Supporting Information for

Benchmark Study of the SCC-DFTB Approach for a Biomolecular Proton Channel

Ruibin Liang, Jessica M. J. Swanson, and Gregory A. Voth*
Department of Chemistry, Institute for Biophysical Dynamics, James Franck Institute, and Computation Institute, University of Chicago, 5735 S. Ellis Ave., Chicago, Illinois 60637, USA

*Corresponding author: Gregory A. Voth.
E-mail: gavoth@uchicago.edu
Fax: 773-795-9106
Phone: 773-702-9092

Section S1: Comparison to multi-state empirical valence bond simulations

The multi-state empirical valence bond method (MS-EVB) has been applied to studying proton transport in the LS2 channel.[1, 2] This method describes the excess proton Grotthuss shuttling and charge delocalization explicitly by evolving the system on a reactive potential energy surface defined by a linear combination of multiple diabatic basis states.[3-7] In addition to the LS2 channel, the MS-EVB method has been widely applied to studying a variety of other biomolecular PT processes, as reviewed in ref 7. In this section of the supplementary information, we compare the results of the MS-EVB3 method[5] to the SCC-DFTB and DFT methods. Comments on these comparisons are given in the table and figure captions.

System setup

The MS-EVB3 systems were equilibrated in the same way as in SCC-DFTB and DFT simulations, except that the SPC/Fw water model,[8] which is consistent with the MS-EVB3 method, was used instead of the TIP3P water model. For the production runs, the MS-EVB3 model was used to explicitly describe excess proton Grotthuss shuttling and charge delocalization. For the LS2 channel, the production run was simulated in the constant NVT ensemble for 300 ps at 300 K with the temperature maintained by a Nose-Hoover thermostat. The cutoff radius for LJ and real space electrostatic interactions was 12 Å employing a switching function starting at 10 Å for the LJ interactions. The long-range electrostatics were treated by Particle-Particle Particle-Mesh (PPPM) method[9] with an accuracy threshold of 10^{-4}. For the CNT, the production run was simulated in the constant NVT ensemble for 200 ps at 300 K using Nose Hoover thermostat. The cutoffs for LJ and real space electrostatic interactions were 10 Å, employing switching function starting at 8 Å for the LJ interactions. Long-range electrostatics were treated by Ewald summation with a relative accuracy threshold of 10^{-6}. The integration time step was 1 fs. All of the MS-EVB simulations were carried out by the RAPTOR software.[10]

Table S1. Average hydrogen bond relaxation times for hydrogen bonds within 4 Å of the excess proton CEC in protonated LS2 channel and CNT. Unlike SCC-DFTB, the MS-EVB3 method predicts slower hydrogen bonding dynamics around the excess proton CEC. The MS-EVB3 hydrogen bond relaxation time for the CNT is comparable to the DFT methods, while that in the LS2 channel is slower due to an apparent over-attraction between the 1st solvation shell around the excess proton CEC and the serine side chains.

Method	LS2 hydrogen bond relaxation time (ps)	CNT hydrogen bond relaxation time (ps)
BLYP-D	12.3	13.7
HCTH	20.4	10.1
B3LYP-D	10.7	NA
SCC-DFTB	1.73	1.4
MS-EVB3	64.9	8.1

Supplementary Figures:

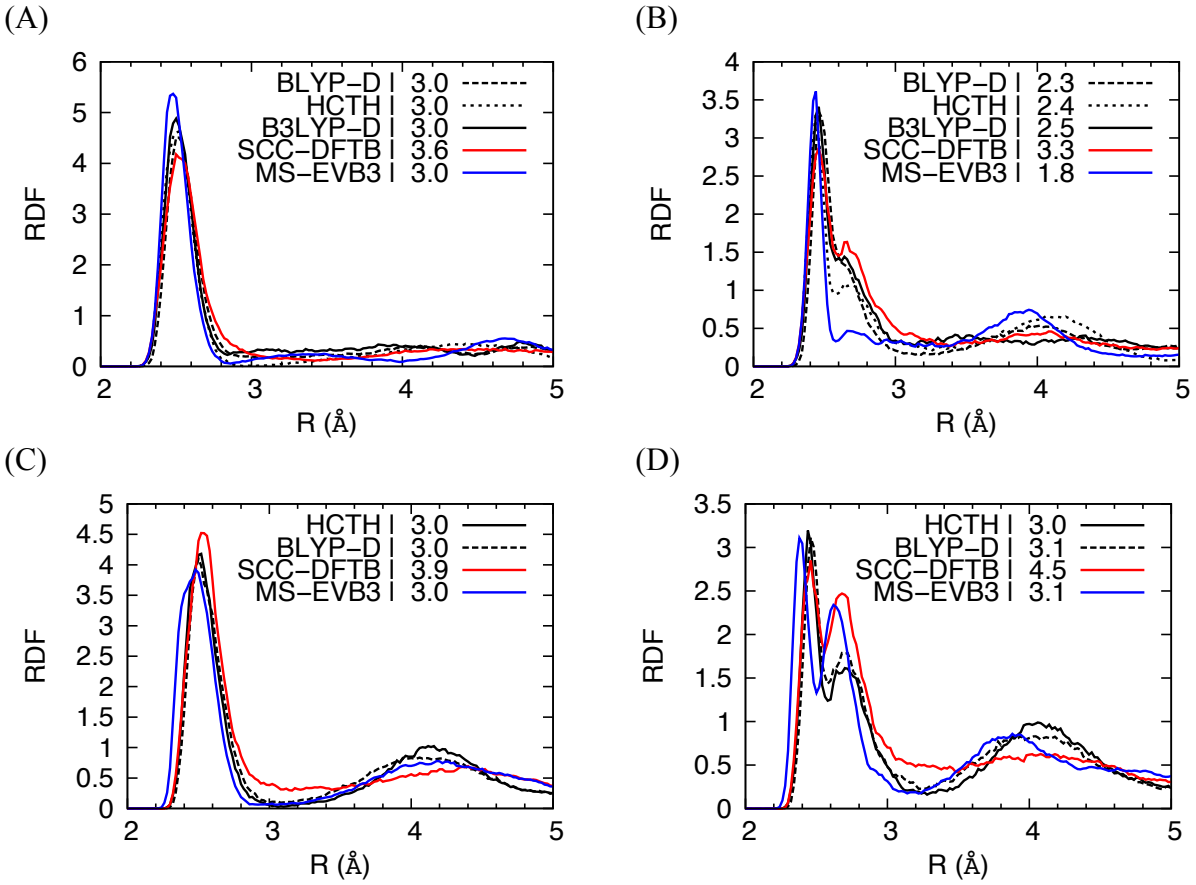

Figure S1. RDFs of (A) O*-O_w and (B) O_{1x}-O, where O includes both O* and O_w, in the protonated LS2 channel. RDFs of (C) O*-O_w and (D) O_{1x}-O, where O includes both O* and O_w, in protonated CNT. Coordination numbers are indicated in the legend. The MS-EVB3 method produces a more structured solvation shell compared to SCC-DFTB and is in better agreement with the DFT methods. The MS-EVB3 coordination numbers also agree better with the DFT methods, suggesting that the MS-EVB3 does not have the overcoordination issue demonstrated for SCC-DFTB. However, the MS-EVB3 O_{1x}-O RDF in LS2 shows an early drop of the first peak around 2.6 Å. This is caused by an apparent over-attraction between the first solvation shell of the hydronium and serine hydroxyl groups. The oxygen atom in serine side chain mostly replaces one of the water molecules coordinating O_{1x}. The RDF of O_{1x}-O including both O_w and O_s (not shown here) recovers the complete solvation shell of O_{1x}. It should also be noted that the QM/MM boundary may cause errors in the water-serine interactions in the DFT and SCC-DFTB results that are excluded in the MSEVB3 results where the entire system is treated on the same MM footing.

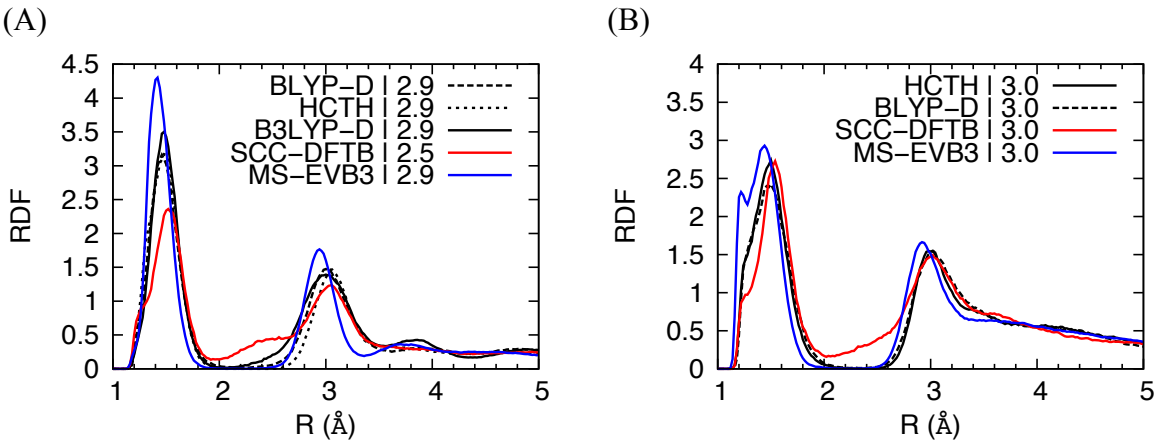

Figure S2. RDFs of H*-O_w in the protonated (A) LS2 channel and (B) CNT channels, with coordination numbers indicated in the legend. The MS-EVB3 method produces a density depletion region after the first peak in both systems and agrees better with the DFT methods than SCC-DFTB does, as demonstrated by the coordination numbers. The more narrow first solvation peak for MS-EVB3 in the LS2 system is again caused by the over-attraction between the hydronium first solvation shell and serine hydroxyl groups. The bimodal first peak the MS-EVB3 H*-O_w CNT RDF implies that the Zundel structure, in which two water molecules equally share an excess proton, occurs more frequently in MS-EVB3 simulation than in DFT. Note that it occurs much less frequently in the SCC-DFTB simulations. It should also be noted that many aspects of the MS-EVB3 model, including excess proton hydrated cluster energies and internal proton transfer barriers, were parameterized to a higher level of electronic structure than DFT (MP2). It is therefore not clear if these MS-EVB3 results are less accurate than the DFT ones.

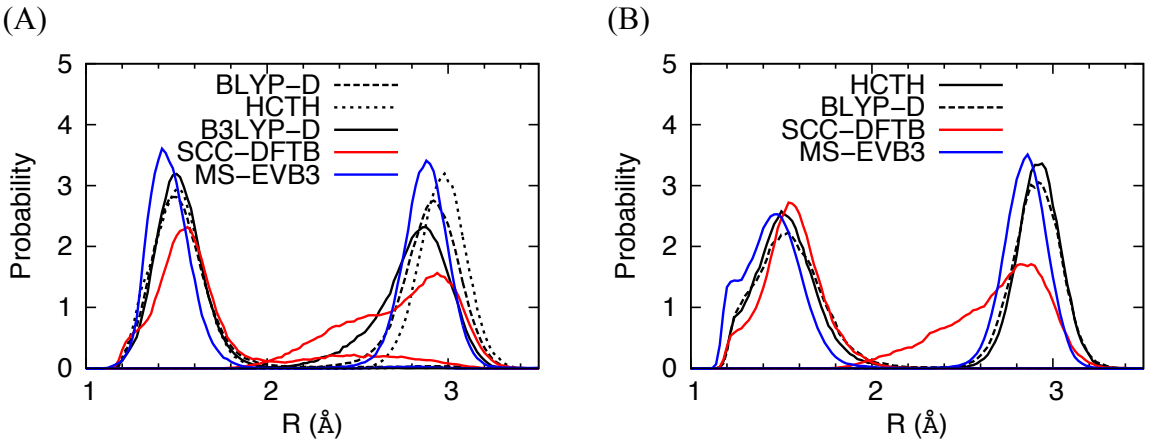

Figure S3. Distribution of H*-O_w distance in the protonated (A) LS2 and (B) CNT channels. The first and second peaks correspond to the closest and second closest O_w to H*. The MS-EVB3 method shows a similar distribution pattern to those described by the DFT methods: two distinct peaks without overlap in the middle. This indicates the absence of bifurcated hydronium hydrogen bonds and is in contrast to the results from SCC-DFTB in which the two peaks have obvious overlap. Again, the MS-EVB3 method predicts a splitting of the first main peak in the CNT system due to more frequent Zundel structures. See also the comment at the end of Fig. S2 caption.

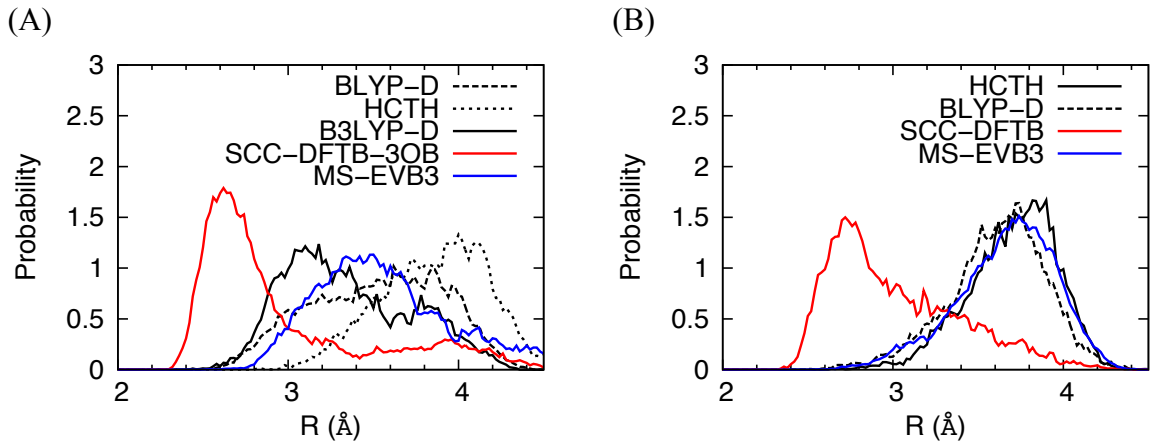

Figure S4. Distribution of the O*-O distance, where O is the closest atom not hydrogen bonded to O*, for the protonated (A) LS2 and (B) CNT channels. The MS-EVB3 method predicts similar distribution patterns to the DFT methods, indicating that for the majority of simulation the water molecules within the first solvation shell of excess proton are hydrogen bonded to the hydronium O*. This is in contrast to SCC-DFTB result, where the water molecules in the first solvation shell are not hydrogen bonded to O* for a significant amount of time.

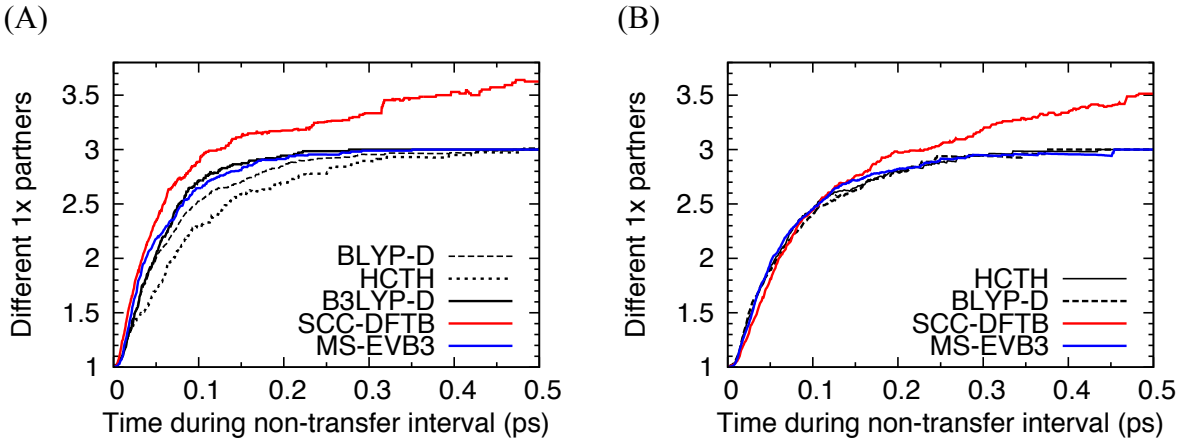

Figure S5. Average number of different O_{1x} partners to O* as a function of time during non-transfer intervals for the protonated (A) LS2 and (B) CNT channels. The MS-EVB3 method predicts a "special-pair dance" behavior[11] similar to the DFT results, with the special partner alternating among 3 nearby water molecules. This is in contrast with SCC-DFTB results, which predict more than 3 water molecules as special partners due to overcoordination of the excess proton.

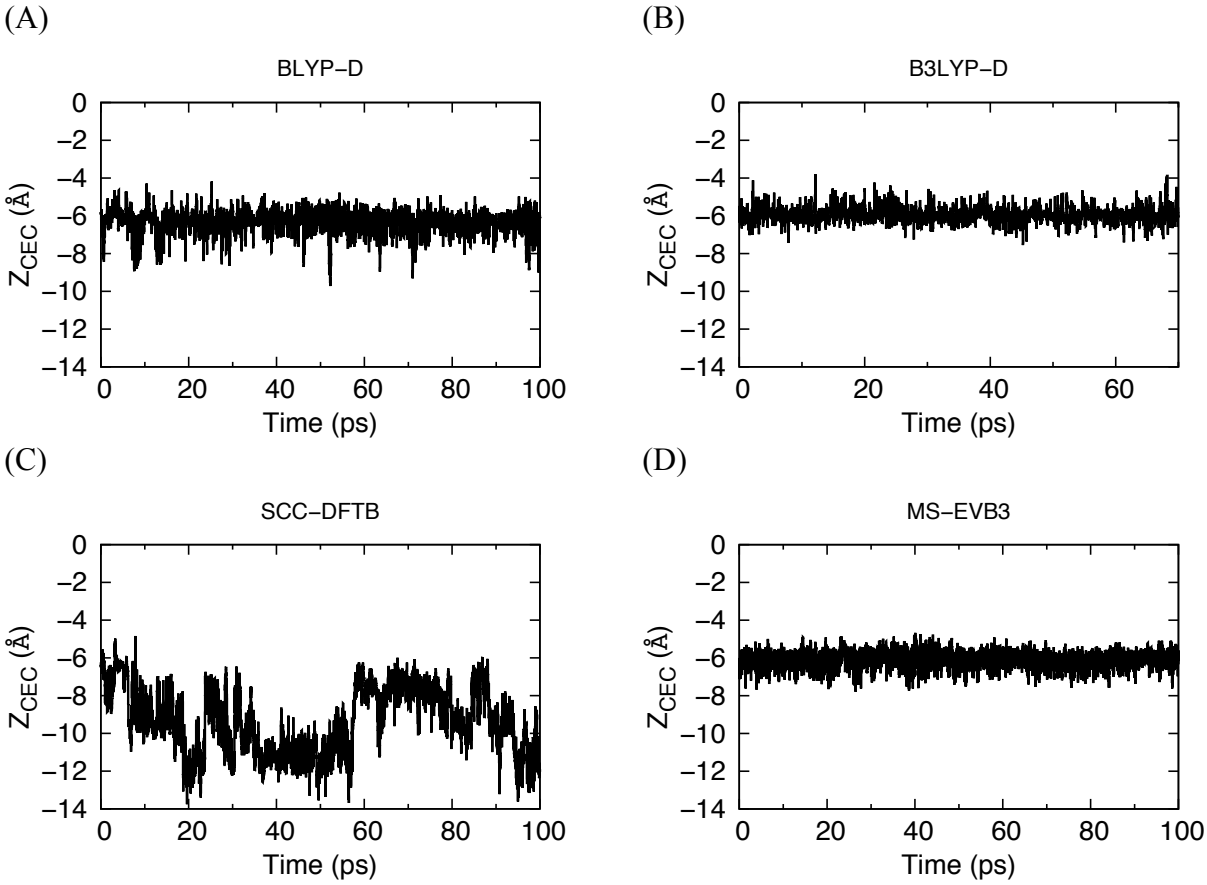

Figure S6. Z coordinate of the proton CEC as a function of time for unconstrained simulations using the (A) BLYP-D (B) B3LYP-D (C) SCC-DFTB and (D) MS-EVB3 methods. The excess proton CEC definition used in the MS-EVB3 simulation was the same as that used in the DFT and SCC-DFTB simulations, and the unconstrained simulation was run in the same way as described for DFT and SCC-DFTB in the main text. For MS-EVB3, the z coordinate fluctuates slightly around z = -6 Å, in agreement with the DFT simulations and with the previous MS-EVB free energy profile[2] that shows a distinct local minimum at z = -6 Å. In contrast, for SCC-DFTB the z coordinate deviates as far as 6 Å from z = -6 Å.

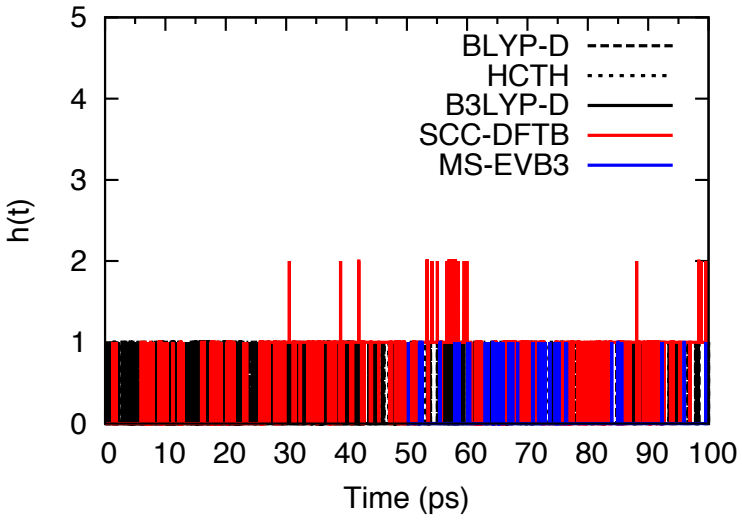

Figure S7. Proton hopping function[5] for the free excess proton simulation initiated from the constrained case where the center of excess charge is trapped at a wide region. In the MS-EVB3 simulation the hopping function fluctuates between 0 and 1, in agreement with the DFT results and in contrast to the SCC-DFTB result.

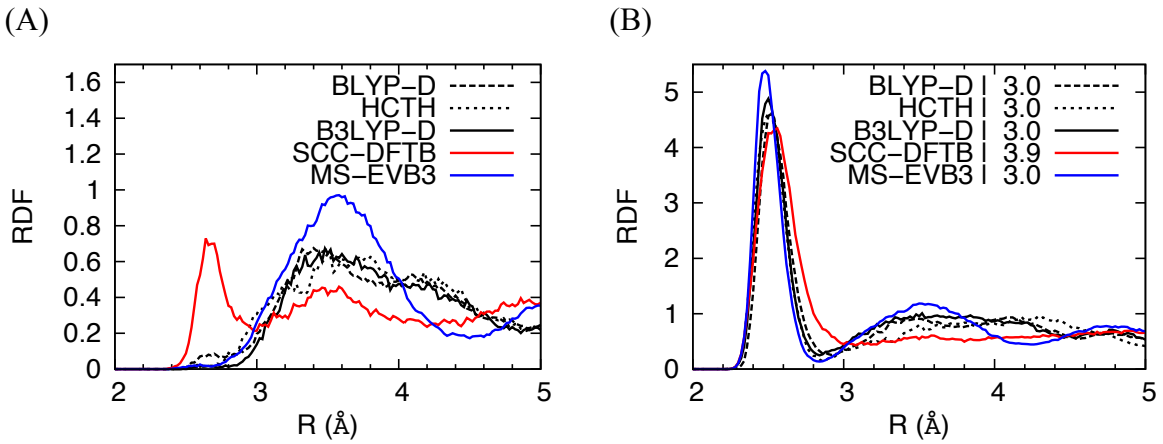

Figure S1. (A) RDF of O*-O_s in the protonated LS2 channel. The location of first peak predicted by MS-EVB3 agrees better with the DFT methods than the SCC-DFTB peak does, which indicates that the hydronium O* is not directly hydrogen bonded to serine O_s. However, the MS-EVB3 method predicts a more pronounced first peak than the DFT methods, which suggests an over-attraction between O* and O_s. (B) The RDF of O*-O, where O includes both O_w and O_s. The coordination number predicted by MS-EVB3 agrees with the DFT methods whereas that predicted by SCC-DFTB is larger.

Section S2: Additional results for proton solvation and transport in a narrower region of the LS2 channel

In this section, additional results for the proton solvation and transport in a narrow region of the LS2 channel (z = - 3.5 Å) are summarized. The results are in line with the conclusions for the wide region of the LS2 channel mentioned above and in the main text, i.e., the SCC-DFTB method yields an overcoordinated excess proton, weak hydrogen bonding around the excess proton CEC, and PT dynamics at odds with the DFT results, whereas the MS-EVB method yields similar results to the DFT methods.

Table S3. Average hydrogen bond relaxation time for hydrogen bonds within 4 Å of the excess proton CEC in the narrow region of the protonated LS2 channel.

Method	LS2 HB relaxation time (ps)
BLYP-D	21
HCTH	22
B3LYP-D	9
SCC-DFTB	0.8
MS-EVB3	6

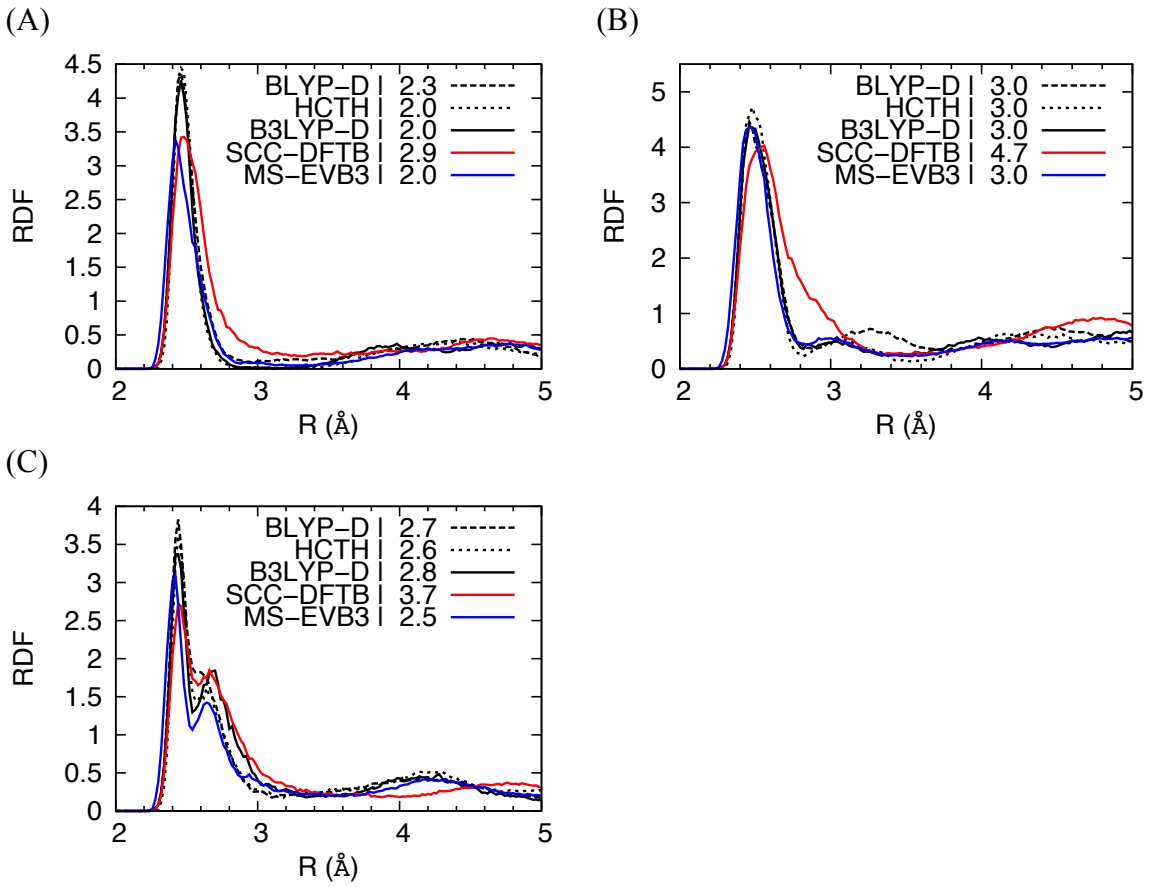

Figure S9. RDFs of (A) O*-O_w, (B) O*-O , and (C) O_{1x}-O , where O includes both O_w and O_s, in a narrow region of the protonated LS2 channel.

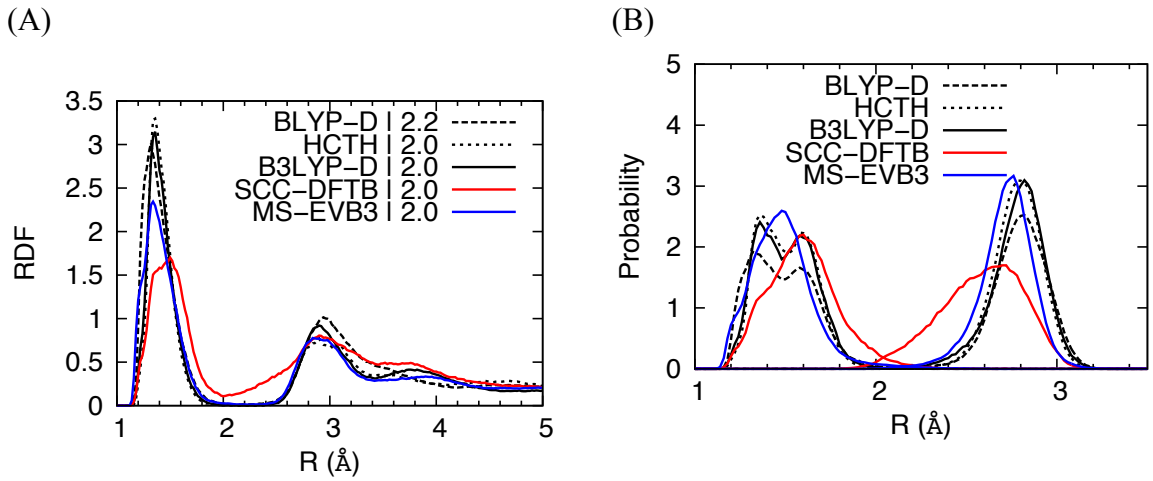

Figure S10. (A) RDF of H*-O_w in the narrow region of the protonated LS2 channel. (B) Distribution of H*-O distances, where O includes both O_w and O_s, in the narrow region of the protonated LS2 channel.

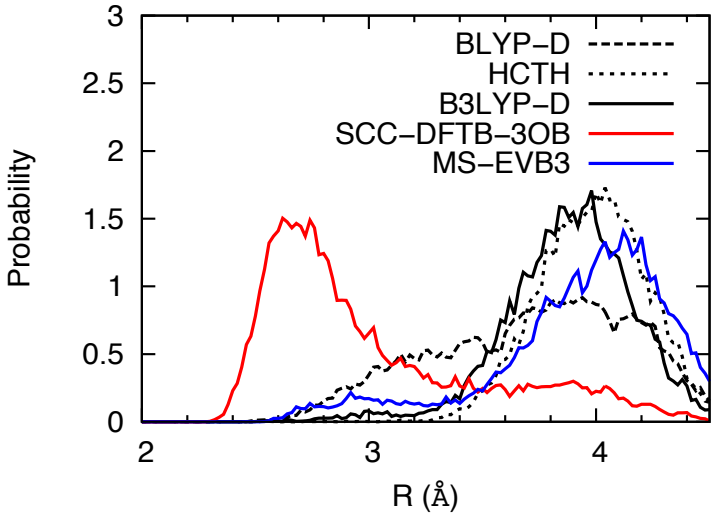

Figure S11. Distribution of the O*-O distance, where O is the closest atom not hydrogen bonded to O*, for a narrow region of the protonated LS2 channel.

Section S3: B3LYP-D/GTH-DZVP improves description of forces and energies over BLYP-D/GTH-TZV2P

To illustrate the advantage of employing B3LYP-D/GTH-DZVP over BLYP-D/GTH-TZV2P, 10 uncorrelated frames were chosen from 80 ps of the BLYP-D/MM simulation, and the forces on each QM oxygen atom was calculated with BLYP-D/GTH-TZV2P and B3LYP-D/GTH-DZVP and plotted against that calculated with B3LYP-D and a large basis set (6-31++G**; Figure S12). The B3LYP-D/GTH-DZVP leads to a root mean squared error (RMSE) that is 6.56 kcal/(mol Å) less than that of BLYP-D/GTH-TZV2P, suggesting that B3LYP-D/GTH-DZVP does indeed offer an improved description of the forces on QM atoms over BLYP-D/GTH-TZV2P. In addition, using B3LYP-D/GTH-TZV2P as benchmark we calculated the energy error for BLYP-D/GTH-TZV2P and B3LYP-D/GTH-DZVP, averaged over the same set of configurations used in the force comparison. The B3LYP-D/GTH-DZVP reduces the energy error of BLYP-D/GTH-TZV2P from 0.65 hartree to 0.27 hartree, and therefore provides significant improvement over the latter. The TZV2P basis set used in the benchmark was shown to predict converged structural properties of bulk water. [12]

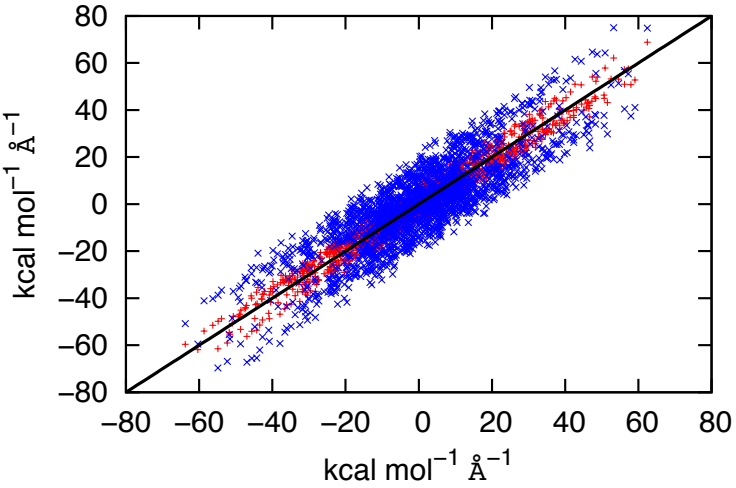

Figure S12. Scatter plot of forces from B3LYP-D/6-31++G** vs B3LYP-D/GTH-DZVP (red) and BLYP-D/GTH-TZV2P (blue). The B3LYP-D/GTH-DZVP provides more accurate forces than BLYP-D/GTH-TZV2P.

Supplementary References

1. Wu, Y. J.; Voth, G. A., A computer simulation study of the hydrated proton in a synthetic proton channel. *Biophys. J.* **2003,** *85*, 864-875.
2. Wu, Y. J.; Ilan, B.; Voth, G. A., Charge delocalization in proton channels, II: The synthetic LS2 channel and proton selectivity. *Biophys. J.* **2007,** *92*, 61-69.
3. Schmitt, U. W.; Voth, G. A., The computer simulation of proton transport in water. *J. Chem. Phys.* **1999,** *111*, 9361-9381.
4. Day, T. J. F.; Soudackov, A. V.; Cuma, M.; Schmitt, U. W.; Voth, G. A., A second generation multistate empirical valence bond model for proton transport in aqueous systems. *J. Chem. Phys.* **2002,** *117*, 5839-5849.
5. Wu, Y. J.; Chen, H. N.; Wang, F.; Paesani, F.; Voth, G. A., An improved multistate empirical valence bond model for aqueous proton solvation and transport. *J. Phys. Chem. B* **2008,** *112*, 467-482.
6. Voth, G. A., Computer simulation of proton solvation and transport in aqueous and biomolecular systems. *Acc. Chem. Res.* **2006,** *39*, 143-150.
7. Swanson, J. M. J.; Maupin, C. M.; Chen, H. N.; Petersen, M. K.; Xu, J. C.; Wu, Y. J.; Voth, G. A., Proton solvation and transport in aqueous and biomolecular systems: Insights from computer simulations. *J. Phys. Chem. B* **2007,** *111*, 4300-4314.
8. Wu, Y. J.; Tepper, H. L.; Voth, G. A., Flexible simple point-charge water model with improved liquid-state properties. *J. Chem. Phys.* **2006,** *124*.
9. Hockney, R. W.; Eastwood, J. W. j. a., *Computer simulation using particles*. McGraw-Hill International Book Co.: New York, 1981; p xix, 540 p.
10. Yamashita, T.; Peng, Y. X.; Knight, C.; Voth, G. A., Computationally Efficient Multiconfigurational Reactive Molecular Dynamics. *J. Chem. Theory. Comput.* **2012,** *8*, 4863-4875.
11. Markovitch, O.; Chen, H.; Izvekov, S.; Paesani, F.; Voth, G. A.; Agmon, N., Special pair dance and partner selection: Elementary steps in proton transport in liquid water. *J. Phys. Chem. B* **2008,** *112*, 9456-9466.
12. VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M., The influence of temperature and density functional models in ab initio molecular dynamics simulation of liquid water. *J. Chem. Phys.* **2005,** *122*.