Supplementary information: Local initiation conditions for water autoionization

Mahmoud Moqadam*[1], Anders Lervik†[1], Enrico Riccardi‡[1], Vishwesh Venkatraman§[1], Bjørn Kåre Alsberg¶[1], and Titus S. van Erp‖[1,2]

[1]Department of Chemistry, Norwegian University of Science and Technology (NTNU), N-7491 Trondheim, Norway
[2]Center for Molecular Modeling (CMM), Ghent University, Technologiepark 903, B9000 Ghent, Belgium

April 18, 2018

This supplementary information contains additional results from the replica exchange transition interface simulations, the predictive power analysis and the machine learning analysis. We refer to the main article for the interpretation of these results.

*mahmoud.moqadam@ntnu.no
†anders.lervik@ntnu.no
‡enrico.riccardi@ntnu.no
§vishwesh.venkatraman@ntnu.no
¶deceased 27.12.2017
‖titus.van.erp@ntnu.no

www.pnas.org/cgi/doi/10.1073/pnas.1714070115

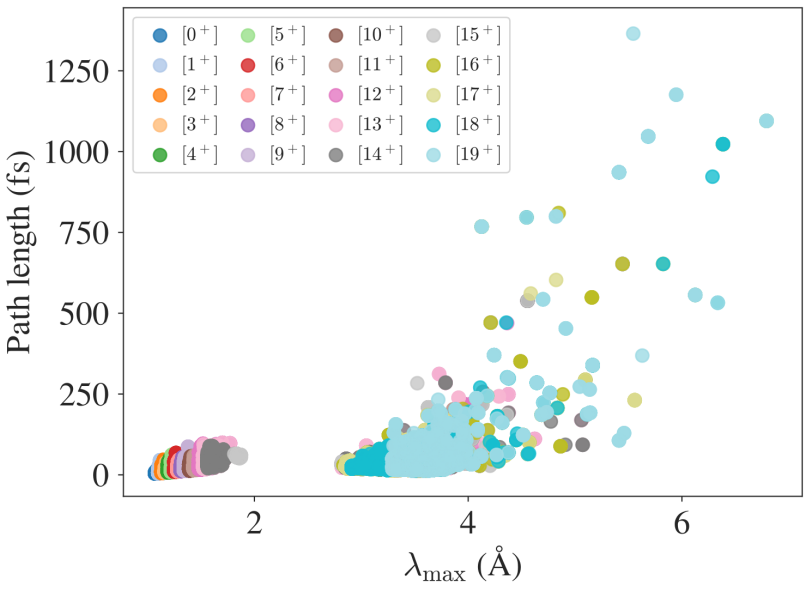

Figure S1: Length of accepted trajectories as function of the maximum order parameter obtained. The trajectories are labeled according to the ensemble they belong to. Note that ionic separation distances up to ca. 4 Å still allow for almost instantaneous recombination.

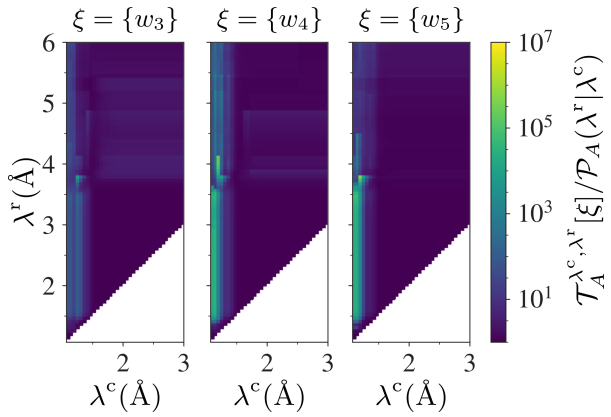

Figure S2: Comparing water wires of different lengths. Here, we compare the predictive abilities ($\mathcal{T}_A^{\lambda^c,\lambda^r}[\xi]$) normalized by the crossing probability ($\mathcal{P}_A(\lambda^r|\lambda^c)$) for water wires consisting of 3, 4 or 5 water molecules. The length of the water wires w_3, w_4 and w_5 respectively were obtained as described in the main text.

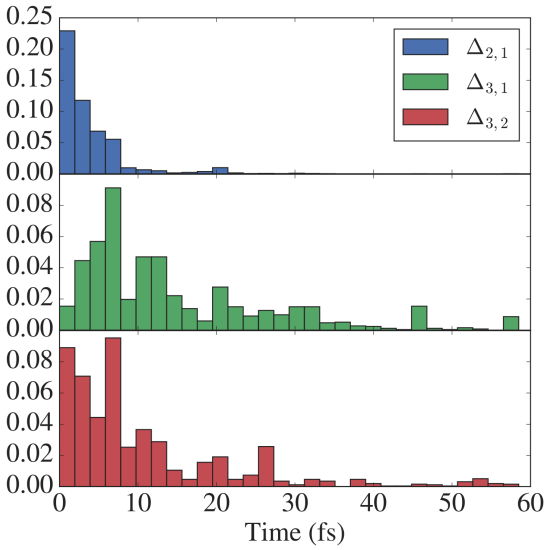

Figure S3: Distribution of waiting times between initiation of hydrogen transfer events. From top to bottom we show the waiting time between initiation of the first and second hydrogen transfer event ($\Delta_{2,1}$), the first and third ($\Delta_{3,1}$), and the second and third event ($\Delta_{3,2}$). The results were obtained by considering the trajectories in the final path ensemble.

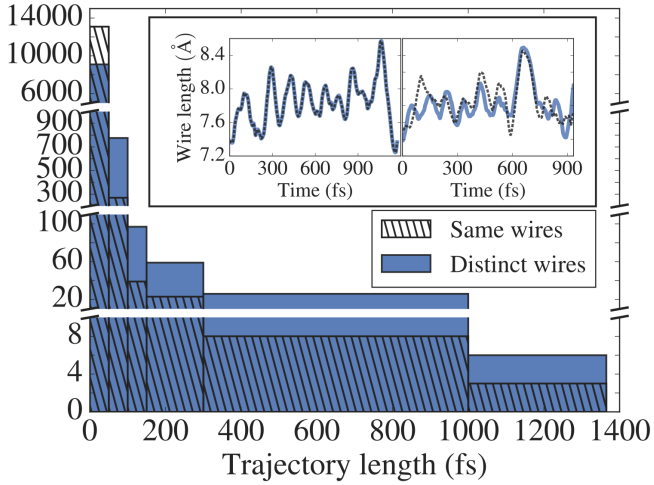

Figure S4: Distribution of trajectory lengths for the final path ensemble. The trajectories have been classified according to the identified hydrogen bond wire in the forward and time-reversed directions as: (1) The same wire is identified in both directions (white color and striped pattern: "Same wires") or (2) different wires are identified in the two directions (color blue: "Distinct wires"). The inset show two examples for the length of the four-membered wires for the two classes (forward direction in solid blue, time-reversed direction in dashed black lines).

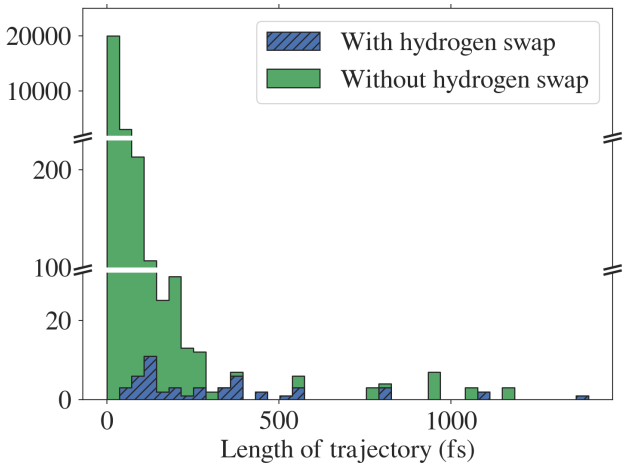

Figure S5: Distribution of trajectory lengths for the final path ensemble. The trajectories have been classified according to whether the hydrogen atoms are bound to the same oxygen atom in the initial and final frames or not (green or blue hatched bars, respectively). The average length of the trajectories with proton swap is 326.73 fs, while the average length for the trajectories without swap is 26.93 fs.

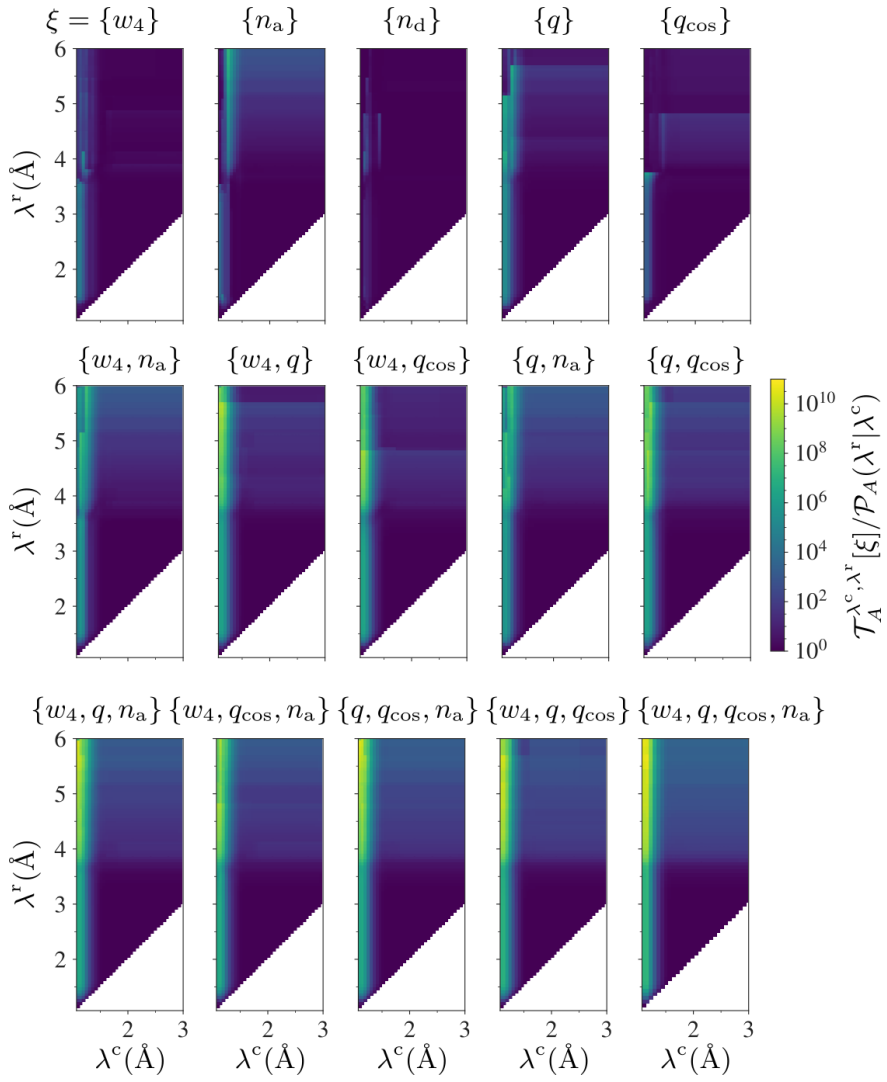

Figure S6: Predictive abilities for different combinations of collective variables. Comparison of the predictive abilities ($\mathcal{T}_A^{\lambda^\mathrm{c},\lambda^\mathrm{r}}[\xi]$) normalized by the crossing probability ($\mathcal{P}_A(\lambda^\mathrm{r}|\lambda^\mathrm{c})$) for different combinations of collective variables (ξ).

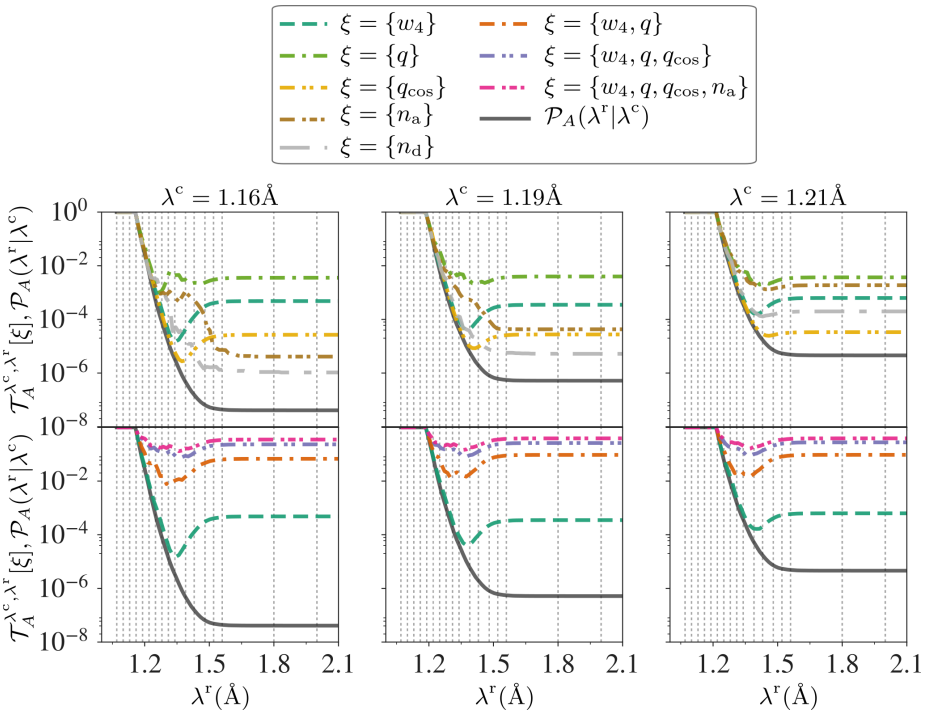

Figure S7: Comparison of predictive powers and the crossing probaibility as a function of λ^r for $\lambda^c = 1.16, 1.19, 1.21$ Å and different combinations of collective variables (ξ). Due to the threshold criterion for defining the wires (see the main text), the probability is shifted so that $\mathcal{P}_A = 1$ for $\lambda < 1.15$ Å.

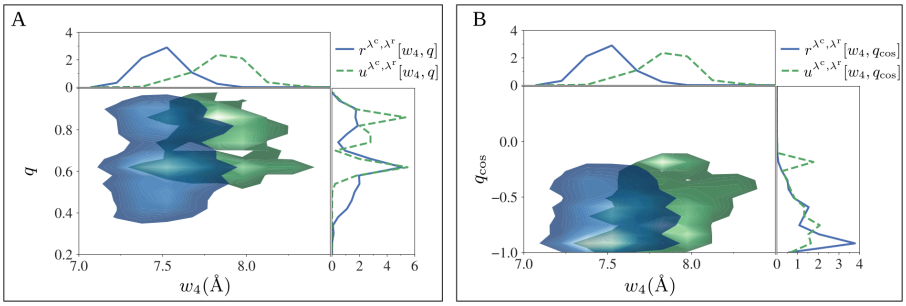

Figure S8: Reactive ($r^{\lambda^c,\lambda^r}(\xi)$) and nonreactive ($u^{\lambda^c,\lambda^r}(\xi)$) distributions for $\xi = \{w_4, q\}$ (panel A) and $\xi = \{w_4, q_{\cos}\}$ (panel B). The depicted distributions have been normalized. The top and right insets show the one-dimensional projections of the distributions.

Table S1: Rules obtained by the Classification and regression tree analysis.

Rule number	Class	Textual representation
5	1	$w_4 \geq 7.6$ and $\lambda_2 \geq 1.1$
7	1	$w_4 < 7.6$ and $w_4 < 7.5$
13	1	$w_4 < 7.6$ and $w_4 \geq 7.5$ and $d_7 < 3.9$
16	0	$w_4 \geq 7.6$ and $\lambda_2 < 1.1$ and $\overline{d}_2 \geq 2.5$ and $d_6 \geq 3.7$
18	0	$w_4 \geq 7.6$ and $\lambda_2 < 1.1$ and $\overline{d}_2 < 2.5$ and $d_5 < 3.3$
19	1	$w_4 \geq 7.6$ and $\lambda_2 < 1.1$ and $\overline{d}_2 < 2.5$ and $d_5 \geq 3.3$
25	1	$w_4 < 7.6$ and $w_4 \geq 7.5$ and $d_7 \geq 3.9$ and $n\_texta < 2.5$
34	0	$w_4 \geq 7.6$ and $\lambda_2 < 1.1$ and $\overline{d}_2 \geq 2.5$ and $d_6 < 3.7$ and $w_4 \geq 7.7$
49	1	$w_4 < 7.6$ and $w_4 \geq 7.5$ and $d_7 \geq 3.9$ and $n_{\mathrm{a}} \geq 2.5$ and $d_{25} \geq 6.1$
71	1	$w_4 \geq 7.6$ and $\lambda_2 < 1.1$ and $\overline{d}_2 \geq 2.5$ and $d_6 < 3.7$ and $w_4 < 7.7$ and $q_4 < 0.23$
96	0	$w_4 < 7.6$ and $w_4 \geq 7.5$ and $d_7 \geq 3.9$ and $n_{\mathrm{a}} \geq 2.5$ and $d_{25} < 6.1$ and $d_2 < 2.7$
97	1	$w_4 < 7.6$ and $w_4 \geq 7.5$ and $d_7 \geq 3.9$ and $n_{\mathrm{a}} \geq 2.5$ and $d_{25} < 6.1$ and $d_2 \geq 2.7$
140	0	$w_4 \geq 7.6$ and $\lambda_2 < 1.1$ and $\overline{d}_2 \geq 2.5$ and $d_6 < 3.7$ and $w_4 < 7.7$ and $q_4 \geq 0.23$ and $\overline{d}_3 < 2.7$
141	1	$w_4 \geq 7.6$ and $\lambda_2 < 1.1$ and $\overline{d}_2 \geq 2.5$ and $d_6 < 3.7$ and $w_4 < 7.7$ and $q_4 \geq 0.23$ and $\overline{d}_3 \geq 2.7$

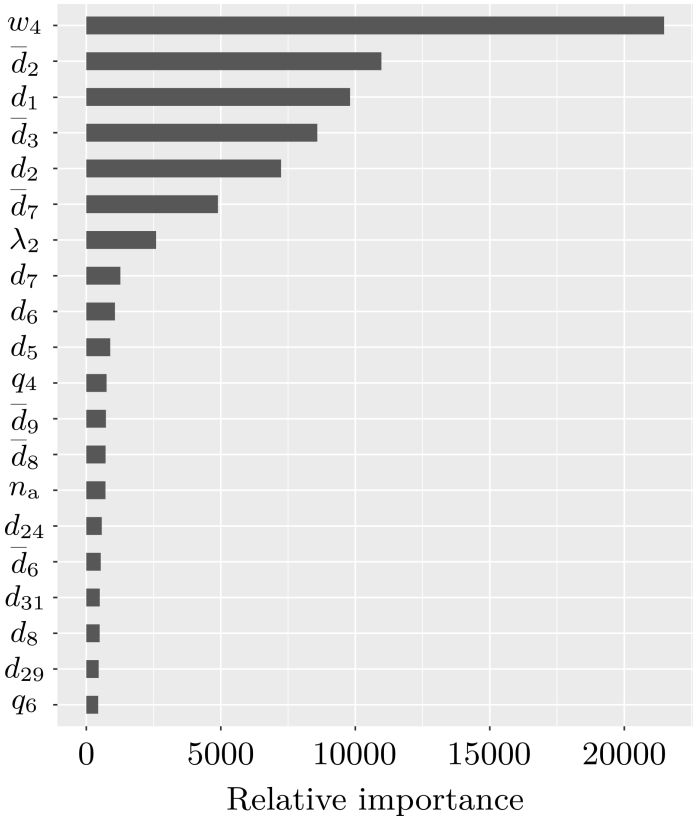

Figure S9: Relative importance of the 20 most important variables from the machine learning analysis. The variable importance is calculated with respect to the reduction in the classification error attributed to each variable at each split in the decision tree. The values are then scaled to sum to 100.