

Supplementary Information

Targeted high-resolution sensing of volatile organic compounds by covalent nanopore detection

Lauren E McGivern^{1,†}, Zhong Hui Lim^{1,†}, Yizhi Yuan¹, Zonghua Bo¹, Guangqi Wu¹, Hagan Bayley^{1*}, Yujia Qing^{1*}

¹Department of Chemistry, Chemistry Research Laboratory, University of Oxford, Oxford, UK, OX1 3TA.

[†]These authors contributed equally: Lauren E McGivern, Zhong Hui Lim

*Email: hagan.bayley@chem.ox.ac.uk ; yujia.qing@chem.ox.ac.uk

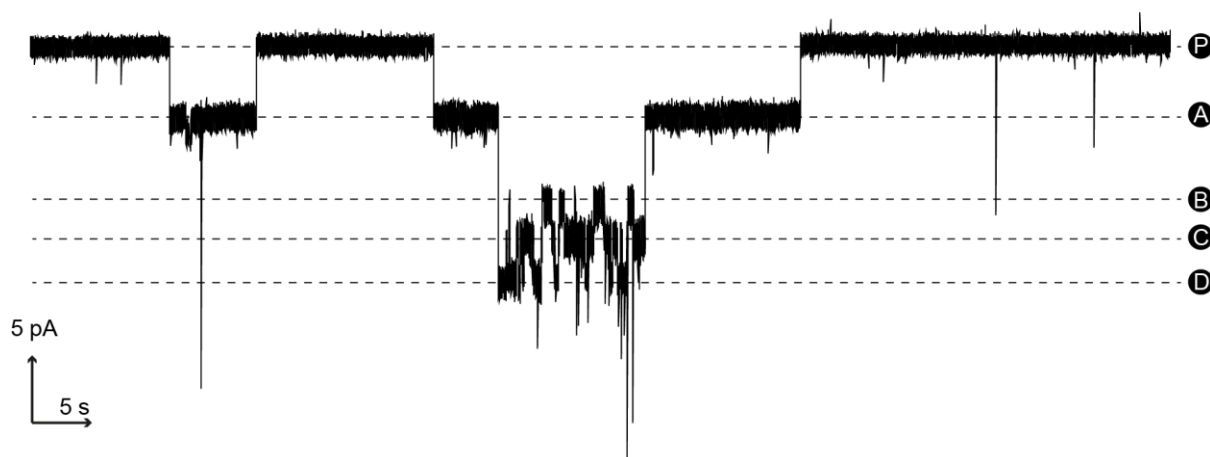
Table of Contents

1	Rationale for using α -hemolysin nanopores with AG background.....	3
2	Concentration-dependence of hemithioacetal formation: Kinetic studies	4
3	Kinetic rate constants for the reversible aldehyde-thiol chemistry and the percentage residual current levels for the hemithioacetal adducts.....	10
4	Hydration constants and solubilities of aldehyde substrates	12
5	Machine learning	15
6	Ratiometric detection of aldehyde mixtures	17
7	Rationale for improved diastereomer discrimination in different nanopores.....	18
8	Improved structural resolution in engineered nanopore mutants	20
9	Experimental methods	23
	General	23
	Plasmid preparation	23
	Nanopore preparation	26
	<i>Arthrobacter cholorphenolicus</i> choline oxidase (AcCO6) preparation.....	27
	Differential sensing of aldehydes and alcohols	27
	Single-channel electrical recording.....	28
	References.....	29

1 Rationale for using α -hemolysin nanopores with AG background

Background reactions were observed between benzaldehyde (20 mM, trans) and the wild-type α -hemolysin (α HL) pore ((WT- α HL)₇), resulting in current levels A, B, C, and D in addition to the open pore level (level P) (Supplementary Fig. 1). These events were eliminated after replacing lysine residues at the cis entrance (Lys-8) and inward-facing lysine residues within the β barrel (Lys-131 and Lys-147). Additionally, the inward-facing methionine residues within the β barrel (Met-113) were replaced to eliminate any potential metal-binding sites within the nanopore.¹ These mutations (i.e., K8A, M113G, K131G and K147G) were collectively called AG mutations. Further studies to identify the adducts producing the undesirable current levels were not conducted.

No detectable reactivity was observed with the (AG)₇ homoheptamer (i.e., AG = WT-K8A-M113G-K131G-K147G) when up to 30 mM of benzaldehyde was introduced into the trans compartment. Hence, events observed when the (AG)₆(AG-T115C) nanopore was used were attributed to the presence of the Cys-115 residue, specifically reversible hemithioacetal formation.

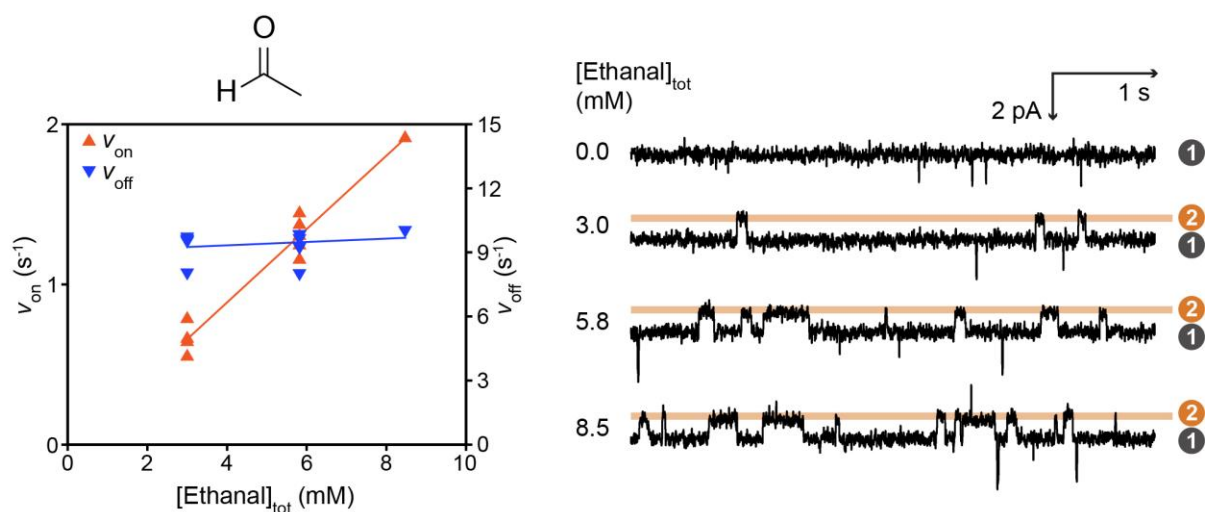


Supplementary Fig. 1: Background reactions between benzaldehyde and the WT α HL pore ((WT- α HL)₇). Levels A, B, C, and D, which occurred in addition to the open pore current level (level P), were attributed to benzaldehyde reacting with lysine residues lining the β barrel and the cis entrance. Conditions: 20 mM benzaldehyde (trans), (WT- α HL)₇ nanopore, 2 M KCl, 200 mM HEPBS (pH 8.5), 20 μ M EDTA, recorded at +50 mV (trans) and 21 ± 1 °C. Signals were low-pass filtered at 10 kHz and sampled at 50 kHz.

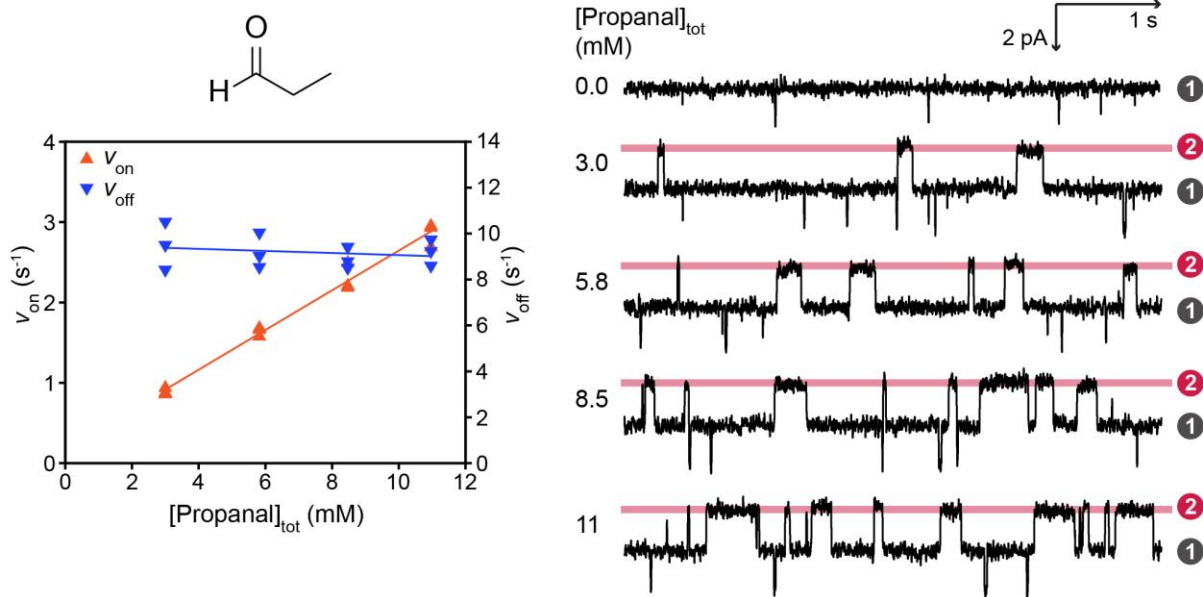
2 Concentration-dependence of hemithioacetal formation: Kinetic studies

Hemithioacetal formation was observed for the reaction between thiols and various aldehydes under aqueous conditions and at neutral pH (e.g., pH 6.5 – 7.2) using ^1H NMR.^{2,3} Additionally, we detected the hemithioacetal product formed from the reaction between hexanal and N-acetyl cysteine methyl ester at pH 6.8 with high resolution mass spectrometry (HRMS) (i.e., HRMS found $[\text{M}+\text{Na}]^+ = 300.1242$; $\text{C}_{12}\text{H}_{23}\text{NNaO}_4\text{S}$ requires 300.1240.)

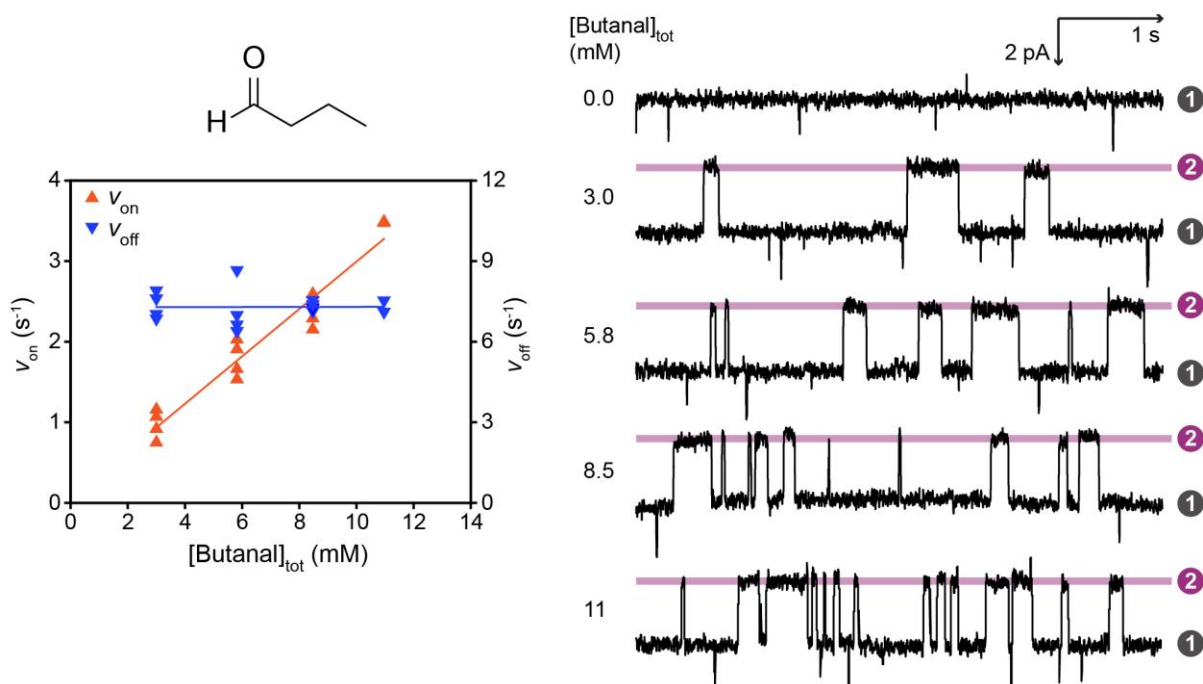
Single-channel recordings were carried out at -50 mV (trans) with 2 M KCl, 200 mM PIPES and 20 μM EDTA at pH 6.8. Signals were low-pass filtered at 10 kHz and sampled at 50 kHz. Traces were further filtered at 100 Hz for display. Errors are standard deviations of rates at each concentration.



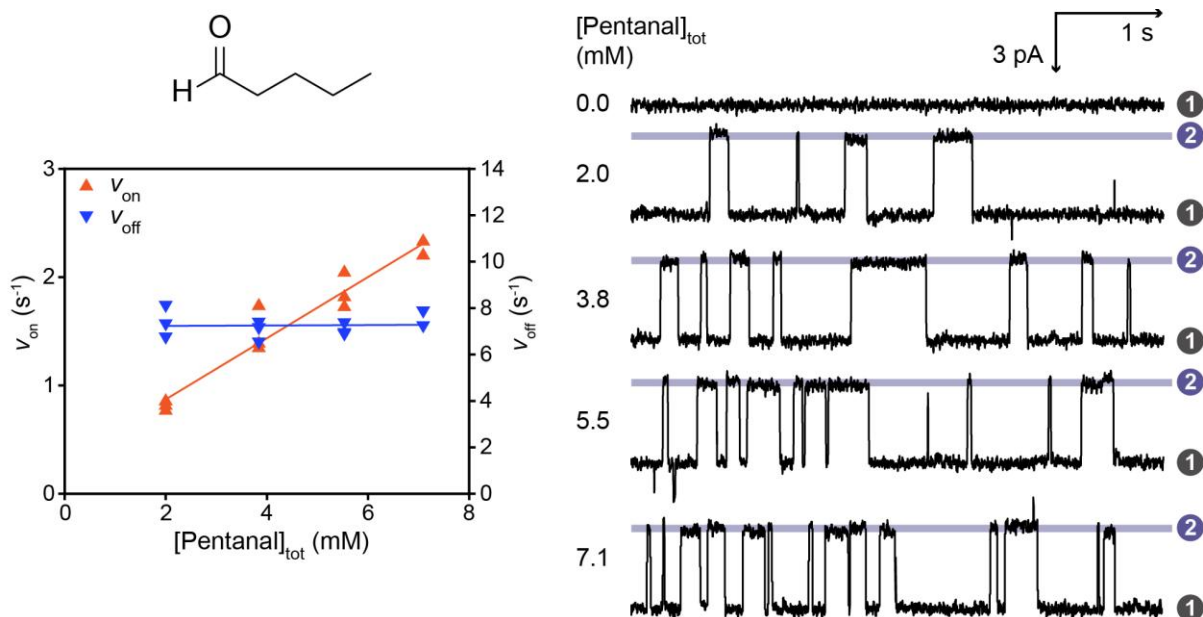
Supplementary Fig. 2: Kinetic analysis of ethanal with $(\text{AG})_6(\text{AG-T115C})$. Left: Rates of adduct formation (v_{on}) and dissociation (v_{off}) plotted against total ethanal concentration. Errors are standard deviations of rates at each concentration. Right: Single-channel recordings with 0.0, 3.0, 5.8 and 8.5 mM total ethanal (trans). Current levels corresponded to the hemithioacetal adducts (2) and the unoccupied nanopore (1).



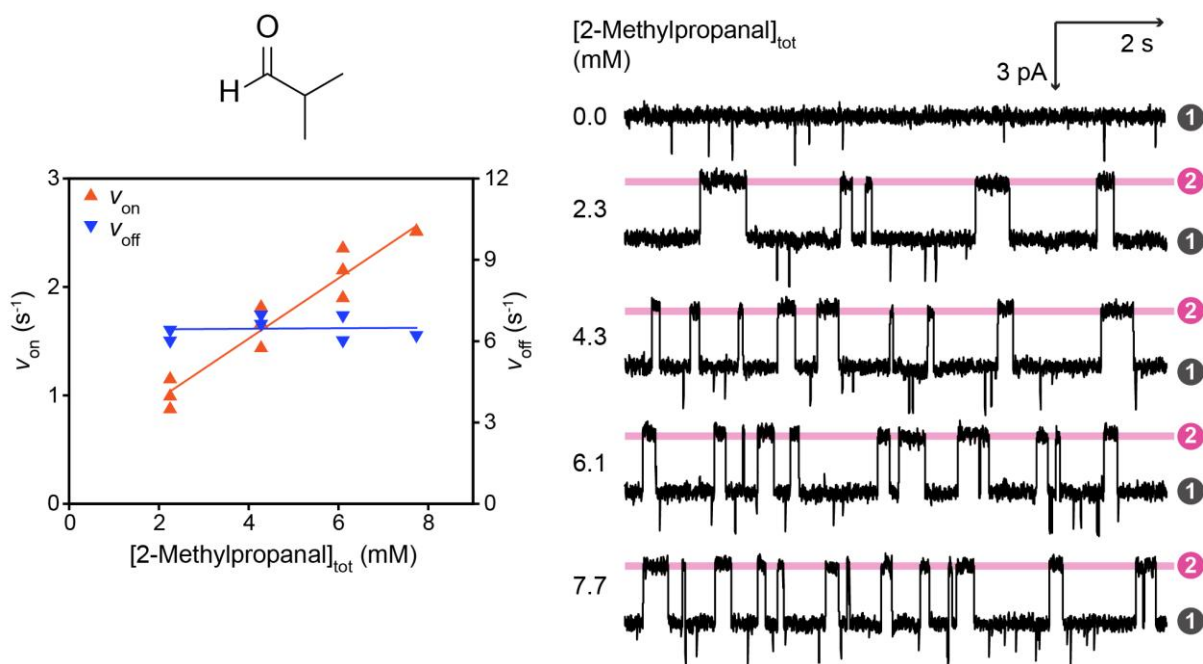
Supplementary Fig. 3: Kinetic analysis of propanal with (AG)₆(AG-T115C). Left: Rates of adduct formation (v_{on}) and dissociation (v_{off}) plotted against total propanal concentration. Errors are standard deviations of rates at each concentration. Right: Single-channel recordings with 0.0, 3.0, 5.8, 8.5 and 11 mM total propanal (trans). Current levels corresponded to the hemithioacetal adducts (2) and the unoccupied nanopore (1).



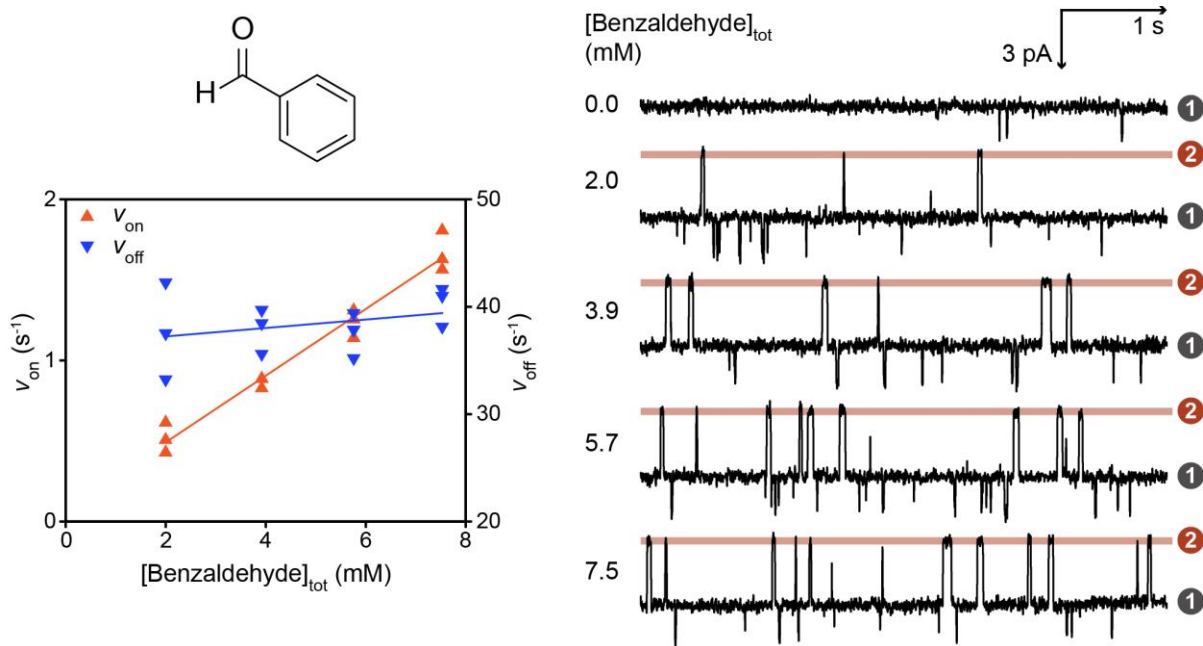
Supplementary Fig. 4: Kinetic analysis of butanal with (AG)₆(AG-T115C). Left: Rates of adduct formation (v_{on}) and dissociation (v_{off}) plotted against total butanal concentration. Errors are standard deviations of rates at each concentration. Right: Single-channel recordings with 0.0, 3.0, 5.8, 8.5 and 11 mM total butanal (trans). Current levels corresponded to the hemithioacetal adducts (2) and the unoccupied nanopore (1).



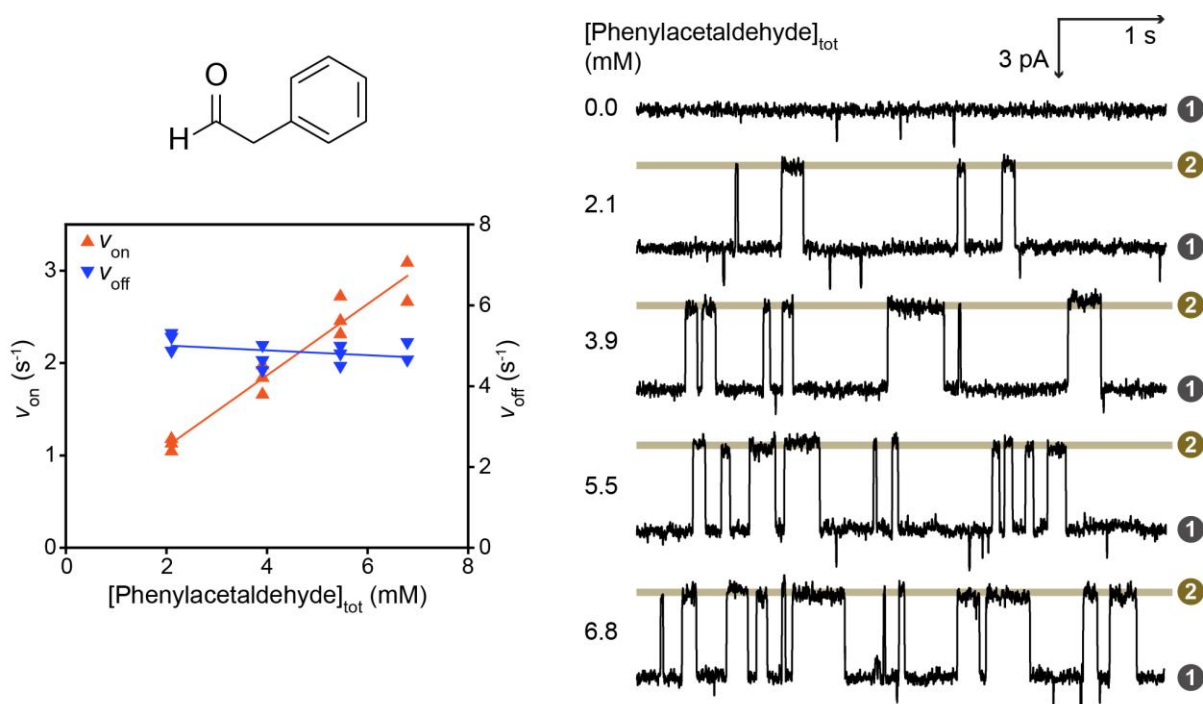
Supplementary Fig. 5: Kinetic analysis of pentanal with (AG)₆(AG-T115C). Left: Rates of adduct formation (v_{on}) and dissociation (v_{off}) plotted against total pentanal concentration. Errors are standard deviations of rates at each concentration. Right: Single-channel recordings with 0.0, 2.0, 3.8, 5.5 and 7.1 mM total pentanal (trans). Current levels corresponded to the hemithioacetal adducts (2) and the unoccupied nanopore (1).



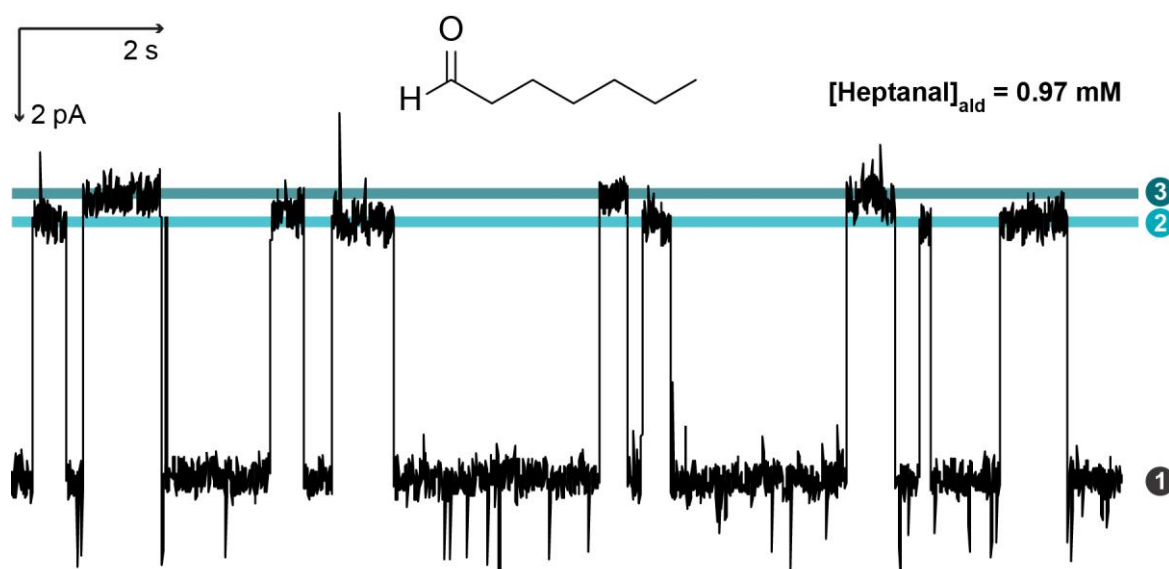
Supplementary Fig. 6: Kinetic analysis of 2-methylpropanal with (AG)₆(AG-T115C). Left: Rates of adduct formation (v_{on}) and dissociation (v_{off}) plotted against total 2-methylpropanal concentration. Errors are standard deviations of rates at each concentration. Right: Single-channel recordings with 0.0, 2.3, 4.3, 6.1 and 7.7 mM total 2-methylpropanal (trans). Current levels corresponded to the hemithioacetal adducts (2) and the unoccupied nanopore (1).



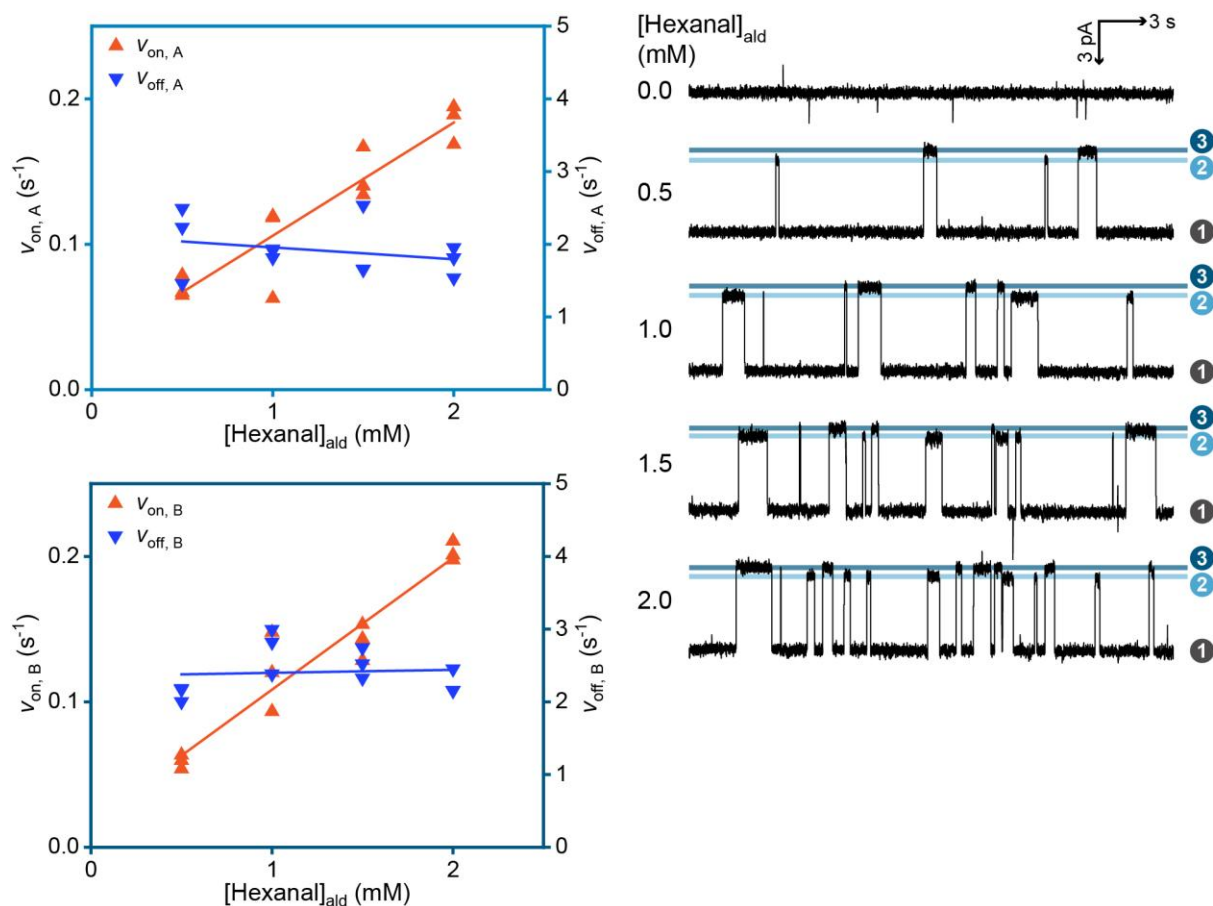
Supplementary Fig. 7: Kinetic analysis of benzaldehyde with (AG)₆(AG-T115C). Left: Rates of adduct formation (v_{on}) and dissociation (v_{off}) plotted against total benzaldehyde concentration. Errors are standard deviations of rates at each concentration. Right: Single-channel recordings with 0.0, 2.0, 3.9, 5.7 and 7.5 mM total benzaldehyde (trans). Current levels corresponded to the hemithioacetal adducts (2) and the unoccupied nanopore (1).



Supplementary Fig. 8: Kinetic analysis of phenylacetaldehyde with (AG)₆(AG-T115C). Left: Rates of adduct formation (v_{on}) and dissociation (v_{off}) plotted against total phenylacetaldehyde concentration. Errors are standard deviations of rates at each concentration. Right: Single-channel recordings with 0.0, 2.1, 3.9, 5.5 and 6.8 mM total phenylacetaldehyde (trans). Current levels corresponded to the hemithioacetal adducts (2) and the unoccupied nanopore (1).



Supplementary Fig. 9: Diastereomeric hemithioacetal adducts from heptanal have different conductance levels with (AG)₆(AG-T115C). Single-channel recordings with 0.97 mM unhydrated heptanal (i.e., concentrations determined from ¹H NMR, Supplementary Section 4). Current levels corresponded to the two possible diastereomeric adducts, diastereomer A (2) and diastereomer B (3) and the unoccupied nanopore (1).



Supplementary Fig. 10: Kinetic analysis of hexanal with (MK)₆(MK-T115C). Left: Rates of adduct formation (v_{on}) and dissociation (v_{off}) for diastereomeric adducts A and B plotted against the concentration of unhydrated hexanal (determined from 1H NMR, Supplementary Section 4). Errors are standard deviations of rates at each concentration. Right: Single-channel recordings with 0.0, 0.5, 1.0, 1.5 and 2.0 mM unhydrated hexanal (trans). Current levels corresponded to the two possible diastereomeric adducts, diastereomer A (2) and diastereomer B (3) and the unoccupied nanopore (1).

3 Kinetic rate constants for the reversible aldehyde-thiol chemistry and the percentage residual current levels for the hemithioacetal adducts

Rates of hemithioacetal adduct formation (v_{on}) were calculated from the maximum interval likelihood algorithm of QuB (see Experimental methods).⁴

Observed rate constants of adduct formation (k_{on}) were calculated from plots of rates of hemithioacetal adduct formation (v_{on}) against the total aldehyde concentration ($[aldehyde]_{tot}$, uncorrected for aldehyde correction, see Supplementary Section 4) being of the form:

$$v_{on} = k_{on} [aldehyde]_{tot}$$

Supplementary Table 1: Observed rate constants of formation (k_{on}) and dissociation (k_{off}), and percentage residual current levels ($I_{res\%}$) of hemithioacetal adducts.^[a,b,c]

Aldehyde		k_{on} (mM ⁻¹ s ⁻¹)	k_{off} (s ⁻¹)	$I_{res\%}$
Ethanal		0.22 ± 0.01	9.4 ± 0.7	99.4 ± 0.1
Propanal		0.25 ± 0.02	9.2 ± 0.6	98.7 ± 0.1
Butanal		0.27 ± 0.02	7.3 ± 0.4	97.9 ± 0.1
Pentanal		0.27 ± 0.02	7.2 ± 0.4	96.9 ± 0.1
Hexanal		0.22 ± 0.01	5.2 ± 0.1	96.1 ± 0.1
Heptanal	(A)	0.10 ± 0.02	3.6 ± 0.3	95.7 ± 0.1
	(B)	0.09 ± 0.02	5.8 ± 0.6	95.4 ± 0.1
Octanal ^[d]	(A)	-	4.0 ± 0.3	95.1 ± 0.1
	(B)	-	5.7 ± 1.0	94.7 ± 0.1
2-Methylpropanal		0.27 ± 0.01	6.5 ± 0.1	98.1 ± 0.1
Benzaldehyde		0.21 ± 0.02	38 ± 2	97.5 ± 0.1
Phenylacetaldehyde		0.39 ± 0.03	4.8 ± 0.1	96.6 ± 0.1

^[a] Observed rate constants have not been corrected to account for aldehyde hydration in solution.

^[b] Conditions: (AG)₆(AG-T115C) with 2 M KCl, 20 μM EDTA, 200 mM PIPES (pH 6.8), recorded at -50 mV (trans) and 22 ± 1 °C.

^[c] Errors are standard deviations of the mean values from N >3 nanopores, excluding octanal (N = 2 pores).

^[d] The low solubility of octanal precluded the accurate determination of adduct formation rate constants.

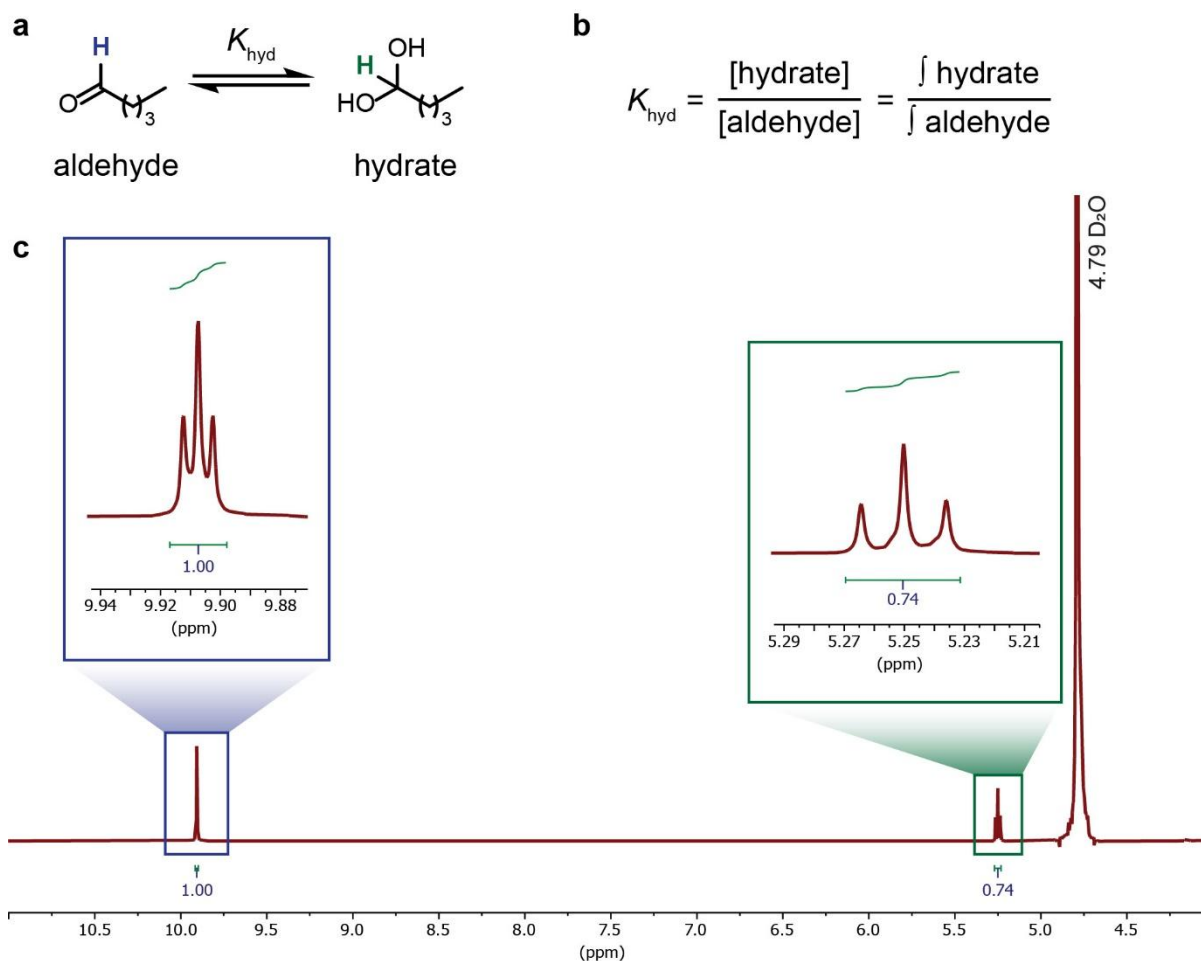
The lower activation energies (higher k_{off} values) for the dissociation of the hemithioacetals formed from benzaldehydes can be attributed to extra resonance stabilisation in the transition state arising from conjugation of the incipient aldehyde group with the aromatic phenyl ring. Higher k_{off} values for benzaldehyde, relative to aliphatic aldehydes, were also observed in the (AG)₆(AG-G137C-Ala3) nanopore (i.e., $k_{off} = 6.4 \pm 0.1$ s⁻¹ and 21 ± 1 s⁻¹ for benzaldehyde (N = 3 pores), 0.82 ± 0.07 s⁻¹ and 1.8 ± 0.2 s⁻¹ for butanal (N = 3 pores), 0.86 ± 0.07 s⁻¹ and 2.5 ± 0.4 s⁻¹ for 2-methylpropanal (N = 3 pores), Supplementary Table 6).

We note that disulfides, if present in the breath, will react with thiols. Because the disulfide reaction is irreversible, aldehydes and disulfides are easily distinguished. Nanopores that have reacted with disulfides would no longer detect aldehydes, and would thereafter be disregarded in a parallel recording device.

4 Hydration constants and solubilities of aldehyde substrates

Hydration constants of aldehyde substrates were determined by using ^1H NMR. Solutions of ethanal, propanal, butanal, pentanal, 2-methylpropanal, benzaldehyde, and phenylacetaldehyde were prepared as follows. The aldehyde (50 mM for ethanal, propanal and butanal; 25 mM for pentanal and benzaldehyde; 22.5 mM for 2-methylpropanal; 15 mM for phenylacetaldehyde) was dissolved in a D_2O solution of 200 mM PIPES and 2 M KCl at pH 6.4 (corresponding to a pD of 6.8),⁵ and left for 10 minutes to equilibrate before ^1H NMR measurement. Hydration constants were calculated from the areas under the characteristic peaks of the aldehydes (J_{aldehyde} , aldehydic proton colored blue in Supplementary Fig. 11) and those of the corresponding hydrates (J_{hydrate} , gem-diol proton colored green in Supplementary Fig. 11). The aldehyde protons were found between chemical shifts δ 9.5-10.5, while the hydrate gem-diol protons were found between chemical shifts δ 4.5-5.5. An example spectrum for pentanal is shown (Supplementary Fig. 11).

For hexanal, heptanal and octanal, solutions were prepared as follows. In a D_2O solution containing 200 mM PIPES and 2 M KCl at pH 6.4 (corresponding to a pD of 6.8), 5 mM calcium formate and 5 mM maleic acid were added as internal standards. To 1 mL of the D_2O solution, 200 μL of aldehyde was added. The undissolved aldehyde formed a layer above the denser D_2O layer. The biphasic mixture was stirred vigorously overnight for the D_2O solution to become saturated with aldehyde. The saturated D_2O layer was then used for ^1H NMR measurement. Hydration constants were calculated from the areas under the characteristic peaks of aldehydes and those of the corresponding hydrates (Supplementary Table 2). Hydration constants for octanal were not determined due to the small area of the hydrate peak precluding accurate calculations. Concentrations of the unhydrated aldehydes in the saturated D_2O layers were determined by comparing the areas of aldehyde signals to the areas of the internal standards (Supplementary Table 3).



Supplementary Fig. 11: ^1H NMR of a pentanal solution. **a**, Pentanal is hydrated in aqueous solution, reducing the concentration of free aldehyde available for hemithioacetal formation. **b**, The hydration equilibrium constant K_{hyd} is defined as the concentration of hydrate over the concentration of aldehyde. K_{hyd} can be calculated from the ratio of the integration of the peaks for the hydrate gem-diol proton (\int_{hydrate} , gem-diol proton colored green) and that of the peaks for the aldehyde proton (\int_{aldehyde} , aldehydic proton colored blue). **c**, ^1H NMR of 25 mM pentanal in a D_2O solution containing 2 M KCl and 200 mM PIPES, adjusted to pH 6.4. Integrations of the peaks for the aldehyde proton (boxed in blue) and the peaks for the gem-diol proton (boxed in green) were used to calculate K_{hyd} .

Supplementary Table 2: Hydration constants of aldehyde analytes derived from NMR studies.

Aldehyde	\int_{aldehyde}	\int_{hydrate}	K_{hyd} ^[a]
Ethanal	1.00	1.30	1.30
Propanal	1.00	1.18	1.18
Butanal	1.00	0.87	0.87
Pentanal	1.00	0.74	0.74
Hexanal	1.00	0.72	0.72
Heptanal	1.00	0.61	0.61
2-Methylpropanal	1.00	0.79	0.79
Benzaldehyde	1.00	0	0.00
Phenylacetaldehyde	1.00	3.02	3.02

^[a] K_{hyd} was calculated using $\int_{\text{hydrate}}/\int_{\text{aldehyde}}$.

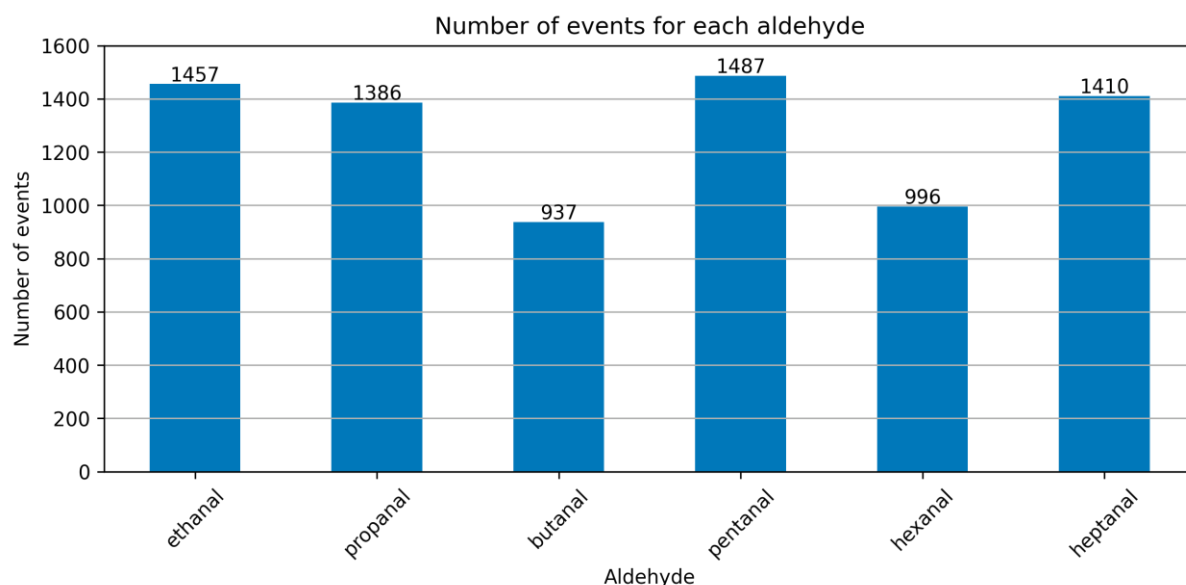
Supplementary Table 3: Concentrations of unhydrated aldehyde in saturated solutions determined by ^1H NMR.

Aldehyde	\int_{aldehyde}	$\int_{\text{calcium formate}}$	$\int_{\text{maleic acid}}$	Saturation concentration (mM) ^[a]
Hexanal	1.00	0.52	0.53	19
Heptanal	1.00	2.23	2.44	4.3
Octanal	1.00	6.74	8.74	1.3

^[a] Saturation concentrations of each aldehyde were calculated by:
 $0.5 \times (2 \times \int_{\text{aldehyde}}/\int_{\text{calcium formate}} + 2 \times \int_{\text{aldehyde}}/\int_{\text{maleic acid}}) \times 5 \text{ mM}$

5 Machine learning

Events were detected using the single-channel-search function in Clampfit 10.7. Using the start and end times for each event, the event features were extracted using an in-house python script. The dataset was curated from individual events from separate traces of each aldehyde. A total of 9732 events were collected with >900 events per aldehyde (Supplementary Fig. 12).

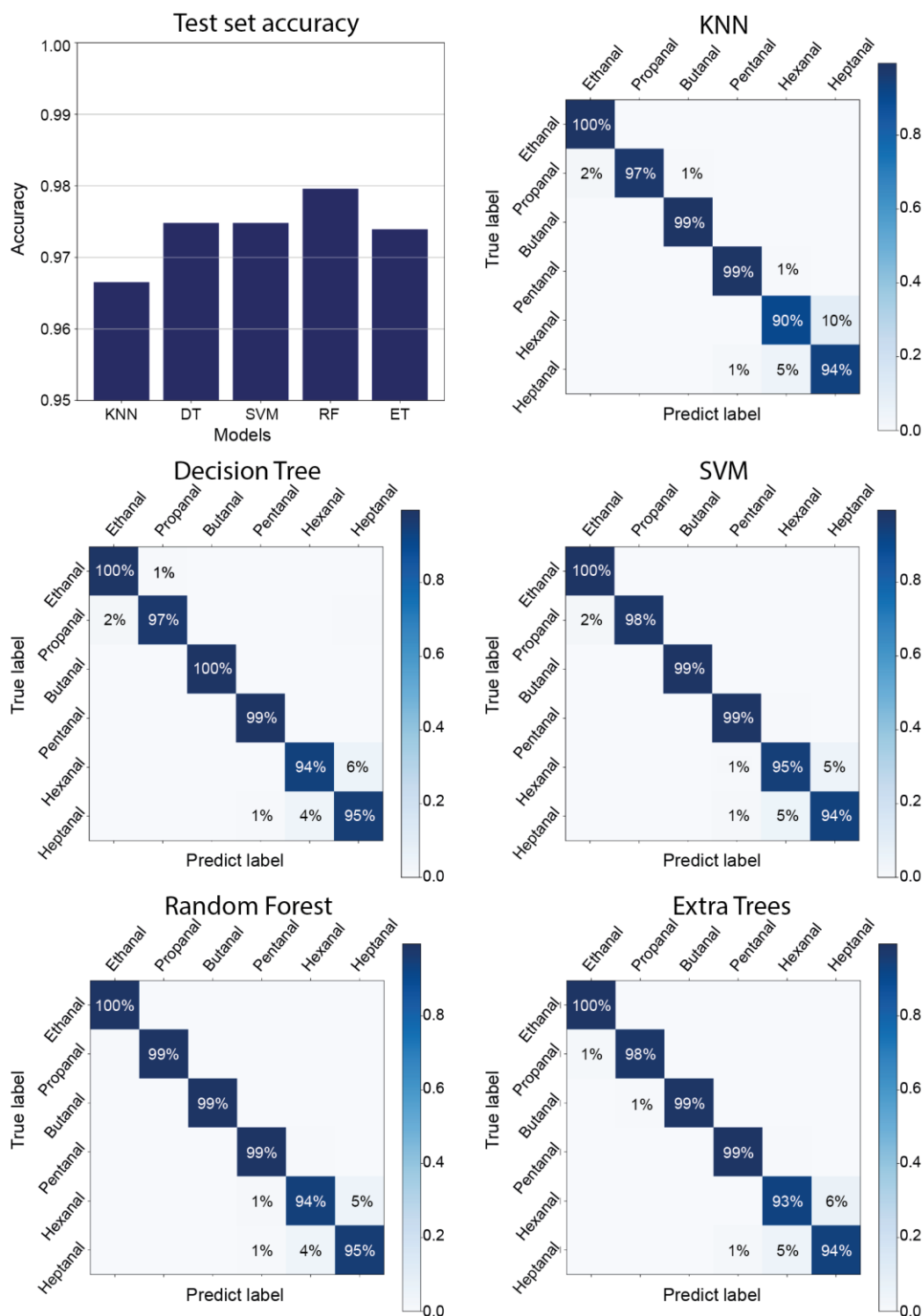


Supplementary Fig. 12: Events collected for each aldehyde for dataset.

Three features were extracted for each event, namely the percentage residual current ($I_{res}\%$), dwell time, and noise. $I_{res}\%$ of each event was calculated as the event's residual current as a percentage of the preceding open pore current ($I_{res}\% = I_{res}/I_P \times 100\%$). The event dwell time was calculated as the difference between the start and end times of each event. The event noise was calculated as the root mean squared error of the current amplitudes of each event.

A random split stratified on aldehyde labels was performed; 70% of the events were used as training data and the other 30% of events were saved as a test set to test the generality of the model. Five classifiers were tested, including k-nearest neighbours (KNN), decision tree, support vector machine (SVM), random forest, and extra trees (Supplementary Fig. 13). Accuracies were calculated as the ratio of correctly classified event over all events in the dataset. Based on 10-fold cross validation, the random forest model achieved the highest accuracy of 98%. We tested this model on the test set, which also yielded an accuracy of 98%. Source data and codes needed to reproduce the results can be found at:

https://github.com/allybo/aldehyde_nanopore_sensing.



Supplementary Fig. 13: Model performance of the five classifiers on the test set. The classifiers tested were k-nearest neighbours (KNN), decision tree, support vector machine (SVM), random forest, and extra trees. Accuracies were calculated as the ratio of correctly classified datapoints over all datapoints in the dataset.

6 Ratiometric detection of aldehyde mixtures

Ratiometric detection of pentanal and butanal in mixtures of different ratios was conducted (Fig. 3b,c). Ratios of aldehydes were altered mid-experiment by removing pentanal and adding butanal. Butanal:pentanal ratios of 1:1, 2:1 and 4:1 were obtained as follows:

1. In the initial single-channel recording, 2 mM of pentanal and 2 mM butanal were present in 500 μ L of trans recording buffer (2 M KCl, 200 mM PIPES (pH 6.8), 20 μ M EDTA).
This is a butanal:pentanal ratio of 1:1.
2. Subsequently, 37 μ L of trans recording buffer was removed and replaced with 37 μ L of 25 mM butanal (in 2 M KCl, 200 mM PIPES (pH 6.8), 20 μ M EDTA). The resulting solution in the trans chamber then had concentrations of 1.9 mM pentanal and 3.7 mM butanal.
This is a butanal:pentanal ratio of 2:1.
3. Finally, 65 μ L of trans recording buffer was removed and replaced with 65 μ L of 25 mM butanal. The resultant solution in the trans chamber then had concentrations of 1.6 mM pentanal and 6.5 mM butanal.
This is a butanal:pentanal ratio of 4:1.

Rates (v_{on}) of hemithioacetal adduct formation for each aldehyde in a mixture were first calculated from the maximum interval likelihood algorithm of QuB using the ratios of adduct formation events (see Experimental methods).² As hemithioacetal adduct formation is a bimolecular process, concentrations of aldehyde were calculated by dividing the rates of adduct formation with their respective observed rate constants (k_{on}) (Supplementary Table 1):

$$\text{Measured [aldehyde]}_{\text{tot}} = v_{on} / k_{on}$$

Measured concentrations were in good agreement with expected concentrations (i.e., <10% difference between measured and expected concentrations) (Fig. 3c). For disease diagnosis and product quality control applications, a more practical, and equally informative approach, is to directly use the ratios of v_{on} as fingerprints of relative concentrations of aldehydes in samples. For example, for aldehydes A and B, v_{on}^A / v_{on}^B is linearly proportional to $[A]_{\text{tot}} / [B]_{\text{tot}}$.

The time needed for accurate ratiometric measurements is dependent on the total concentration of the aldehydes present, as well as the ratio of aldehydes. We found that approximately 100 events per aldehyde was sufficient to provide accurate ratiometric measurements (i.e., measured concentration within <10% of the expected concentration). This was achieved with a recording time of <10 minutes for aldehyde concentrations above 1.5 mM. However, a longer time would be needed for aldehyde mixtures at lower concentrations, or if the ratio of aldehydes were higher (e.g., 10:1). These issues could be addressed by improving the limits of detection as discussed in the Conclusions.

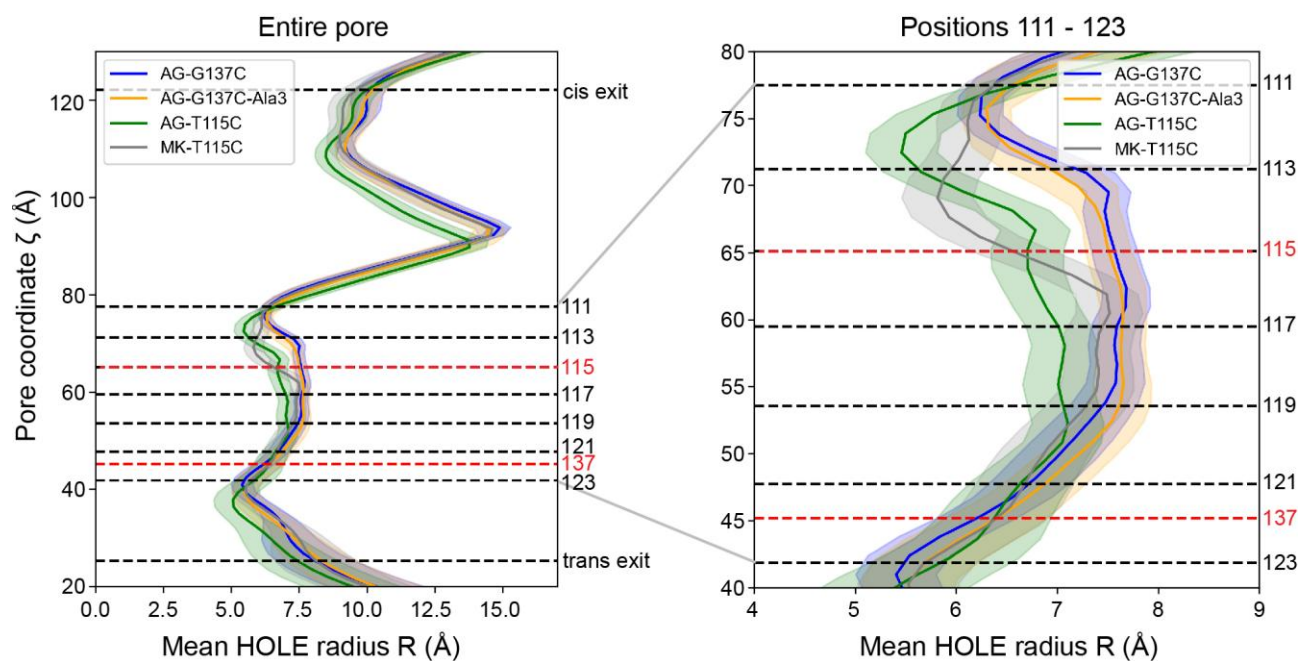
7 Rationale for improved diastereomer discrimination in different nanopores

Nanopores with a cysteine at the position 137 produced larger differences in $I_{res}\%$ between structural isomers and diastereomers (compared to nanopores with a cysteine at the 115 position). The (MK)₆(MK-T115C) nanopore produced slightly larger differences in $I_{res}\%$ between diastereomers than the (AG)₆(AG-T115C) nanopore. We performed molecular dynamics (MD) simulations of the (AG)₆(AG-T115C), (MK)₆(MK-T115C), (AG)₆(AG-G137C) and (AG)₆(AG-G137C-Ala3) nanopores to explore the improved diastereomer discrimination in these nanopores.

MD simulations were performed with GROMACS (v2019.2).⁶ For each system, the nanopore was modeled with the AMBER99SB-ILDN force field and inserted into a cubic box of TIP3P water molecules.^{7,8} The system was neutralized and 0.15 M NaCl was added. Energy minimization was performed with the steepest descent algorithm until the maximum force was below 1000 kJ mol⁻¹ nm⁻¹. A positional restraint of 1000 kJ mol⁻¹ nm² was applied on the protein backbone atoms. Three independent repeats were performed with random initial velocities generated according to the Maxwell–Boltzmann distribution at 298 K. The system was equilibrated under a constant volume and temperature (NVT) ensemble at 298 K (100 ps, 2 fs timestep). This was followed by an equilibration under a constant pressure and temperature (NPT) ensemble at 1 bar and 298 K (100 ps, 2 fs timestep). Production simulations were performed in the NPT ensemble at 1 bar and 298 K (200 ns, 2 fs timestep). The temperature of the system was maintained at 298 K using the V-rescale thermostat.⁹ The pressure was controlled by the Parrinello-Rahman barostat at 1.0 bar, with an isothermal compressibility of 4.5×10^{-5} bar⁻¹.¹⁰ All simulations were performed with three-dimensional periodic boundary conditions. Long-range electrostatics was described with the Particle Mesh Ewald (PME) algorithm.^{11,12} All bond lengths involving hydrogen atoms were constrained using the LINCS algorithm.¹³ All three replicas of 50 ns MD trajectories for a given mutant were concatenated.

We assessed the internal pore radius of each mutant using the HOLE program with the default parameters.^{14,15} The narrowest constriction in the wild-type α HL nanopore is formed from the residues of Glu-111, Met-113 and Lys-147. After mutations at these positions (i.e., M113G and K147G) in the (AG)₆(AG-T115C), (AG)₆(AG-G137C) and (AG)₆(AG-G137C-Ala3) nanopore mutants, we observed the removal of the original constriction. Instead, the narrowest part of the (AG)₆(AG-T115C), (AG)₆(AG-G137C) and (AG)₆(AG-G137C-Ala3) nanopore is formed by a ring of leucine residues at position 135. Improved diastereomer discrimination at the 137 position is hypothesized to arise from the location's proximity to the narrowest constriction site (Supplementary Fig. 14).

In the (MK)₆(MK-T115C) nanopore, the residue at the 113 position is reverted back to methionine. The internal diameter of the nanopore close to the 115 position is reduced, resulting in the slightly improved diastereomer discrimination (Supplementary Fig. 14).



Supplementary Fig. 14: Model of the internal radius of the nanopore. Profile of the pore radius along the $(AG)_6(AG-T115C)$, $(MK)_6(MK-T115C)$, $(AG)_6(AG-G137C)$ and $(AG)_6(AG-G137C-Ala3)$ nanopores along the z -coordinate. For each mutant, the solid line represents the mean radius, and the shaded region represents the standard deviation.

8 Improved structural resolution in engineered nanopore mutants

Two additional nanopores were designed to address the challenge in distinguishing between structural isomers butanal and 2-methylpropanal. Using the (MK)₆(MK-T115C) and (AG)₆(AG-G137C) nanopores, hemithioacetal adducts from ethanal to octanal were easily separable (Fig. 4 and Supplementary Tables 4-5).

Supplementary Table 4: Percentage residual current of hemithioacetal adducts of straight-chain aldehydes in (MK)₆(MK-T115C) nanopore.^[a,b]

Aldehyde		I _{res} %
Ethanal	(A)	99.5 ± 0.1
	(B)	99.2 ± 0.1
Propanal	(A)	98.9 ± 0.1
	(B)	98.5 ± 0.1
2-Methylpropanal	(A)	98.0 ± 0.1
	(B)	97.8 ± 0.1
Butanal	(A)	98.0 ± 0.1
	(B)	97.6 ± 0.1
Pentanal	(A)	96.8 ± 0.1
	(B)	96.5 ± 0.1
Hexanal	(A)	95.7 ± 0.1
	(B)	95.1 ± 0.1
Heptanal	(A)	94.7 ± 0.1
	(B)	93.7 ± 0.1
Octanal	(A)	93.2 ± 0.1
	(B)	92.4 ± 0.1

^[a] Conditions: (MK)₆(MK-T115C) with 2 M KCl, 20 μM EDTA, 200 mM PIPES (pH 6.8), recorded at -50 mV (trans) and 22 ± 1 °C.

^[b] Errors are standard deviations of the mean values from N > 2 nanopores.

Supplementary Table 5: Percentage residual current of hemithioacetal adducts of straight-chain aldehydes in (AG)₆(AG-G137C) nanopore.^[a,b]

Aldehyde		I _{res} %
Ethanal	(A)	98.6 ± 0.1
	(B)	98.2 ± 0.1
Propanal	(A)	97.9 ± 0.1
	(B)	97.2 ± 0.1
2-Methylpropanal	(A)	96.8 ± 0.1
	(B)	96.2 ± 0.1
Butanal	(A)	96.8 ± 0.1
	(B)	95.9 ± 0.1
Pentanal	(A)	95.8 ± 0.1
	(B)	94.9 ± 0.1
Hexanal	(A)	95.5 ± 0.1
	(B)	93.8 ± 0.1
Heptanal	(A)	95.1 ± 0.1
	(B)	93.5 ± 0.1
Octanal	(A)	94.0 ± 0.2
	(B)	92.6 ± 0.1

^[a] Conditions: (AG)₆(AG-G137C) with 2 M KCl, 20 μM EDTA, 200 mM PIPES (pH 6.8), recorded at -50 mV (trans) and 22 ± 1 °C.

^[b] Errors are standard deviations of the mean values from N = 3 nanopores, excluding octanal (N = 2 pores).

Using the (AG)₆(AG-G137C) nanopore, butanal and 2-methylpropanal can be distinguished using diastereomers B (i.e., $\Delta I_{res}\%$ = 0.25 ± 0.02 % for diastereomers B of butanal and 2-methylpropanal, N = 3 pores, >35 events for each aldehyde) (Fig. 4f).

Further engineering was performed on the (AG)₆(AG-G137C) nanopore to give the (AG)₆(AG-G137C-Ala3) nanopore (i.e., AG-G137C-Ala3 = WT-K8A-M113G-N121A-N123A-K131G-G137C-N139A-K147G). In the (AG)₆(AG-G137C-Ala3) nanopore, butanal and 2-methylpropanal could be distinguished using diastereomers A (i.e., $\Delta I_{res}\%$ = 0.42 ± 0.03 % for diastereomers A of butanal and 2-methylpropanal, N = 4 pores, >60 events for each aldehyde) (Fig. 4f and Supplementary Table 6).

Supplementary Table 6: Observed rate constants of formation (k_{on}) and dissociation (k_{off}) of hemithioacetal adducts for select aldehydes in (AG)₆(AG-G137C-Ala3).^[a,b]

Aldehyde		k_{on} (mM ⁻¹ s ⁻¹)	k_{off} (s ⁻¹)	I _{res} %
Butanal	(A)	0.052 ± 0.008	0.82 ± 0.07	98.6 ± 0.1
	(B)	0.030 ± 0.002	1.8 ± 0.2	97.2 ± 0.2
2-Methylpropanal	(A)	0.016 ± 0.002	0.86 ± 0.07	98.2 ± 0.1
	(B)	0.018 ± 0.002	2.5 ± 0.4	97.1 ± 0.1
Benzaldehyde	(A)	0.22 ± 0.01	6.4 ± 0.1	98.0 ± 0.1
	(B)	0.088 ± 0.004	21 ± 1	96.3 ± 0.1

^[a] Conditions: (AG)₆(AG-G137C-Ala3) with 2 M KCl, 20 μM EDTA, 200 mM PIPES (pH 6.8), recorded at -50 mV (trans) and 22 ± 1 °C.

^[b] Errors are standard deviations of the mean values from N = 3 nanopores.

9 Experimental methods

General

All reagents were purchased from Sigma-Aldrich and used without further purification. Catalogue numbers together with purities of aldehydes are listed as follows: Ethanal ($\geq 99.5\%$ purity, Sigma Aldrich catalogue no. 402788); propanal ($\geq 98.0\%$ purity, Sigma Aldrich catalogue no. 64409); butanal ($\geq 99.5\%$ purity, Sigma Aldrich catalogue no. 418102); pentanal ($\geq 98.0\%$ purity, Sigma Aldrich catalogue no. 8.08504); hexanal ($\geq 95.0\%$ purity, Sigma Aldrich catalogue no. 18109); heptanal ($\geq 97.0\%$ purity, Sigma Aldrich catalogue no. 61696); octanal ($\geq 98.0\%$ purity, Sigma Aldrich catalogue no. 52466); 2-methylpropanal ($\geq 99.0\%$ purity, Sigma Aldrich catalogue no. 240788); benzaldehyde ($\geq 99.0\%$ purity, Sigma Aldrich catalogue no. B1334); phenylacetaldehyde ($\geq 95.0\%$ purity, Sigma Aldrich catalogue no. W287407). 1,2-Diphytanoyl-sn-glycerol-3-phosphocholine (DPhPC) used for single-channel electrical recordings was purchased from Avanti Polar Lipids.

Plasmid preparation

The construction of plasmids encoding the AG (α HL-K8A-M113G-K131G-K147G) and AG-D8 (α HL-K8A-M113G-K131G-K147G-D8) monomers were previously described.¹

Plasmids encoding the AG-T115C (α HL-K8A-M113G-T115C-K131G-K147G-D8) monomer were prepared from AG-D8 plasmids by a single round of site-directed mutagenesis (Agilent QuikChange II XL Site-Directed Mutagenesis Kit), as follows. Polymerase chain reactions (PCR) (25 μ L) were set up by mixing the following reagents in the order listed: 10 \times reaction buffer (2.5 μ L), nuclease-free water (to make up the final volume of the reaction to 25 μ L), double-stranded plasmid DNA template (AG-D8 plasmid, 5 μ L, 1 ng/ μ L), mutagenic primers (sense primer: 5' GAGTATGGGAGTIGTTTAACTTATGGATTCAACGGTAATG 3'; antisense primer: 5' CATTACCGTTGAATCCATAAGTTAAACAACTCCCATACTC 3'; 62.5 ng of each primer, target mutations are underlined and highlighted in red), dNTP mix (0.5 μ L), Quik Solution (1.5 μ L) and Pfu Ultra HF DNA polymerase (0.5 μ L, 2.5 U/ μ L). The PCRs were carried out with the following program: 95 $^{\circ}$ C for 5 min, 18 cycles of: 95 $^{\circ}$ C (50 s), 53 $^{\circ}$ C (50 s), 68 $^{\circ}$ C (5 min), followed by a final extension at 68 $^{\circ}$ C for 7 min. The PCRs were then cooled on ice for 2 min and the template was digested with DpnI (0.5 μ L, supplied with the kit) at 37 $^{\circ}$ C for 1 h. Plasmids containing the mutant genes were produced from *Escherichia coli* (*E. coli*) XL-10 Gold ultracompetent cells transformed with the PCR product. The DNA sequences of the genes were verified (Source BioScience).

A plasmid encoding the AG-G137C (α HL-K8A-M113G-K131G-G137C-K147G-D8) monomer was prepared from AG-D8 plasmids using a single round of site-directed mutagenesis (Agilent QuikChange II XL Site-Directed Mutagenesis Kit) as follows. A PCR (25 μ L) was set up by mixing the following reagents in the order listed: 10 \times reaction buffer (2.5 μ L), nuclease-free water (to make up the final reaction volume to 25 μ L), double-stranded plasmid DNA template (AG-D8 plasmid, 5 μ L, 1 ng/ μ L), mutagenic primers (sense primer: 5' GAATTGGCGGCCTTATTIGTGCAAATGTTTC 3'; antisense primer: 5' GAAACATTTGCACAAATAAGGCCGCCAATTC 3'; 62.5 ng of each primer, target mutations are underlined and highlighted in red), dNTP mix (0.5 μ L), QuikSolution (1.5 μ L) and PfuUltra HF DNA polymerase (0.5 μ L, 2.5 U/ μ L). PCR

was carried out with the following program: 95 °C for 5 min, 18 cycles of 95 °C (50 s), 60 °C (50 s), 68 °C (5 min), followed by a final extension at 68 °C for 7 min. The PCR was then cooled on ice for 2 min and then the template DNA was digested with DpnI (0.5 µL, supplied with the kit) at 37 °C for 1 h. Plasmids containing the mutant genes were produced from *E. coli* XL-10 Gold ultracompetent cells transformed with the PCR products. DNA sequences of the genes were verified (Source BioScience).

A plasmid encoding the AG-G137C-Ala3 (α HL-K8A-M113G-N121A-N123A-K131G-G137C-N139A-K147G-D8) monomer was prepared from AG-D8 plasmids by four successive rounds of site-directed mutagenesis (Agilent QuikChange II XL Site-Directed Mutagenesis Kit), as follows. Each PCR (25 µL) was set up by mixing the following reagents in the order listed: 10× reaction buffer (2.5 µL), nuclease-free water (to make up the final reaction volume to 25 µL), double-stranded plasmid DNA template (5 µL, 1 ng/µL), mutagenic primers (62.5 ng of each primer, see below), dNTP mix (0.5 µL), QuikSolution (1.5 µL) and PfuUltra HF DNA polymerase (0.5 µL, 2.5 U/µL). PCRs were carried out with the following program: 95 °C for 5 min, 18 cycles of 95 °C (50 s), 60 °C (50 s for all mutants), 68 °C (5 min), followed by a final extension at 68 °C for 7 min. The PCRs were then cooled on ice for 2 min and then the template DNA was digested with DpnI (0.5 µL, supplied with the kit) at 37 °C for 1 h. Plasmids containing the mutant genes were produced from *E. coli* XL-10 Gold ultracompetent cells transformed with the PCR products. DNA sequences of the genes were verified (Source BioScience).

Mutants, primers and DNA templates for each site-directed mutagenesis reaction are shown below. Codons underlined and highlighted in red represent the target mutations in each site-directed mutagenesis PCR.

- **α HL-K8A-M113G-K131G-G137C-K147G-D8** (DNA template: AG-D8; sense primer: 5' GAATTGGCGGCCTTATTIGTGCAAATGTTTC 3'; antisense primer: 5' GAAACATTTGCACAAATAAGGCCGCCAATTC 3')
- **α HL-K8A-M113G-N121A-K131G-G137C-K147G-D8** (DNA template: α HL-K8A-M113G-K131G-G137C-K147G-D8; sense primer: 5' CTTATGGATTCGCCGGTAATGTTACTGGTGATGATACAGG 3'; antisense primer: 5' CCTGTATCATCACCAGTAACATTACCGGCGAATCCATAAG 3'),
- **α HL-K8A-M113G-N121A-K131G-G137C-N139A-K147G-D8** (DNA template: α HL-K8A-M113G-N121A-K131G-G137C-K147G-D8; sense primer: 5' GCGGCCTTATTTGTGCAGCTGTTTCGATTGGTCATAC 3'; antisense primer: 5' GTATGACCAATCGAAACAGCTGCACAAATAAGGCCGC 3'), and
- **AG-G137C-Ala3** (DNA template: α HL-K8A-M113G-N121A-K131G-G137C-N139A-K147G-D8; sense primer: 5' GGATTCGCCGGTGCTGTTACTGGTGATGATACAGG 3'; antisense primer: 5' CCTGTATCATCACCAGTAACAGCACCGGCGAATCC 3')

Plasmids encoding the MK monomer were prepared from AG plasmids using two parallel rounds of site-directed mutagenesis (Phusion), followed by homologous recombination. The first PCR (25 µL) was set up by mixing the following reagents in the order listed: DMSO (0.75 µL), double-stranded plasmid DNA template (AG, 1 µL, 1 ng/µL), sense primer 5' GTTTAACTTTAAGAAGGAGATATACATATGGCAGATTCTGATATTAATATTGC 3' (1.25 µL, 10 µM) and antisense primer 5' CCGTTGAATCCGTACGTTAACGTACTCATATACTCTTTTGTATCAATCGAATTCC

G 3' (1.25 μ L, 10 μ M) (target mutations in primers are underlined and highlighted in red), nuclease-free water (8.25 μ L), 2 \times Phusion Master Mix (12.5 μ L). The PCR was carried out with the following program: 98 $^{\circ}$ C for 30 s, 25 cycles of: 98 $^{\circ}$ C (10 s), 45 $^{\circ}$ C (30 s), 72 $^{\circ}$ C (30 s), followed by a final extension at 72 $^{\circ}$ C for 7 min. The second PCR (25 μ L) was set up by mixing the following reagents in the order listed: DMSO (0.75 μ L), double-stranded plasmid DNA template (AG, 1 μ L, 1 ng/ μ L), sense primer 5'

CGGAATTCGATTGATACAAAAGAGTATATGAGTACGTTAACGTACGGATTCAAC
GG 3' (1.25 μ L, 10 μ M) and antisense primer 5'
GCCATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTC 3' (1.25 μ L, 10 μ M)
(target mutations in primers are underlined and highlighted in red), nuclease-free water (8.25 μ L), 2 \times Phusion Master Mix (12.5 μ L). The PCR was carried out with the following program: 98 $^{\circ}$ C for 30 s, 25 cycles of: 98 $^{\circ}$ C (10 s), 45 $^{\circ}$ C (30 s), 72 $^{\circ}$ C (3 min), followed by a final extension at 72 $^{\circ}$ C for 7 min. PCR mixtures were then cooled on ice for 2 min and then the template DNA was digested with DpnI (1 μ L, 20 U/ μ L, NEB Catalog #R0176S) at 37 $^{\circ}$ C for 1 h. Plasmids containing the mutant genes were produced by transforming *E. coli* XL-10 Gold ultracompetent cells (45 μ L) with 2 μ L of PCR product from each PCR. The DNA sequences of the genes were verified (Source BioScience).

Plasmids encoding the MK-D8 monomer were prepared from AG-D8 plasmids using two rounds of site-directed mutagenesis (Phusion), followed by homologous recombination. The first PCR (25 μ L) was set up by mixing the following reagents in the order listed: DMSO (0.75 μ L), double-stranded plasmid DNA template (AG-D8, 1 μ L, 1 ng/ μ L), sense primer 5'
GTTTAACTTTAAGAAGGAGATATACATATGGCAGATTCTGATATTAATATTGC 3'
(1.25 μ L, 10 μ M) and antisense primer 5'
CCGTTGAATCCGTACGTTAACGTACTCATATACTCTTTTGTATCAATCGAATTCC
G 3' (1.25 μ L, 10 μ M) (target mutations in primers are underlined and highlighted in red), nuclease-free water (8.25 μ L), 2 \times Phusion Master Mix (12.5 μ L). The PCR was carried out with the following program: 98 $^{\circ}$ C for 30 s, 25 cycles of: 98 $^{\circ}$ C (10 s), 45 $^{\circ}$ C (30 s), 72 $^{\circ}$ C (30 s), followed by a final extension at 72 $^{\circ}$ C for 7 min. The second PCR (25 μ L) was set up by mixing the following reagents in the order listed: DMSO (0.75 μ L), double-stranded plasmid DNA template (AG-D8, 1 μ L, 1 ng/ μ L), sense primer 5'

CGGAATTCGATTGATACAAAAGAGTATATGAGTACGTTAACGTACGGATTCAAC
GG 3' (1.25 μ L, 10 μ M) and antisense primer 5'
GCCATATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTC 3' (1.25 μ L, 10 μ M)
(target mutations in primers are underlined and highlighted in red), nuclease-free water (8.25 μ L), 2 \times Phusion Master Mix (12.5 μ L). The PCR was carried out with the following program: 98 $^{\circ}$ C for 30 s, 25 cycles of: 98 $^{\circ}$ C (10 s), 45 $^{\circ}$ C (30 s), 72 $^{\circ}$ C (3 min), followed by a final extension at 72 $^{\circ}$ C for 7 min. PCR mixtures were then cooled on ice for 2 min and then the template DNA was digested with DpnI (1 μ L, 20 U/ μ L, NEB Catalog #R0176S) at 37 $^{\circ}$ C for 1 h. Plasmids containing the mutant genes were produced by transforming *E. coli* XL-10 Gold ultracompetent cells (45 μ L) with 2 μ L of PCR product from each PCR. The DNA sequences of the genes were verified (Source BioScience).

A plasmid encoding the MK-T115C (α HL-K8A-T115C-K131G-K147G-D8; Minus-K) monomer were prepared from MK-D8 plasmids via a single round of site-directed

mutagenesis (Agilent QuikChange II XL Site-Directed Mutagenesis Kit), as follows. Polymerase chain reactions (PCR) (25 μ L) were set up by mixing the following reagents in the order listed: 10 \times reaction buffer (2.5 μ L), nuclease-free water (to make up the final volume of the reaction to 25 μ L), double-stranded plasmid DNA template (MK-D8, 5 μ L, 1 ng/ μ L), mutagenic primers (sense primer: 5' CCAAGAAATTCGATTGATACAAAAGAGTATATGAGTIGTTTAACTTATGGATTCA ACGG 3'; antisense primer: 5' CCGTTGAATCCATAAGTTAAACAACTCATATACTCTTTTGTATCAATCGAATTTCT TGG 3'; 62.5 ng of each primer, target mutations are underlined and highlighted in red), dNTP mix (0.5 μ L), Quik Solution (1.5 μ L) and Pfu Ultra HF DNA polymerase (0.5 μ L, 2.5 U/ μ L). The PCRs were carried out with the following program: 95 $^{\circ}$ C for 1 min, 18 cycles of: 95 $^{\circ}$ C (50 s), 55 $^{\circ}$ C (50 s), 68 $^{\circ}$ C (5 min 30 s), followed by a final extension at 68 $^{\circ}$ C for 7 min. The PCRs were then cooled on ice for 2 min and the template was digested with DpnI (0.5 μ L, supplied with the kit) at 37 $^{\circ}$ C for 1 h. Plasmids containing the mutant genes were produced from *Escherichia coli* (*E. coli*) XL-10 Gold ultracompetent cells transformed with the PCR product. The DNA sequences of the genes were verified (Source BioScience).

Nanopore preparation

(AG)₆(AG-T115C), (MK)₆(MK-T115C), (AG)₆(AG-G137C) and (AG)₆(AG-G137C-Ala3) heteroheptamers were prepared according to the procedure below, which is a modification of a method previously reported.¹⁶

α HL monomers were prepared with an *E. coli in vitro* transcription and translation (IVTT) system (*E. coli* T7 S30 Extract System for Circular DNA, Cat #L1130, Promega). Prior to use, the T7 S30 extract provided in the kit was treated with 1 μ L rifampicin (1 μ g/mL, final concentration) to suppress transcription by *E. coli* RNA polymerase. A standard reaction comprised: DNA plasmid mixture (<4 μ g, plasmids encoding cysteine-free and cysteine-containing subunits were in an 8:1 ratio), amino acid mixture without methionine (5 μ L, as supplied in the kit), S30 premix without amino acids (20 μ L, as supplied in the kit), [³⁵S]methionine (2 μ L, 1200 Ci/mmol, 15 mCi/mL, MP Biomedicals), rabbit red blood cell membranes (2 μ L, ~ 1 mg protein/mL), and T7 S30 extract for circular DNA (15 μ L, as supplied in the kit). The reaction mixture was incubated at 37 $^{\circ}$ C for 2 h.

Next, MBSA buffer (1 mL; 3-morpholinopropane-1-sulfonic acid (10 mM), NaCl (150 mM), bovine serum albumin (1 mg/mL), pH 7.4) was added to the reaction mixture. The mixture was centrifuged at 12000 RCF for 10 min at 4 $^{\circ}$ C. The supernatant was then removed, and the pellet solubilized at room temperature with 2 \times Laemmli sample buffer (25 μ L) which contains 10% 2-mercaptoethanol. Before loading the samples, a 5.5 % SDS/PAGE gel was pre-run with 1 \times Tris-Glycine SDS running buffer containing dithiothreitol (2 mM) and sodium thioglycolate (1 mM). The resuspended pellets were then loaded onto the gel and electrophoresed at +70 V overnight (13 h).

α HL heptamers containing different numbers of mutant subunits were separated in the gel based on their electrophoretic mobilities which were determined by the number of octa-aspartate (D8) tails present (i.e., each cysteine-bearing mutant subunit contained a D8 tail). Hence, the top band corresponded to homoheptamers bearing no octa-aspartate tail (i.e., (AG)₇ or (MK)₇), the second band corresponded to heteroheptamers bearing a single octa-aspartate tail (i.e., (mutant-D8)₁(AG)₆ or (mutant-D8)₁(MK)₆) and

so on, with consecutive bands having an aspartate-tail-free subunit replaced with a mutant subunit bearing an octa-aspartate tail. In this work, the second band from the top contained the desired protein pore with a single cysteine residue.

To extract the protein pores, the gel was dried under vacuum onto Whatman 3MM filter paper for 5 h at 50 °C. The dried gel was then exposed to photographic film (Kodak Bio Max MR autoradiography film) for 8 h and the developed film was used to locate the target protein bands in the gel. The desired protein bands were excised and rehydrated in TE buffer (300 µL; Tris·HCl (10 mM), DTT (0.5 mM), EDTA (1 mM), pH 8.0) for 1 h at room temperature. The paper was then removed, and the gel crushed with a plastic pestle. The resulting suspension was filtered through a 0.2 µm hydrophilic membrane filter (Proteus Mini Clarification Spin Column, Genex). The filtrate was stored in 10 µL aliquots at –80 °C.

***Arthrobacter cholorphenolicus* choline oxidase (AcCO6) preparation**

The plasmid pET28a-AcCO6 encoding an engineered *Arthrobacter cholorphenolicus* choline oxidase (AcCO6) was a kind gift from Nicholas J. Turner.¹⁷ The plasmid was transformed in *E. coli* BL21(DE3) competent cells for expression, which were cultured in 10 mL LB medium containing 50 µg/mL kanamycin at 37 °C, 220 rpm overnight. The overnight culture was transferred to 1 L autoinduction media (LB based, Formedium, #AIMLB0110) containing 50 µg/mL kanamycin and grown at 220 rpm for 16 hours at 20 °C. Cells were harvested by centrifugation at 5000 rpm for 20 min at 4 °C. The pellet was resuspended in 50 mL pre-chilled Buffer A (100 mM KPi pH 7.8, 300 mM KCl, 20 mM imidazole) supplemented with 250 U nucleases (Pierce™ Universal Nuclease for Cell Lysis, #88700). Cells were lysed by sonication (10 s pulse, 10 s pause, 3 min duration) on ice, followed by centrifugation at 40000 g for 30 min at 4 °C. The supernatant was filtered through a 0.45 µm syringe filter and applied for purification using a 5 mL HisTrap™ FF column (Buffer A: 100 mM KPi pH 7.8, 300 mM KCl, 20 mM imidazole; Buffer B: 100 mM KPi pH 7.8, 300 mM KCl, 1,000 mM imidazole). Fractions containing His-tagged AcCO6 were pooled and dialysed against a buffer containing 100 mM KPi pH 7.8 at 4 °C overnight. The His-tagged AcCO6 were concentrated, aliquoted, flash frozen with liquid N₂, and stored at –80 °C.

Differential sensing of aldehydes and alcohols

The control aldehyde-alcohol mixture was prepared as follows: propanal, butanal, 1-pentanol, 1-hexanol (5 mM each) were dissolved in oxygen-saturated KPi buffer (100 mM, pH 7.8). The treated mixture was prepared as follows: To the control aldehyde-alcohol mixture was added AcCO6 (0.1 mg/mL for each mM of alcohol) and allowed to shake overnight at 300 RPM at 30 °C. After incubation, the control or treated mixture was adjusted to pH 6.8, and diluted 4-fold with buffer (2.66 M KCl, 200 mM PIPES at pH 6.8, 20 µM EDTA) (i.e., final concentration of 2 M KCl).

For the enzyme-nanopore coupled experiments, either the control (100 µL) or treated mixture (100 µL) was introduced into the trans chamber. As the control mixture was not treated with AcCO6, only propanal and butanal could be detected. Single-channel recordings with the treated mixture revealed additional events with conductance levels corresponding to hemithioacetal adducts derived from pentanal and hexanal. This confirmed the oxidation of alcohols to aldehydes after AcCO6 treatment.

Single-channel electrical recording

Single-channel recordings were carried out in a planar bilayer apparatus as previously described.¹⁸ A single α HL pore was allowed to insert into the bilayer. Aldehyde substrates were introduced from the trans compartment. Experiments were conducted using recording buffer containing 2 M KCl, 200 mM PIPES and 20 μ M EDTA titrated to pH 6.8. Aldehyde solutions were prepared with recording buffer and titrated to pH 6.8. Single-channel recordings were conducted with a coverslip placed atop the recording chamber.

Ionic currents were recorded by using a patch clamp amplifier (Axopatch 200B, Axon Instruments), and filtered with a low-pass Bessel filter (80 dB/decade) with a corner frequency of 10 kHz. Signals were digitized with a Digidata 1320A digitizer (Molecular Devices) at an acquisition frequency of 50 kHz. The current traces were processed with Clampfit 10.7 (Molecular Devices). Current traces were idealized by using Clampfit 10.7 (Molecular Devices). The idealized data were analyzed with QuB 2.0 software (www.qub.buffalo.edu).¹⁹ Dwell time analysis and rate constant determinations were performed by using the maximum interval likelihood algorithm of QuB.⁴

References

1. Choi, L. S., Mach, T. & Bayley, H. Rates and Stoichiometries of Metal Ion Probes of Cysteine Residues within Ion Channels. *Biophys. J.* **105**, 356-364 (2013).
2. Bracchi, M. E. & Fulton, D. A. Orthogonal Breaking and Forming of Dynamic Covalent Imine and Disulfide Bonds in Aqueous Solution. *Chem. Commun.* **51**, 11052-11055 (2015). DOI: 10.1039/c5cc02716k.
3. Caraballo, R., Dong, H., Ribeiro, J. P., Jiménez-Barber, J. & Ramström, O. Direct STD NMR Identification of β -Galactosidase Inhibitors from a Virtual Dynamic Hemithioacetal System. *Angew. Chem. Int. Ed.* **49**, 589-593 (2010). DOI: 10.1002/anie.200903920.
4. Qin, F., Auerbach, A. & Sachs, F. Estimating Single-Channel Kinetic Parameters from Idealized Patch-Clamp Data Containing Missed Events. *Biophys. J.* **70**, 264–280 (1996).
5. Krezel, A. & Bal, W. A Formula for Correlating pK_a Values Determined in D_2O and H_2O . *J. Inorg. Biochem.* **98**, 161-166 (2004).
6. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. & Lindahl, E. GROMACS: High Performance Molecular Simulations Through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **1-2**, 19-25 (2015).
7. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Bioinf.* **78**, 1950-1958 (2010).
8. Price, D. J. & Brooks, C. L. A Modified TIP3P Water Potential for Simulation with Ewald Summation. *J. Chem. Phys.* **121**, 10096–10103 (2004).
9. Bussi, G., Donadio, D. & Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
10. Parrinello, M. & Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
11. Darden, T., York, D. & Pedersen, L. Particle Mesh Ewald: An $N \cdot \log(N)$ Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
12. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H. & Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
13. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **18**, 1463-1472 (1998).
14. Smart, O. S., Goodfellow, J. M. & Wallace, B. A. The Pore Dimensions of Gramicidin A. *Biophys J.* **65**, 2455–2460 (1993).
15. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: A Program for the Analysis of the Pore Dimensions of Ion Channel Structural Models. *J. Mol. Graph.* **14**, 354-360 (1996).

16. Braha, O., Walker, B., Cheley, S., Kasianowicz, J. J., Song, L., Gouaux, J.E. & Bayley, H. Designed Protein Pores as Components for Biosensors. *Chem. Biol.* **4**, 497-505 (1997).
17. Heath, R. S., Birmingham, W. R., Thompson, M. P., Taglieber, A., Daviet, L. & Turner, N. J. An Engineered Alcohol Oxidase for the Oxidation of Primary Alcohols. *ChemBioChem* **20**, 276-281 (2019).
18. Montal, M. & Mueller, P. Formation of Bimolecular Membranes from Lipid Monolayers and a Study of their Electrical Properties. *Proc. Natl. Acad. Sci. USA* **69**, 3561-3566 (1972).
19. Nicolai, C. & Sachs, F. Solving Ion Channel Kinetics with the QuB Software. *Biophys. Rev. Lett.* **08**, 191–211 (2013).