Supplementary Information to

Proton transport in water is doubly gated by sequential hydrogen-bond exchanges

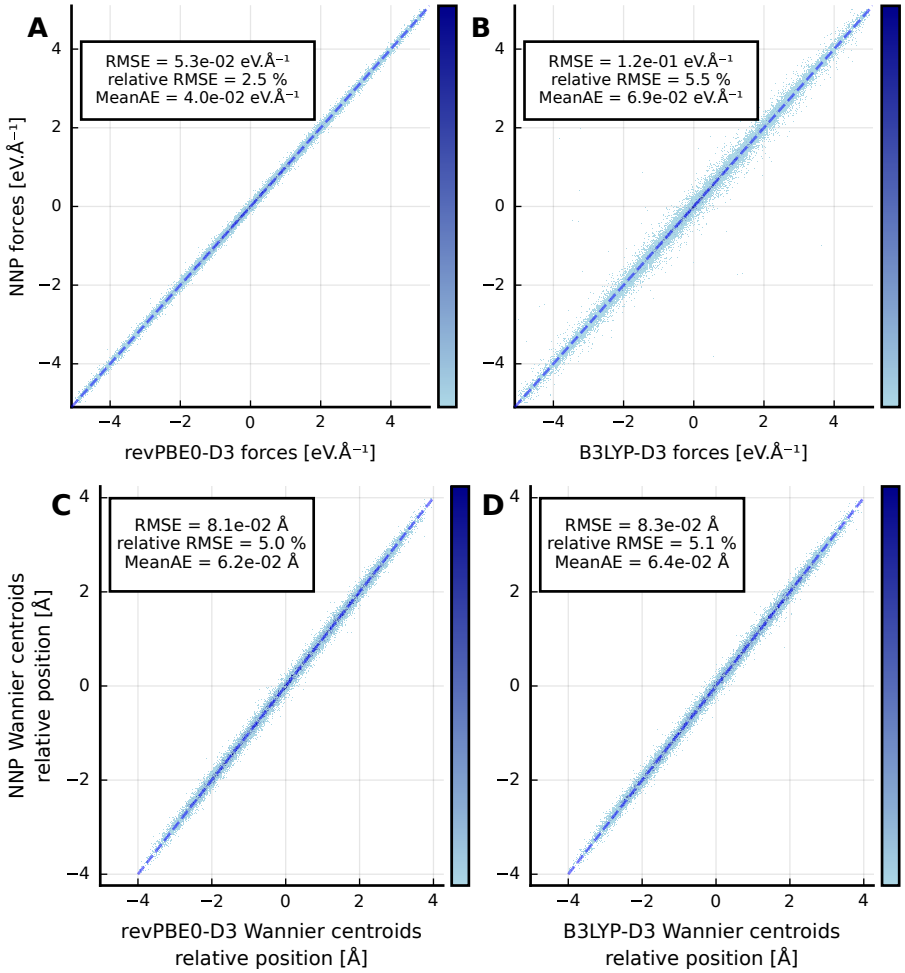
Axel Gomez[1], Ward H. Thompson[2] and Damien Laage[1*]

[1]PASTEUR, Department of chemistry, Ecole normale supérieure, PSL university, Sorbonne university, CNRS; Paris, France.
[2]Department of chemistry, University of Kansas; Lawrence, KS 66045, USA.
*Corresponding author. Email: damien.laage@ens.psl.eu.

S-1 Validation of neural network potentials

The neural network potentials (NNP) used in this work as reactive force-fields are tested against reference DFT calculations. An independent set of 40 configurations, labelled both at the revPBE0-D3 and at the B3LYP-D3 levels of theory, serves for validation. Figure S1A-B shows the correlation of atomic forces from the NNP and from the reference DFT calculations. This shows that the NNPs developed in this work are able to reproduce the reference DFT forces with a high accuracy. The insets of Figure S1A-B provide the root mean squared error (RMSE), the RMSE scaled by the standard deviation of reference data (relative RMSE) and the mean absolute error (MeanAE). The NNPs reproduce the reference DFT forces with a relative RMSE of 2.5% and 5.5% for the revPBE0-D3 and B3LYP-D3 levels respectively.

For infrared spectra calculations, as described in the Methods section, we train deep Wannier neural networks (DWNN) that predict the displacement of Wannier centroids with respect to the closest oxygen atom. Figure S1C-D represent the correlation between DFT-computed Wannier centroids and the DWNN predictions for the two hybrid functionals employed. The correlation is excellent and the predictions can be used to compute infrared spectra with a high accuracy.

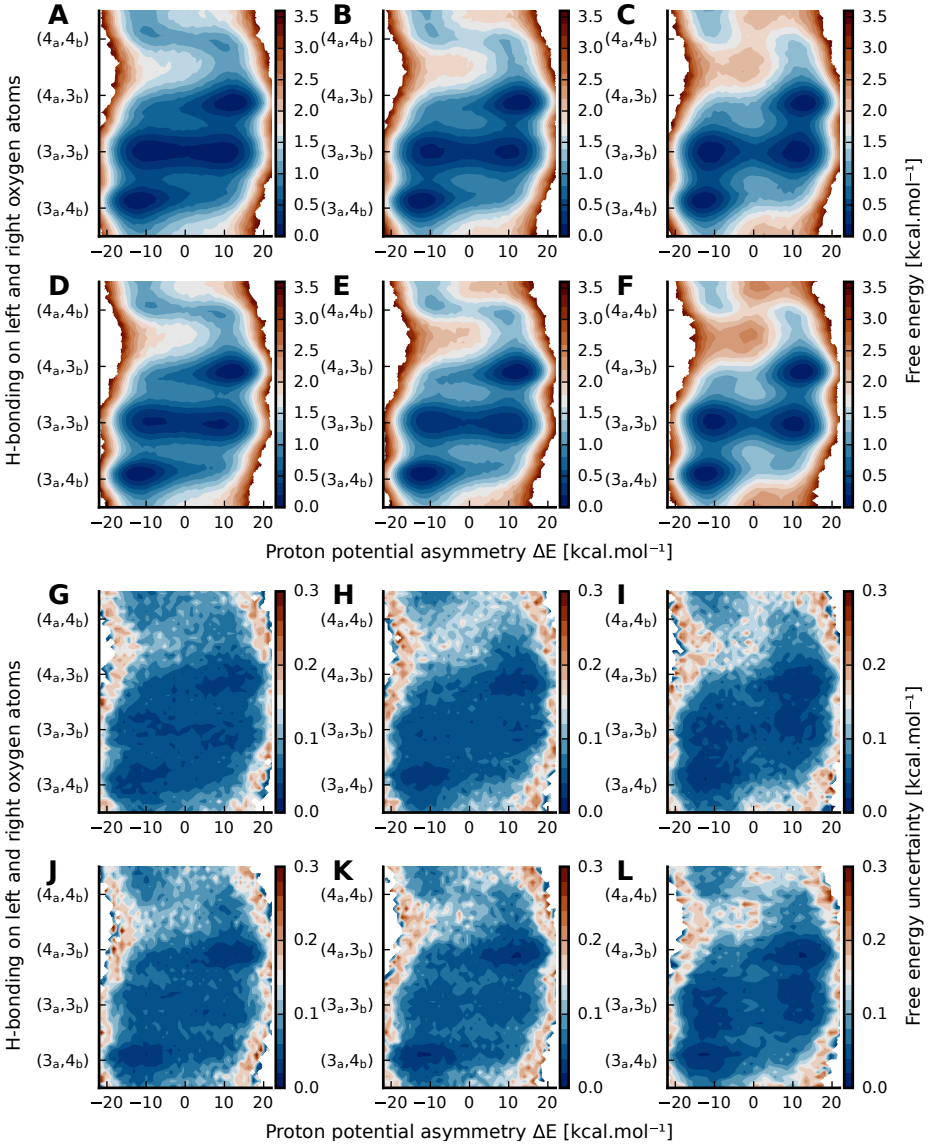


Supplementary Figure S1: **Validation of neural networks.** Upper panels: Correlation between atomic forces at the revPBE0-D3 (**A**) and B3LYP-D3 (**B**) levels from DFT calculations and from our NNP. The insets provide the root mean squared error (RMSE), the RMSE scaled by the standard deviation of reference data (relative RMSE) and the mean absolute error (MeanAE). Lower panels: Correlation between the (x, y, z) components of the relative positions of Wannier centroids with respect to their closest oxygen atoms at revPBE0-D3 (**C**) and B3LYP-D3 (**D**) levels. Insets provide the root mean squared error (RMSE), the RMSE scaled by the standard deviation of reference data (relative RMSE) and the mean absolute error (MeanAE).

S-2 Comparison of proton transfer free energy surfaces at different levels of description

To confirm the robustness of the proposed proton transfer mechanism at ambient temperature, we have verified that the free energy surface along the proton potential asymmetry and H-bond coordinates are not significantly affected by the choice of the exchange-correlation functional and by the inclusion of nuclear quantum effects. Figure S2 shows the free energy surfaces computed at the revPBE0-D3 and B3LYP-D3 levels of theory from TRPMD simulations in H_2O (Figure S2A-B) and in D_2O (Figure S2D-E) and from classical MD simulations (Figure S2C-F). The free energy surfaces are very similar for the two hybrid functionals, for the two isotopes with quantum nuclei and for classical nuclei. These surfaces are integrated to obtain the 1D free energy profiles for all levels of description (Extended Data Figure 3 and Figure 1C of the main text). Following the typical minimum free energy path shown in Figure 1B, the 1D free energy is obtained by numerically integrating the free energy along the transverse coordinate. The integration range changes with the nature of the transverse coordinate and is defined along three consecutive intervals. On the first segment, from the $(3_a, 4_b)$ to the $(3_a, 3_b)$ basins, the profile evolves along the H-bond coordinate and the solvent coordinate is integrated over $\Delta E < 0$. On the second segment, from left to right in the Zundel-like $(3_a, 3_b)$ configuration, the H-bond coordinate is integrated over the $(3_a, 3_b)$ state. On the final segment, from the $(3_a, 3_b)$ to the $(4_a, 3_b)$ basins, the solvent coordinate is integrated over $\Delta E > 0$.

The uncertainty on the computed free energy is defined as the half-width of the 95% confidence interval calculated from the 20 independent trajectories. It is shown in the lower panels of Figure S2 for each bin of the 2D free energy surfaces. Free energy uncertainties uniformly averaged on the -20 kcal.mol^{-1} $< \Delta E <$ 20 kcal.mol^{-1} range and over the (3_a,4_b), (3_a,3_b) and (4_a,3_b) states are reported in Table S1. They are approximately one order of magnitude smaller than the thermal energy and much smaller than the barriers along the reaction path, confirming that the free energy surfaces are very well converged.



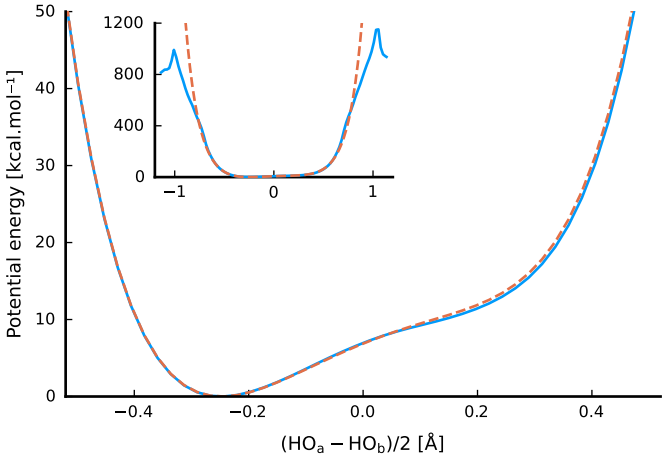
Supplementary Figure S2: **Proton transfer free energy surfaces at 300 K.** Upper panels: free energy surfaces computed at the revPBE0-D3 (top) and B3LYP-D3 (bottom) levels from TRPMD simulations in H_2O (**A**, **D**), in D_2O (**B**, **E**) and from classical MD simulations (**C**,**F**). Lower panels: Free energy uncertainty, defined as the half-width of the 95% confidence interval calculated from the 20 independent trajectories at the same levels of theory as in the upper panels.

Simulation	Free energy uncertainty [kcal.mol^{-1}]
revPBE0-D3 TRPMD H^+,H_2O	0.065
revPBE0-D3 TRPMD D^+,D_2O	0.073
revPBE0-D3 classical H^+,H_2O	0.070
B3LYP-D3 TRPMD H^+,H_2O	0.069
B3LYP-D3 TRPMD D^+,D_2O	0.072
B3LYP-D3 classical H^+,H_2O	0.087

Supplementary Table S1: Free energy uncertainty uniformly averaged on the -20 kcal.mol^{-1} $< \Delta E <$ 20 kcal.mol^{-1} range and over the (3_a,4_b), (3_a,3_b) and (4_a,3_b) states from our simulated systems

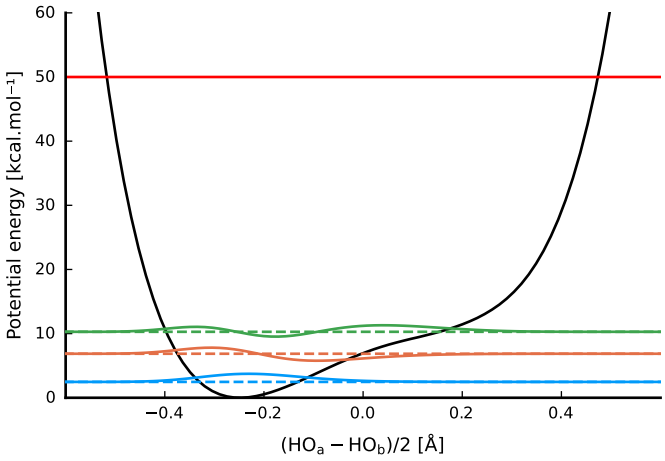
S-3 Calculation of vibrational energy levels

In addition to the validation of the NNP, we confirm here that the 1D potential energy profiles used for DVR calculations are accurate. Figure S3 shows a typical 1D potential energy profile along the proton transfer coordinate from our NNP trained at the revPBE0-D3 level and from reference revPBE0-D3 DFT calculations. The agreement is excellent and the potential energy profile is very well reproduced by the NNP. This validates the use of the NNP for the quantum energy levels calculations and therefore for the computed spectroscopy observables.



Supplementary Figure S3: **Validation of NNP for 1D potential energy profiles.** One-dimensional potential energy profile along the proton transfer coordinate from our NNP trained at the revPBE0-D3 level (blue solid line) and from reference revPBE0-D3 DFT calculations (orange dashes). The inset shows the same energy profile with a broader range of values for both axes.

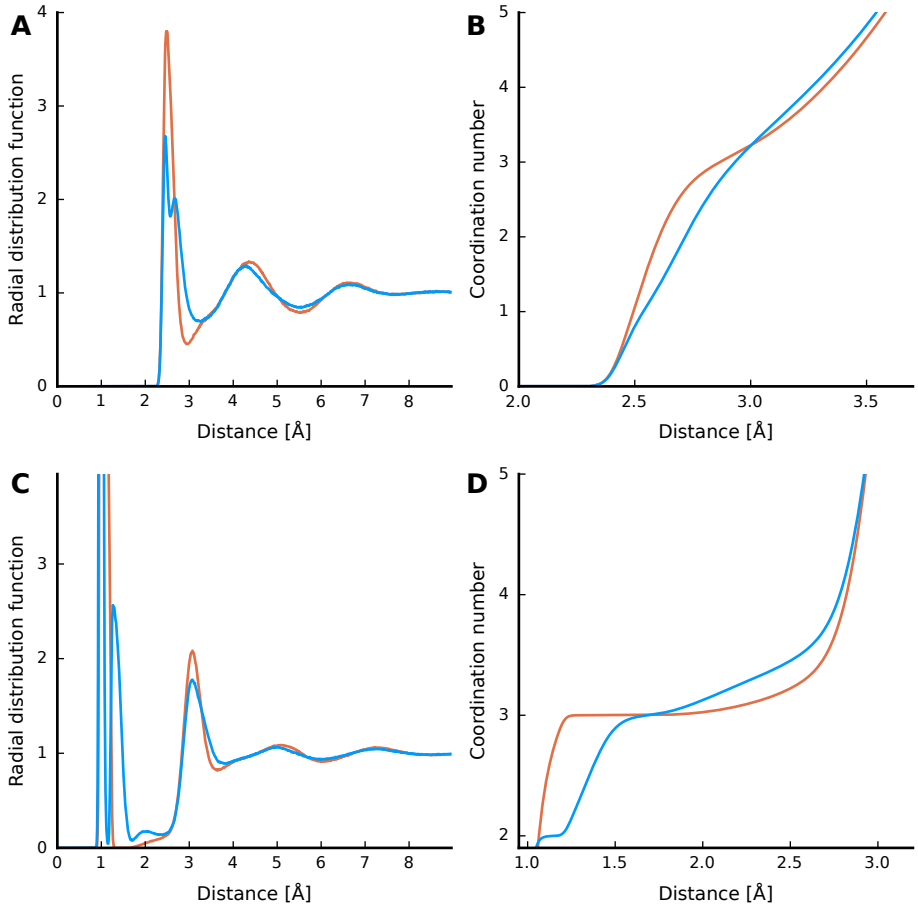
For the same profile, we also represent in Figure S4 the quantum vibrational energy levels and wavefunctions obtained from one-dimensional discrete variable representation (DVR) calculations. This one-dimensional quantization approach has been extensively used to calculate nonlinear vibrational spectra of water in quantitative agreement with experimental measurements (see, *e.g.*, a review in J.L. Skinner, B.M. Auer, Y.-S. Lin, *Adv. Chem. Phys.* **142**, 59-103 (2009)).



Supplementary Figure S4: **Quantum vibrational energy levels calculation.** Typical potential energy profile (black) along the proton transfer mode and quantum vibrational energy levels (dashes) and wavefunctions (solid lines) for the $\nu = 0$ (blue), $\nu = 1$ (orange) and $\nu = 2$ (green) levels, and energy cutoff (red).

S-4 Structural characterization of the excess proton

The structure of an excess proton in liquid water at ambient temperature is studied via the hydrogen atom radial distribution function around the oxygen atom O_a carrying 3 hydrogen atoms and around its best proton acceptor O_b, *i.e.*, its H-bonded neighbor with the smallest $|\Delta E|$. Figure S5 clearly shows that O_a does not accept a fourth H-bond, whereas O_b can accept one. This therefore justifies the definition of the H-bond coordinate in this work.

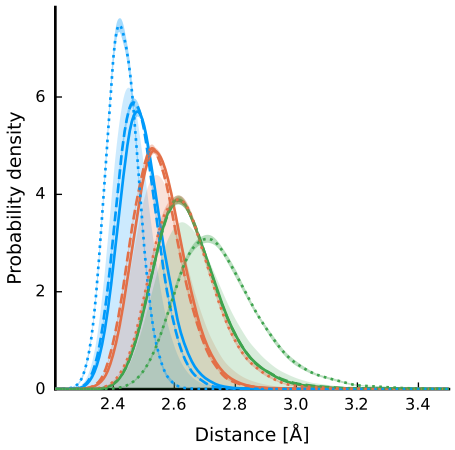


Supplementary Figure S5: **Local structure around the excess proton.** (**A**) Oxygen radial distribution functions around O_a (orange) and O_b (blue) and (**B**) associated oxygen coordination numbers, based on TRPMD simulations at the revPBE0-D3 level in H_2O at 300 K. (**C**) and (**D**) idem for the hydrogen coordination.

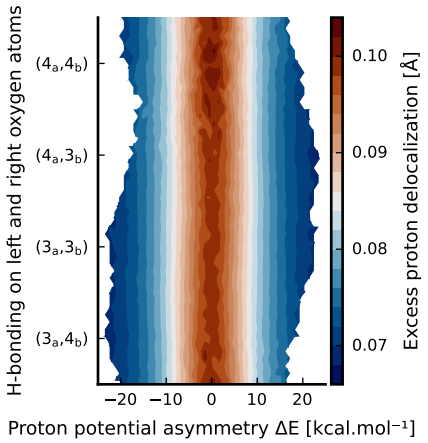
The local environment of the excess proton is further characterized by the distances between O_a and its three H-bonded water neighbors. Figure S6 shows their distribution. The ordering is based on the ΔE solvent coordinate and thus reflects the strength of the H-bond and the shape of the proton potential. Our results are consistent with a recent X-ray absorption study on protonated water complexes[37].

State	Proton potential asymmetry values [eV]	H-bond coordinate values
1: Eigen-like	$-0.7 < \Delta E < -0.4$	$-\frac{3\pi}{8} < \gamma_{HB} < -\frac{\pi}{8}$
2: Zundel-like	$-0.7 < \Delta E < -0.4$	$\frac{\pi}{8} < \gamma_{HB} < \frac{3\pi}{8}$
3: Zundel	$-0.1 < \Delta E < 0.1$	$\frac{\pi}{8} < \gamma_{HB} < \frac{3\pi}{8}$

Supplementary Table S2: Definition of the three states **1**, **2** and **3** (see Figure 1) that are used to calculate state-specific oxygen-oxygen distances (Figure S6) based on the two proton transfer reaction coordinates.



Supplementary Figure S6: **Oxygen-oxygen distances around the excess proton.** (Left) Oxygen-oxygen distances around the hydronium with its best excess proton acceptor (blue), second best (orange) and least favorable (green), as determined by the ΔE coordinate, based on TRPMD simulations at the revPBE0-D3 level in H_2O at 300 K. Light filled distributions correspond to the average structure in the simulation, and the solid, dashed and dotted lines correspond to states **1**, **2** and **3** respectively (see Figure 1). The definition of the three states as a function of the reaction coordinates is in Table S2.



Supplementary Figure S7: **Proton delocalization.** Excess proton delocalization measured as the standard deviation of the squared proton wavefunction as a function of the proton potential asymmetry and of the H-bond coordinate.

S-5 Proton diffusion

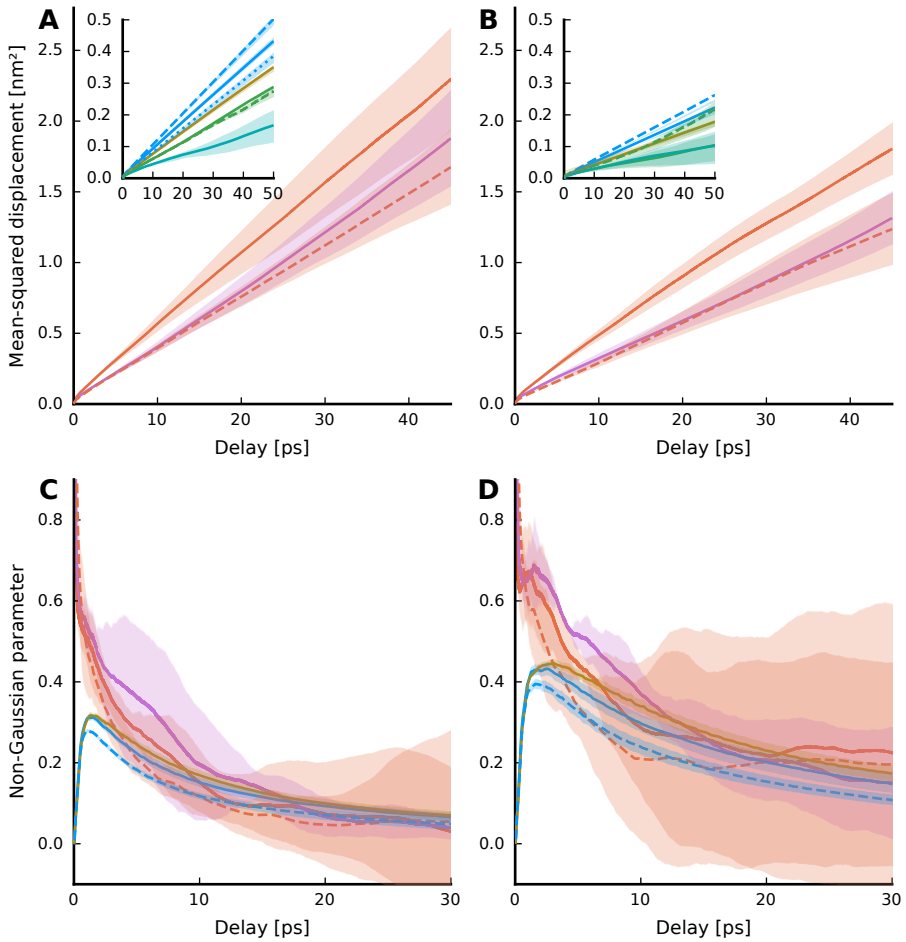
S-5.1 Proton diffusion for different levels of description

Using the trajectories propagated with the revPBE0-D3 and B3LYP-D3 NNPs, with and without nuclear quantum effects, we have computed the mean squared displacements (MSD) of the defect-carrying oxygen atom (Figure S8A-B). Our long simulations allowing attaining the linear regime, which can then be used to determine the diffusion coefficients for the different levels of description at ambient temperature (Table S3). The values are reported in Table S3 and used for comparison with the random walk modeling in Figure 3B of the main text.

To investigate the potentially anomalous character of proton diffusion, we have computed the non-Gaussian parameter $\alpha_2(t) = \frac{3}{5} \times \frac{\langle \Delta r(t)^4 \rangle}{\langle \Delta r(t)^2 \rangle^2} - 1$ (Figure S8C-D) for the same systems. Diffusion is shown to be as normal as water diffusion for all descriptions, which confirms that proton transport can be described as a normal diffusion after a few picoseconds.

S-5.2 Finite-size effect on diffusion

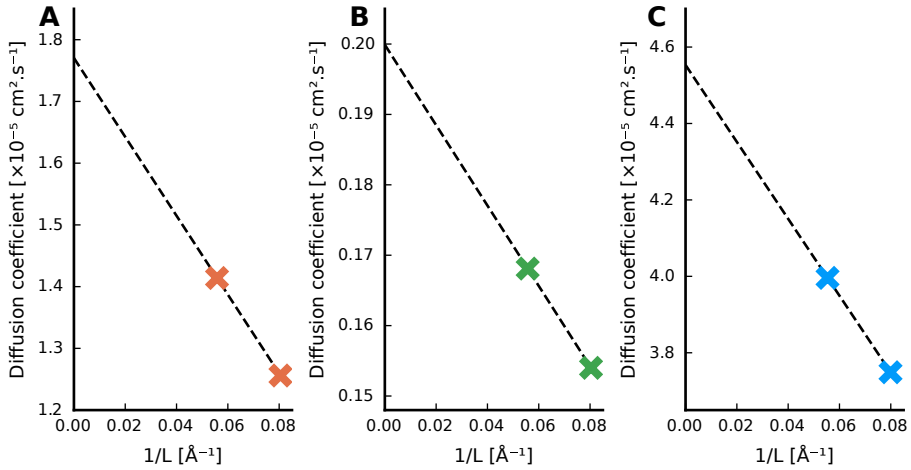
We have computed the finite-size correction to the water diffusion coefficient from simulations with two different box lengths, 192 and 64 water molecules in a cubic box with a side of 17.927 Å and 12.42 Å respectively (Yeh, I. C. & Hummer, G. System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions. *J. Phys. Chem. B* **108**, 15873–15879 (2004)). Figure S9 shows the box length effect on the diffusion coefficient at three temperatures. Because this hydrodynamic correction is independent of the solute, it is the same for water and for the excess proton. Finite-size corrected diffusion coefficients are used for comparisons with experiments.



Supplementary Figure S8: **Excess proton diffusion at 300 K.** Upper panels: Mean squared displacements of the defect-carrying oxygen atom with (**A**) revPBE0-D3 and (**B**) B3LYP-D3 NNPs respectively in H_2O with nuclear quantum effects (solid orange line), H_2O with classical nuclei (orange dashes) and D_2O with nuclear quantum effects (solid purple line). The insets show the water oxygen atom mean squared displacements, respectively in H_2O with nuclear quantum effects (solid blue lines), H_2O with nuclear quantum effects in a smaller simulation box containing one excess proton and 64 H_2O (blue dots), H_2O with classical nuclear dynamics (blue dashes) and D_2O with nuclear quantum effects (solid gold lines), and the mean squared displacement of a constrained hydronium cation whose proton transfers are prevented (solid green line for H_3O^+ with nuclear quantum effects, green dashes for H_3O^+ with classical dynamics, turquoise line for D_3O^+ with nuclear quantum effects). Lower panel: Non-Gaussian parameter $\alpha_2(t) = \frac{3}{5} \times \frac{\langle \Delta r(t)^4 \rangle}{\langle \Delta r(t)^2 \rangle^2} - 1$ for the defect-carrying oxygen atom in (**C**) revPBE0-D3 and (**D**) B3LYP-D3 simulations, respectively in H_2O with nuclear quantum effects (solid orange line), H_2O with classical nuclei (orange dashes) and D_2O with nuclear quantum effects (solid purple line). $\alpha_2(t)$ for water oxygen atom distributions are shown for comparison, respectively in H_2O with nuclear quantum effects (solid blue lines), in H_2O with classical nuclear dynamics (blue dashes) and in D_2O with nuclear quantum effects (solid gold lines).

Species	Method	D_{MD} [$\times 10^{-5}$ cm^2.s^{-1}]
H^+	revPBE0-D3 TRPMD	(8.21 ± 0.05)
H_2O	revPBE0-D3 TRPMD	(1.414 ± 0.002)
H_3O^+ vehicular	revPBE0-D3 TRPMD	(0.96 ± 0.01)
H_2O 64 molecules	revPBE0-D3 TRPMD	(1.256 ± 0.002)
D^+	revPBE0-D3 TRPMD	(6.90 ± 0.04)
D_2O	revPBE0-D3 TRPMD	(1.150 ± 0.002)
D_3O^+ vehicular	revPBE0-D3 TRPMD	(0.465 ± 0.004)
H^+	revPBE0-D3 classical	(6.05 ± 0.03)
H_2O	revPBE0-D3 classical	(1.646 ± 0.002)
H_3O^+ vehicular	revPBE0-D3 classical	(0.845 ± 0.004)
H^+	B3LYP-D3 TRPMD	(6.28 ± 0.03)
H_2O	B3LYP-D3 TRPMD	(0.719 ± 0.001)
H_3O^+ vehicular	B3LYP-D3 TRPMD	(0.286 ± 0.004)
D^+	B3LYP-D3 TRPMD	(4.61 ± 0.03)
D_2O	B3LYP-D3 TRPMD	(0.568 ± 0.001)
D_3O^+ vehicular	B3LYP-D3 TRPMD	(0.313 ± 0.004)
H^+	B3LYP-D3 classical	(4.63 ± 0.03)
H_2O	B3LYP-D3 classical	(0.847 ± 0.001)
H_3O^+ vehicular	B3LYP-D3 classical	(0.670 ± 0.006)

Supplementary Table S3: Diffusion coefficients at 300 K obtained from the mean squared displacements (Figure S8A-B). The reported uncertainties correspond to the 95% confidence interval calculated on 20 trajectories of 200 ps (except for H_3O^+ and D_3O^+ vehicular diffusion coefficients which are computed from 5 independent simulations).



Supplementary Figure S9: **Simulation box size dependence of the water diffusion coefficient.** (A) Crosses show the self-diffusion coefficient of water for the two simulation boxes in this work (192 and 64 water molecules) with the revPBE0-D3 NNP at 300 K. The difference from finite-size calculated values to infinite-size diffusion coefficient (origin of the dashed line) gives the finite-size correction to the diffusion coefficient. For the larger simulation box, this correction is $(0.356 \pm 0.006) \times 10^{-5}$ cm^2.s^{-1}. (B) Same at 250 K where the finite-size correction for the larger simulation box is $(0.032 \pm 0.004) \times 10^{-5}$ cm^2.s^{-1} and (C) at 350 K where the correction is $(0.56 \pm 0.02) \times 10^{-5}$ cm^2.s^{-1}.

S-5.3 Temperature dependence of the diffusion coefficient and activation energy

The temperature dependence of the diffusion coefficient shown in Figure 3C of the main text is fitted with a Speedy-Angell power law (Speedy, R. J. & Angell, C. A. Isothermal compressibility of supercooled water and evidence for a thermodynamic singularity at -45°C. *J. Chem. Phys.* **65**, 851–858 (1976)):

$$D(T) = D_0 \left(\frac{T}{T_S} - 1 \right)^{\gamma}$$

The fitting parameters are reported in Table S4 for TRPMD simulations at the revPBE0-D3 level of theory for the excess proton diffusion and water self-diffusion.

Parameter	H^+	H_2O
D_0 [$\times 10^{-5}$ cm^2.s^{-1}]	34.8	14.2
T_S [K]	227.8	229.6
γ	1.22	1.76

Supplementary Table S4: Fitting parameters for the Speedy-Angell power law describing the temperature-dependence of the diffusion coefficient for TRPMD simulations at the revPBE0-D3 level of theory.

The activation energy for diffusion is computed from the derivative of the power law as

$$E_{\mathrm{a}}(T) = -\frac{\partial \ln D(T)}{\partial \beta} = \frac{\gamma k_{\mathrm{B}} T^2}{T - T_S}$$

where $\beta = \frac{1}{k_{\mathrm{B}} T}$.

S-6 Proton hopping times and hopping amplitudes for random walk modeling of proton transport

The proton hopping times extracted from the time correlation function analysis for all descriptions are reported in Table S5 and used for the random walk modeling.

System	Method	τ [ps]
H_2O	revPBE0-D3 TRPMD	1.47 ± 0.05
D_2O	revPBE0-D3 TRPMD	2.03 ± 0.09
H_2O	revPBE0-D3 classical	1.84 ± 0.09
H_2O	B3LYP-D3 TRPMD	1.91 ± 0.15
D_2O	B3LYP-D3 TRPMD	2.67 ± 0.17
H_2O	B3LYP-D3 classical	2.55 ± 0.14

Supplementary Table S5: Proton hopping times at 300 K for each type of simulation from time correlation function analysis (Extended Data Figure 6A).

The average $O_aO_b^2$ squared distances for all descriptions are reported in Table S6.

System	$\langle O_aO_b^2 \rangle$ [Å^2]
revPBE0-D3 TRPMD H^+,H_2O	6.089 ± 0.003
revPBE0-D3 TRPMD D^+,D_2O	6.102 ± 0.004
revPBE0-D3 classical H^+,H_2O	6.147 ± 0.003
B3LYP-D3 TRPMD H^+,H_2O	6.136 ± 0.003
B3LYP-D3 TRPMD D^+,D_2O	6.153 ± 0.003
B3LYP-D3 classical H^+,H_2O	6.206 ± 0.004

Supplementary Table S6: Average $O_aO_b^2$ squared distance at 300 K for each simulation (Extended Data Figure 6C).