Probing the role of acid site distribution on water structure in aluminosilicate zeolites: insights from molecular dynamics

*Mingze Zheng and Brandon C. Bukowski**

Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, Maryland 21218, United States

ABSTRACT: Water plays a pivotal role in numerous chemical processes, especially in the production of fuels and fine chemicals derived from bio-based feedstocks. Zeolites are porous catalysts used extensively due to their shape-selective adsorption and confinement interactions; However the kinetics of zeolite-catalyzed reactions are significantly impacted by the presence of water, which may affect product selectivity and intrinsic rate constants depending on transition state polarity. In this study, we employed machine learning force fields (MLFFs) to accelerate *ab initio* molecular dynamics (AIMD) simulations and enhance the phase space exploration of water configurations in a model Brønsted acid zeolite, H-AFI. We interrogated the structure of adsorbed water based on the Si/Al ratio and acid site distribution to disentangle the impact of acid site density and distribution on water matrix organization as a function of water loading. We integrated adsorption thermodynamics, vibrational spectroscopy simulations, and local density maps to interrogate the spatial orientation of adsorbed water clusters and their degree of hydrogen bonding. Our analysis unveiled the intricate interplay between zeolite structure, Brønsted acid site location, and water where spatially disparate acid sites nucleate extended clusters that span siliceous regions of the zeolite. We found that the length scale of ordered water regions is directly related to the Si/Al ratio and spatial distribution of Al sites. These findings provide insights into the molecular-level structure of water in microporous aluminosilicate micropores and demonstrate how acid sites can be used to control water activity which has applications to heterogeneous catalysis and adsorptive separations.

KEYWORDS: Zeolite, porous material, water, machine learning force fields, density functional theory, molecular dynamics

INTRODUCTION

Zeolites are nanoporous solids characterized by high crystallinity and hydrothermal stability that have been studied extensively due in part to their role as catalysts and sorbents for the petrochemical industry,[1–5] environmental remediation,[6–11], and sustainable chemical synthesis.[12–14] Notably, aluminosilicate zeolites possess anionic lattice charges resulting from aluminum heteroatom substitutions that are balanced by cations such as protons, which can function as catalytic active sites. The distinct coordination and confinement environments of aluminum atoms within the zeolite framework allow for cations and guest adsorbates to be accommodated in diverse void spaces, which are determined by the underlying pore architecture. The catalytic reactivity of a framework proton is determined by the stability of a proton on a bridging oxygen, which acts as a conjugate base (AlO_4^-).[15]

The distribution and location of intracrystalline aluminum sites within zeolite frameworks impact active site accessibility and product selectivity.[16] The strategic placement of aluminum enables targeted chemical transformations and shape-selectivity for the formation of desired products.[17,18] Understanding the intricate relationship between acid site location and reactivity is pivotal for designing and optimizing zeolite catalysts. This understanding is facilitated by advanced characterization techniques and computational modeling, leading to highly efficient catalyst development for diverse catalytic processes such as methanol-to-olefins (MTO)[19,20] and Fischer–Tropsch synthesis.[21–23] Furthermore, acid site location also significantly affects the structure and energetics of guest molecules, including solvents, within the zeolite framework.

Water often has a prominent role in zeolites, especially in hydrophilic aluminosilicate zeolites, due to its large heat of adsorption and its frequent formation as a product in oxygenate chemistries.[12] Additionally, water acts as a solvent, which can directly affect the stability of confined transition states, especially those involving renewable feedstocks and polar reaction intermediates within the zeolite.[24–26] The protonation of water molecules can give rise to various structures, including Zundel and Eigen structures and the interplay between these structures further contributes to the rich behavior of water in zeolites.[27,28] However, the presence of water can also hinder reaction rates.[29] Water molecules tend to form clusters, limiting the accessibility of reactants to active sites and consequently reducing the reaction rate.[30] Moreover, coadsorbed water may participate in transition states, affecting intrinsic kinetics.[31] Despite its ubiquitous role, fundamental insights into water structuring in acid zeolites remain challenging due to the formation of hydrogen bonding networks and complex cluster structures.[27,32–34] These characteristics also complicate experimental characterization techniques and contribute to the complexity of understanding water's behavior in zeolites.

Simulation methods play a critical role in this endeavor. Two commonly used approaches are classical parameterized force fields (FF) and Density Functional Theory (DFT), the latter explicitly includes electron interactions. While it is true that FF methods are primarily employed for tasks like diffusion or adsorption, reactive force fields have been developed for various classes of reactions.[35–37] These specialized force fields aim to bridge the gap between computational efficiency, as seen in traditional FF methods, and the ability to model bond-breaking and bond-formation processes, characteristic of catalytic reactions. On the other hand, DFT-based first-principles (FP) calculations provide an electron-level description of electronic structure and energetics, making them a crucial tool for studying

catalysis. However, it is important to acknowledge that DFT calculations are computationally more demanding.

Efficiently predicting interatomic potentials is crucial to understanding and modeling complex molecular systems. Traditional methods for fitting interatomic potentials, such as quantum chemical calculations including DFT are not only computationally intensive but demand substantial time and resources. Machine learning force fields (MLFFs) describe the potential energy as a function of atomic structure descriptors, and their parameters are optimized to reproduce training data obtained from FP calculations. MLFFs have been successfully applied to predict the properties of various systems, including phase transition[38–40], interfaces,[41] adsorption,[42] and reactions[43].

Machine learning potentials have also been developed to understand interactions between water and silica polymorphs including zeolites. Roy et al. developed reactive potentials for silicates using an equivariant neural network that captured water deprotonation and silica dimerization reactions.[44] Neural network potentials for aluminosilicate zeolites have been developed by Erlebach et al.[45] which included interactions between water molecules and Brønsted acid sites using ten different zeolite structures. Saha et al. used these potential to quantify how Si/Al ratio affects the proton solvation and water dynamics in faujasite (FAU) zeolites.[46] By understanding the molecular interactions between the solvent and zeolites, simulations can open new avenues for optimizing reaction kinetics to enable the development of more efficient and selective catalysts.

In this work, we examined adsorbed water behavior within AFI aluminosilicate zeolites based on their acid site distribution and Si/Al ratio (acid site density) as a model zeolite framework to determine the impact of these factors on water organization and hydrogen bonding interactions as a function of water loading in the framework. Our investigation focused on developing fundamental insights into how aluminum density and distribution affect water organization as a function of water loading in the framework. To enhance the sampling efficiency for the large configurational space of possible water structures, we adopted a strategy that leverages MLFFs to accelerate *ab initio* molecular dynamics (AIMD) simulations. The results unveil the role of specific aluminum next-nearest neighbor interactions on both the thermodynamics of water adsorption and the distribution of adsorbed water clusters. The findings underscore the relationship Brønsted acid site locations and protic solvents, which are relevant for liquid phase catalysis as well as oxygenate chemistries that co-produce water. These results have broad applications in designing catalysts with tunable control over adsorbed water matrices and consequently improved selectivities and kinetics for aqueous phase reactions or those that co-produce water such as dehydration from bio-based feedstocks.

METHODOLOGY

Model Construction. A CIF for zeolite AFI was obtained from the International Zeolite Association (IZA) structure database.[47] One and two aluminum-substituted AFI zeolites (denoted here as H-AFI and 2H-AFI, labeled based on the number of aluminum heteroatoms and protons per unit cell) were built using the atomic simulation environment (ASE) which was also used to place adsorbate water molecules at their initial positions.[48] For 2H-AFI, we constructed four different configurations based on the different aluminum locations and noted them as 2H-AFI-1, 2H-AFI-2, 2H-AFI-3, and 2H-AFI-4. AFI 1×1×2 supercells with Si/Al ratios of 11 and 23, which correspond to 4 and 2 aluminum per unit cell (denoted u.c.$^{-1}$), respectively, were constructed by replicating the optimized 2H-AFI-1 unit cell.

First Principles Calculations. All density functional theory (DFT) based ab initio molecular dynamic (AIMD) simulations were performed by Vienna Ab initio Simulation Package (VASP)[49–52] with exchange correlation function of Perdew-Burke-Ernzerhof (PBE)[53] in conjunction with D3 correction method,[54] and Becke-Johnson (BJ) was used as the damping function. The cutoff energy was set to 500 eV, while the tolerance for the self-consistent field (SCF) was set to 1×10^{-5} eV. The time step was set to 0.5 fs. A Nosé-Hoover thermostat[55] was used to simulate the NVT ensemble at a temperature of 300 K. Example INCAR files are included in the Supporting Information.

The original AFI CIF obtained from IZA was re-optimized using VASP. Force criteria were set to 0.03 eV/Å with all other parameters the same as described above. The cell constants, a, b, and c were adjusted in VASP with the optimized values for a and b being 13.68 Å and c being 8.60 Å and were kept constant in subsequent simulations. The volume deviation of these optimized cell constants from the original a=b=13.83 Å and c=8.58 Å is 2.07%. Aluminum atoms replaced silicon atoms to form acid sites, and protons were added to the oxygen atom next to each aluminum atom in the 12-membered ring to balance the charge and expose the proton to adsorbates. We calculated the saturation loading of water using the accessible volume (199.88 Å^3) of zeolite AFI which accounts for 14.07% of the total volume (1420.60 Å^3), leading to an approximate equilibrium loading of 6 H_2O u.c.$^{-1}$. Once we determined the water loading, the 2H-AFI configurations with water molecules were optimized by VASP using the same simulation conditions to obtain the initial configurations for the following MD simulations. POSCAR-formatted structures are included in the Supporting Information.

On-the-fly Machine Learning Force Fields. Investigating the behavior of water in confined environments requires simulations conducted over large timescales that include electronic interactions due to proton shuttling that occurs at Brønsted acid sites. Simulations based on FP calculations, such as AIMD, explicitly calculate the forces acting on each ion based on their electronic structure, resulting in slower computational speeds. Conversely, simulation methods based on empirical parameters like force field require solid fundamental knowledge of the given system to investigate the relevant interactions at an atomic level. The VASP on-the-fly MLFF[38,56,57] is a force field method that derives its parameters from first-principles calculations and employs a Bayesian active learning algorithm. The VASP on-the-fly MLFF incorporated radial and angular descriptors and basis functions,[56] utilizing the Gaussian Approximation Potentials (GAP) method[58] and Smooth Overlap of Atomic Positions (SOAP) descriptors.[59]

After obtaining the 0 K optimized aqueous 2H-AFI structures, we employed VASP on-the-fly MLFF simulations to 2H-AFI-1, 2H-AFI-2, and 2H-AFI-3 with all other simulation conditions the same as we used in the AIMD simulations. The total duration of each MLFF simulation was 215 ps.

The accuracy of a trained machine learning force field is related to the frequency of FP calculations. More frequent FP calculations result in higher accuracy but significantly increase the computational demand. FP calculations were only carried out when the Bayesian error for the forces and spilling factors exceeded the threshold.[57,60] In this work, the MLFFs underwent

refinement after each training step. Refinement occurred when the Bayesian error of the force exceeded the average maximum Bayesian error of the force observed over the last 10 training steps and the relative standard deviation was less than 0.2 in normalized units.[57] Example force field files and INCARs for our training are included in the Supporting Information.

We assessed the accuracy of our MLFF simulations by both visually inspecting the obtained structures as well as quantifying the error in energies and forces. In each on-the-fly MLFF simulation, we conducted a thorough examination of the system geometry which included the AFI framework and water configurations. This ensured that our models did not exhibit unphysical bond formation or bond cleavage events. We calculated the mean absolute error (MAE) for each aluminum distribution and water loading to quantify the energy and force error of our potential. The MAE for energies across all models was 24 meV/atom and the MAE for forces across all models averaged 0.81 eV/Å. Further comprehensive data is available in **Figure S6** and **Figure S7** for reference.

We performed 10 ps AIMD simulations of both 2H-AFI-2 at 2 H_2O u.c.$^{-1}$ and 2H-AFI-1 at 5 H_2O u.c.$^{-1}$ to compare the distribution of adsorbed water molecules compared to those predicted by our MLFF models. These are shown in **Figure S4**. The AIMD and MLFF models show similar adsorbed structures indicating that the MLFF models correctly predict the adsorbed water configuration in addition to their energies.

We conducted on-the-fly MLFF training simulations on three different aluminum distributions (2H-AFI-1, 2H-AFI-2, and 2H-AFI-3) while varying the water loading from 1-6 H_2O u.c.$^{-1}$ Our calculations indicated that the MLFF model of 2H-AFI-2 at a loading of 6 H_2O u.c.$^{-1}$ exhibited the lowest energy and force errors (with MAEs of 7.8 meV/atom and 0.79 eV/Å, respectively). We then used this MLFF model for H-AFI and 2H-AFI-4.

Aluminum Incorporation Energy. The total aluminum incorporation energy was calculated by the following equation (1):

$$E_{inc,total} = (E_{xH-AFI} + xE_{Si(OH)_4}) - \left(E_{Si-AFI} + xE_{Al(OH)_3} + xE_{H_2O,g}\right) \quad \text{(eq. 1)}$$

Here x denotes the number of substituted aluminums, E_{xH-AFI} and E_{Si-AFI} represent the potential energy of aluminum incorporated AFI and siliceous AFI, respectively. $E_{Si(OH)_4}$, $E_{Al(OH)_3}$, and E_{H_2O} denote the potential energy of $Si(OH)_4$, $Al(OH)_3$, and H_2O in the gas phase. The configurations of AFI, H-AFI, and 2H-AFI were obtained from the geometry optimizations described above.

Adsorption Thermodynamics. The water adsorption energy per water molecule at 300 K averaged across the 200 ps simulation production run is presented in **Figure S1**. The average water adsorption energy per water molecule in 2H-AFI-1, 2H-AFI-2, and 2H-AFI-3 by equation (2) and shown in **Figure S2a**:

$$\langle E_{ads} \rangle = \langle E_{nW+Z} \rangle - \langle E_Z \rangle - n_{water}\,\overline{\langle E_W \rangle} \quad \text{(eq.2)}$$

Here, $\langle E_{nW+Z} \rangle$ is the average total energy of the zeolite with n adsorbed water over the production run. $\langle E_Z \rangle$ is the average potential energy of the framework, obtained from 10 ps MLFF simulations using the predicted force field. The force field we used here was based on the equilibrated system of the corresponding system with adsorbed water. $\overline{\langle E_W \rangle}$ represents the average intensive potential energy of a bulk water system comprising 45 water molecules, obtained from a 10 ps AIMD simulation in units of energy H_2O^{-1}. The potential energy of each water molecule was normalized and calculated using equation (3) to account for various interactions between water molecules, including Coulomb and van der Waals interactions.[61]

$$\overline{\langle E_W \rangle} = \frac{\langle E_{W,\ bulk,\ aimd,\ n=45} \rangle}{45} \quad \text{(eq.3)}$$

To make meaningful comparisons for adsorption energies, it is essential to establish uniform reference states. We performed additional 10 ps MD simulations from the equilibrated structures using our trained potentials for all water loadings. The reference energy of water was also calculated by averaging a 10 ps prediction simulation utilizing the same model as above to ensure a consistent thermodynamic reference state using the same trained potentials. These simulations used the same thermostat, temperature, and timestep as described previously.

The adsorption energy per water molecule at 0 K was also computed for a comparison shown in **Figure S2b** by equation (4):

$$\overline{E_{ads,0\ K}} = \left(E_{W+Z,0\ K} - E_{Z,0\ K} - E_{W,0\ K}\right)/n_{water} \quad \text{(eq.4)}$$

$E_{Z,0\ K}$ and $E_{W,0\ K}$ here are the potential energy of 2H-AFI and a single water molecule at 0 K from geometry optimizations. The calculated water adsorption energies at 0 K shown in **Figure S2b** differ from 300 K due to the finite-temperature reorganization of the water molecules. To determine the 0 K ground state adsorption configuration, we obtained at least 6 water configurations from each MD simulation with the lowest potential energy and geometry optimized each of them. The configuration with the lowest geometry-optimized potential energy was then used as the corresponding $E_{W+Z,0\ K}$ for that system.

Radial Distribution Functions. Aluminum-water radial distribution functions (RDFs) were computed by obtaining the interatomic distances between the aluminum and the oxygen atoms belonging to water molecules.

Water Density Plots. We generated time-averaged heatmaps that illustrate water density distribution by slicing perpendicular to the AFI center pore which corresponds to the [001] plane. Higher brightness in the heatmaps indicates a higher probability of water molecules visiting that region. Each heatmap encompasses simulation data spanning from 15 ps to 215 ps. To improve clarity, we centered the AFI pore, making it easier to visualize the water distribution. In these heatmaps, water positions are represented by the oxygen atoms within the water molecules.

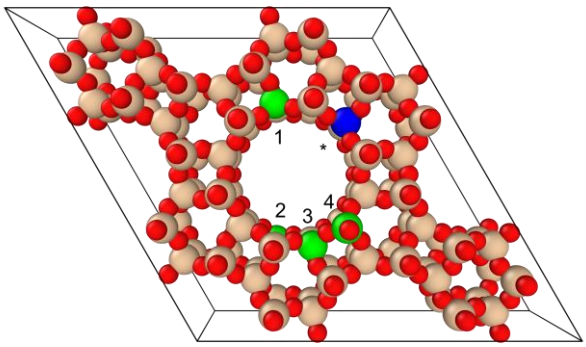

Figure 1. A 2x2x1 supercell of zeolite AFI with O colored red, and Si colored tan. Plausible locations for the next-nearest neighbor aluminum (marked as 1, 2, and 3) and the third nearest neighbor aluminum (marked as 4) substitutions are labeled in green relative to the initial aluminum substitution labeled in blue.

RESULTS AND DISCUSSION

Zeolite AFI Structure. AFI is a one-dimensional framework structure characterized by identical T-sites. **Figure 1** depicts a 2 by 2 by 1 supercell to illustrate aluminum atom positions, although all simulations were conducted using a single unit cell. A potential aluminum atom location in H-AFI is colored blue and labeled by a star. A second aluminum atom was placed relative to the position of the first aluminum. Aluminum distribution in zeolites must obey Löwenstein's rule to avoid Al-O-Al bonds, which could lead to highly unstable structures.[62] We first explored three aluminum locations (represented by green atoms and labeled 1 to 3) for the second aluminum atom, which are next-nearest neighbors to the first aluminum substitution. These configurations were selected to represent unique combinations of aluminum positions relative to the anchoring initial aluminum. We note that the AFI unit cell can be divided into two SiO_2 layers along the z-axis. Specifically, the first and the third positions of the second aluminum atom align with the same layer as the anchor aluminum, while the second position, although having a similar Al-Al closest geometric distance as the third position, sits in a different layer.

To increase the range of Al-Al distances while maintaining Löwenstein's rule, we explored a fourth aluminum position. Here the second aluminum site was chosen to be the third nearest neighbor to the anchor aluminum and situated in the same layer as the first and third positions. **Table 1** presents the distances between two aluminum atoms as both the minimum geometric distance, and distance across the AFI center pore. For all further analysis, we used the Al-Al distance across the pore as that measures the interaction length for water molecules within a single pore.

The single T-site characteristic of AFI makes it a model material to interrogate site distribution effects by removing the local structure differences around each Al site. We calculated Al incorporation energies and found that they differ by at most 0.14 eV (**Table 1**) indicating that the thermodynamics of heteroatom substitution do not depend on which Si atom we choose.

Table 1. Aluminum-aluminum distances and their incorporation energy calculated for each acid site distribution considered.

Al distribution	Al-Al direct distance (Å)	Al-Al distance across the pore (Å)	Incorporation energy (eV)
2H-AFI-1	5.55	5.55	-2.38
2H-AFI-2	5.00	11.50	-2.51
2H-AFI-3	4.98	9.63	-2.49
2H-AFI-4	8.01	8.20	-2.37

Water adsorption at different acid site distributions. We found that the water adsorption energy varies significantly with the water loading and acid site distribution, which can be attributed to the acid site distribution modulating water molecule reorganization. We first interrogated the role of aluminum distribution by calculating the time-averaged adsorption energies from our trained machine learning potentials. **Figure 2a** shows the variation of the water adsorption energy per water molecule as a function of the Al-Al distance distance measured across the pore. Our results indicate that water adsorption becomes less favorable as the Al-Al distance increases for intermediate water loadings between 2 - 4 H_2O u.c.$^{-1}$ (~33-66% of the saturation loading). At the lowest loading of 1 H_2O u.c.$^{-1}$, the Al distribution is irrelevant as only one acid site can be populated by a given water molecule. As a result, we see that the adsorption energy for 1 H_2O u.c.$^{-1}$ is invariant with the Al-Al distance. Close to the saturation loading of 5 – 6 H_2O u.c.$^{-1}$ the adsorption energy per water molecule is also invariant with Al-Al distance. This is reasonable as at saturation there are sufficient water molecules to stabilize protonated clusters regardless of the Al proximity. At intermediate water loadings (2 – 4 H_2O u.c.$^{-1}$) we see that the Al-Al distance plays a role that implicates the importance of site proximity in stabilizing water clusters.

We observed that the magnitude of the water adsorption energy decreased as the water loading increased. We hypothesize that two primary factors significantly influence the water adsorption energy: the direct adsorption of water at the acid site and hydrogen bond formation. In the case of 1 - 3 H_2O u.c.$^{-1}$, adsorption is dominated by electrostatic interactions with the framework proton. From 4 - 6 H_2O u.c.$^{-1}$, the formation of a hydrogen bond network among water molecules contributes to stabilizing different adsorbate configurations, diminishing the primary role of direct adsorption when normalized per water molecule. At higher water loadings, the dominant factor in adsorption energy is the shift from direct interactions of water molecules with the acid site to intermolecular interactions, such as the formation of hydrogen bonds, proton transfer, and interatomic interactions.[28] This observation is further supported by **Figure 2b**, which includes a snapshot of 2H-AFI-2 at the saturation loading where we observed extended hydrogen bonding structures.

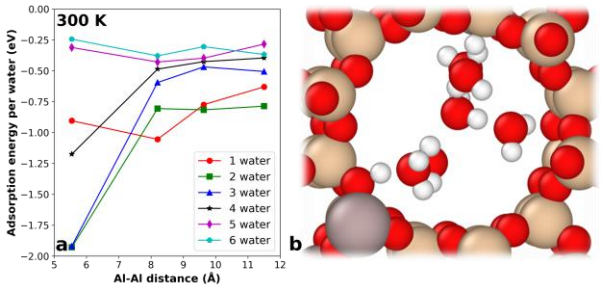

Figure 2. (**a**) Water adsorption energy per water molecule with the increasing Al-Al distance across the AFI center pore. (**b**) Snapshot of the water cluster in AFI center pore from 2H-AFI-2 with 6 H_2O u.c.$^{-1}$.

Adsorbed water structure. To complement our thermodynamic analysis, we tracked the position of water molecules over time in our simulations. The behavior of water molecules within the AFI center pore was studied as a function of water loading. **Figure 3** shows heatmaps that illustrate the local density of water within the AFI center pore. A subset of water conditions are shown in **Figure 3**, with the complete set of water density maps included in **Figure S3**. Brighter regions of the density map indicate where water molecules are more frequently located, suggesting the presence of favorable adsorbed water structures within those regions. For water loadings below 3 H_2O u.c.$^{-1}$, we find that water molecules cluster around the acid sites throughout the duration of the simulation. As the water loading increases, the water distribution extends throughout the central pore. At the saturation loading, we observe that water molecules access the full volume of the AFI center pore space regardless of the aluminum distribution.

To illustrate the effect of changing water loading, we chose a representative aluminum distribution (2H-AFI-2) and plotted the density heat maps from 1-6 H_2O u.c.$^{-1}$ (**Figure 3a-d**). We see that a single water can adsorb at a single site with a single bright region localized near the Al heteroatom indicated by the

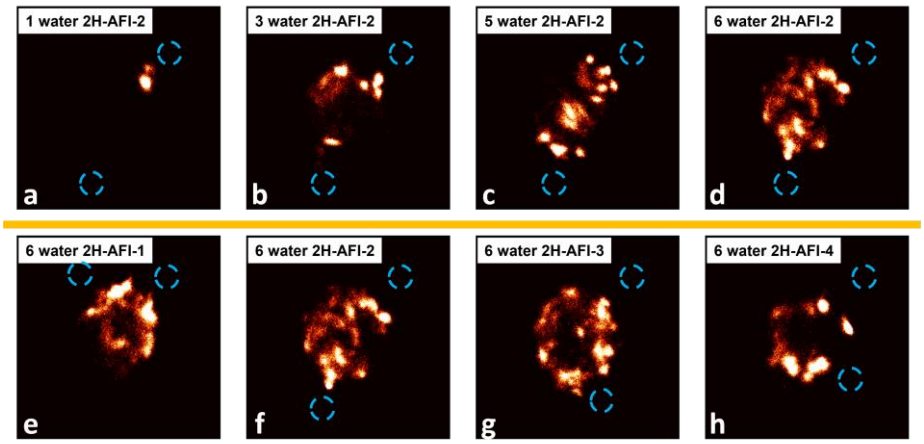

Figure 3. Time-averaged heatmaps of the water density distribution in AFI center pore viewed along the [001] plane. The zeolite framework is not shown. (a) to (d) are 2H-AFI-2 with one, three, five and six H_2O u.c.$^{-1}$ and (e) to (h) are 2H-AFI-1 to 2H-AFI-4 at the saturation water loading. Blue circles indicate the position of Al substituted heteroatoms. A complete set of density maps at all water loadings and aluminum distributions are included in **Supplementary Figure S3**.

blue circle. At 3 H_2O u.c.$^{-1}$ water molecules populate both Al sites as well as the central pore. At 5 H_2O u.c.$^{-1}$, an extended water cluster that spans both Al sites emerges which then saturates the accessible pore volume at 6 H_2O u.c.$^{-1}$.

At the saturation loading, we observed differences in the adsorbed water structure depending on the Al distribution. We observed that larger Al-Al spacing promoted the formation of broad water clusters, contributing to the stabilization of the water network within the zeolite AFI center pore (**Figure 3e-h**). For the shortest Al-Al distance (2H-AFI-1) the local density of water is more concentrated proximal to the two Al sites with minimal exploration of the siliceous region along the outer pore wall. This differs from the other distributions where distinct water clustering is observed which encompasses the full pore volume. The observed differences in equilibrium water density in conjunction with adsorption thermodynamics indicate that Al distribution can modulate both the thermodynamic stability and structure of adsorbed water.

Radial distribution functions of adsorbed water. We calculated radial distribution functions (RDFs) for the distribution of oxygen atoms in water molecules to Al heteroatoms as an additional measure of water clustering. RDFs for each water loading and Al distribution are provided in **Figure S5**. The RDFs support the density heatmap conclusions that water molecules occupy the central pore at high water loadings, and that the proximal Al distribution (2H-AFI-1) nucleates a localized water cluster as compared to the disparate Al distributions (2H-AFI-2,3,4). We also compared RDFs for a single Al heteroatom which exhibits a greater degree of peak broadening than the paired Al sites indicating weakly nucleated water structures. We calculated the first peak position for each RDF (**Figure 4a**) which portrays an overall increasing trend as the water loading increases.

We also compared RDF peak positions as a function of Al-Al distance. **Figure 4b** shows that at the saturation loading of 6 H_2O u.c.$^{-1}$, the position of the first RDF peak is invariant with the Al-Al distance. Below the saturation loading we observe that the peak positions exhibit a minimum at intermediate Al-Al distances between 8-10 Å. This observation supports our hypothesis that there exists a critical distance between two aluminum acid sites that facilitates water adsorption and the formation of clusters around acid sites. At higher water loadings (above 5 H_2O u.c.$^{-1}$), water molecules are evenly distributed within the AFI pore, as seen in **Figure 3**, stabilized by hydrogen bonding and other interactions within the water cluster.

Proton solubility. In addition to tracking the position of water molecules, we tracked the position of protons to determine their solubility in the pore centers. Previous studies from Mei, Rousseau, Lercher, and coworkers have shown that proton solubility depends on the number of coadsorbed waters.[63,64] Since we cannot *a priori* label protons in our system due to proton-shuttling events, we calculated proton identities dynamically over the simulation. To do this, we generated neighbor lists for each oxygen and hydrogen atom excluding framework oxygens. The protons were determined by pairing hydrogen atoms with oxygen atoms based on neighbor lists and then determining which

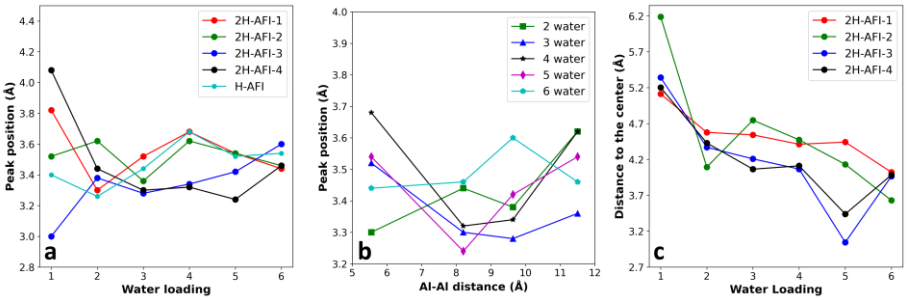

Figure 4. (a) The first RDF peak position for all aluminum distributions and water loadings (**b**) The first peak positions of RDFs in 2H-AFI with water loading from two to five u.c.$^{-1}$ (**c**) Average distances between framework protons and the AFI pore center.

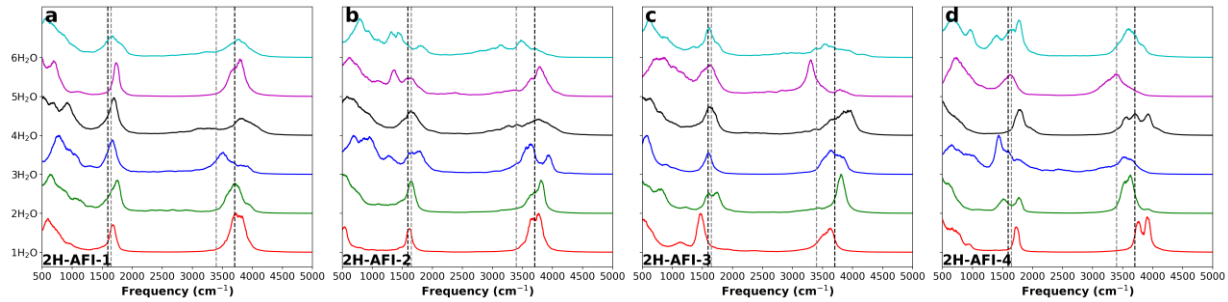

Figure 5. Vibrational density of state (vDOS) of water with different water loading in (**a**) 2H-AFI-1, (**b**) 2H-AFI-2, (**c**) 2H-AFI-3, and (**d**) 2H-AFI-4. Black dashed lines are reference gas water bending and stretching frequency, and grey dashed lines are reference liquid water bending and stretching frequency.[34]

hydrogens were weakly bound. These hydrogens were then labeled in that frame as a proton.

We sought to determine whether Al distributions affected the proton solubilities. **Figure 4c** illustrates the average distances of the framework protons to the AFI pore center in all four 2H-AFI configurations (i.e. a distance to the center of 0 Å would be the center of the pore). At 1 H_2O u.c.$^{-1}$, we note large distance values indicating protons remain close to the framework. We also find that single waters do not appreciably protonate. As the water loading increases, the average distance in all configurations decreases, indicating protonation of adsorbed water and solvation into the cluster. This observation aligns with previous literature at high Si/Al ratios,[30] and work in faujasite[46] which reported similar trends of proton solvation assisted by coadsorbed water. We noticed that all Al distributions exhibit a similar trend in proton distance with near-saturation loadings having average proton locations closest to the pore center. These data would indicate that protons solvate near the center of the pore more frequently on average when there are sufficient water molecules adsorbed. Solubility does not appear to depend on the relative position of the acid sites under these conditions, but this opens questions into whether different pore topologies or temperatures could be used to change the framework-proton distance.

Vibrational spectra of adsorbed water. We interrogated the hydrogen bond strength as a function of water loading by comparing water bending and stretching modes. These modes were compared by calculating the vibrational density of state (vDOS) which is shown in **Figure 5**. The dashed lines denote literature values for the vibrational frequencies of both liquid water and water vapor.[34] We observed shifts in the bending and stretching modes as water loading increased and as we changed aluminum distribution. As the water loading increased, both the frequencies associated with bending and stretching modes exhibited broader peak profiles, while remaining within a reasonable range that would be expected for water. The observed range of bending and stretching vibrational frequencies further indicates that our MLFF strategy captures the intermolecular forces present in water. At the saturation loading, we observed a shift in the stretching peak from 3746 cm^{-1} to 3448 cm^{-1} as the Al-Al distance increased. A decrease in the stretching peak has been shown to describe stronger hydrogen bonding.[65] These data indicate that disparate Al sites that nucleate extended water clusters form stronger hydrogen bonding networks than proximal Al sites.

Variation in Si/Al ratio. All data thus far have focused on a fixed Si/Al ratio with variations in the Al distribution. We applied our MLFF strategy to probe how changes to the Si/Al ratio can affect the nucleation of water clusters. We chose a single Al distribution (2H-AFI-2) and then expanded the simulation box by doubling its size along the AFI center pore. **Figure 6a-b** illustrates the water density distribution heat map along the zeolite AFI center cage which corresponds to the [010] plane. Here, **Figure 6a** illustrates the saturation loading of water at Si/Al 11 and subplot b illustrates the saturation loading of water at Si/Al 23. A Si/Al 11 corresponds to four aluminum atoms per supercell and Si/Al 23 corresponds to two aluminum atoms per supercell. We observed that Si/Al 23 exhibits a higher degree of clustering around the aluminum sites than the Si/Al 11 configuration which has water molecules adsorbed near both sets of Al sites marked in blue circles. To quantify the water structure around these acid sites, we calculated RDFs which are shown in **Figure 6c**. Both Si/Al 11 and Si/Al 23 RDFs exhibited a similar position for the first peak, indicating a consistent average distance between the first coordination shell of water molecules around acid sites. Si/Al 11 has a sharper RDF beyond the first peak as compared to Si/Al 23.

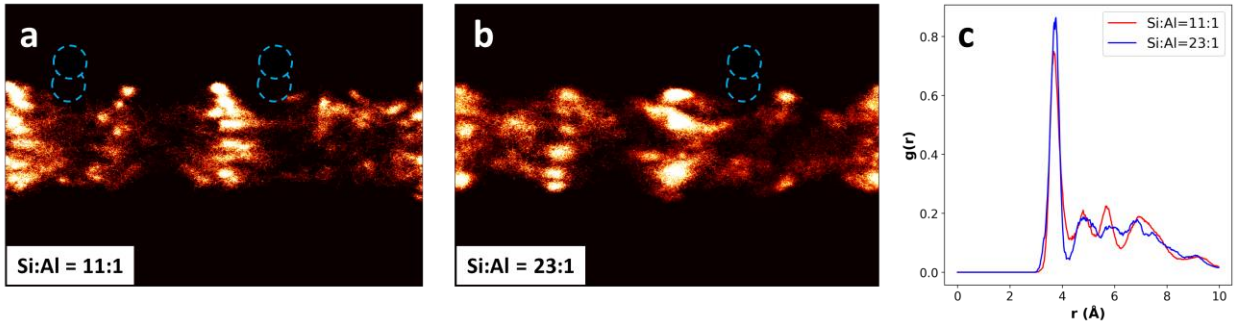

Figure 6. Heatmaps of the water density distribution along the AFI center pore in supercells with (**a**) Si/Al 11, (**b**) Si/Al 23, and (**c**) water RDFs for both Si/Al ratios.

This suggests that Si/Al 11 produces a more structured water environment than the high Si/Al ratio which exhibited broader peaks indicating that intraporous water is more evenly distributed water along the center pore in comparison. At higher Si/Al ratio we observe clustering around acid sites, but more disordered liquid-like water fills the siliceous intraporous region. Lower Si/Al ratio leads to extended water clusters where a greater degree of water structuring was observed indicating that acid site proximity plays a role in the spatial extent of water clustering.

CONCLUSION

Our investigation into water structuring provides insights into the intricate relationship between aluminum density and distribution within AFI aluminosilicate zeolites and the organization of adsorbed water molecules as a function of water loading. Using MLFFs to expedite AIMD simulations and improve the sampling of diverse water configurations, we have gained a deeper understanding of the pivotal role played by specific aluminum distributions in the thermodynamics of water adsorption and the spatial arrangement and persistence of water clusters. Notably, our findings indicate a consistent weakening in water adsorption energies as the Al-Al distance increases across loading. Additionally, our examination of water distribution within the AFI center pore revealed a propensity for water clusters to form at higher water loadings, where framework protons become integral to cluster stability. These findings were corroborated by independent computational characterization methods including density maps, radial distribution functions, proton locations, and vibrational spectra. The union of these data indicates that spatially disparate Al sites better stabilize extended hydrogen bonding networks resulting in denser water clusters. At high Si/Al ratios, water clusters nucleate in isolation, with intraporous disordered water filling the siliceous regions. As the Si/Al ratio decreases, water clusters interact which leads to larger clusters with strong hydrogen bonding.

These findings underscore the interplay between zeolite structure, Brønsted acid site location, and water adsorption, providing evidence that precise control over Si/Al ratio and distribution will directly influence the structure and stability of adsorbed water matrices. This may play a role in liquid-phase catalysis and oxygenate chemistries involving water co-generation where adsorption, diffusion, and transition state stability are affected by the water activity. Our error validation and data characterization have demonstrated the suitability of MLFF for studying such complex systems. This approach has key implications for future studies that link the cluster geometries observed here with the kinetics of reactions catalyzed by zeolites or other porous materials. These findings provide evidence that controlling the microenvironments around catalytic active sites can tailor water activity.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental procedures and simulation details (PDF). ZIP folder containing simulation settings, structures for bulk water, H-AFI, 2H-AFI, supercells, and MLFF model for 2H-AFI-2.

AUTHOR INFORMATION

Corresponding Author

* Brandon C. Bukowski – Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, Maryland 21218, United States; E-mail: bbukows1@jhu.edu

All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENTS

MZ and BCB acknowledge the Advanced Research Computing at Hopkins (ARCH) core facility (rockfish.jhu.edu), which is supported by the National Science Foundation (NSF) grant number OAC1920103. MZ and BCB acknowledge support from the Ralph E. Powe Junior Faculty Enhancement Award from Oak Ridge Associated Universities to perform this work.

REFERENCES

(1) Hemelsoet, K.; Van der Mynsbrugge, J.; De Wispelaere, K.; Waroquier, M.; Van Speybroeck, V. Unraveling the Reaction Mechanisms Governing Methanol-to-Olefins Catalysis by Theory and Experiment. *ChemPhysChem* **2013**, *14* (8), 1526–1545. https://doi.org/10.1002/cphc.201201023.

(2) Göltl, F.; Michel, C.; Andrikopoulos, P. C.; Love, A. M.; Hafner, J.; Hermans, I.; Sautet, P. Computationally Exploring Confinement Effects in the Methane-to-Methanol Conversion Over Iron-Oxo Centers in Zeolites. *ACS Catal.* **2016**, *6* (12), 8404–8409. https://doi.org/10.1021/acscatal.6b02640.

(3) Lercher, J. A.; van Santen, R. A.; Vinek, H. Carbonium Ion Formation in Zeolite Catalysis. *Catal Lett* **1994**, *27* (1), 91–96. https://doi.org/10.1007/BF00806981.

(4) Corma, A.; Huber, G. W.; Sauvanaud, L.; O'Connor, P. Processing Biomass-Derived Oxygenates in the Oil Refinery: Catalytic Cracking (FCC) Reaction Pathways and Role of Catalyst. *Journal of Catalysis* **2007**, *247* (2), 307–327. https://doi.org/10.1016/j.jcat.2007.01.023.

(5) Corma, A.; Domine, M. E.; Nemeth, L.; Valencia, S. Al-Free Sn-Beta Zeolite as a Catalyst for the Selective Reduction of Carbonyl Compounds (Meerwein−Ponndorf−Verley Reaction). *J. Am. Chem. Soc.* **2002**, *124* (13), 3194–3195. https://doi.org/10.1021/ja012297m.

(6) Hudson, M. R.; Queen, W. L.; Mason, J. A.; Fickel, D. W.; Lobo, R. F.; Brown, C. M. Unconventional, Highly Selective CO2 Adsorption in Zeolite SSZ-13. *J. Am. Chem. Soc.* **2012**, *134* (4), 1970–1973. https://doi.org/10.1021/ja210580b.

(7) Guo, P.; Shin, J.; Greenaway, A. G.; Min, J. G.; Su, J.; Choi, H. J.; Liu, L.; Cox, P. A.; Hong, S. B.; Wright, P. A.; Zou, X. A Zeolite Family with Expanding Structural Complexity and Embedded Isoreticular Structures. *Nature* **2015**, *524* (7563), 74–78. https://doi.org/10.1038/nature14575.

(8) Kim, C.; Cho, H. S.; Chang, S.; Cho, S. J.; Choi, M. An Ethylenediamine-Grafted y Zeolite: A Highly Regenerable Carbon Dioxide Adsorbent: Via Temperature Swing Adsorption without Urea Formation. *Energy and Environmental Science* **2016**, *9* (5), 1803–1811. https://doi.org/10.1039/c6ee00601a.

(9) Ayodele, O. B. Eliminating Reverse Water Gas Shift Reaction in CO 2 Hydrogenation to Primary Oxygenates

over MFI-Type Zeolite Supported Cu/ZnO Nanocatalysts. *Journal of CO2 Utilization* **2017**, *20*, 368–377. https://doi.org/10.1016/j.jcou.2017.06.015.

(10) Kwak, J. H.; Tonkyn, R. G.; Kim, D. H.; Szanyi, J.; Peden, C. H. F. Excellent Activity and Selectivity of Cu-SSZ-13 in the Selective Catalytic Reduction of NOx with NH3. *Journal of Catalysis* **2010**, *275* (2), 187–190. https://doi.org/10.1016/j.jcat.2010.07.031.

(11) Xu, Y.; Zheng, M.; Musgrave, C. B. I.; Zhang, L.; Goddard, W. A. I.; Bukowski, B. C.; Liu, Y. Assessing the Kinetics of Quinone–CO2 Adduct Formation for Electrochemically Mediated Carbon Capture. *ACS Sustainable Chem. Eng.* **2023**, *11* (30), 11333–11341. https://doi.org/10.1021/acssuschemeng.3c03321.

(12) Zhang, J.; Wegener, E. C.; Samad, N. R.; Harris, J. W.; Unocic, K. A.; Allard, L. F.; Purdy, S.; Adhikari, S.; Cordon, M. J.; Miller, J. T.; Krause, T. R.; Cheng, S.; Liu, D.; Li, M.; Jiang, X.; Wu, Z.; Li, Z. Isolated Metal Sites in Cu–Zn–Y/Beta for Direct and Selective Butene-Rich C_{3+} Olefin Formation from Ethanol. *ACS Catal.* **2021**, *11* (15), 9885–9897. https://doi.org/10.1021/acscatal.1c02177.

(13) Muñoz-Guerra, S.; Lavilla, C.; Japu, C.; Ilarduya, A. M. de. Renewable Terephthalate Polyesters from Carbohydrate-Based Bicyclic Monomers. *Green Chem.* **2014**, *16* (4), 1716–1739. https://doi.org/10.1039/C3GC42394H.

(14) Silva, C. X. A. da; Gonçalves, V. L. C.; Mota, C. J. A. Water-Tolerant Zeolite Catalyst for the Acetalisation of Glycerol. *Green Chem.* **2009**, *11* (1), 38–41. https://doi.org/10.1039/B813564A.

(15) Jin, M.; Ravi, M.; Lei, C.; Heard, C. J.; Brivio, F.; Tošner, Z.; Grajciar, L.; van Bokhoven, J. A.; Nachtigall, P. Dynamical Equilibrium between Brønsted and Lewis Sites in Zeolites: Framework-Associated Octahedral Aluminum. *Angewandte Chemie International Edition n/a* (n/a), e202306183. https://doi.org/10.1002/anie.202306183.

(16) Tang, X.; Chen, W.; Dong, W.; Liu, Z.; Yuan, J.; Xia, H.; Yi, X.; Zheng, A. Framework Aluminum Distribution in ZSM-5 Zeolite Directed by Organic Structure-Directing Agents: A Theoretical Investigation. *Catalysis Today* **2022**, *405–406*, 101–110. https://doi.org/10.1016/j.cattod.2022.06.027.

(17) Csicsery, S. M. Shape-Selective Catalysis in Zeolites. *Zeolites* **1984**, *4* (3), 202–213. https://doi.org/10.1016/0144-2449(84)90024-1.

(18) Liu, X.; Liu, S.; Yan, T.; Shang, N.; Li, H.; Wang, Z.; Xu, H.; Wu, P. Tin Active Sites Confined in Zeolite Framework as a Promising Shape-Selective Catalyst for Ethylene Oxide Hydration. *Chemistry – A European Journal* **2023**, *29* (16), e202203696. https://doi.org/10.1002/chem.202203696.

(19) Svelle, S.; Olsbye, U.; Joensen, F.; Bjørgen, M. Conversion of Methanol to Alkenes over Medium- and Large-Pore Acidic Zeolites: Steric Manipulation of the Reaction Intermediates Governs the Ethene/Propene Product Selectivity. *J. Phys. Chem. C* **2007**, *111* (49), 17981–17984. https://doi.org/10.1021/jp077331j.

(20) Brogaard, R. Y.; Wang, C.-M.; Studt, F. Methanol–Alkene Reactions in Zeotype Acid Catalysts: Insights from a Descriptor-Based Approach and Microkinetic Modeling. *ACS Catal.* **2014**, *4* (12), 4504–4509. https://doi.org/10.1021/cs5014267.

(21) Iglesia, E. Design, Synthesis, and Use of Cobalt-Based Fischer-Tropsch Synthesis Catalysts. *Applied Catalysis A: General* **1997**, *161* (1), 59–78. https://doi.org/10.1016/S0926-860X(97)00186-5.

(22) Jacobs, G.; Ma, W.; Davis, B. H. Influence of Reduction Promoters on Stability of Cobalt/g-Alumina Fischer-Tropsch Synthesis Catalysts. *Catalysts* **2014**, *4* (1), 49–76. https://doi.org/10.3390/catal4010049.

(23) Li, G.; Wang, B.; Resasco, D. E. Water-Mediated Heterogeneously Catalyzed Reactions. *ACS Catal.* **2020**, *10* (2), 1294–1309. https://doi.org/10.1021/acscatal.9b04637.

(24) Chen, K.; Kelsey, J.; White, J. L.; Zhang, L.; Resasco, D. Water Interactions in Zeolite Catalysts and Their Hydrophobically Modified Analogues. *ACS Catal.* **2015**, *5* (12), 7480–7487. https://doi.org/10.1021/acscatal.5b02040.

(25) Zapata, P. A.; Faria, J.; Ruiz, M. P.; Jentoft, R. E.; Resasco, D. E. Hydrophobic Zeolites for Biofuel Upgrading Reactions at the Liquid–Liquid Interface in Water/Oil Emulsions. *J. Am. Chem. Soc.* **2012**, *134* (20), 8570–8578. https://doi.org/10.1021/ja3015082.

(26) Bregante, D. T.; Johnson, A. M.; Patel, A. Y.; Ayla, E. Z.; Cordon, M. J.; Bukowski, B. C.; Greeley, J.; Gounder, R.; Flaherty, D. W. Cooperative Effects between Hydrophilic Pores and Solvents: Catalytic Consequences of Hydrogen Bonding on Alkene Epoxidation in Zeolites. *J. Am. Chem. Soc.* **2019**, *141* (18), 7302–7319. https://doi.org/10.1021/jacs.8b12861.

(27) Yu, Q.; Bowman, J. M. High-Level Quantum Calculations of the IR Spectra of the Eigen, Zundel, and Ring Isomers of $H^+(H_2O)_4$ Find a Single Match to Experiment. *J. Am. Chem. Soc.* **2017**, *139* (32), 10984–10987. https://doi.org/10.1021/jacs.7b05459.

(28) Mei, D.; Lercher, J. A. Effects of Local Water Concentrations on Cyclohexanol Dehydration in H-BEA Zeolites. *J. Phys. Chem. C* **2019**, *123* (41), 25255–25266. https://doi.org/10.1021/acs.jpcc.9b07738.

(29) Cordon, M. J.; Harris, J. W.; Vega-Vila, J. C.; Bates, J. S.; Kaur, S.; Gupta, M.; Witzke, M. E.; Wegener, E. C.; Miller, J. T.; Flaherty, D. W.; Hibbitts, D. D.; Gounder, R. Dominant Role of Entropy in Stabilizing Sugar Isomerization Transition States within Hydrophobic Zeolite Pores. *J. Am. Chem. Soc.* **2018**, *140* (43), 14244–14266. https://doi.org/10.1021/jacs.8b08336.

(30) Grifoni, E.; Piccini, G.; Lercher, J. A.; Glezakou, V.-A.; Rousseau, R.; Parrinello, M. Confinement Effects and Acid Strength in Zeolites. *Nat Commun* **2021**, *12* (1), 2630. https://doi.org/10.1038/s41467-021-22936-0.

(31) Bates, J. S.; Bukowski, B. C.; Greeley, J.; Gounder, R. Structure and Solvation of Confined Water and Water–Ethanol Clusters within Microporous Brønsted Acids and Their Effects on Ethanol Dehydration Catalysis. *Chem. Sci.* **2020**, *11* (27), 7102–7122. https://doi.org/10.1039/D0SC02589E.

(32) Berens, P. H.; Mackay, D. H. J.; White, G. M.; Wilson, K. R. Thermodynamics and Quantum Corrections from Molecular Dynamics for Liquid Water. *The Journal of Chemical Physics* **1983**, *79* (5), 2375–2389. https://doi.org/10.1063/1.446044.

(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The*

Journal of Chemical Physics **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.

(34) Seki, T.; Chiang, K.-Y.; Yu, C.-C.; Yu, X.; Okuno, M.; Hunger, J.; Nagata, Y.; Bonn, M. The Bending Mode of Water: A Powerful Probe for Hydrogen Bond Structure of Aqueous Systems. *J. Phys. Chem. Lett.* **2020**, *11* (19), 8459–8469. https://doi.org/10.1021/acs.jpclett.0c01259.

(35) van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A* **2001**, *105* (41), 9396–9409. https://doi.org/10.1021/jp004368u.

(36) van Duin, A. C. T.; Strachan, A.; Stewman, S.; Zhang, Q.; Xu, X.; Goddard, W. A. ReaxFFSiO Reactive Force Field for Silicon and Silicon Oxide Systems. *J. Phys. Chem. A* **2003**, *107* (19), 3803–3811. https://doi.org/10.1021/jp0276303.

(37) Naserifar, S.; Chen, Y.; Kwon, S.; Xiao, H.; Goddard, W. A. Artificial Intelligence and QM/MM with a Polarizable Reactive Force Field for Next-Generation Electrocatalysts. *Matter* **2021**, *4* (1), 195–216. https://doi.org/10.1016/j.matt.2020.11.010.

(38) Jinnouchi, R.; Karsai, F.; Kresse, G. On-the-Fly Machine Learning Force Field Generation: Application to Melting Points. *Phys. Rev. B* **2019**, *100* (1), 014105. https://doi.org/10.1103/PhysRevB.100.014105.

(39) Liu, P.; Verdi, C.; Karsai, F.; Kresse, G. $\ensuremath{\alpha}\text{\ensuremath{-}}\ensuremath{\beta}$ Phase Transition of Zirconium Predicted by on-the-Fly Machine-Learned Force Field. *Phys. Rev. Mater.* **2021**, *5* (5), 053804. https://doi.org/10.1103/PhysRevMaterials.5.053804.

(40) Verdi, C.; Karsai, F.; Liu, P.; Jinnouchi, R.; Kresse, G. Thermal Transport and Phase Transitions of Zirconia by On-the-Fly Machine-Learned Interatomic Potentials. *npj Comput Mater* **2021**, *7* (1), 1–9. https://doi.org/10.1038/s41524-021-00630-5.

(41) Zhang, G.; Liu, W.; Hu, T.; Shuai, S.; Chen, C.; Xu, S.; Ren, W.; Wang, J.; Ren, Z. On-the-Fly Machine Learning Force Field Study of Liquid-Al/α-Al2O3 Interface. *Applied Surface Science* **2023**, 158141. https://doi.org/10.1016/j.apsusc.2023.158141.

(42) Liu, P.; Wang, J.; Avargues, N.; Verdi, C.; Singraber, A.; Karsai, F.; Chen, X.-Q.; Kresse, G. Combining Machine Learning and Many-Body Calculations: Coverage-Dependent Adsorption of CO on Rh(111). *Phys. Rev. Lett.* **2023**, *130* (7), 078001. https://doi.org/10.1103/PhysRevLett.130.078001.

(43) Xia, D.; Chen, J.; Xie, H.-B.; Zhong, J.; Francisco, J. S. Counterintuitive Oxidation of Alcohols at Air–Water Interfaces. *J. Am. Chem. Soc.* **2023**, *145* (8), 4791–4799. https://doi.org/10.1021/jacs.2c13661.

(44) Roy, S.; Dürholt, J. P.; Asche, T. S.; Zipoli, F.; Gómez-Bombarelli, R. Learning a Reactive Potential for Silica-Water through Uncertainty Attribution. arXiv July 4, 2023. https://doi.org/10.48550/arXiv.2307.01705.

(45) Erlebach, A.; Šípka, M.; Saha, I.; Nachtigall, P.; Heard, C. J.; Grajciar, L. A Reactive Neural Network Framework for Water-Loaded Acidic Zeolites. arXiv July 13, 2023. https://doi.org/10.48550/arXiv.2307.00911.

(46) Saha, I.; Erlebach, A.; Nachtigall, P.; Heard, C. J.; Grajciar, L. Quantifying the Effect of Si/Al Ratio on Proton Solvation and Water Diffusion in H-FAU Using Reactive Neural Network Potential. ChemRxiv July 12, 2023. https://doi.org/10.26434/chemrxiv-2022-d1sj9-v3.

(47) *Database of Zeolite Structures*. http://www.iza-structure.org/databases/ (accessed 2023-04-30).

(48) Hjorth Larsen, A.; Jørgen Mortensen, J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Bjerre Jensen, P.; Kermode, J.; Kitchin, J. R.; Leonhard Kolsbjerg, E.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Bergmann Maronsson, J.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schütt, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. W. The Atomic Simulation Environment—a Python Library for Working with Atoms. *J. Phys.: Condens. Matter* **2017**, *29* (27), 273002. https://doi.org/10.1088/1361-648X/aa680e.

(49) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal--Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49* (20), 14251–14269. https://doi.org/10.1103/PhysRevB.49.14251.

(50) Kresse, G.; Hafner, J. *Ab Initio* Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47* (1), 558–561. https://doi.org/10.1103/PhysRevB.47.558.

(51) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169–11186. https://doi.org/10.1103/PhysRevB.54.11169.

(52) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Computational Materials Science* **1996**, *6* (1), 15–50. https://doi.org/10.1016/0927-0256(96)00008-0.

(53) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(54) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465. https://doi.org/10.1002/jcc.21759.

(55) Evans, D. J.; Holian, B. L. The Nose–Hoover Thermostat. *The Journal of Chemical Physics* **1985**, *83* (8), 4069–4074. https://doi.org/10.1063/1.449071.

(56) Jinnouchi, R.; Karsai, F.; Verdi, C.; Asahi, R.; Kresse, G. Descriptors Representing Two- and Three-Body Atomic Distributions and Their Effects on the Accuracy of Machine-Learned Inter-Atomic Potentials. *J. Chem. Phys.* **2020**, *152* (23), 234102. https://doi.org/10.1063/5.0009491.

(57) Jinnouchi, R.; Lahnsteiner, J.; Karsai, F.; Kresse, G.; Bokdam, M. Phase Transitions of Hybrid Perovskites Simulated by Machine-Learning Force Fields Trained on the Fly with Bayesian Inference. *Phys. Rev. Lett.* **2019**, *122* (22), 225701. https://doi.org/10.1103/PhysRevLett.122.225701.

(58) Bartók, A. P.; Csányi, G. Gaussian Approximation Potentials: A Brief Tutorial Introduction. arXiv February 5, 2020. http://arxiv.org/abs/1502.01366 (accessed 2023-04-29).

(59) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B* **2013**, *87* (18), 184115. https://doi.org/10.1103/PhysRevB.87.184115.
(60) Miwa, K.; Ohno, H. Molecular Dynamics Study on β -Phase Vanadium Monohydride with Machine Learning Potential. *Phys. Rev. B* **2016**, *94* (18), 184109. https://doi.org/10.1103/PhysRevB.94.184109.
(61) Bukowski, B. C.; Bates, J. S.; Gounder, R.; Greeley, J. Defect-Mediated Ordering of Condensed Water Structures in Microporous Zeolites. *Angew. Chem. Int. Ed.* **2019**, *58* (46), 16422–16426. https://doi.org/10.1002/anie.201908151.
(62) Loewenstein, W. The Distribution of Aluminum in the Tetrahedra of Silicates and Aluminates. *American Mineralogist* **1954**, *39* (1–2), 92–96.
(63) Mei, D.; Lercher, J. A. Mechanistic Insights into Aqueous Phase Propanol Dehydration in H-ZSM-5 Zeolite. *AIChE Journal* **2017**, *63* (1), 172–184. https://doi.org/10.1002/aic.15517.
(64) Wang, M.; Jaegers, N. R.; Lee, M.-S.; Wan, C.; Hu, J. Z.; Shi, H.; Mei, D.; Burton, S. D.; Camaioni, D. M.; Gutiérrez, O. Y.; Glezakou, V.-A.; Rousseau, R.; Wang, Y.; Lercher, J. A. Genesis and Stability of Hydronium Ions in Zeolite Channels. *J. Am. Chem. Soc.* **2019**, *141* (8), 3444–3455. https://doi.org/10.1021/jacs.8b07969.
(65) Dalla Bernardina, S.; Paineau, E.; Brubach, J.-B.; Judeinstein, P.; Rouzière, S.; Launois, P.; Roy, P. Water in Carbon Nanotubes: The Peculiar Hydrogen Bond Network Revealed by Infrared Spectroscopy. *J. Am. Chem. Soc.* **2016**, *138* (33), 10437–10443. https://doi.org/10.1021/jacs.6b02635.