



HAL
open science

Identification and Detection of a Peptide Biomarker and Its Enantiomer by Nanopore

Laura Ratinho, Laurent Bacri, Bénédicte Thiebot, Benjamin Cressiot, Juan Pelta

► **To cite this version:**

Laura Ratinho, Laurent Bacri, Bénédicte Thiebot, Benjamin Cressiot, Juan Pelta. Identification and Detection of a Peptide Biomarker and Its Enantiomer by Nanopore. ACS Central Science, In press, <10.1021/acscentsci.4c00020>. <hal-04575965>

HAL Id: hal-04575965

<https://hal.science/hal-04575965v1>

Submitted on 15 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Identification and Detection of a Peptide Biomarker and Its Enantiomer by Nanopore

Laura Ratinho, Laurent Bacri, Bénédicte Thiebot,* Benjamin Cressiot,* and Juan Pelta*



Cite This: <https://doi.org/10.1021/acscentsci.4c00020>



Read Online

ACCESS |



Metrics & More

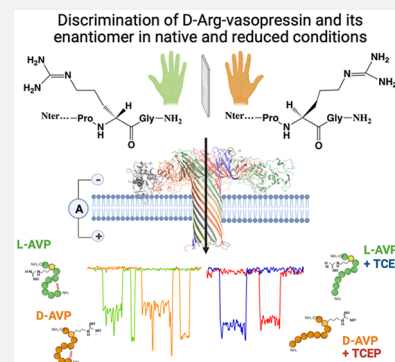


Article Recommendations



Supporting Information

ABSTRACT: Until now, no fast, low-cost, and direct technique exists to identify and detect protein/peptide enantiomers, because their mass and charge are identical. They are essential since L- and D-protein enantiomers have different biological activities due to their unique conformations. Enantiomers have potential for diagnostic purposes for several diseases or normal bodily functions but have yet to be utilized. This work uses an aerolysin nanopore and electrical detection to identify vasopressin enantiomers, L-AVP and D-AVP, associated with different biological processes and pathologies. We show their identification according to their conformations, in either native or reducing conditions, using their specific electrical signature. To improve their identification, we used a principal component analysis approach to define the most relevant electrical parameters for their identification. Finally, we used the Monte Carlo prediction to assign each event type to a specific L- or D-AVP enantiomer.



INTRODUCTION

L-Amino acid enantiomers are the predominant forms found in animal tissues. However, although D-amino acids are present in small quantities, they can be detected in either their free forms or polypeptide chains.¹ In humans, the free D-amino acids, such as D-Asp, D-Ser, and D-Ala, play significant roles in various physiological functions, such as the development and functioning of the central nervous system, reproduction, and hormone release.^{2–4} Altered regulation production of these isomers can have profound effects on organisms. For example, during the natural aging process, there is a decline in the level of D-serine.⁵ Conversely, in pathological aging conditions, such as Alzheimer's disease, there is excessive activation of serine racemase, resulting in hyperstimulation of the NDMA receptor due to an excess of D-Ser.⁵ Notably, levels of D-Arg levels are decreased.⁶

Additionally, several D-amino acids containing peptides (DAACPs) have been identified in various pathological conditions associated with diseases or natural aging processes, as evidenced by the presence of D-β-Asp-containing peptides in elastic fibers of sun-damaged skin,⁷ D-Asp in αA-crystallin from the lenses of individuals with cataracts,⁸ and the β-amyloid peptide of Alzheimer's patients.⁹ These DAACPs are mainly formed following the post-translational modification of L-amino acids, either by enzymatic racemization of L-amino acids into their D-form or spontaneously, despite a very slow process, within long-lived proteins and tissues.¹ It has been established that oxidative stress and free radicals can induce such modifications in proteins.^{10,11} Notably, the alteration of a specific residue from L- to D-configuration often leads to a modification of the biological activity of the peptide and could

be developed as therapeutic molecules.¹² Modifications of peptide functional properties are likely due to a change in the higher-order structure of a protein.^{13,14} The significant increase of DAACPs discovered in diseases has encouraged researchers to consider them as potential biomarkers,¹ and any new technique capable of detecting/quantifying or finding new DAACPs, even at low concentrations, would represent a major advancement in early diagnosis.

As mentioned above, there is a clear need to develop fast, low-cost, and direct techniques to identify chiral amino acids in a polypeptide chain without prechemical/enzymatic treatment. Due to the difference between L- and D-amino acids, we expect the polypeptide chains to adopt different conformations. Therefore, developing a technique that can sense the protein conformational landscape is necessary, while just a few studies have already been published using nanopores.^{15–17} To the best of our knowledge, the classical methods used, such as mass spectrometry, liquid chromatography, or enzymatic assays, for the detection of L- and D-amino acids do not allow the simultaneous detection of proteins, their conformation, and their chiral amino-acid composition.¹⁴

The nanopore electrical detection technique offers a unique multiscale analytical tool for peptide and protein enantiomer biomarker detection. Up to now, nanopores have been used to

Received: January 4, 2024

Revised: April 24, 2024

Accepted: April 25, 2024

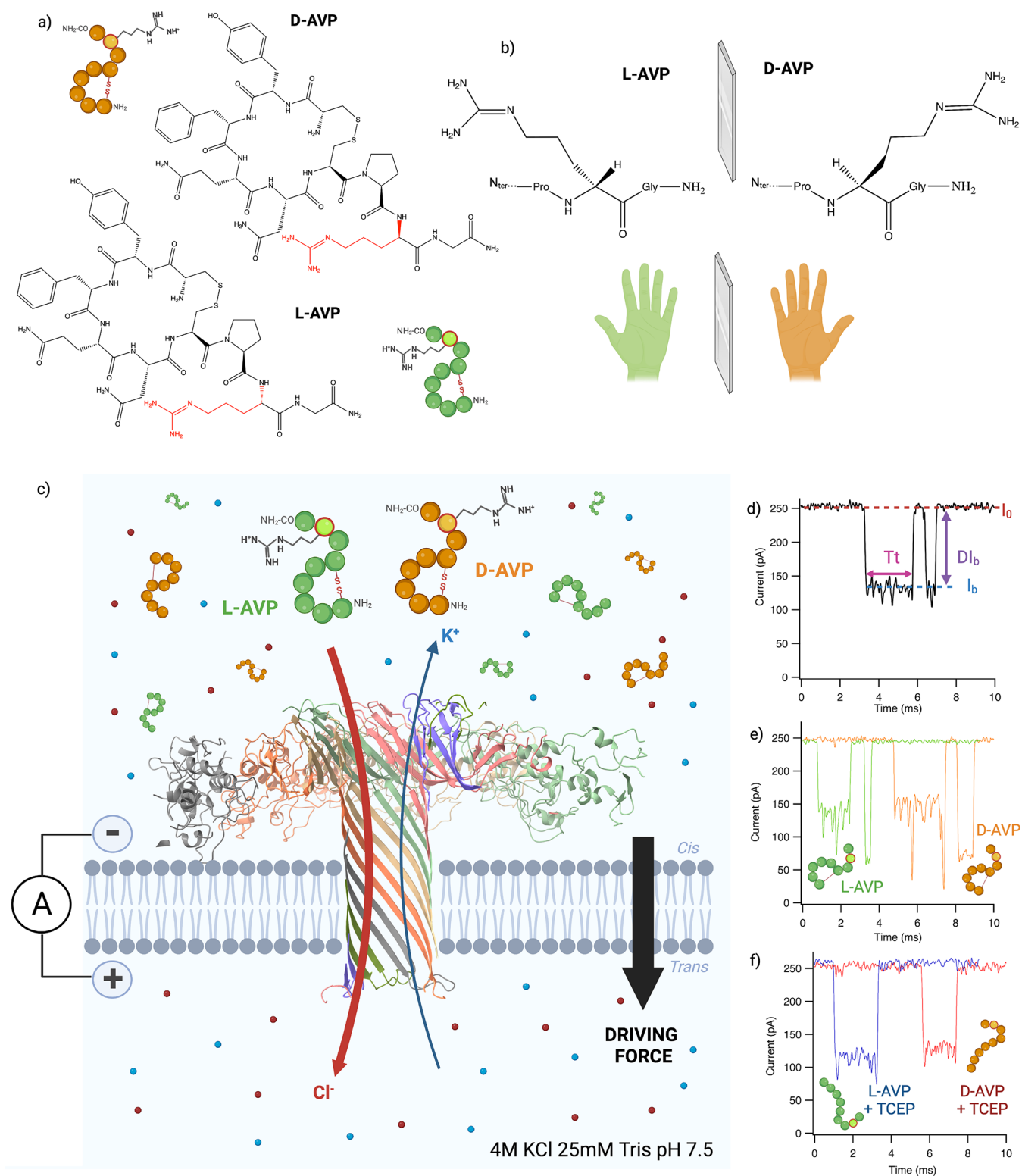


Figure 1. (a) Isomeric representation of L- and D-AVP (chemical structures created with ChemDraw). (b) Chemical representation of the chirality of vasopressin Arg8. (c) Representation of experimental conditions for the analysis of L-AVP (green) and D-AVP (orange) using a wild-type aerolysin nanopore (ribbon representation of PDB 5JZT using ChimeraX), inserted in a lipid bilayer in 4 M KCl and 25 mM Tris, pH 7.5. With an applied voltage, Cl^- and K^+ ions go through the pore toward the oppositely charged electrodes, creating an ionic current (I_0). (Created using BioRender.com.) (d) Example of an event showing the open pore current (I_0) and blockade current (I_b) resulting in a blockade level (DI_b) over a dwell time (T_t), characteristic of a peptide at 110 mV. (e) Example of the most representative types of events of L-AVP (green) and D-AVP (orange) at 110 mV. (f) Example of the most representative type of event of L-AVP (blue) and D-AVP (red) in the presence of 5 mM TCEP at 110 mV. (Image generated using Biorender.com.)

discriminate between chiral forms of free amino acids,¹⁸ as well as detect and identify protein/peptide biomarkers.^{19–29} More recently, research has focused on detecting post-translational modifications (PTMs)^{21,30–37} in peptide biomarkers that could help understand their implications in pathologies.^{38–40} Machine learning is now used to validate the data from the sequencing of individual amino acids³⁷ to the identification of biomolecules^{41,42} and biomarkers.⁴³

While the first study on detecting individual chiral amino acids (Trp, Phe, Tyr, Cys, and Asp) using an engineered Cu²⁺ phenanthroline alpha-hemolysin channel was published in 2012 by Bayley's group,¹⁸ only a few studies have been published since. Two other groups showed the ability to discriminate individual L- and D-amino acids using cyclodextrin adapters⁴⁴ inside a covalent organic framework-(His) or an alpha-hemolysin mutant⁴⁵ (Phe, Trp, Tyr). Mubarak et al. developed a nanopipette system to detect single amino-acid enantiomers (Tyr, Trp, and Phe) using a polymeric conical nanopore functionalized with BSA.⁴⁶ They used current rectification to detect each kind of enantiomer. The first example of detecting chiral amino acids within a polypeptide biomarker was described by Luchian's group, where they identified, using Cu²⁺ chiral recognition, the presence of histidine enantiomers.⁴⁷ A recent study used mutant OmpF to control the side chain orientation of a β -amyloid peptide inside the nanopore due to its lateral electric field.⁴⁸ They were able to discriminate between D-Ser and D-Asp isoforms and mutants. Finally, Maglia's group studied the detection of D-Ala and D-Leu in enkephalin peptides using FraC and CytK mutants.⁴⁹

Until now, no studies have been conducted to detect enantiomers within a biologically relevant peptide that has secondary and tertiary structures such as disulfide bonds. Furthermore, most previous studies needed an adapter or several pore mutations to detect chiral amino acids in the polypeptide sequence. To prove the ability to sense enantiomers containing elements in secondary and tertiary structures under native conditions, we used vasopressin as a model peptide. Vasopressin (L-Arg-AVP, later named L-AVP), a hormone produced in mammals, is involved in the regulation of water balance and blood pressure, as well as having an influence on social behavior, memory, and the cardiovascular system through its interaction with V1a, V1b, and V2 receptors.⁵⁰ Vasopressin quantification was performed to assess hyper- or hyposecretion pathologies. In the case of diabetes insipidus,⁵¹ quantifying vasopressin levels allows to determine the cause of the disease, which can be AVP receptor insensitivity or a decrease in hormone production. Notably, the preferred technique for detecting L-AVP in biomedical analyses has shifted to HPLC-MS/MS, sidelining traditional immunological techniques, partly due to its low molecular weight. This choice, however, requires essential preliminary steps for extracting and purifying the peptide of interest. The vasopressin nonapeptide possesses two cysteine residues at positions 1 and 6, forming a constrained peptide loop structure via a disulfide bridge. NMR analysis has revealed saddle and open conformations within the cyclic part of the molecule, while the noncyclic part demonstrates greater flexibility.⁵⁰ The D-Arg-AVP (later named D-AVP) enantiomer represents a synthetic derivative of L-AVP, primarily employed for research purposes.⁵²

Conversely, the C-terminally deaminated form of D-AVP, known as desmopressin, is used for therapeutic applications.

Indeed, desmopressin selectively binds to a single receptor, V2, unlike the natural hormone L-AVP, limiting its physiological effects on water retention. Including D-Arg in the peptide chain makes it more resistant to proteolysis and enhances the molecule's lifespan in the bloodstream.⁵³

This work proves that a wild-type (WT) aerolysin nanopore can discriminate, at the single-molecule level, enantiomeric peptides, L-Arg8 or D-Arg8, contained in the native peptide chain. The L- and D-AVP can also be identified in a mixture. Interestingly, we can detect multiple AVP conformations, such as open and saddle states already observed by NMR. Furthermore, we show that after using a reducing agent several conformations observed in native conditions disappear, and the nanopore is still sensitive enough to identify both L- and D-forms in single and mixed experiments. By changing the relative ratio of each component in the mixture, we demonstrate the identification of each enantiomeric state according to its blockage level or volume. To identify each kind of event population and attribute them to different conformations, we used a PCA approach to first determine the best electrical parameters for their discrimination. We also developed a Monte Carlo prediction approach to identify each enantiomer conformation.

RESULTS AND DISCUSSION

Electrical Enantiomer Characterization in Native Conditions. L- and D-AVP are two analogous peptides of 9 amino acids differing from a single chiral amino acid Arg8 (Figure 1a,b). Chirality can be represented as the non-superposable mirror image of an object. Therefore, the main difference between L- and D-AVP is the orientation of the lateral chain of Arg8 (Figure 1b). The two peptides form a disulfide bond between Cys1 and Cys6, creating a looped peptide structure (Figure 1a). As well as a disulfide bond, it has been found that L-AVP can possess one or two beta-turns stabilized by hydrogen bonds:⁵⁰ the first one within the loop created by the disulfide bond creates the saddle conformation, and the other is formed across Pro7. L- and D-AVP structures are similar,⁵⁴ so we can hypothesize that D-AVP can also form these two beta-turns. Discriminating L- and D-AVP with a WT aerolysin nanopore would enable us to show that we can detect a slight difference in conformation with a single chiral amino acid in structurally complex peptides. By reducing the disulfide bonds, we want to show if we can detect a conformational difference between the native and reduced states and determine the best conditions to discriminate between L- and D-AVP.

Using L- and D-AVP as model peptides, we want to show if with a wild-type aerolysin nanopore we can discriminate and identify a single amino acid enantiomer in both native and reducing conditions (Figure 1c). We performed nanopore experiments with L- and D-AVP, independently or in a mix, in the presence or absence of tris(2-carboxyethyl)phosphine (TCEP), a reducing agent. A WT aerolysin nanopore is inserted into a lipid bilayer separating two compartments (*cis* and *trans*) filled with an electrolyte solution (4 M KCl, 25 mM Tris, pH 7.5) (Figure 1c). A potential difference is applied by using two electrodes in each compartment. Ions in the electrolyte solutions flow through the pore, and we can measure an ionic current of 250 pA at 110 mV in our experimental conditions (I_0). In the presence of an analyte at the pore entry, we observe current blockades (I_b), characterized by a blockade level ($DI_b = I_0 - I_b$) and by a dwell time

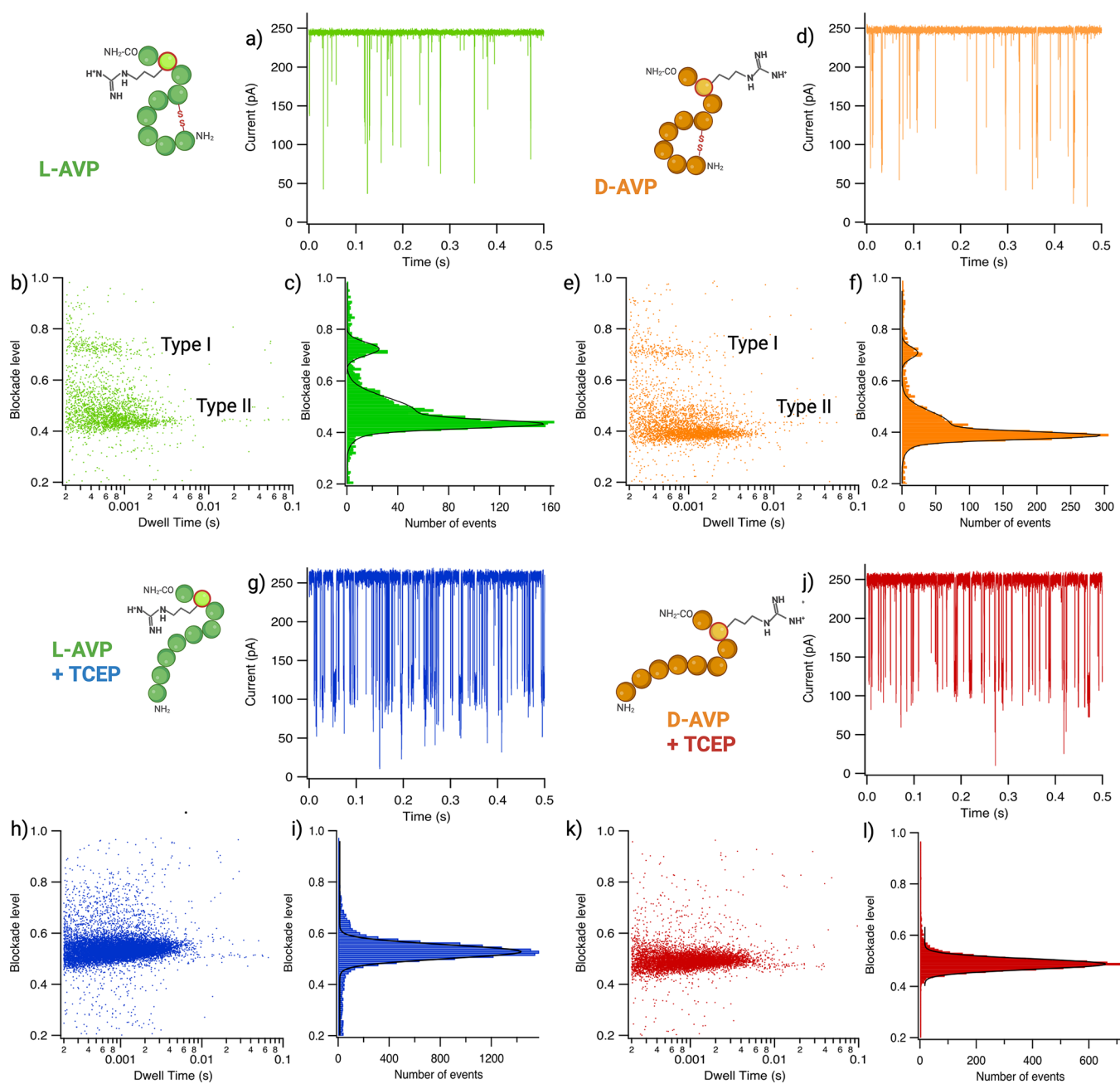


Figure 2. Experimental result of the independent analysis of 10 μM L- and D-AVP with or without 5 mM TCEP using an aerolysin nanopore in 4 M KCl 25 and mM Tris pH 7.5. (a, d, g, j) Example of current traces (filtered at 5 kHz) for each independent experimental condition at 110 mV: L-AVP (green), D-AVP (orange), L-AVP+TCEP (blue), and D-AVP+TCEP (red). Open pore currents, $I_0 = 243.82 \pm 1.77$ pA for L-AVP (a), $I_0 = 246.78 \pm 2.04$ pA for D-AVP (d), $I_0 = 259.22 \pm 3.33$ pA for L-AVP+TCEP (g), and $I_0 = 250.16 \pm 3.46$ pA for D-AVP+TCEP (j). (b, e, h, k) Scatter plot representing the normalized average blockade level (DI_b) against the dwell time of each event longer than 200 μs between 0.2 and 1.0 in the blockade level representing the interacting population of each peptide in the absence and presence of TCEP. (Raw scatter plot can be found in Supporting Information S3 and S4.) (b, e) Scatter plot showing populations of events for L- and D-AVP in the absence of TCEP: Type I between 0.6 and 0.8 blockade level and Type II between 0.38 and 0.6 blockade level. (c, f, i, l) Histograms of normalized blockade levels as a function of the number of events fitted with a Gaussian function (black lines) to determine the most probable blockade level for Type I and a bi-Gaussian for Type II: L-AVP: Type I, 0.72 ± 0.01 ; Type IIa, 0.43 ± 0.01 for the population with the most events; Type IIb, 0.48 ± 0.01 ; and D-AVP: Type I, 0.71 ± 0.01 ; Type IIa, 0.39 ± 0.01 ; Type IIb, 0.43 ± 0.01 . With TCEP average blockade levels fitted by a Gaussian for L-AVP + TCEP: 0.53 ± 0.01 and D-AVP + TCEP: 0.49 ± 0.01 . Data shown are from a single recording, with fitted values being mean and standard deviation for three independent fits. $N_{\text{L-AVP}} = 2812$ events; $N_{\text{D-AVP}} = 3961$ events; $N_{\text{L-AVP+TCEP}} = 13383$ events; $N_{\text{D-AVP+TCEP}} = 7644$ events. The number of events was calculated by selecting the population with a dwell time superior to 200 μs , depending on their blockade level. (Image generated using Biorender.com.)

(T_i) (Figure 1d). In native conditions, we observe two types of events for both L- and D-AVP: Type I with short dwell times of ~ 440 and ~ 390 μs , respectively, and a lower blockade current

of ~ 60 and ~ 70 pA; and Type II with longer dwell times: for L-AVP ~ 820 μs and D-AVP ~ 1140 μs and a higher blockade current of ~ 150 pA (Figure 1e) (Supporting Information S1,

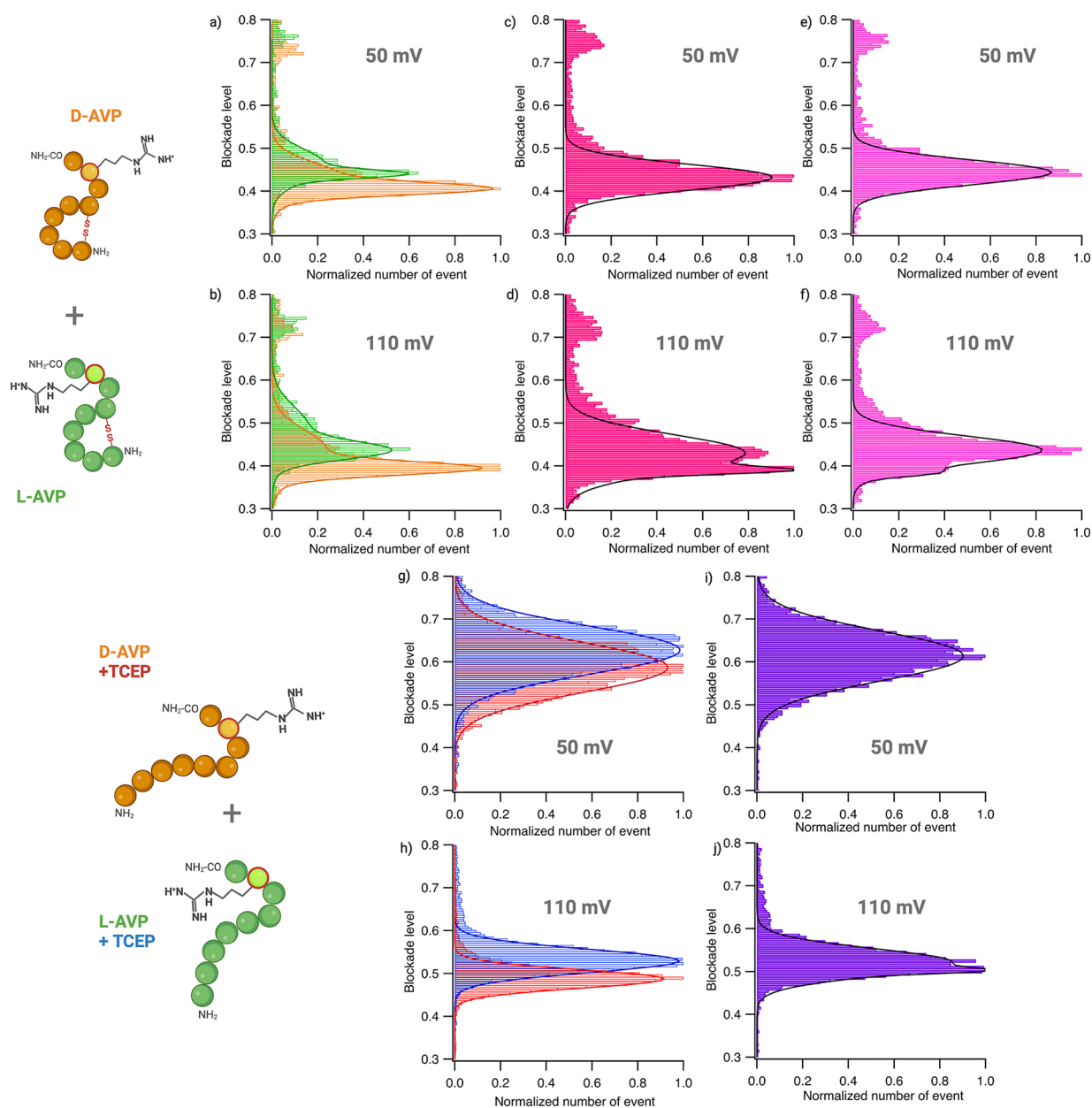


Figure 3. L- and D-AVP discrimination in a mix with and without TCEP using a wild-type aerolysin nanopore at 50 and 110 mV. Histograms of blockade levels from 0.3 to 0.8 as a function of the normalized total number of events fitted with a Gaussian or bi-Gaussian function (lines) to determine each population's most probable blockade level. (Raw scatter plot can be found in [Supporting Information S3 and S4](#).) (a, b) Superposition of L-AVP in green and D-AVP in orange (independent experiments at 10 μM). (c, d) 10 μM equimolar mix of L- and D-AVP (dark pink). (e, f) 7.5 μM L-AVP and 2.5 μM D-AVP (light pink). (g, i) Superposition of L-AVP with TCEP in blue and D-AVP with TCEP in red (independent experiments at 10 μM). (h, j) 5 μM equimolar mix of L- and D-AVP with 5 mM TCEP. (a, c, e, g, h) Experiments conducted at 50 mV. (b, d, f, i, j) Experiments conducted at 110 mV. (a) Independent blockade levels of Type IIa L-AVP = 0.43 ± 0.01 and Type IIb = 0.46 ± 0.01 ; Type IIa D-AVP = 0.40 ± 0.01 and Type IIb = 0.43 ± 0.01 at 50 mV. (b) Independent blockade levels of Type IIa L-AVP = 0.43 ± 0.01 and Type IIb = 0.48 ± 0.01 ; Type IIa D-AVP = 0.39 ± 0.01 and Type IIb = 0.43 ± 0.01 at 110 mV. (c) Blockade level measured: 0.43 ± 0.01 in an equimolar mix of L- and D-AVP at 50 mV. (d) Blockade levels measured: 0.39 ± 0.01 and 0.43 ± 0.01 representing Type IIa most probable mean blockade level of L- and D-AVP in an equimolar mix at 110 mV. (e) Blockade level measured: 0.44 ± 0.01 in a 75/25 concentration ratio mix of L- and D-AVP at 50 mV. (f) Blockade levels measured: 0.38 ± 0.01 and 0.44 ± 0.01 representing Type IIa most probable mean L- and D-AVP blockade levels in a 75/25 concentration ratio at 110 mV. (g) Independent blockade levels of L-AVP+TCEP = 0.62 ± 0.01 and D-AVP+TCEP = 0.58 ± 0.01 at 50 mV. (h) Independent blockade levels L-AVP+TCEP = 0.53 ± 0.01 and D-AVP+TCEP = 0.49 ± 0.01 at 110 mV. (i) Blockade level measured: 0.61 ± 0.01 in an equimolar mix of L- and D-AVP+TCEP at 50 mV. (j) Blockade level measured: 0.50 ± 0.01 and 0.52 ± 0.01 in an equimolar mix of L- and D-AVP+TCEP representing their respective blockade levels at 110 mV. Data shown are from a single recording with fitted values being mean and standard deviation for three independent fits. $N_{\text{L-AVP}}_{50 \text{ mV}} = 2099$ events; $N_{\text{D-AVP}}_{50 \text{ mV}} = 1461$ events; $N_{\text{L-AVP+TCEP}}_{50 \text{ mV}} = 7849$ events; $N_{\text{D-AVP+TCEP}}_{50 \text{ mV}} = 5231$ events; $N_{\text{L-AVP}}_{110 \text{ mV}} = 2756$ events; $N_{\text{D-AVP}}_{110 \text{ mV}} = 3927$ events; $N_{\text{L-AVP+TCEP}}_{110 \text{ mV}} = 13227$ events; $N_{\text{D-AVP+TCEP}}_{110 \text{ mV}} = 7523$ events; $N_{\text{L-D-AVP}_{\text{eq}}}_{50 \text{ mV}} = 3832$ events; $N_{\text{L-D-AVP}_{75/25}}_{50 \text{ mV}} = 3226$ events; $N_{\text{L-D-AVP}_{\text{eq}}}_{110 \text{ mV}} = 5460$ events; $N_{\text{L-D-AVP}_{\text{eq}}}_{110 \text{ mV}} = 4630$ events; $N_{\text{L-D-AVP}_{75/25}}_{110 \text{ mV}} = 5907$ events; $N_{\text{L-D-AVP}_{\text{eq}}}_{110 \text{ mV}} = 4460$ events. The number of events was calculated by selecting the population with a dwell time superior to 200 μs , depending on their blockade level. (Image generated using [Biorender.com](#).)

S2). In the presence of TCEP, a reducing agent, we observe only one type of event for each peptide characterized by long dwell times with a medium blockade current of ~ 130 pA (D-AVP) and ~ 120 pA (L-AVP) (Figure 1f). Interestingly, for these individual events, we can already observe a difference in the average blockade current for L- and D-AVP in either native or reducing conditions (Figure 1e and f, respectively).

For experiments in native conditions, the characteristic parameters of dwell time and average blockade level for each event were extracted from the current traces and plotted as a bidimensional cloud (Figure 2a,b,d,e,g,h,j,k). Only the events of longer dwell time (>200 μ s) are shown here, corresponding to the entrance of the peptides in the pore with a well-characterized blockade level as opposed to events with short dwell time (<200 μ s) known as bumping events.^{55–57} We observed the two main types of events characterized by at least two defined current blockades, 0.43 ± 0.01 (Type IIa) and 0.72 ± 0.01 (Type I) for L-AVP, while 0.39 ± 0.01 (Type IIa) and 0.71 ± 0.01 (Type I) for D-AVP (Figure 2c,f) (Supporting Information S1). These results show that we can discriminate and identify L- and D-AVP in native conditions using the Type II events and their average current blockade. Unfortunately, the less frequent Type I events for L- and D-AVP cannot be discriminated. These events could be attributed to the entrance of the cycle formed by the disulfide bound. In fact, due to the volume of this cycle, we can suppose it is entropically unfavorable for entry into the pore. On the other hand, it could explain the highest blockade level observed. Furthermore, we observe a discrete subgroup in the Type II population that blocks the pore slightly more (Figure 2b,c,e,f). For L-AVP, we can measure a blockade level of 0.48 ± 0.01 (Type IIb) for this population and 0.43 ± 0.01 for D-AVP (Type IIb) (Supporting Information S1). Blockade levels depend on the volume of the chain interacting with or passing through the pore.^{36,49,58,59} Several studies showed that L-AVP could adopt multiple conformations. NMR studies^{50,60} showed that L-AVP can be either in a “saddle” or “open” conformation with a ratio of 70:30, respectively (Supporting Information S5). These two populations with similar observed blockade levels (Types IIa and IIb) could be due to this conformation change, with the most probable population observed being the saddle conformation, or Type IIa, and the less probable population being the open conformation, or Type IIb. These two conformations have different values for molecule volume.⁵⁰ To better discriminate and identify each population, we will discuss this later in the paper by performing semisupervised classification to assign types of events to specific conformations.

Electrical Enantiomer Characterization in Reducing Conditions. In experiments repeated in the presence of 5 mM TCEP, the bidimensional clouds for dwell time and blockade level display a single population of events with an average current blockade of 0.53 ± 0.01 for L-AVP and 0.49 ± 0.01 for D-AVP (Figure 2h,k; Supporting Information S1, S4). This single population in the presence of a reducing agent shows that we indeed have a change or loss of conformation. We confirmed that this single population is not due to TCEP affecting the detection of the peptides (Supporting Information S6). By comparing the previous experiments in native conditions, we can attribute events Type I and II to the disulfide bond. In the presence of the reducing agent, we observe that for both peptides the current blockade for Type II events increases, from 0.39 ± 0.01 for D-AVP to 0.49 ± 0.01

for D-AVP+TCEP and from 0.43 ± 0.01 for L-AVP to 0.53 ± 0.01 for L-AVP+TCEP (Figure 2i,j) (Supporting Information S1). The reduction of the disulfide bond releases the constraint on the peptide conformation, conferring more flexibility. To confirm these results, we analyzed and compared the event frequency in native and reducing conditions (Supporting Information S7). We can observe that event frequency at similar concentrations of peptide and applied voltage increased drastically after treatment with TCEP, from 19.1 ± 1.0 Hz to 153.3 ± 8.5 Hz for L-AVP at 110 mV in native and reducing conditions and 28.1 ± 0.9 and 107.2 ± 3.6 Hz for D-AVP, respectively. This result shows a reduced energy barrier for the entry of peptides treated with TCEP inside the nanopore. The reduction of the disulfide bond allows the peptide chain to adopt different conformations, which makes the peptide more flexible. This result is confirmed by an increase in the blockade level between the native and the reduced peptides (Figure 2c,f,i,l).

Enantiomer Discrimination in a Mix. Since we observed a significant difference in blockade level between L- and D-AVP in native and reducing conditions, we performed equimolar mixes in both conditions to determine (1) if we can discriminate them in a mix and (2) what are the best experimental conditions for their discrimination (native or reducing conditions, applied voltage).

We compared the superposition of the histograms of each independent experiment and the mixes (Figure 3a,b,g,h). In the absence of TCEP, we focused on the Type II events that allowed discrimination between the peptides. We varied the applied voltage (Supporting Information S8) and salt concentration (Supporting Information S9) to determine the best experimental conditions, with data at 50 and 110 mV in 4 M KCl displayed in Figure 3. At 50 mV, the blockade level histograms of L- and D-AVP in native or reducing conditions overlap (Figure 3a,c,e,g,i). Indeed, in a mix, we could not discriminate between each peptide population at this voltage. By increasing the voltage to 110 mV, the blockade levels for each peptide were better resolved (Figure 3b,d,f,h,j).

We measured the most probable mean blockade level for both mixes (Supporting Information S10). In native conditions, blockade levels of 0.39 ± 0.01 and 0.43 ± 0.01 were measured, while with TCEP they were found to be 0.49 ± 0.01 and 0.53 ± 0.01 (Figure 3d,j) (Supporting Information S11), confirming that we can discriminate L- and D-AVP in an equimolar mix in both conditions. To verify that each population observed is attributed to L- or D-AVP, we changed the concentration ratio to 75% L-AVP and 25% D-AVP (Figure 3e,f). We observed a decrease in the number of events for the population attributed to D-AVP compared to L-AVP. To further confirm this, we changed the ratio to 25% L-AVP and 75% D-AVP and showed the same ability to discriminate between the peptides (Supporting Information S12). Here, we demonstrated that we could discriminate and identify L- and D-AVP, two native or reduced peptides differing only by one chiral amino acid in a mix using a WT aerolysin nanopore at 110 mV.

Identification by Principal Component Analysis and Monte Carlo Approaches. In order to accurately identify each enantiomeric peptide and their different conformations, we performed a principal component analysis (PCA). For the experimental data analysis, we considered just the blockade level DI_b to characterize each blockade. We can go further by better characterizing the shape of the blockades using

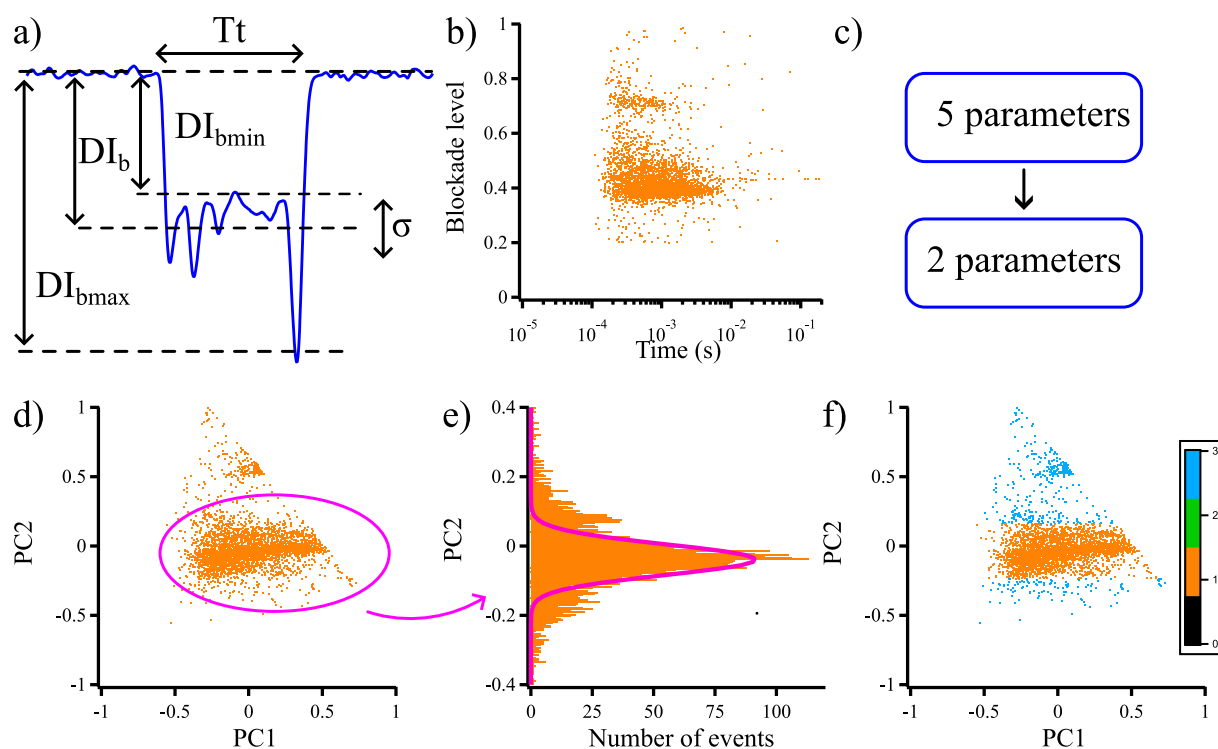


Figure 4. Principal component analysis (PCA). (a) Parameters used to characterize each blockade minimum, average, maximum blockades (DI_{bmin} , DI_b , DI_{bmax}), standard deviation (σ), and duration (T_t) of each blockade. (b) Scatter plot of the average blockade of each D-AVP blockade according to its duration. (c) PCA strategy to decrease the parameter numbers. (d) Second principal component (PC2) according to the first one (PC1). (e) PC2 histogram fitted by a Gaussian distribution ($\mu_D = 0.0474$, $\sigma_D = 0.0348$, describing the most probable value and the standard deviation of the distribution, respectively). (f) Selection of the relevant blockades in the range $[\mu_D \pm \sigma_D]$ in $10 \mu\text{M}$ D-AVP using an aerolysin nanopore in 4 M KCl and 25 mM Tris , $\text{pH } 7.5$, $V = 110 \text{ mV}$.

additional parameters.^{41–43} Then, we can consider four other parameters defined in Figure 4a: maximum and minimum blockades (DI_{bmax} , DI_{bmin} , respectively), duration (T_t), and standard deviation (σ). We removed the blockades characterized by a standard deviation smaller than 1 pA , which were due to bumping, or by a blockade level smaller than 0.2 (see above). These data are plotted in Figure 4b from the results obtained with $10 \mu\text{M}$ D-AVP and show two main clusters: a small one characterized by a high blockade level of around 0.7 and a large one around 0.4 .

If we want to take the five parameters together into account, it is necessary to perform a PCA to reduce the number of relevant parameters. Then, we consider the two most relevant parameters: the first and second principal components (PC1 and PC2, respectively; Figure 4e). We observe two clusters similar to those in Figure 4b: a small cluster and a large one. We are particularly interested in the second principal component (PC2) and its Gaussian distribution for the larger population, which is similar to the blockade levels (Figure 2c,f). The two peptides have very similar structures, so the corresponding distributions are very close. These two distributions overlap, making classical clustering analyses with well-separated domains impossible (Supporting Information S13).

To overcome this difficulty, we considered the results obtained from experiments with individual peptides. As a first approximation, each distribution is fitted by a Gaussian centered at μ_D (respectively μ_L) with a standard deviation of σ_D (respectively σ_L) for the D-AVP (respectively L-AVP) peptides. Blockades belonging to the domain $[\mu - 3\sigma, \mu + 3\sigma]$ are attributed to the corresponding peptide (D-AVP or L-AVP),

allowing the classification of 99% of blockades (Figure 4e,f and Supporting Information S13). Fitting the histogram of the main PC2 population for the D-AVP like this determines $\mu_D = -0.039$ and $\sigma_D = 0.043$ (Figure 4e). The classification is performed by taking into account all of the blockades in the range $[\mu_D - 3\sigma_D, \mu_D + 3\sigma_D]$. These data are colored in orange. The remaining data are not labeled and are in cyan (Figure 4f).

We followed the same approach with the L-AVP experimental data using the correlation matrix previously calculated from D-AVP data. We can observe a similar behavior: a PC2 distribution fitted by a Gaussian function as before, where $\mu_L = 0.047$ and $\sigma_L = 0.035$ are the fitted values for L-AVP. All the blockades in the range $[\mu_L - 3\sigma_L, \mu_L + 3\sigma_L]$ are attributed to L-AVP peptides (in green). The other blockades are unassigned (in cyan) (Supporting Information S13).

We used these criteria to discriminate between D-AVP and L-AVP in a mixture. First, we combined data from D-AVP and L-AVP current traces to define the training data (Figure 5a). These classifications led to labeling the previous data combined (Figure 5a). The PCA is computed using the correlation matrix previously calculated from D-AVP data. The PC2 distribution shows the two Gaussian peaks previously observed for each peptide (Figure 5b). As both distributions are overlaid, the logistic regression approach to define clusters is irrelevant (Supporting Information S14). We follow a Monte Carlo approach to discriminate the two peptides and label them (Figure 5c), followed by a comparison of our prediction with the original labeling of D-AVP and L-AVP by calculating a confusion matrix (Figure 5e). The success rates are 71% for D-AVP and 75% for L-AVP recognition. Nevertheless, the false-

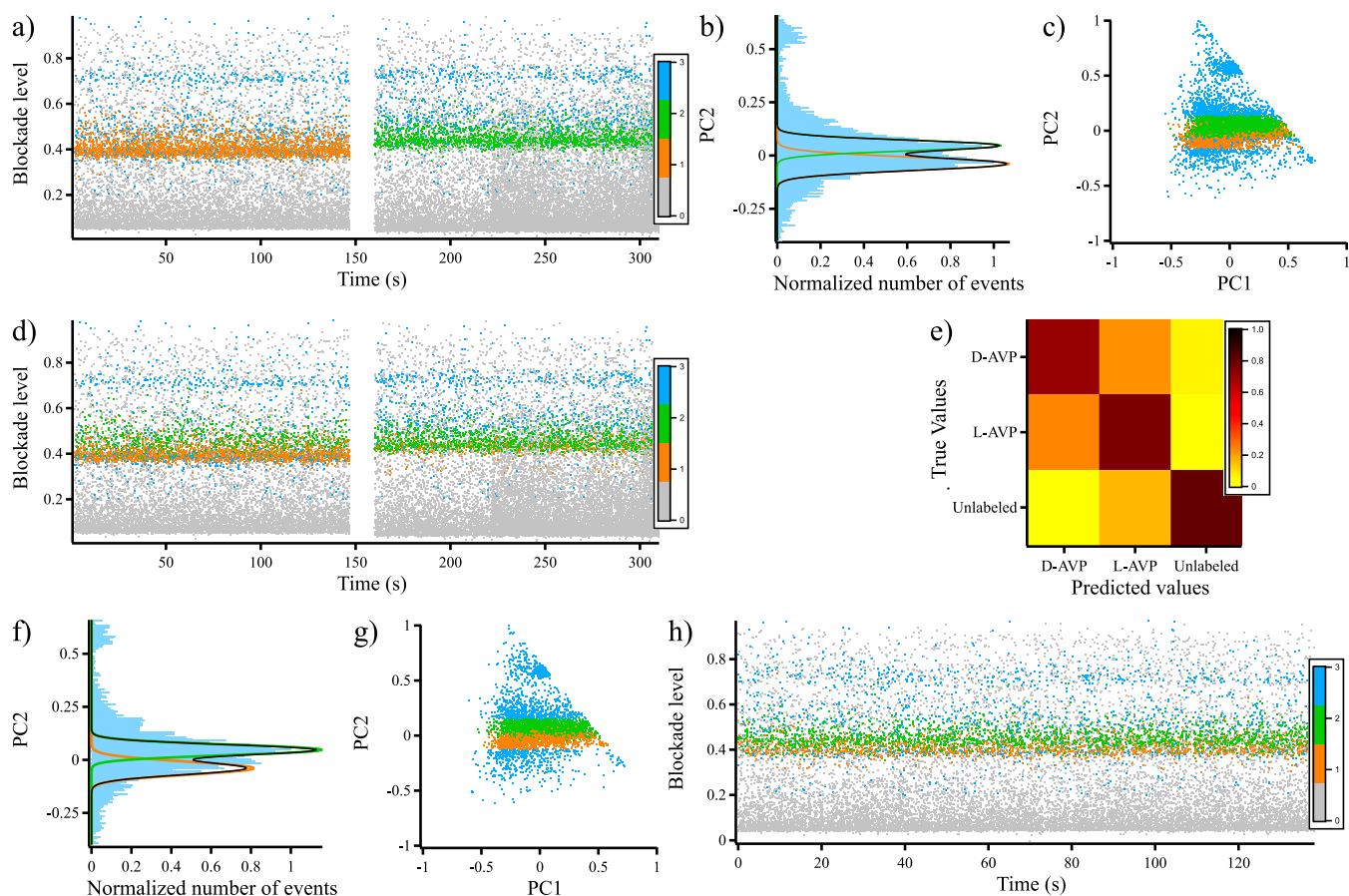


Figure 5. Prediction of D- and L-AVP mixtures. (a) Combined data from blockade traces obtained for D- (orange) or L-AVP (green), respectively. The gray events correspond to the ones with $\sigma < 1$ pA (10 μM D- or L-AVP using an aerolysin nanopore in 4 M KCl and 25 mM Tris, pH 7.5, $V = 110$ mV). (b) Distribution of the second principal component (PC2), fitted by bi-Gaussian (gray) or Gaussian (orange or green) curves. (c) Scatter plot of the two principal components. The area colored in orange or green corresponds to the bands centered in μ_D (μ_L) and with a width of $\pm 3\sigma_D$ (σ_L). It is labeled according to the nature of each isomer. (d) Evaluation of Monte Carlo prediction. Predicted blockade trace according to four types of labeling obtained from the data used in (a). (e) Confusion matrix calculated from the Monte Carlo prediction. (f) Equimolar mixture of 10 μM D- and L-AVP distribution of the second principal component (PC2), fitted by bi-Gaussian (gray) or Gaussian (orange or green) curves. $\Delta V = 110$ mV. (g) Scatter plot of the two principal components. The area colored in orange or green is calculated according to a Monte Carlo algorithm. (h) Blockade trace according to four types of labeling: $\sigma < 1$ (gray), D-AVP (orange), L-AVP (green), and not attributed (cyan).

positive rates are similar for both peptides (around 22%) due to the overlap in the two distributions.

We applied this approach to data collected from an equimolar mixture (10 μM D-AVP and L-AVP). The PC2 distribution (Figure 5f) is similar to the one observed for the combined data from D-AVP and L-AVP in Figure 5b. We used the same criteria as used previously to label the blockades: 1 = D-AVP (orange), 2 = L-AVP (green), and 3 = unlabeled data (cyan). Figure 5g shows the scatter plot of the two first-principal components according to this labeling, which leads to the blockade trace represented in Figure 5h.

We applied this method to nonequimolar mixtures (2.5 μM L-AVP/7.5 μM D-AVP and 7.5 μM L-AVP/2.5 μM D-AVP), leading to the ratio calculation of the number of L-AVP blockades divided by that of D-AVP blockades (Supporting Information S15). The correlation between this ratio and the relative composition of the mixtures (1:3, 1:1, and 1:3 L-AVP:D-AVP, respectively) shows the accuracy of our approach.

Using five parameters provides a more comprehensive view of the blockade phenomena and allows the determination of the most relevant parameters. The correlation matrix calculated from the D-AVP data was systematically used for each analysis. The scatter plots representing the first two

principal components show a structure with several domains (a large and a small) comparable to classic representations of blockade levels versus durations, whether for D-AVP or L-AVP peptides.

In the case of data collected from a mixture, we followed a Monte Carlo approach to assign the blockades to one peptide or the other (Figure 5f). Our approach was validated using the combined data sets from experiments performed with the L-AVP and D-AVP peptides separately. The corresponding confusion matrix shows a success rate of 73%, which is explained by the overlapping between the two distributions. This result is interesting and demonstrates the relevance of this approach in the statistical analysis of peptides by nanopores.

Furthermore, these blockade level distributions (Figure 2c,f) highlight a shoulder attributed to the open conformation.⁵⁰ The PC2 distributions clearly show two symmetrical shoulders for D-AVP (Figure 4e), whereas there is only one for L-AVP (Supporting Information S13). Then, we could also discriminate the two peptides from the PC2 shape distribution. Principal component analyses are very promising for further analysis of the different conformations of peptides at the single-molecule level.

CONCLUSION

The world's population is increasing, and people worldwide are living longer. The degradation of the environment also impacts health. In this context, we expect an increase in chronic diseases, cancers, and neurodegenerative diseases.⁴⁰ To answer this health challenge, we need a powerful, sensitive, and specific tool to perform early disease detection. Many molecules, including established peptide and protein clinical biomarkers and yet-to-be-identified biomarkers, remain underutilized or unexploited for clinical applications. This is due to the absence of techniques capable of discerning different conformations, post-translational modifications, and D/L amino acids. This study paves the way for the ability to identify, characterize, and quantify free or contained D-amino acids in peptides or proteins, which will be crucial for the early detection of human diseases. From a chemical and pharmaceutical point of view, this technique would also allow the scientific community to discriminate isomers⁶¹ or follow the conversion of molecule chirality⁶² and explain nonenzymatic racemization pathways.⁶³ While this study focuses on a specific peptide model, there is no doubt that this technique can be extended to other enantiomer peptides (lanthipeptide and A β peptides), as shown in two recent publications using different protein nanopores such as OmpF,⁴⁸ CytK, and FraC.⁴⁹

We showed that we can detect and discriminate biologically relevant peptides differing by a single amino-acid enantiomer, L- or D-Arg. The peptides can be identified individually or as a mixture. We detected two different peptide conformations, saddle (Type IIa) and open (Type IIb), which NMR already described. Due to their unique conformation, we could identify each peptide using two electrical parameters: dwell time and current blockade. Using a reducing agent, we did not detect any more conformational variants observed under native conditions. On the other hand, we can still identify L- and D-AVP that adopt different conformations. We also used a PCA approach to confirm our experimental data analysis and define the best electrical parameters to discriminate each population of events. This method is up-and-coming and will make it possible to study the conformations of biomarkers in more detail at the single-molecule level.

METHODS

Peptides. Vasopressin peptides, L-Arg vasopressin (L-AVP), Nter-C₁-Y₂-F₃-Q₄-N₅-C₆-P_{7(L)}-R₈-G₉-CO-NH₂, and D-Arg vasopressin (D-AVP), Nter-C₁-Y₂-F₃-Q₄-N₅-C₆-P_{7(D)}-R₈-G₉-CO-NH₂, were synthesized and purified by HPLC (Proteogenics, Schiltigheim, France) with a disulfide bond between C₁-C₆, dissolved in 25 mM Tris pH 7.5 to either 1 or 10 mM, and aliquoted and stored at -20 °C until use. The purity analysis of the L-AVP (Supporting Information S16) and D-AVP (Supporting Information S17) is checked by mass spectrometry and liquid chromatography.

Chemicals. Tris(2-carboxyethyl)phosphine (ref: T2556), hydrochloride (TCEP) was synthesized by Invitrogen (Thermo Fisher Scientific) and dissolved in 100 mM Tris pH 7.5 to 1 M and aliquoted and stored at -20 °C until use.

Protein Production. Proaerolysin was produced by DreamPore SAS (Cergy, France) as described previously.²¹ Briefly, C-terminally His-tagged protein was expressed into the periplasm of BL21 Ros 2 cells (MilliporeSigma), harvested by osmotic shock, before further purification by nickel affinity and

buffer exchange chromatography (Cytiva, Marlborough MA, USA) in standard buffers containing 350 mM NaCl. Purified protein was stored at 4 °C until further use, when it was activated using trypsin-bound beads (Thermo Scientific, Waltham, MA, USA). It was then used in nanopore experiments at a 0.5–1 nM final concentration of aerolysin monomers.

Nanopore Experiments. Nanopore experiments were performed using a vertical planar lipid bilayer setup (Warner Instruments, Hamden CT, USA) with a 150 nm aperture, as described previously.⁵⁷ A 10 mg/mL stock of diphytanoyl-phosphatidyl-choline (DphPC, Avanti Polar Lipids, Alabaster, AL, USA), dissolved in decane, was used to create a planar lipid bilayer separating compartments containing 1 mL of electrolyte solution, 4 M KCl, and 25 mM Tris, pH 7.4. Once a single aerolysin pore was inserted, peptide analytes (individually or in a mix) were introduced at different concentrations before a 110 or 50 mV voltage differential was applied through Ag/AgCl electrodes.

Data Acquisition and Analysis. Electrical measurements were performed by using an Axopatch 200B and a Digidata 1440 digitizer. Data were recorded at 250 kHz or 4 μ s sampling time and filtered at 5 kHz using Clampex software (Axon Instruments, Union City, CA, USA). Three-minute recordings taken at 50 or 110 mV were cleaned and analyzed with Igor Pro (Wavemetrics, Portland, OR, USA), using in-house algorithms to detect events and extract their characteristic parameters (current baseline level I_0 , values for the average (I_b), maximum (I_{bmax}), and minimum (I_{bmin}) current blockade level within the event, event noise or standard deviation of the current values within the event (I_{bs}), dwell time, and blockade time location in the data set. The average open pore current (I_0) and standard deviation (σ) for each recording were determined statistically,⁶⁴ and a threshold of $I_0 - 7\sigma$ was used to define events for the target peptides in the characterization experiments (Figures 2, 3). Extracted parameters from multiple recordings were concatenated for low-frequency events to provide sufficient information for robust statistics. To remove bias introduced by short bumping events with a high blockade fraction, events with dwell times equal to or greater than 200 μ s were used for further analysis. This is a reasonable approach given the overall distribution of average blockade fraction vs dwell time data (Supporting Information S2, S3), relative length of event dwell times, and the use of a 5 kHz filter. Histograms of the average blockade fraction for each event (DI_b/I_0 , where $DI_b = I_0 - I_b$; binned at 0.005) were fit with Gaussian or bi-Gaussian functions to determine the average blockade fraction for each population of events. Semilog histograms of the number of events against dwell time, with 30 bins per decade, were fit with an exponential function to determine the most probable dwell time (Supporting Information S2). Semilog distributions of interevent time, binned at 2 ms for global frequency and 4 ms for population frequency, were fit with single-exponential functions to determine the event frequency (Supporting Information S7). Each experiment contained at least 1700 events of more than 200 μ s in dwell time. Depending on their blockade level distribution, populations were selected to determine each mean dwell time value. For Type I (L- and D-AVP), the chosen points were found between 0.7 and 0.8 blockade level, and for Type II, between 0.35 and 0.55. In the presence of TCEP, points were selected between 0.4 and 0.8. The average and standard deviation determined from 3

independent fits were used for each fit to account for fit robustness.

Semisupervised Clustering Using Principal Component Analysis and Logistic Regression Classification.

The five parameters (T_v , DI_{bmin} , DI_{bmax} , DI_b , and I_{bs}) were determined (Figure 5b) using $DI_{bmin} = I_0 - I_{bmax}$, $DI_{bmax} = I_0 - I_{bmin}$, and $DI_b = I_0 - I_b$, respectively. PCA was performed using the PCA module of the scikit-learn library in Python.^{65,66} Briefly, the covariance matrix between these previously normalized five parameters was calculated to determine this matrix's eigenvalues and vectors. Each parameter is projected onto each eigenvector, with only the eigenvectors characterized by the most significant eigenvalues (PC1 and PC2) used for further analysis. The calculations with L-AVP or mixtures were performed using the correlation matrix already calculated with D-AVP. The logistic regression classification⁶⁷ was performed using the corresponding module in the scikit-learn Python library. The training is performed with 20% of the data set composed by concatenating data from D-AVP and L-AVP experiments (without or with TCEP). We can also follow a Monte Carlo type approach. We first used the histograms of PC2 obtained with each D-AVP or L-AVP peptide. These histograms are fitted by distinct Gaussian distributions centered in μ with a standard sigma deviation. We used these two Gaussian distributions to classify the blockages of a mixture according to a Monte Carlo approach. This method is evaluated using data obtained after the concatenation of data obtained previously with D-AVP and L-AVP peptides. We only take into account the blockages included in the band $[\mu - 3\sigma, \mu + 3\sigma]$ of each Gaussian distribution. The predictions are compared to the true initial values to evaluate this method and calculate the corresponding confusion matrix using the corresponding sci-kit-learn library in Python.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscentsci.4c00020>.

Additional experimental details and results as well as tables regrouping population characteristics in each condition (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Bénédicte Thiebot – Université Paris-Saclay, Univ Evry, CY Cergy Paris Université, CNRS, LAMBE, 95000 Cergy, France; Email: benedicte.thiebot@cyu.fr

Benjamin Cressiot – Université Paris-Saclay, Univ Evry, CY Cergy Paris Université, CNRS, LAMBE, 95000 Cergy, France; orcid.org/0000-0001-9319-3152; Email: benjamin.cressiot@cyu.fr

Juan Pelta – Université Paris-Saclay, Univ Evry, CY Cergy Paris Université, CNRS, LAMBE, 91025 Evry-Courcouronnes, France; orcid.org/0000-0002-1896-4137; Email: juan.pelta@univ-evry.fr

Authors

Laura Ratinho – Université Paris-Saclay, Univ Evry, CY Cergy Paris Université, CNRS, LAMBE, 95000 Cergy, France; orcid.org/0009-0006-1113-5949

Laurent Bacri – Université Paris-Saclay, Univ Evry, CY Cergy Paris Université, CNRS, LAMBE, 91025 Evry-Courcouronnes, France; orcid.org/0000-0002-1763-2456

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acscentsci.4c00020>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by CY Initiative of Excellence Investissements d'Avenir No. ANR-16-IDEX-0008, by ANR JJCJ “QuaBioNH_y” No. ANR-23-CE44-0013, and by Direction Générale de l'Armement (DGA) and Agence d'Innovation Défense (AID). We acknowledge Dreamport for providing aerolysin. The authors thank S. J. Greive for careful reading and for improving the English in the manuscript.

■ REFERENCES

- (1) Abdulbagi, M.; Wang, L.; Siddig, O.; Di, B.; Li, B. D-Amino Acids and D-Amino Acid-Containing Peptides: Potential Disease Biomarkers and Therapeutic Targets? *Biomolecules* **2021**, *11* (11), 1716.
- (2) Mothet, J. P. Physiological Relevance of Endogenous Free D-Serine in the Mammalian Brain: Are Scientists on a Royal Road for the Treatment of Glutamatergic-Related Brain Disorders? *Pathol. Biol.* **2001**, *49* (8), 655–659.
- (3) Ploux, E.; Freret, T.; Billard, J.-M. D-Serine in Physiological and Pathological Brain Aging. *Biochim. Biophys. Acta (BBA) - Proteins Proteom.* **2021**, *1869* (1), 140542.
- (4) Li, Y.; Han, H.; Yin, J.; Li, T.; Yin, Y. Role of D-Aspartate on Biosynthesis, Racemization, and Potential Functions: A Mini-Review. *Anim. Nutr.* **2018**, *4* (3), 311–315.
- (5) Orzylowski, M.; Fujiwara, E.; Mousseau, D. D.; Baker, G. B. An Overview of the Involvement of D-Serine in Cognitive Impairment in Normal Aging and Dementia. *Front. Psychiatry* **2021**, *12*, 754032.
- (6) Ota, N.; Shi, T.; Sweedler, J. V. D-Aspartate Acts as a Signaling Molecule in Nervous and Neuroendocrine Systems. *Amino Acids* **2012**, *43* (5), 1873–1886.
- (7) Fujii, N.; Tajima, S.; Tanaka, N.; Fujimoto, N.; Takata, T.; Shimo-Oka, T. The Presence of D-β-Aspartic Acid-Containing Peptides in Elastic Fibers of Sun-Damaged Skin: A Potent Marker for Ultraviolet-Induced Skin Aging. *Biochem. Biophys. Res. Commun.* **2002**, *294* (5), 1047–1051.
- (8) Fujii, N.; Satoh, K.; Harada, K.; Ishibashi, Y. Simultaneous Stereo-inversion and Isomerization at Specific Aspartic Acid Residues in AA-Crystallin from Human Lens. *J. Biochem.* **1994**, *116* (3), 663–669.
- (9) Roher, A. E.; Lowenson, J. D.; Clarke, S.; Wolkow, C.; Wang, R.; Cotter, R. J.; Reardon, I. M.; Zürcher-Neely, H. A.; Heinrikson, R. L.; Ball, M. J. Structural Alterations in the Peptide Backbone of Beta-Amyloid Core Protein May Account for Its Deposition and Stability in Alzheimer's Disease. *J. Biol. Chem.* **1993**, *268* (5), 3072–3083.
- (10) Sheykhkarimli, D.; Choo, K.-L.; Owen, M.; Fiser, B.; Jójárt, B.; Csizmadia, I. G.; Viskolcz, B. Molecular Ageing: Free Radical Initiated Epimerization of Thymopentin - A Case Study. *J. Chem. Phys.* **2014**, *140* (20), 205102.
- (11) Kuge, K.; Kitamura, K.; Nakaoji, K.; Hamada, K.; Fujii, N.; Saito, T.; Fujii, N. Oxidative Stress Induces the Formation of D-Aspartyl Residues in the Elastin Mimic Peptides. *Chem. Biodivers.* **2010**, *7* (6), 1408–1412.
- (12) Yan, J.; Yao, Y.; Yan, S.; Gao, R.; Lu, W.; He, W. Chiral Protein Supraparticles for Tumor Suppression and Synergistic Immunotherapy: An Enabling Strategy for Bioactive Supramolecular Chirality Construction. *Nano Lett.* **2020**, *20* (8), 5844–5852.

- (13) Fujii, N.; Takata, T.; Fujii, N.; Aki, K.; Sakaue, H. D-Amino Acids in Protein: The Mirror of Life as a Molecular Index of Aging. *Biochim. Biophys. Acta (BBA) - Proteins Proteom.* **2018**, *1866* (7), 840–847.
- (14) Lander, A. J.; Jin, Y.; Luk, L. Y. P. D-Peptide and D-Protein Technology: Recent Advances, Challenges, and Opportunities**. *ChemBioChem.* **2023**, *24* (4), No. e202200537.
- (15) Oukhaled, G.; Mathe, J.; Bianca, A. L.; Bacri, L.; Betton, J.-M.; Lairez, D.; Pelta, J.; Auvray, L. Unfolding of Proteins and Long Transient Conformations Detected by Single Nanopore Recording. *Phys. Rev. Lett.* **2007**, *98* (15), 158101–158101.
- (16) Merstorf, C.; Cressiot, B.; Pastoriza-Gallego, M.; Oukhaled, A.; Betton, J.-M.; Auvray, L.; Pelta, J. Wild Type, Mutant Protein Unfolding and Phase Transition Detected by Single-Nanopore Recording. *ACS Chem. Biol.* **2012**, *7* (4), 652–658.
- (17) Galenkamp, N. S.; Biesemans, A.; Maglia, G. Directional Conformer Exchange in Dihydrofolate Reductase Revealed by Single-Molecule Nanopore Recordings. *Nat. Chem.* **2020**, *167*, 378–8.
- (18) Boersma, A. J. A.; Bayley, H. H. Continuous Stochastic Detection of Amino Acid Enantiomers with a Protein Nanopore. *Angew. Chem., Int. Ed. Engl.* **2012**, *51* (38), 9606–9609.
- (19) He, L.; Tessier, D. R.; Briggs, K.; Tsangaris, M.; Charron, M.; McConnell, E. M.; Lomovtsev, D.; Tabard-Cossa, V. Digital Immunoassay for Biomarker Concentration Quantification Using Solid-State Nanopores. *Nat. Commun.* **2021**, *12* (1), 5348.
- (20) Chuah, K.; Wu, Y.; Vivekchand, S. R. C.; Gaus, K.; Reece, P. J.; Micolich, A. P.; Gooding, J. J. Nanopore Blockade Sensors for Ultrasensitive Detection of Proteins in Complex Biological Samples. *Nat. Commun.* **2019**, *10* (1), 2109.
- (21) Stierlen, A.; Greive, S. J.; Bacri, L.; Manivet, P.; Cressiot, B.; Pelta, J. Nanopore Discrimination of Coagulation Biomarker Derivatives and Characterization of a Post-Translational Modification. *Acs Central Sci.* **2023**, *9* (2), 228–238.
- (22) Lucas, F. L. R.; Versloot, R. C. A.; Yakovlieva, L.; Walvoort, M. T. C.; Maglia, G. Protein Identification by Nanopore Peptide Profiling. *Nat. Commun.* **2021**, *12* (1), 5795.
- (23) Huang, G.; Voorspoels, A.; Versloot, R. C. A.; Heide, N. J.; Carlon, E.; Willems, K.; Maglia, G. PlyAB Nanopores Detect Single Amino Acid Differences in Folded Haemoglobin from Blood**. *Angewandte Chemie Int. Ed* **2022**, *61* (34), No. e202206227.
- (24) Ahmad, M.; Ha, J.-H.; Mayse, L. A.; Presti, M. F.; Wolfe, A. J.; Moody, K. J.; Loh, S. N.; Movileanu, L. A Generalizable Nanopore Sensor for Highly Specific Protein Detection at Single-Molecule Precision. *Nat. Commun.* **2023**, *14* (1), 1374.
- (25) Huang, G.; Willems, K.; Soskine, M.; Wloka, C.; Maglia, G. Electro-Osmotic Capture and Ionic Discrimination of Peptide and Protein Biomarkers with FraC Nanopores. *Nat. Commun.* **2017**, *8* (1), 935.
- (26) Galenkamp, N. S.; Soskine, M.; Hermans, J.; Wloka, C.; Maglia, G. Direct Electrical Quantification of Glucose and Asparagine from Bodily Fluids Using Nanopores. *Nat. Commun.* **2018**, *9* (1), 4085.
- (27) Jiang, J.; Li, M.-Y.; Wu, X.-Y.; Ying, Y.-L.; Han, H.-X.; Long, Y.-T. Protein Nanopore Reveals the Renin-Angiotensin System Crosstalk with Single-Amino-Acid Resolution. *Nat. Chem.* **2023**, 1–9.
- (28) Greive, S. J.; Bacri, L.; Cressiot, B.; Pelta, J. Identification of Conformational Variants for Bradykinin Biomarker Peptides from a Biofluid Using a Nanopore and Machine Learning. *ACS Nano* **2024**, *18* (1), 539–550.
- (29) Zhang, M.; Tang, C.; Wang, Z.; Chen, S.; Zhang, D.; Li, K.; Sun, K.; Zhao, C.; Wang, Y.; Xu, M. J.; et al. Real-Time Detection of 20 Amino Acids and Discrimination of Pathologically Relevant Peptides with Functionalized Nanopore. *Nat. Methods* **2024**, 1–10.
- (30) Restrepo-Pérez, L.; Wong, C. H.; Maglia, G.; Dekker, C.; Joo, C. Label-Free Detection of Post-Translational Modifications with a Nanopore. *Nano Lett.* **2019**, *19* (11), 7957–7964.
- (31) Versloot, R. C. A.; Lucas, F. L. R.; Yakovlieva, L.; Tadema, M. J.; Zhang, Y.; Wood, T. M.; Martin, N. I.; Marrink, S. J.; Walvoort, M. T. C.; Maglia, G. Quantification of Protein Glycosylation Using Nanopores. *Nano Lett.* **2022**, *22* (13), 5357–5364.
- (32) Restrepo-Pérez, L.; Huang, G.; Bohländer, P. R.; Worp, N.; Eelkema, R.; Maglia, G.; Joo, C.; Dekker, C. Resolving Chemical Modifications to a Single Amino Acid within a Peptide Using a Biological Nanopore. *ACS Nano* **2019**, *13* (12), 13668–13676.
- (33) Huo, M.; Hu, Z.; Ying, Y.; Long, Y. Enhanced Identification of Tau Acetylation and Phosphorylation with an Engineered Aerolysin Nanopore. *Proteomics* **2022**, *22* (5–6), 2100041.
- (34) Meng, F.-N.; Ying, Y.-L.; Yang, J.; Long, Y.-T. A Wild-Type Nanopore Sensor for Protein Kinase Activity. *Anal. Chem.* **2019**, *91* (15), 9910–9915.
- (35) Ying, Y.-L.; Yang, J.; Meng, F.-N.; Li, S.; Li, M.-Y.; Long, Y.-T. A Nanopore Phosphorylation Sensor for Single Oligonucleotides and Peptides. *Research (Wash D C)* **2019**, *2019* (8), 1050735–1050738.
- (36) Ouldali, H.; Sarthak, K.; Ensslen, T.; Pigué, F.; Manivet, P.; Pelta, J.; Behrends, J. C.; Aksimentiev, A.; Oukhaled, A. Electrical Recognition of the Twenty Proteinogenic Amino Acids Using an Aerolysin Nanopore. *Nat. Biotechnol.* **2019**, *31*, 247–6.
- (37) Wang, K.; Zhang, S.; Zhou, X.; Yang, X.; Li, X.; Wang, Y.; Fan, P.; Xiao, Y.; Sun, W.; Zhang, P.; et al. Unambiguous Discrimination of All 20 Proteinogenic Amino Acids and Their Modifications by Nanopore. *Nat. Methods* **2023**, 1–10.
- (38) Cressiot, B.; Bacri, L.; Pelta, J. The Promise of Nanopore Technology: Advances in the Discrimination of Protein Sequences and Chemical Modifications. *Small Methods* **2020**, *4* (11), 2000090.
- (39) Ying, Y.-L.; Hu, Z.-L.; Zhang, S.; Qing, Y.; Fragasso, A.; Maglia, G.; Meller, A.; Bayley, H.; Dekker, C.; Long, Y.-T. Nanopore-Based Technologies beyond DNA Sequencing. *Nat. Nanotechnol.* **2022**, *17* (11), 1136–1146.
- (40) Tanimoto, I. M. F.; Cressiot, B.; Greive, S. J.; Pioufle, B. L.; Bacri, L.; Pelta, J. Focus on Using Nanopore Technology for Societal Health, Environmental, and Energy Challenges. *Nano Res.* **2022**, 1–15.
- (41) Meyer, N.; Janot, J.-M.; Lepoitevin, M.; Smietana, M.; Vasseur, J.-J.; Torrent, J.; Balme, S. Machine Learning to Improve the Sensing of Biomolecules by Conical Track-Etched Nanopore. *Biosensors* **2020**, *10* (10), 140.
- (42) Reynaud, L.; Bouchet-Spinelli, A.; Janot, J.-M.; Buhot, A.; Balme, S.; Raillon, C. Discrimination of A-Thrombin and Γ -Thrombin Using Aptamer-Functionalized Nanopore Sensing. *Anal. Chem.* **2021**, *93* (22), 7889–7897.
- (43) Zhang, J.-H.; Liu, X.-L.; Hu, Z.-L.; Ying, Y.-L.; Long, Y.-T. Intelligent Identification of Multi-Level Nanopore Signatures for Accurate Detection of Cancer Biomarkers. *Chem. Commun.* **2017**, *53* (73), 10176–10179.
- (44) Yuan, C.; Wu, X.; Gao, R.; Han, X.; Liu, Y.; Long, Y.; Cui, Y. Nanochannels of Covalent Organic Frameworks for Chiral Selective Transmembrane Transport of Amino Acids. *J. Am. Chem. Soc.* **2019**, *141* (51), 20187–20197.
- (45) Guo, Y.; Niu, A.; Jian, F.; Wang, Y.; Yao, F.; Wei, Y.; Tian, L.; Kang, X. Metal-Organic Complex-Functionalized Protein Nanopore Sensor for Aromatic Amino Acids Chiral Recognition. *Analyst* **2017**, *142* (7), 1048–1053.
- (46) Ali, M.; Nasir, S.; Ensinger, W. Stereoselective Detection of Amino Acids with Protein-Modified Single Asymmetric Nanopores. *Electrochim. Acta* **2016**, *215*, 231–237.
- (47) Schiopu, I.; Iftemi, S.; Luchian, T. Nanopore Investigation of the Stereoselective Interactions between Cu²⁺ and d,L-Histidine Amino Acids Engineered into an Amyloidic Fragment Analogue. *Langmuir* **2015**, *31* (1), 387–396.
- (48) Wang, J.; Prajapati, J. D.; Gao, F.; Ying, Y.-L.; Kleinekathöfer, U.; Winterhalter, M.; Long, Y.-T. Identification of Single Amino Acid Chiral and Positional Isomers Using an Electrostatically Asymmetric Nanopore. *J. Am. Chem. Soc.* **2022**, *144* (33), 15072–15078.
- (49) Versloot, R. C. A.; Arias-Orozco, P.; Tadema, M. J.; Lucas, F. L. R.; Zhao, X.; Marrink, S. J.; Kuipers, O. P.; Maglia, G. Seeing the Invisibles: Detection of Peptide Enantiomers, Diastereomers, and Isobaric Ring Formation in Lanthipeptides Using Nanopores. *J. Am. Chem. Soc.* **2023**, *145* (33), 18355–18365.

(50) Haensele, E.; Banting, L.; Whitley, D. C.; Clark, T. Conformation and Dynamics of 8-Arg-Vasopressin in Solution. *J. Mol. Model.* **2014**, *20* (11), 2485.

(51) Christ-Crain, M.; Gaisl, O. Diabetes Insipidus. *Press. Médicale* **2021**, *50* (4), 104093.

(52) Manning, M.; Balaspiri, L.; Moehring, J.; Haldar, J.; Sawyer, W. H. Synthesis and Some Pharmacological Properties of Deamino[4-Threonine,8-D-Arginine]Vasopressin and Deamino[8-D-Arginine]-Vasopressin, Highly Potent and Specific Antidiuretic Peptides, and [8-D-Arginine]Vasopressin and Deamino-Arginine-Vasopressin. *J. Med. Chem.* **1976**, *19* (6), 842–845.

(53) Saito, M.; Tahara, A.; Sugimoto, T. 1-Desamino-8-d-Arginine Vasopressin (DDAVP) as an Agonist on V1b Vasopressin Receptor. *Biochem. Pharmacol.* **1997**, *53* (11), 1711–1717.

(54) Lubecka, E. A.; Sikorska, E.; Sobolewski, D.; Prah, A.; Slaninová, J.; Ciarkowski, J. Arginine-, d-Arginine-Vasopressin, and Their Inverso Analogues in Micellar and Liposomic Models of Cell Membrane: CD, NMR, and Molecular Dynamics Studies. *Eur. Biophys. J.* **2015**, *44* (8), 727–743.

(55) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (24), 13770–13773.

(56) Brun, L.; Pastoriza-Gallego, M.; Oukhaled, G.; Mathe, J.; Bacri, L.; Auvray, L.; Pelta, J. Dynamics of Polyelectrolyte Transport through a Protein Channel as a Function of Applied Voltage. *Phys. Rev. Lett.* **2008**, *100* (15), 158302.

(57) Cressiot, B.; Braselmann, E.; Oukhaled, A.; Elcock, A. H.; Pelta, J.; Clark, P. L. Dynamics and Energy Contributions for Transport of Unfolded Pertactin through a Protein Nanopore. *ACS Nano* **2015**, *9* (9), 9050–9061.

(58) Bétermier, F.; Cressiot, B.; Muccio, G. D.; Jarroux, N.; Bacri, L.; della Rocca, B. M.; Chinappi, M.; Pelta, J.; Tarascon, J.-M. Single-Sulfur Atom Discrimination of Polysulfides with a Protein Nanopore for Improved Batteries. *Commun. Mater.* **2020**, *1* (1), 59.

(59) Robertson, J. W. F.; Rodrigues, C. G.; Stanford, V. M.; Rubinson, K. A.; Krasilnikov, O. V.; Kasianowicz, J. J. Single-Molecule Mass Spectrometry in Solution Using a Solitary Nanopore. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (20), 8207–8211.

(60) Haensele, E.; Saleh, N.; Read, C. M.; Banting, L.; Whitley, D. C.; Clark, T. Can Simulations and Modeling Decipher NMR Data for Conformational Equilibria? Arginine-Vasopressin. *J. Chem. Inf. Model.* **2016**, *56* (9), 1798–1807.

(61) Jia, W.; Hu, C.; Wang, Y.; Gu, Y.; Qian, G.; Du, X.; Wang, L.; Liu, Y.; Cao, J.; Zhang, S.; et al. Programmable Nano-Reactors for Stochastic Sensing. *Nat. Commun.* **2021**, *12* (1), 5811.

(62) Du, X.; Zhang, S.; Wang, L.; Wang, Y.; Fan, P.; Jia, W.; Zhang, P.; Huang, S. Single-Molecule Interconversion between Chiral Configurations of Boronate Esters Observed in a Nanoreactor. *ACS Nano* **2023**, *17* (3), 2881–2892.

(63) Bitchagno, G. T. M.; Nchiozem-Ngnitedem, V.-A.; Melchert, D.; Fobofou, S. A. Demystifying Racemic Natural Products in the Homochiral World. *Nat. Rev. Chem.* **2022**, *6* (11), 806–822.

(64) Oukhaled, A.; Bacri, L.; Pastoriza-Gallego, M.; Betton, J.-M.; Pelta, J. Sensing Proteins through Nanopores: Fundamental to Applications. *ACS Chem. Biol.* **2012**, *7* (12), 1935–1949.

(65) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Müller, A.; Nothman, J.; Louppe, G.; et al. Scikit-Learn: Machine Learning in Python. *arXiv* **2012**, DOI: 10.48550/arxiv.1201.0490.

(66) Lever, J.; Krzywinski, M.; Altman, N. Principal Component Analysis. *Nat. Methods* **2017**, *14* (7), 641–642.

(67) Hilbe, J. M. *International Encyclopedia of Statistical Science*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp 755–758.