Electronic Supplementary Material (ESI) for Energy & Environmental Science.

Supporting Information

Uncovering Fast Solid-Acid Proton Conductors Based on Dynamics of Polyanion Groups and Proton Bonding Strength

Pjotrs Žguns[1*], Konstantin Klyukin[1,3*], Louis S. Wang[4], Grace Xiong[4], Ju Li[1,2], Sossina M. Haile[4], Bilge Yildiz[1,2,#]

[1] *Department of Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*

[2] *Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*

[3] *Department of Materials Engineering, Auburn University, Auburn, Alabama 36849, USA*

[4] *Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA*

* *These authors contributed equally to this work*

\# byildiz@mit.edu

Table of Contents

Section S1. Computational Methods

Initial and Final Datasets

Initial dataset to identify strong descriptors of proton conductivity: We created an initial dataset to identify correlations between the proposed descriptors and the calculated proton diffusivities (see AIMD simulations below). These compounds were selected from the Materials Project database, with criteria imposed on (*I*) chemical composition, (*II*) electronic band gap, (*III*) synthesizability/phase stability, and *(IV*) structure (hydrogen bonding), as described below (see Screening workflow). To keep the computational cost of AIMD simulations manageable, we applied tight criteria in steps (*II*) and (*III*), reducing the pool of materials to about a hundred. The great majority of thus selected compounds contained oxygen; the small fraction of compounds that had no oxygen (e.g., ZrF_4NH_3, AlF_4NH_4) was excluded for the sake of dataset coherence. Further, we excluded compounds that had no cations (for the list of cations, see step (*VI*)).

The initial dataset consisted of 88 compounds (Table S1). The most of compounds (49) were solid acids with SO_4, PO_4, ClO_4, SeO_3 or other polyanion groups; these included well-known solid-acids such as $CsHSO_4$ and $CsH_5(PO_4)_2$, as well less-known compounds such as $CaHSO_4(H_2SO_4)$, $Er(H_2O)(ClO_4)_3$, $(NH_4)(ReO_4)$, $Mg(H_2O)_6(BrO_3)$, etc. The initial dataset also included compounds (28) that do not have any monomeric groups but rather have oligomeric or polymeric groups such as P_2O_7 (e.g., $CaP_2H_2O_7$) or B_nO_m (e.g., $Ba_3B_6H_2O_{13}$), and compounds (11) that consist solely of water, fluorine or hydroxyl groups (e.g., $Ba(H_2O)(OH)_2$, $ZrF_4(H_2O)_3$). Since such compounds do not have groups that could actively rotate, we excluded them from the analysis of group rotation (see footnotes in Table S1), however we retained them for the analysis of proton transfer as it may still occur. These 39 compounds did not show any significant proton diffusion unless they were completely melted in AIMD simulations, and we therefore excluded such compounds in the final screening (see Final dataset and Screening workflow).

To assess the proton conductivity of the selected materials, we carried out AIMD simulations at elevated temperatures to compute proton diffusivities. Molten compounds (e.g., those having a substantial diffusion of ions other than protons) were excluded from analysis (see Table S1). For 15 of the selected compounds, the mean-square displacement (MSD) analysis of the AIMD trajectories at 700 K revealed a significant proton diffusion characterized by a near-linear MSD growth, while other investigated materials have not demonstrated appreciable diffusion.

Final dataset to screen for fast proton conductors: To identify potential proton conductors based on our identified descriptors, we expanded the number of materials considered by loosening the criteria in steps (*I*), (*II*), and (*III*), while maintaining step (*IV*). We then screened this larger set of candidates by applying criteria determined from the descriptors of proton transfer and group rotation in steps (*V*) and (*VI*), respectively, as described in detail below. The final dataset included 143 solid acids (Table S2). We further explored the proton diffusivities of these materials by running MD simulations as follows. First, in step (*VII*) we ran machine learning force-fields (MLFF) which were trained on-the-fly during AIMD simulations (MLFF–AIMD), as implemented

in the VASP code. MLFF–AIMD simulations were run at 650 K to filter out poor conductors. Finally, we ran AIMD (or MLFF–AIMD) for several selected compounds and explored their proton diffusivities at various temperatures. For this analysis, we specifically selected solid acids with less common cations such as Ag, Sr, Ba, Er, and omitted more common solid acids with Cs, Rb, K cations to reduce the computational cost. However, for the sake of comparison we calculated proton conductivity of $CsHSO_4$ (mp-1192419) to compare the calculated conductivities of non-Cs compounds with that of $CsHSO_4$.

Screening workflow

I. Chemical composition. We considered only materials that already have hydrogen in their chemical formula. For the initial dataset, we excluded compounds with other small cations (Li^+, Na^+, K^+) that could lower the proton transference number, however this restriction was not applied for the final screening. For the initial dataset, we also excluded compounds that do not have oxygen in their chemical formula at this step (for the final dataset, we applied similar criterion in step (*V*)).

II. Electronic band gap. To minimize plausible electronic conductivity, we filtered out materials with a low band gap. We used $E_g > 4$ eV criterion for the initial dataset to keep the number of materials bearable for AIMD. We used $E_g > 2$ eV criterion for the final dataset to include more potential conductors. The band gaps were retrieved from the Materials Project database[1].

III. Synthesizability/*Phase stability*. The initial and final datasets consisted of synthesizable compounds that are reported in the ICSD database[53]. The respective ICSD identifiers were retrieved from the Materials Project database[1]. For the initial dataset, we also applied criterion on the phase stability, i.e., the convex hull energy ($E_{hull} \leq 5$ meV/atom), which was also retrieved from the Materials Project database[1].

IV. Structure. Only compounds possessing D–H…A hydrogen bonds (at least one) were included in the datasets (we applied $d_{H...A} < 2$ Å cut-off). For simplicity, and since polyanion groups may rotate, we set no further restrictions on the number of bonds or their topology.

V. Descriptors of proton transfer. We used the maximal O–H covalent bond length observed in compound, d_{OH}, as a descriptor (it was calculated for the structure as retrieved from the Material Project database). We included only compounds that have (at least one) $d_{OH} > 1.007$ Å belonging to the O–H…O fragments, where either or both O atoms belong to the monomeric oxyanion group, $H_zXO_n^{m-}$; the cut-off of 2 Å was used to identify XO bonds and thus XO_n groups. This allowed us to exclude compounds that have only groups with low rotational flexibility such as P_2O_7, but to include compounds with H_2O/H_3O^+ as units in the hydrogen bonded network. In this step we also excluded compounds that have H_2 molecules or lone H atoms in their structures as retrieved from the Materials Project[1] database. The justification of the chosen cut-off value for d_{OH} is provided in the Results and Discussion section.

VI. Descriptors of group rotation. We used the average cation phonon band center, ω_{cat}, as a descriptor. For the final dataset, we selected compounds that have cations with $\omega_{cat} < 3.5$ THz (the justification for choosing this cut-off value is provided in the Results and Discussion section). For

compounds with multiple cations, we selected the compound if this criterion was satisfied by any of the cations. For the screening purposes, we broadly defined cation as any element excluding the following: halogens, O, S, Se, Te, Po, N, P, As, Sb, C, Si, Ge, B, H, and noble gases. Compounds that have no cations were also excluded.

VII. MD Screening. We computed proton diffusivities of 143 compounds (final dataset) at 650 K, by running MLFF–AIMD simulations (see the AIMD subsection below).

VIII. MD Simulations. We computed proton diffusivities of selected compounds at various temperatures by AIMD (or MLFF–AIMD) simulations (see the AIMD subsection below).

AIMD simulations of proton diffusion

Proton diffusivities were assessed with *Ab Initio* Molecular Dynamics (AIMD) simulations using the Projector Augmented Wave method[2,3] and Perdew–Burke–Ernzerhof (PBE) exchange-correlation functional[4,5] implemented in the Vienna Ab Initio Simulation Package (VASP)[6–9]. Since the initial dataset included diverse materials, we consistently applied the D3 dispersion correction[10] to account for van der Walls interactions which may be sizable in some compounds. Spin-polarization was used on-demand for compounds that might have unpaired electrons. The Hubbard U correction[11] was applied for transition metals with U values as used in the Materials Project database[1]. The plane-wave cut-off was set to 400 eV. The 2s2p oxygen electrons were treated as valence states. The supercell size was at least 6 Å in each crystallographic direction; a typical supercell consisted of about 50–120 atoms (the number of hydrogen atoms was on the order of ~10 to ~50). Brillouin Zone integration was performed in the Γ point only. The crystal structures were retrieved from the Materials Project database[1]. The AIMD simulations were run in the isothermal–isochoric (*NVT*) ensemble with the Nosé–Hoover thermostat[12,13] using fixed cell parameters, as obtained from the database. To reduce the computational time, we used deuterons (2 Da) with 1 fs time step. To speed up some of the calculations, we employed machine learning force-fields (MLFF) which were trained on-the-fly during AIMD simulations (MLFF–AIMD), as implemented in the VASP code[14–16]. The MLFF–AIMD method is robust and executes MD with nearly *ab initio* accuracy. It is particularly reliable since it resorts to *ab initio* calculation of forces whenever the MLFF needs to be retrained. Our results support the credibility of the MLFF method. Thus, the AIMD conductivity of $CsHSO_4$ (mp-1192419, Fig. 7) is 0.0083 S/cm at 650 K with a confidence interval of approximately 0.0025 S/cm to 0.0141 S/cm. The MLFF–AIMD conductivity is 0.0051 S/cm, which is well within the limits and the difference can be attributed to the natural variability of independent MD runs. The analysis of the force-field errors further confirms that MLFF–AIMD simulations are reliable. The typical root-mean-square errors of energies and forces are 0.001 eV/atom and 0.114 eV/Å, respectively ($CsHSO_4$ @650 K). These errors are relatively small for MD and are comparable to other state-of-the-art neural network fore-fields such as CHGNet (0.033 eV/atom and 0.079 eV/Å)[17]. For other compounds, e.g., $Sr(HSO_4)_2(H_2SO_4)$ and $Ba(H_2PO_4)_2$ in Fig. 7, the typical errors are comparable, on the order of 0.001 eV/atom and 0.1 eV/Å. These figures of merit, the good agreement between AIMD and

MLFF–AIMD and the small errors in energies and forces, support the credibility of our MLFF–AIMD simulations.

Few words should be said regarding the *NVT* ensemble employed in this work. Because the simulations were performed in the *NVT* ensemble, structural transitions to superprotonic phases were not directly captured. Nevertheless, even with the confines of the unit cells defined by the room temperature, non-superprotonic phases, our elevated temperature AIMD simulations achieve thermal activation of the polyanion group rotation and disordering of hydrogen bonds. Given this crystallographic constraint, we expect that the so-calculated diffusivities provide a lower bound of the true values that would be obtained with temperature-dependent lattice parameters. (At present, the computational cost of explicit modelling of superprotonic transitions in AIMD is prohibitive.)

Initial dataset: Proton diffusivities (D_H) of compounds in the initial dataset were assessed by AIMD runs using 25 ps long trajectories at 700 K. $D_H = \lim_{t\to\infty} \langle R_H^2 \rangle / 6t$, where $\langle R_H^2 \rangle$ is the mean-squared displacement of protons (R_H) and t is time. We calculated the D_H as the fit of $\langle R_H^2 \rangle$ vs. t obtained from AIMD. The calculated D_H values were used to test the proposed descriptors. To differentiate between fast and poor proton conductors in the initial dataset, we applied the cut-off for D_H that was set to 10^{-6} cm^2/s. This was necessary to filter out trajectories that did not result in any long-range diffusion. Proton excursions from the local equilibrium can produce mean-squared displacement of about 1 Å^2, resulting in an apparent D_H of about [1 Å^2 / (6×simulation time)], even in the absence of any diffusion. For 25 ps trajectory, this corresponds to about 10^{-6} cm^2/s. In other words, this cut-off allowed us to differentiate between diffusion versus noise in the AIMD trajectories, as shown in Fig. S0 below. (This cut-off corresponds to about 0.02 S/cm conductivity at 700 K, which is within the range of experimental diffusivities[18] extrapolated by us to 700 K, i.e., from about 0.003 S/cm to 0.3 S/cm.) In addition to D_H, we also computed the metrics for each of the two steps of the Grotthuss mechanism (see Local metrics of proton transfer and polyanion rotation rates subsection below), providing further credibility to the analysis of descriptors. Further, the strong correlation between the calculated D_H and local metrics of proton transfer in Fig. S1b indicates that Grotthuss diffusion is a dominant mechanism in solid acids in good agreement with previous studies[19–21]. While the possibility of vehicle diffusion (common in many liquid electrolytes[22] and contributing to proton conductivity at oxides' grain boundaries through water interlayers[23]) cannot be entirely excluded, its contribution is considered negligible for the majority of the proton-conducting compounds in both the initial and final datasets.

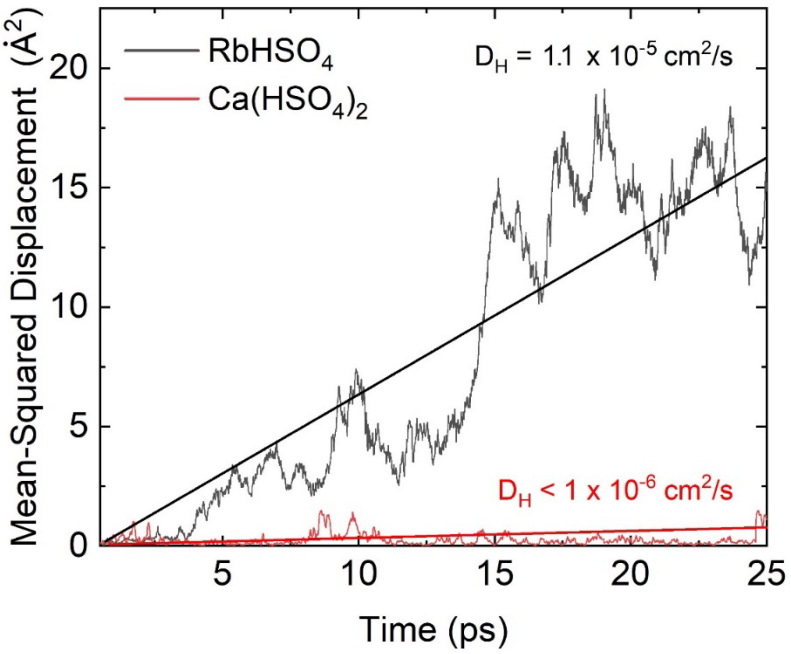

Figure S0. Mean-squared displacement of H^+ in two representative compounds: $RbHSO_4$ (fast proton diffusion) and $Ca(HSO_4)_2$ (negligible proton motion). The diffusion coefficients are calculated as a slope of the mean-squared displacement.

Final dataset (step VII): Proton diffusivities of compounds in the final dataset were assessed from the MLFF–AIMD runs at 650 K. A typical trajectory was 150 ps long. D_H was computed as outlined in *step VIII* below.

AIMD simulations at various temperatures (step VIII). For a few most promising compounds, we calculated D_H at lower temperatures (starting from 625 K and, depending on compound, down to 350 K). For each temperature, D_H was estimated from the MSD at the end of the run, $D_H \approx \langle R_{H,end}^2 \rangle / 6t_{end}$, and was averaged over multiple runs (the runs that did not show appreciable diffusivity, which was checked by the displacement of protons, were considered to have zero diffusivity for the averaging purposes). This approach was more reliable than a linear fit, especially for long runs at low-temperature showing a lower diffusivity (see example in Fig. S6). These runs employed AIMD (or MLFF–AIMD) and were up to 2 ns long. Proton conductivity, σ_H, was calculated by using the Nernst–Einstein relation: $\sigma_H = ((Ze)^2 C_H)/k_B T) D_H$, where Z is the proton charge, e is the elementary charge, $C_H = N_H/V$ is the total number of protons (N_H) per supercell volume (V), and D_H is the proton diffusion coefficient averaged over all protons in the simulation.

Phonon spectra calculations

Phonons were calculated using the finite-difference method as implemented in the VASP code [3,6–9] (without D3 correction). Supercells and atomic positions were fully relaxed, so that forces acting on atoms were below 10^{-3} eV/Å. The plane-wave cut-off energy was set to 520 eV. The phonon

band center (for each atom) was then calculated[24]. We also calculated the power spectra from the AIMD for each atomic species from the respective velocity autocorrelation functions[25].

Assignment of pK_a values to proton donor and proton acceptor groups

The pK_a of an acid group is defined with respect to the K_a, the acid dissociation constant, i.e., the equilibrium constant of the $HA = A^- + H^+$ reaction, and $K_a = [A^-][H^+]/[HA]$, $pK_a = -\log_{10} K_a$ and $\Delta G = -RT \ln K_a$. As the first step in assigning pK_a descriptors, the donor and acceptor sites, as well as corresponding donor and acceptor (polyanion) groups, associated with each proton in the crystal structure were identified by assessing the spatial connectivity of atoms and the respective bond lengths. Each donor and acceptor group was then assigned an acidity constant (pK_a) value from the literature (see Table S3; compounds with groups that have no tabulated pK_a values were excluded from consideration) [26–29]. Here we primarily use pK_a tabulated for acids in water. For the acceptor groups, we used the pK_a value of the conjugated acid (protonated group, i.e., after the proton transfer from the donor site to the acceptor site would have occurred). We then assign $\Delta pK_a = |pK_{a,donor} - pK_{a,acceptor}|$ as a descriptor that shows the difference in acidity between the proton bonded to its initial donor and the proton bonded to its initial acceptor.

Local metrics of proton transfer and polyanion rotation rates

Beyond computing macroscopic diffusivity, we assess individual metrics for each of the two steps of the Grotthuss mechanism. To assess the rate of proton transfer, we calculate the maximal distance between the proton and its initial donor atom encountered with the 25 ps simulation period, $L_{OH,max}$ (Figure 1a). This parameter, which has a value close to d_{OH} at the initiation of the simulation, reveals whether any proton transfer took place during the allocated time (Figure 1b). To assess the rate of polyanion group rotation, we calculate the standard deviations σ_θ and σ_φ of spherical angles θ (polar) and φ (azimuthal) of all bond vectors (e.g., S–O in the SO_4 group, in the internal coordinate system associated with the initial vector orientation) tracking displacement from their initial orientation over time (Figure 1c), and use the larger one of them, i.e., $max(\sigma_\theta, \sigma_\varphi)$ as descriptor. We note that although the proton transfer and rotation of polyanion groups are often correlated[19,30], our analysis shows that these two metrics (viz., $L_{OH,max}$ and maximal standard deviation of the rotation angles) are nevertheless indicative of the rates of the two rate-limiting steps (i.e., we can infer which step is rate limiting, see Figure S1). This assessment therefore allows us to correlate each descriptor with both diffusivity and the metric of respective step, making our analysis more robust (see Results and Discussions).

Section S2. Supporting Figures S1–S8

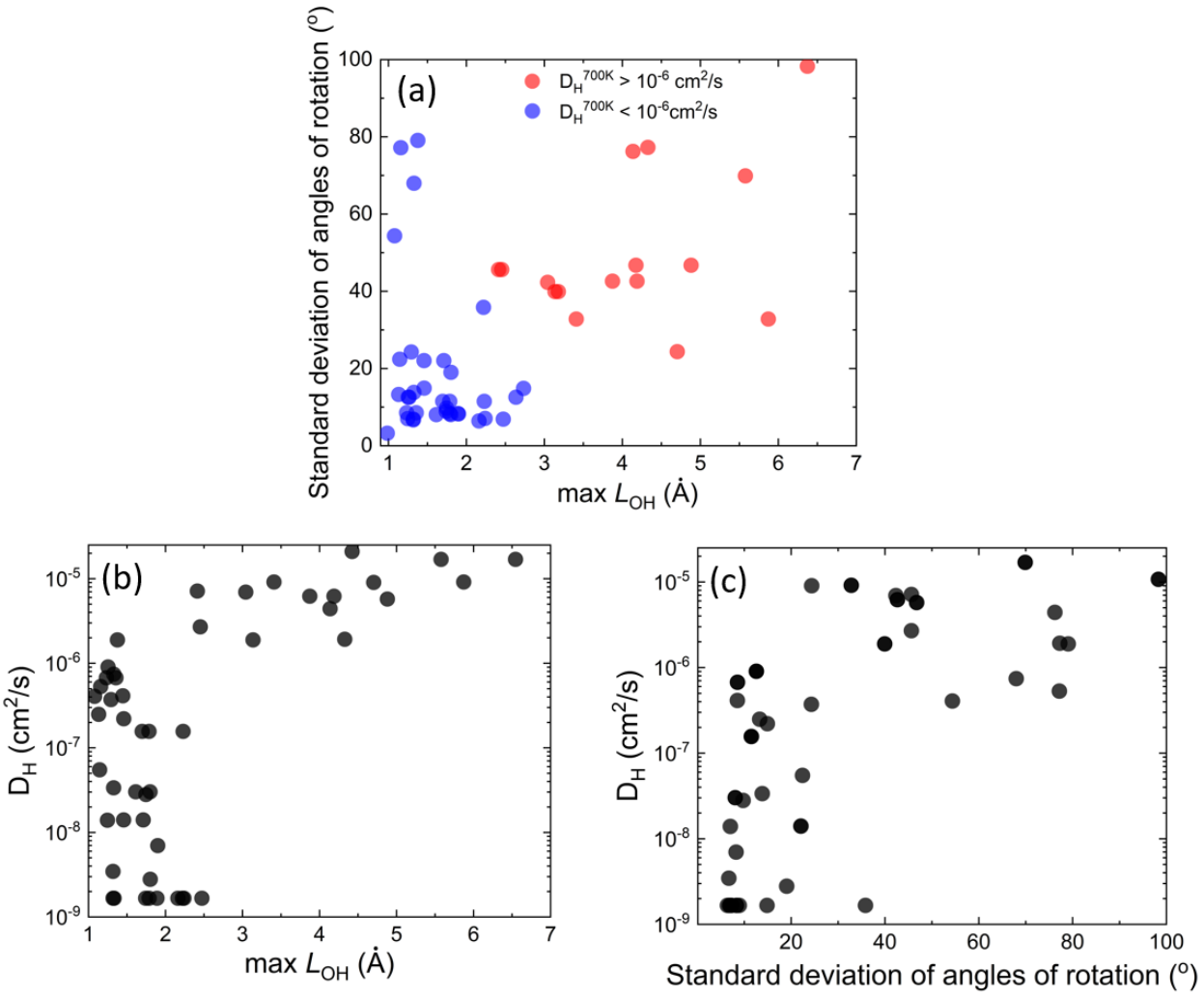

Figure S1. Analysis of limiting factors for two steps of the Grotthuss proton transport. (a) The scatter plot of rotational metric (the largest of the θ and φ standard deviations) vs. maximal distance L_{OH} between the proton and its initial donor atom showing limiting factors of proton diffusion in the preselected set of compounds for group rotations and local proton transfer, correspondingly. Materials demonstrating a long-range hydrogen diffusion are in red. (b) Scatter plot of D_H vs. *max* L_{OH}. (c) Scatter plot of D_H vs. the rotational metric.

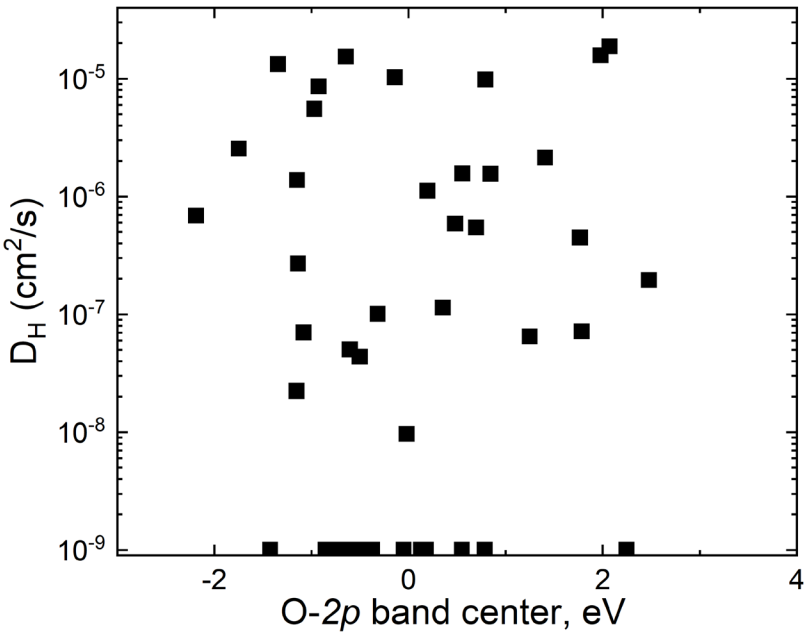

Figure S2. Scatter plot of diffusion coefficient D_H vs. O-2p band center for the subset of materials from the initial dataset. No generalizable correlation is observed.

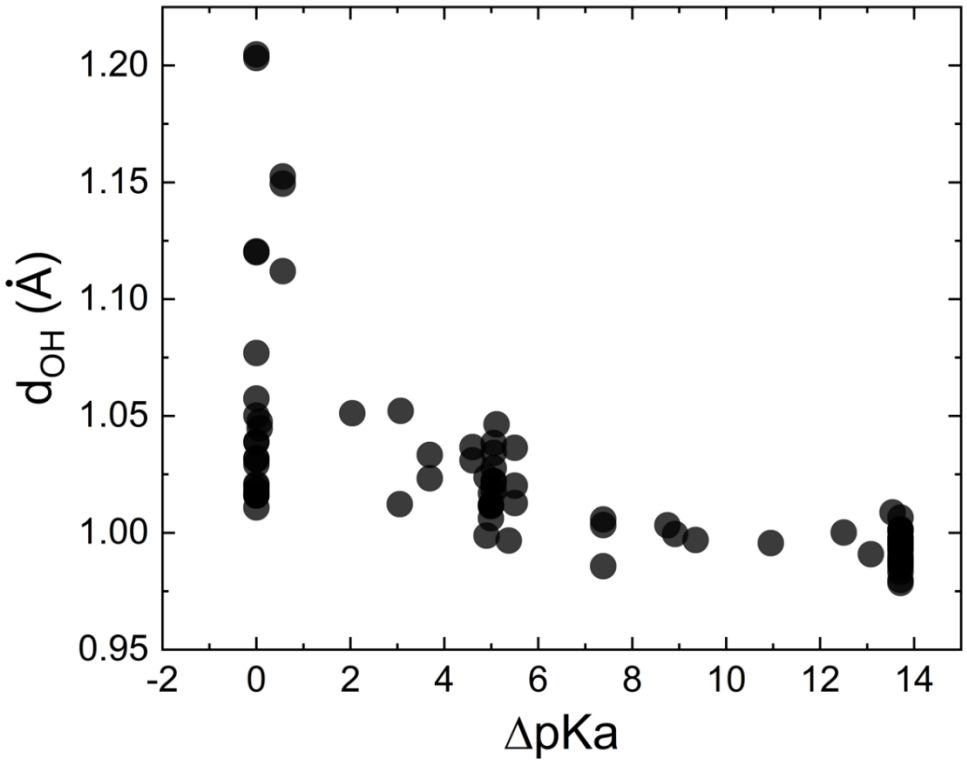

Figure S3. A scatter plot showing maximal initial O–H bond length d_{OH} vs. ΔpK_a.

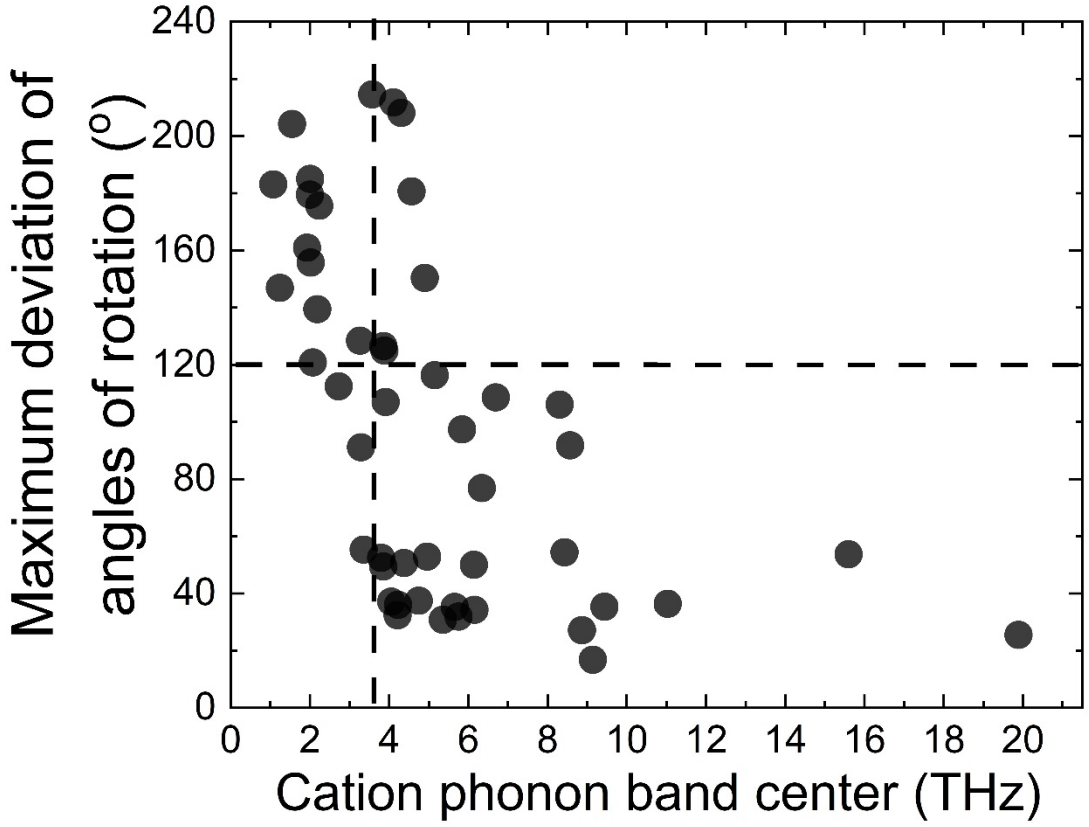

Figure S4. AIMD results from the initial dataset to extract group rotation descriptors: Complementary rotational metric (the maximal of the θ and φ angles as observed in AIMD) vs. cation phonon band center. Compounds that do not possess mono-oxyanion groups were excluded from analysis (Table S1). All data at 700 K.

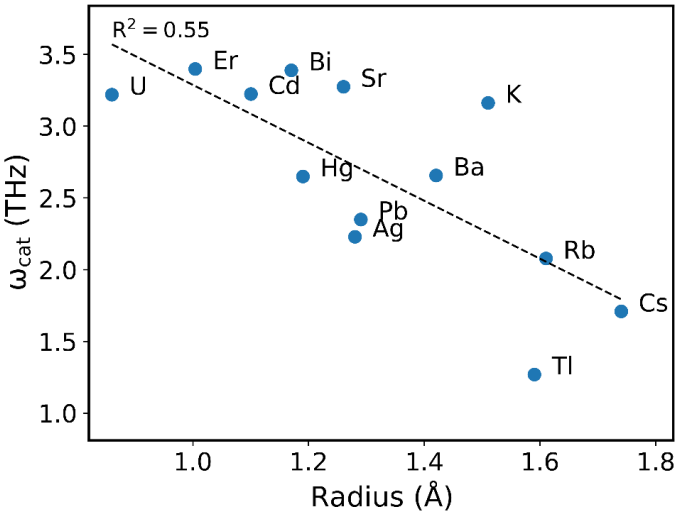

Figure S5. Average cation phonon band center vs. ionic radius[31].

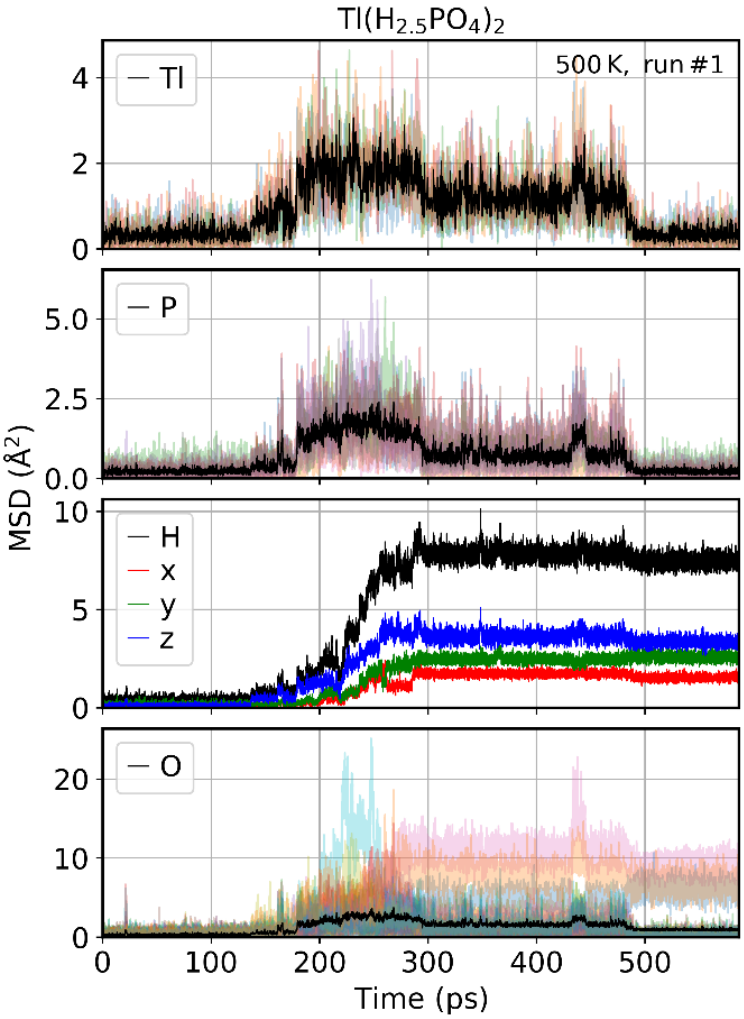

Figure S6. MSD vs. time plot for $Tl(H_{2.5}PO_4)_2$ (mp-696762) at 500 K (*NVT* simulation). Each panel show one type of atoms (Tl, P, H, O). Black lines show the average MSD (for each atom type). For Tl, P, O, also MSD of each individual atom is shown (different colors). For H, only the average MSD is shown (and x, y, z components of the MSD). In about first 150 ps no diffusion events were observed. Diffusion of hydrogen is accompanied by rotations of oxyanions (MSD of O) and displacement of Tl atoms. In fact, the whole crystal changes the structure (see radial distribution function (RDF) in Fig. S7).

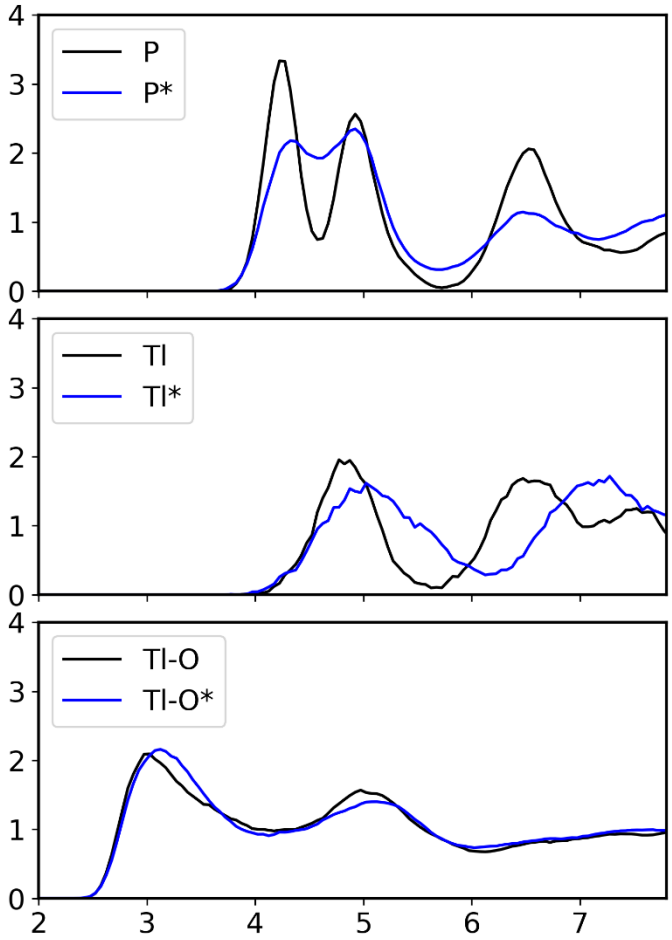

Figure S7. Radial distribution functions (RDFs) of P–P, Tl–Tl, and Tl–O ($Tl(H_{2.5}PO_4)_2$ at 500 K, *NVT* simulation). The black line corresponds to the 0 ps ≤ t < 100 ps interval, where no diffusion happened (see Fig. S6). The blue line corresponds to the 200 ps ≤ t < 300 ps interval, where proton diffuses (see Fig. S6).

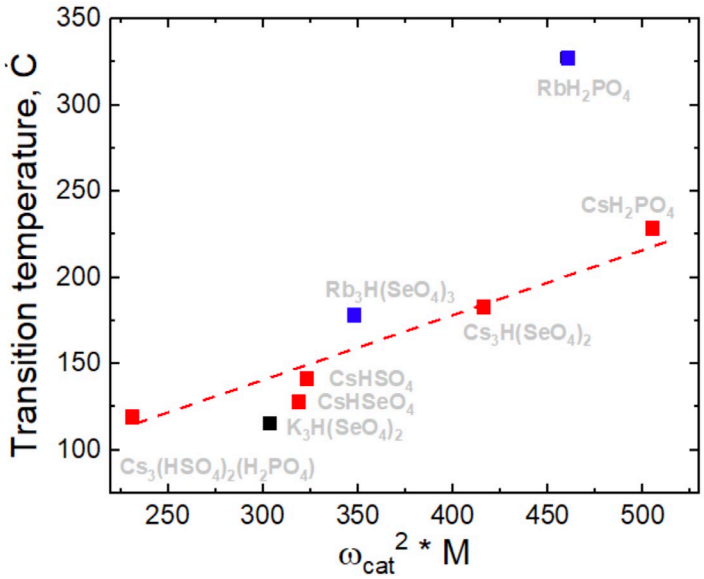

Figure S8. Experimental temperatures of superprotonic transition for Cs, Rb, K-based solid acids taken from references[32–41] vs. force constants associated with cation phonon band center. The line is the guide to the eye.

Section S3. Supporting Tables S1–S3

Table S1. List of compounds in the initial dataset. Compounds demonstrating fast proton diffusion at 700K (AIMD) are highlighted in bold; compounds that melted are marked with * (these compounds were excluded from further analysis); compounds that do not have mono-oxyanion groups (XO_n) are highlighted with † (these compounds were excluded from analysis of group rotation metrics).

MP-ID	Formula
mp-707424	$Al_2H_{12}(SeO_5)_3$ †
mp-850293	$Al_2H_{16}S_3O_{20}$ †
mp-722527	$AlH_{11}SO_{10}$ *
mp-1196038	$Ba_2B_{11}H_7O_{22}$
mp-1199009	$Ba_3B_6H_2O_{13}$ †
mp-23904	BaH_4O_3 †
mp-642844	BaH_5BrO_3 †
mp-642834	BaH_5ClO_3 †
mp-28587	BaH_8O_5 †
mp-626973	$Ba(HO)_2$ †
mp-706400	$BaP_2(HO_2)_4$
mp-720433	$Be_2As(HO)_9$ †
mp-23883	Be_2BHO_4 †
mp-24674	$Be_3P_2(H_2O_5)_2$ †
mp-757836	BeH_2SeO_4 †
mp-23996	BeH_8SO_8 *
mp-24348	BiP_4HO_{12} †
mp-1198782	$Ca_2Al_3H_3O_8$ †
mp-706291	$Ca_2B_8H_2O_{15}$ † *
mp-722262	$CaB_3H_3O_7$ †
mp-23701	$CaB_8H_4O_{15}$ †
mp-703574	$CaH_2(SO_4)_2$
mp-643898	$CaP_2H_2O_7$ †
mp-697657	$CaP_4(HO_6)_2$ †
mp-24390	$CaPHO_4$
mp-1197015	$CdP_2(H_4O_5)_2$
mp-743538	$CrH_{18}(O_3F)_3$ †
mp-758948	$Cs_2H_6CO_6$ *
mp-867975	$CsB_2H_5O_6$ †
mp-23742	$CsB_5(H_2O_3)_4$ †
mp-1192419	**$CsHSO_4$**
mp-1195974	**$CsP_2H_5O_8$**
mp-1197651	**$CsP(HO_2)_2$**
mp-1202664	$ErH_2Cl_3O_{13}$
mp-24640	**$ErH_3(SO_4)_3$**
mp-707317	ErP_4HO_{12} †
mp-643564	FeH_2SO_5
mp-24465	$GaP_3(HO_5)_2$ †
mp-697339	$Ho_2H_4CO_7$
mp-1195109	HoP_4HO_{12} †

mp-560314	$In_2H_{10}S_3O_{17}$
mp-762358	**$LaH_9(SO_6)_2$**
mp-757220	$LuH_6(ClO_5)_3$
mp-699232	$Mg_7P_2(HO_2)_8$
mp-23748	$MgBHO_3$ †
mp-24006	$MgH_{12}(BrO_6)_2$
mp-1202344	$MgH_{12}(ClO_6)_2$ *
mp-504894	$MgH_{12}SeO_9$
mp-24041	MgH_2SO_5
mp-24460	**$MgH_6(SO_4)_4$**
mp-865188	$MgH_8(ClO_6)_2$
mp-24620	$MgMoH_2O_5$
mp-24071	**$MnH_6(SO_4)_4$**
mp-865024	$MnP_2H_2O_7$ †
mp-1203000	**$NdH_{11}S_3O_{16}$**
mp-24574	**$NdH_5(SeO_4)_2$**
mp-757886	$NdHS_3O_{11}$
mp-1181715	$NiH_{10}(SeO_5)_2$
mp-1180537	$NiH_{12}(ClO_6)_2$*
mp-559054	$NiH_{12}SO_{10}$ *
mp-23877	$NiH_{14}SO_{11}$ *
mp-1202294	$NiP_2H_2O_7$ †
mp-698164	$Rb_2B_{10}H_6O_{19}$ †
mp-1197195	$Rb_2B_7H_5O_{14}$ †
mp-23781	$Rb_3H(SO_4)_2$
mp-863420	$Rb_4H_4C_3O_{10}$
mp-23700	RbH_2OF †
mp-28264	RbH_3O_2 †
mp-733612	**$RbH_3(SO_4)_2$**
mp-24022	**$RbH_3(SeO_3)_2$**
mp-1195562	$RbHS_2O_7$ †*
mp-1195896	**$RbHSO_4$**
mp-1199511	**$RbP_2H_5O_8$**
mp-23667	**$RbP(HO_2)_2$**
mp-24128	ReH_4NO_4 *
mp-541015	$ScH_{13}Cl_2O_7$ †*
mp-756748	$ScH_3(ClO_5)_2$
mp-696457	$ScP_3(HO_2)_6$
mp-542237	SmP_4HO_{12} †
mp-1201690	$Sr_3B_6H_2O_{13}$ †
mp-1195189	TbP_4HO_{12} †
mp-696762	**$TlP_2H_5O_8$**
mp-24656	$YH_3(SO_4)_3$
mp-24335	$YH_5(SeO_4)_2$
mp-24502	$ZnH_2(SO_4)_2$
mp-643066	ZnH_2SO_5
mp-706386	$ZrH_6O_3F_4$ †
mp-1192835	$ZrH_6(OF_3)_2$ †

Table S2. List of 143 compounds that passed the screening criteria (plus one extra compound, see †). d_{OH} is the maximal O–H distance in the pristine compound (such that either O donor atom or O acceptor atom, or both, belong to the mono-oxyanion group). ω_{cat} is the cation phonon band center (for compounds with multiple cations, the lowest value is provided). D_H is the hydrogen diffusivity as calculated at 650 K (MLFF—AIMD). The second to last column marks compounds that melted during MD ("melted", i.e., MSD of any cation exceeded 10 Å^2). The part of trajectory where compound melted was discarded and not used for D_H calculation. The last column provides comments and * marks promising compounds that showed fast diffusion and/or diffusional event(s) in MD.

#	MP-ID	Formula	d_{OH} (Å)	ω_{cat} (THz)	D_H (cm^2/s)	Melting	Comments
1	mp-867593	$Ag_2H_4(SO_4)_3$	1.031	2.22	–	melted	*; fast diffusion prior to melting
2	mp-24072	$AgH_5S_2O_9$	1.037	2.15	1.26E-05	melted	*
3	mp-707734	$AsHPb_4(ClO)_4$	1.013	2.06	1.60E-08		
4	mp-504751	$AsHPbO_4$	1.214	2.39	1.51E-08		
5	mp-758007	$Ba_2CaP_4(H_3O_8)_2$	1.036	2.66	1.64E-08		
6	mp-1199629	$Ba_2CdP_4(H_3O_8)_2$	1.028	2.69	1.84E-08		
7	mp-735530	$Ba_2Fe_3P_6HO_{22}$	1.200	2.59	1.03E-08		
8	mp-1194837	$Ba_3As_2H_{34}O_{25}$	1.015	2.78	1.33E-06	melted	*
9	mp-23806	$BaAsH_3O_5$	1.015	2.59	2.03E-08		
10	mp-23810	$BaAsHO_4$	1.045	2.63	1.55E-08		
11	mp-733968	$BaCaP_2(HO_4)_2$	1.058	2.53	1.28E-08		
12	mp-698325	$BaH_4(CO_2)_3$	1.113	2.68	5.71E-08		
13	mp-1199471	$BaNaAlMo_6(H_{13}O_{17})_2$	1.018	2.37	1.22E-05	melted	
14	mp-706400	$BaP_2(HO_2)_4$	1.023	2.69	1.52E-08		
15	mp-706543	$BaP_2(HO_2)_4$	1.210	3.14	1.13E-05	melted	*

16	mp-698163	$BaPHO_4$	1.050	2.53	3.46E-08		
17	mp-1105235	$BH(PbO_2)_2$	1.016	2.60	1.82E-08		
18	mp-1200655	$BiH_{25}C_8(SO_2)_6$	1.012	3.39	9.25E-06		*
19	mp-1202482	$Cd_3H_{18}S_4N_2O_{21}$	1.012	3.19	2.39E-07		
20	mp-1197015	$CdP_2(H_4O_5)_2$	1.022	3.26	5.26E-08		
21	mp-733936	$Cs_2MgH_8(CO_5)_2$	1.016	1.86	2.90E-07		*
22	mp-1181702	$Cs_3H(SeO_4)_2$	1.056	1.66	1.16E-06		*
23	mp-1192732	$Cs_3H(SeO_4)_2$	1.224	1.63	5.93E-07		*
24	mp-23980	$Cs_3H(SeO_4)_2$	1.227	1.70	2.91E-08		*
25	mp-1197340	$Cs_4AsH_5Se_3O_{16}$	1.056	1.54	2.55E-07		
26	mp-1185571	$CsAs(HO_2)_2$	1.056	1.82	3.06E-08		
27	mp-24141	$CsHSeO_4$	1.031	1.52	3.07E-06		*
28	mp-1192419	$CsHSO_4$	1.030	1.82	4.71E-06		*
29	mp-540981	CsHSO4	1.023	1.69	1.53E-05	melted	*
30	mp-557752	$CsHSO_4$	1.212	1.52	3.97E-08		*
31	mp-1195974	$CsP_2H_5O_8$	1.121	1.24	3.33E-07		*
32	mp-1198944	$CsP_2H_5O_8$	1.203	1.41	2.38E-08		
33	mp-1191996	$CsP_2(HO_2)_3$	1.145	1.30	2.30E-08		
34	mp-574928	$CsPH_3O_3F$	1.039	1.67	2.90E-07		*
35	mp-542541	$CsPH_3O_4F$	1.032	2.65	7.32E-07		*

36	mp-1188121	$CsP(HO_2)_2$	1.061	1.96	2.30E-08		
37	mp-1197651	$CsP(HO_2)_2$	1.048	2.05	2.54E-08		
38	mp-542557	$CsScAs_2(HO_4)_2$	1.015	1.72	3.21E-08		
39	mp-24640	$ErH_3(SO_4)_3$	1.008	3.40	1.62E-06		*
40	mp-757360	$HgTe(HO)_7$	1.013	2.65	1.86E-07		*
41	mp-24350	$K_2CoH_8(CO_5)_2$	1.0214	3.28	1.51E-08		
42	mp-604071	$K_2FePH_5(CO_5)2$	1.0207	3.12	7.97E-08		
43	mp-733853	$K_2HI_2ClO_6$	1.0242	3.2	2.58E-08		
44	mp-757963	$K_2MgH_8(CO_5)_2$	1.024	3.28	4.95E-08		
45	mp-850535	$K_2Mn_3H_{10}S_4O_{21}$	1.0147	2.95	3.47E-07		
46	mp-695963	$K_2NaH_{10}IO_{10}$	1.0163	3.29	1.42E-07	melted	
47	mp-505771	$K_2NaH_5(CO_4)_2$	1.0835	3.21	1.30E-07		
48	mp-766427	$K_2NaZn_2H_5(C_2O_7)_2$	1.2168	3.15	1.52E-08		
49	mp-24573	$K_2NiH_8(CO_5)_2$	1.0225	3.36	1.49E-08		
50	mp-721562	$K_2P_3H_5O_{11}$	1.1	3.02	3.18E-08	melted	
51	mp-707096	$K_2P_3H_7O_9$	1.0478	3.23	1.10E-07		*
52	mp-723043	$K_2TeH_6SeO_{10}$	1.0115	2.85	2.42E-07		
53	mp-867184	$K_2TeH_6SeO_{10}$	1.0139	2.9	4.77E-07		
54	mp-697127	$K_3HPdS_2(ClO_3)_2$	1.2044	3.15	2.63E-08		
55	mp-706983	$K_3HPtS_2(ClO_3)_2$	1.2023	3.19	2.28E-08		

56	mp-1200107	$K_3HS_3O_{11}$	1.016	3.1	4.50E-08		
57	mp-23979	$K_3H(SeO_4)_2$	1.2241	3.08	5.29E-06		*
58	mp-23779	$K_3H(SO_4)_2$	1.2177	3.31	4.88E-07		*
59	mp-557941	$K_4AsH_5S_3O_{16}$	1.0373	3.19	3.66E-07		*
60	mp-746688	$K_4MnH_6(S_2O_9)_2$	1.0316	3.29	1.20E-07		
61	mp-697128	$KCa_2P_4H_{11}O_{18}$	1.1958	3.23	8.48E-08		
62	mp-746371	$KCo_2As_2HO_8$	1.0755	3.13	1.80E-08		
63	mp-735586	$KFe_2H(SeO_3)_4$	1.0289	3.19	1.97E-08		
64	mp-23682	$KH_3(CO_2)_2$	1.1421	3.1	2.43E-06	melted	
65	mp-1198088	KH_3CO_6	1.0096	3.02	2.35E-07		
66	mp-1180593	$KH_3(SeO_3)_2$	1.047	3.38	2.29E-07		*
67	mp-706579	$KH_3(SeO_3)_2$	1.2224	3.36	1.31E-07		*
68	mp-1200593	$KH_3(SeO_4)_2$	1.0748	2.6	1.43E-05	melted	*
69	mp-733655	$KH_3(SO_4)_2$	1.0466	3	5.37E-06	melted	*
70	mp-697284	KH_3SO_6	1.0198	3.28	7.59E-07		
71	mp-23724	$KHCO_3$	1.0345	3.11	3.38E-08		
72	mp-634431	$KHCO_3$	1.0368	3.49	2.80E-08		
73	mp-706273	$KH(IO_3)_2$	1.0646	3.41	2.15E-08		
74	mp-720407	$KH(IO_3)_2$	1.011	3.46	1.13E-06	melted	
75	mp-766021	$KH(IO_3)_2$	1.2187	3.42	1.93E-08		

76	mp-24433	$KHSeO_3$	1.0136	3.03	4.48E-08		
77	mp-707536	$KHSeO_4$	1.0376	3.45	1.84E-07		*
78	mp-23800	$KHSO_4$	1.0336	3.19	6.02E-08		*
79	mp-1200428	$KMgAs(H_6O_5)_2$	1.0134	2.89	4.43E-08		
80	mp-23905	$KMgH_9(CO_5)_2$	1.2182	2.76	9.98E-08		
81	mp-721617	$KMgP(H_6O_5)_2$	1.0107	2.58	4.54E-08		
82	mp-1192905	$KNaTe(HO)_6$	1.0076	3.07	1.60E-07	melted	
83	mp-697458	$KNiH_9(CO_5)_2$	1.2168	2.92	4.88E-08		
84	mp-1197171	$KP_2H_5O_8$	1.1654	2.96	2.48E-08	melted	
85	mp-706608	KPH_3O_3F	1.0287	3.49	2.55E-08		
86	mp-1197097	KPH_3O_4	1.0468	3.15	1.34E-07		*
87	mp-24214	KPH_3O_4F	1.0303	2.91	3.51E-08		
88	mp-761185	KPH_3O_4F	1.0339	3.11	2.67E-08	melted	
89	mp-1106168	$KP(HO_2)_2$	1.2088	3.37	1.58E-08		
90	mp-1198928	$KP(HO_2)_2$	1.0783	3.03	7.56E-07		*
91	mp-23959	$KP(HO_2)_2$	1.057	3.22	2.04E-08		
92	mp-24262	$KP(HO_2)_2$	1.2116	3.31	2.04E-08		
93	mp-24263	$KP(HO_2)_2$	1.2128	3.41	1.96E-08		
94	mp-696752	$KP(HO_2)_2$	1.2093	3.23	1.28E-08		
95	mp-699437	$KP(HO_2)_2$	1.0326	3.2	6.85E-06	melted	*

96	mp-757909	$KP(HO_2)_2$	1.0801	3.17	2.92E-08		
97	mp-644015	$KVHSe_2O_7$	1.2162	3.41	1.65E-08		
98	mp-541071	$KZrP_2HO_8$	1.0773	3.28	2.16E-08		
99	mp-556009	$MgTlAs(H_6O_5)_2$	1.013	1.09	5.82E-08		
100	mp-554894	$MgTlP(H_6O_5)_2$	1.011	0.88	3.16E-08		
101	mp-24731	$Ni_3AgP_3(HO_6)_2$	1.084	2.32	1.48E-08		
102	mp-24624	$Rb_2CoH_8(CO_5)_2$	1.022	2.21	1.23E-07		
103	mp-24606	$Rb_2MgH_8(CO_5)_2$	1.018	2.28	2.07E-08		
104	mp-505772	$Rb_2NaH_5(CO_4)_2$	1.0947	2.12	6.43E-06	melted	
105	mp-753854	$Rb_2PH_3SeO_8$	1.073	2.08	4.42E-07		*
106	mp-603414	$Rb_2TeH_6SeO_{10}$	1.032	1.75	8.51E-07		
107	mp-23897	$Rb_3H(SeO_4)_2$	1.224	2.26	1.12E-06		*
108	mp-1197110	$Rb_3H(SO_4)_2$	1.013	2.14	1.26E-06		*
109	mp-23781	$Rb_3H(SO_4)_2$	1.219	2.30	1.06E-06		*
110	mp-863420	$Rb_4H_4C_3O_{10}$	1.050	2.20	4.69E-08		
111	mp-707406	$Rb_6Te_3P_6(H_{13}O_{20})_2$	1.020	1.96	9.91E-08	melted	
112	mp-699453	$RbAs_2H_5O_8$	1.209	1.88	9.86E-07		*
113	mp-1196732	$RbH_2(IO_3)_3$	1.024	2.00	4.25E-07		*
114	mp-23688	$RbH_3(CO_2)_2$	1.146	2.04	–	melted	
115	mp-24022	$RbH_3(SeO_3)_2$	1.057	1.86	1.09E-07		*

116	mp-733612	$RbH_3(SO_4)_2$	1.059	1.93	6.62E-06	melted	*
117	mp-695829	$RbHSeO_4$	1.047	2.44	6.30E-08		*
118	mp-696794	$RbHSeO_4$	1.034	2.05	4.30E-07		*
119	mp-1195896	$RbHSO_4$	1.019	2.20	3.35E-06	melted	
120	mp-1197989	$RbHSO_4$	1.022	2.24	1.31E-06	melted	*
121	mp-707377	$RbHSO_4$	1.013	2.00	–	melted	
122	mp-559463	$RbMgAs(H_6O_5)_2$	1.012	1.73	3.11E-08		
123	mp-604725	$RbMgP(H_6O_5)_2$	1.014	1.86	3.06E-08		
124	mp-1199511	$RbP_2H_5O_8$	1.205	1.97	2.05E-08		
125	mp-1190228	$RbP_2(HO_2)_3$	1.127	1.68	4.12E-08		
126	mp-761252	$RbPH_3O_4F$	1.032	2.24	2.03E-08		
127	mp-703312	$RbPH_4O_5$	1.015	2.02	1.58E-06		*
128	mp-23667	$RbP(HO_2)_2$	1.053	2.17	2.20E-08		
129	mp-642831	$RbP(HO_2)_2$	1.209	2.15	1.47E-08		
130	mp-643791	$RbP(HO_2)_2$	1.210	2.29	2.19E-08		
131	mp-703528	$RbP(HO_2)_2$	1.216	2.15	2.10E-08		
132	mp-722348	$RbP(HO_2)_2$	1.029	1.81	1.76E-06	melted	*
133	mp-867132	$RbZnH_2Se_2BrO_6$	1.023	2.35	2.12E-08		
134	mp-867129	$RbZnH_2Se_2ClO_6$	1.015	2.21	1.16E-07	melted	
135	mp-510709	$SrAsH_3O_5$	1.015	3.13	2.20E-08		

136	mp-757723	$SrH_4(SO_4)_3$	1.038	3.55	2.52E-06		* †
137	mp-863961	$SrNi_2P_2(H_2O_5)_2$	1.055	3.42	1.11E-07		*
138	mp-696762	$TlP_2H_5O_8$	1.203	1.19	9.40E-07		*
139	mp-1196113	$TlP(HO_2)_2$	1.048	1.34	4.87E-08		
140	mp-643701	$TlP(HO_2)_2$	1.212	1.34	2.11E-08		
141	mp-690711	$TlP(HO_2)_2$	1.204	1.76	2.75E-08		*
142	mp-697267	$TlP(HO_2)_2$	1.051	1.30	3.26E-08		
143	mp-768283	$UAs_2H_6O_{11}$	1.016	3.22	6.61E-08		
144	mp-542045	$UTlH(SeO_4)_2$	1.009	1.25	6.03E-08		

† this compound has ω_{cat} = 3.55 THz just above the cut-off, therefore we run MLFF–AIMD to explore the potential diffusivity of this Sr-based compounds

Table S3. List of pK_a values used to calculate descriptors[26–29].

Acid	pK_a
HF	3.1
NH_4	9.24
H_2O	14
H_3O	-1.74
$HAsO_4$	11.3
H_2AsO_4	7.10
H_3PO_4	2.16
H_2PO_4	7.21
HPO_4	12.32
H_2SO_4	-3
HSO_4	2
H_2SeO_3	2.62
$HSeO_3$	8.32
HCO_3	10.3
H_2CO_3	6.3
H_3BO_3	9.2

References

1 Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater* 2013; **1**: 011002.

2 Blöchl PE. Projector augmented-wave method. *Phys Rev B* 1994; **50**: 17953.

3 Joubert D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys Rev B Condens Matter Mater Phys* 1999; **59**: 1758–1775.

4 Perdew JP, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Phys Rev Lett* 1996; **77**: 3865–3868.

5 Perdew JP, Burke K, Ernzerhof M. Erratum: Generalized gradient approximation made simple (Physical Review Letters (1996) 77 (3865)). *Phys Rev Lett* 1997; **78**: 1396.

6 Kresse G, Hafner J. Ab initio molecular dynamics for liquid metals. *Phys Rev B* 1993; **47**: 558–561.

7 Kresse G, Hafner J. Ab initio molecular-dynamics simulation of the liquid-metalamorphous- semiconductor transition in germanium. *Phys Rev B* 1994; **49**: 14251–14269.

8 Kresse G, Furthmüller J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys Rev B Condens Matter Mater Phys* 1996; **54**: 11169–11186.

9 Kresse G, Furthmüller J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput Mater Sci* 1996; **6**: 15–50.

10 Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys* 2010; **132**: 154104.

11 Dudarev SL, Botton GA, Savrasov SY, Humphreys CJ, Sutton AP. Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study. *Phys Rev B* 1998; **57**: 1505.

12 Nosé S. A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys* 1998; **81**: 511.

13 Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A (Coll Park)* 1985; **31**: 1695.

14 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**: 225701.

15 Jinnouchi R, Karsai F, Kresse G. On-the-fly machine learning force field generation: Application to melting points. *Phys Rev B* 2019; **100**: 14105.

16 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**: 234102.

17 Deng B, Zhong P, Jun KJ, Riebesell J, Han K, Bartel CJ *et al.* CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nat Mach Intell* 2023; **5**. doi:10.1038/s42256-023-00716-3.

18 Haile SM, Boysen DA, Chisholm CRI, Merie RB. Solid acids as fuel cell electrolytes. *Nature* 2001; **410**: 910–913.

19 Kreuer K-D. Proton Conductivity: Materials and Applications. *Chemistry of Materials* 1996; **8**: 610–641.
20 Kreuer KD. Fast proton conductivity: A phenomenon between the solid and the liquid state? *Solid State Ion* 1997; **94**: 55–62.
21 Baranov AI. Crystals with disordered hydrogen-bond networks and superprotonic conductivity. Review. *Crystallography Reports* 2003; **48**: 1012–1037.
22 Fischer SA, Dunlap BI, Gunlycke D. Correlated dynamics in aqueous proton diffusion. *Chem Sci* 2018; **9**. doi:10.1039/c8sc01253a.
23 Miyoshi S, Akao Y, Kuwata N, Kawamura J, Oyama Y, Yagi T *et al.* Low-temperature protonic conduction based on surface protonics: An example of nanostructured yttria-doped zirconia. *Chemistry of Materials* 2014; **26**. doi:10.1021/cm5012923.
24 Muy S, Bachman JC, Giordano L, Chang HH, Abernathy DL, Bansal D *et al.* Tuning mobility and stability of lithium ion conductors based on lattice dynamics. *Energy Environ Sci* 2018; **11**: 850–859.
25 Smith JG, Siegel DJ. Low-temperature paddlewheel effect in glassy solid electrolytes. *Nature Communications 2020 11:1* 2020; **11**: 1–11.
26 Evans DA, Ripin DH. pKa's of Inorganic and Oxo-Acids. https://www.wanglab.chem.pitt.edu/wp-content/uploads/2017/09/Evans-D.-A.-Ripin-D.-H.-The-Evans-pKa-Table-Harvard-2006.pdf (accessed 12 Feb2024).
27 Peacock CJ, Nickless G. The Dissociation Constants of some Phosphorus(V) Acids. *Zeitschrift fur Naturforschung - Section A Journal of Physical Sciences* 1969; **24**: 245–247.
28 Engineering ToolBox, (2017). Inorganic Acids and Bases - pKa Values. https://www.engineeringtoolbox.com/pKa-inorganic-acid-base-hydrated-metal-ion-monoprotic-diprotic-triprotic-tetraprotic-d_1950.html (accessed 12 Feb2024).
29 Haynes WM (ed.). *CRC Handbook of Chemistry and Physics*. 95th Edition. Taylor & Francis, 2014.
30 Dreßler C, Sebastiani D. Effect of anion reorientation on proton mobility in the solid acids family CsH: YXO4 (X = S, P, Se, y = 1, 2) from ab initio molecular dynamics simulations. *Physical Chemistry Chemical Physics* 2020; **22**: 10738–10752.
31 Database of Ionic Radii, http://abulafia.mt.ic.ac.uk/shannon/ptable.php. .
32 Jiráak Z, Dlouhá M, Vratislav S, Balagurov AM, Beskrovnyi AI, Gordelii VI *et al.* A neutron diffraction study of the superionic phase in CsHSO4. *Physica Status Solidi (a)* 1987; **100**: K117–K122.
33 Haile SM, Lentz G, Kreuer KD, Maier J. Superprotonic conductivity in Cs3(HSO4)2(H2PO4). *Solid State Ion* 1995; **77**: 128–134.
34 Zakharov MA, Troyanov SI, Kemnitz E. Superprotonic high temperature phase and refinement of the low temperature structure of cshseo4. *Zeitschrift fur Kristallographie* 2001; **216**: 172–175.
35 Boysen DA, Haile SM, Liu H, Secco RA. High-temperature behavior of CsH2PO4 under both ambient and high pressure conditions. *Chemistry of Materials* 2003; **15**: 727–736.
36 Lavrova G V., Burgina EB, Matvienko AA, Ponomareva VG. Bulk and surface properties of ionic salt CsH5(PO4)2. *Solid State Ion* 2006; **177**: 1117–1122.
37 Ikeda A, Haile SM. The thermodynamics and kinetics of the dehydration of CsH2PO4 studied in the presence of SiO2. *Solid State Ion* 2012; **213**: 63–71.

38 Sanghvi S, Haile SM. Crystal structure, conductivity, and phase stability of Cs3(H1.5PO4)2 under controlled humidity. *Solid State Ion* 2020; **349**: 115291.
39 Yi D, Sanghvi S, Kowalski CP, Haile SM. Phase Behavior and Superionic Transport Characteristics of (MxRb1- x)3H(SeO4)2 (M = K or Cs) Solid Solutions. *Chemistry of Materials* 2019; **31**: 9807–9818.
40 Chisholm CRI, Haile SM. Superprotonic behavior of Cs2(HSO4)(H2PO4) - a new solid acid in the CsHSO4-CsH2PO4 system. *Solid State Ion* 2000; **136–137**: 229–241.
41 Chisholm CRI, Merle RB, Boysen DA, Haile SM. Superprotonic phase transition in CsH(PO3H). *Chemistry of Materials* 2002; **14**: 3889–3893.