2016**, *88* (23), 11900-11907.

(33) N. Meyer et al., Machine Learning to Improve the Sensing of Biomolecules by Conical Track-Etched Nanopore. *Biosensors* **2020**, *10* (10), 140.

(34) A. Acharya; J. D. Prajapati; U. Kleinekathöfer, Improved Sampling and Free Energy Estimates for Antibiotic Permeation through Bacterial Porins. *J. Chem. Theory Comput.* **2021**, *17* (7), 4564-4577.

(35) S. Pangeni et al., Large-Peptide Permeation Through a Membrane Channel: Understanding Protamine Translocation Through CymA from Klebsiella Oxytoca. *Angew. Chem. Int. Ed.* **2021**, *60* (15), 8089-8094.

(36) J. D. Prajapati; C. J. F. Solano; M. Winterhalter; U. Kleinekathöfer, Enrofloxacin Permeation Pathways across the Porin OmpC. *J. Phys. Chem. B* **2018**, *122* (4), 1417-1426.

(37) P. Mueller; D. O. Rudin; H. Ti Tien; W. C. Wescott, Reconstitution of cell membrane structure in vitro and its transformation into an excitable system. *Nature* **1962**, *194*, 979-980.

(38) J. P. Rosenbusch, Characterization of the major envelope protein from Escherichia coli regular arrangement on the peptidoglycan and unusual dodecyl sulfate binding. *J. Biol. Chem.* **1974**, *249* (24), 8019-8029.

(39) V.K. Golla; E. Sans-Serramitjana; K.R. Pothula; L. Benier; J.A. Bafna; M. Winterhalter; U. Kleinekathöfer., Fosfomycin permeation through the outer membrane porin OmpF. *Biophys. J.* **2019**, *116* (2), 258-269.

(40) S. A. Ionescu et al., Orientation of the OmpF Porin in Planar Lipid Bilayers. *ChemBioChem* **2017**, *18* (6), 554-562.