Supporting Information

Resolving Atomic Structure of γ-Alumina: A Non-Spinel Phase with Distorted Anion Lattice and Three Adjacent Long Channels

Xiao Yang [a], Cheng Shang[a,*], and Zhi-Pan Liu [a,b,*]

[a] *Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, iChEM, Department of Chemistry, Fudan University, Shanghai 200433, China*
[b] *Key Laboratory of Synthetic and Self-Assembly Chemistry for Organic Functional Molecules, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China*

1. Calculation details

1.1 Construction of global dataset using SSW-NN

Undoubtedly, the dataset used for training the NN determines largely the quality of the potential energy surface (PES) of G-NN. Our previous works have shown that the stochastic surface walking (SSW) global optimization can be used to fast generate a global dataset, which incorporates different structural patterns on the global PES. The SSW PES search is fully automated and does not need a priori knowledge on the system, such as the structure motif (e.g., bonding patterns, symmetry) of materials. The final Al-O-H global dataset in this work is detailed in Table S1. In brief, the SSW-NN method involves three stages for constructing the global dataset, as described in the following.

(i) **The first stage** constructs a raw dataset, which contains the most common atomic environment and serves as the training dataset for building an initial NN PES. This is done by performing density functional theory (DFT) SSW global optimization in a massively parallel way. The DFT calculation is typically with low accuracy setups and restricted to small unit cells to speed up the SSW search. By collecting and screening the structures from the SSW trajectories, a raw dataset is finally obtained.

(ii) **The second stage** trains an NN global PES. This is done by first refining the dataset using first principles calculation with high accuracy setups, followed by the NN training on the accurate global dataset. The NN architecture applied in this stage utilizes a small set of structural descriptors and a small network size.

(iii) **The third stage** iteratively expands the global dataset. It targets to increase the predictive power of NN PES by incorporating more structural patterns into the dataset. This is done by carrying out SSW PES search using the NN PES obtained in the second stage, starting from a variety of initial structures. These initial structures are often randomly configurated and also include large systems with many atoms per unit cell. It is worth noting that the molecular dynamics is performed during SSW exploration at a predefined frequency, e.g., once every 10 SSW steps, to include some configuration from high temperature region of the PES. The structures from all SSW trajectories are collected and filtered to generate the additional dataset. This new dataset is then fed to the global dataset (back to stage 2) to start a new cycle of NN training.

1.2 G-NN Potential Generated from the SSW-NN Method

All simulations based on G-NN potential were carried out using LASP code developed in our group, which implements data generation using SSW global optimization, G-NN training, and potential energy surface (PES) evaluation of G-NN potentials.[1-4] The Al-O-H ternary G-NN potential was trained by self-learning the SSW global potential energy surface dataset that covers a wide range of $(Al_2O_3)_x(H_2O)_y$ compositions with different value of x and y. More than 10^6 structures on Al-O-H global PES were visited by SSW-NN during NN potential generation and the final training dataset of Al-O-H consists of 13610 structures that are selected to represent the global PES. The dataset was calculated using plane-wave DFT calculations[5] as implemented in VASP (Vienna Ab-initio Simulation Package) (see below). The training dataset is available on online.[6]

The G-NN potential has a five-layer (269-80-80-80-5) feed-forward MBNN architecture for each element,[7] in total containing 104895 fitting parameters.[1,8] The details on the training dataset are shown in Supporting information (SI) Table S1. The rot-mean-square- errors (RMSE) for the energy and the force of the G-NN are 4.882 meV/atom and 0.177 eV/Å, respectively. This G-NN potential is now included in the G-NN library of LASP and available online.[9] We have also

benchmarked the G-NN accuracy against the DFT results for important structures, which shows that the RMSE of potential energy is less than 1.66 meV/atom for low energy structures in this work (see SI Table S2). This small error suggests that the G-NN PES is a good approximation to DFT PES and can be utilized to expedite the global structure search and pathway determination. With G-NN to expedite PES calculations, all results reported in this work are finally converged using DFT calculations (see the following calculation setups).

The trained potential is then used for the SSW simulation. The central idea of SSW method comes from the bias-potential-driven dynamics[10] and Metropolis Monte Carlo (MC)[11]. It manipulates smoothly the structural configuration from one minimum to another on PES, and relies on Metropolis MC at a given temperature to decide the acceptance of the move. A series of consecutive minimum structures are generated in SSW simulation forming a continuous trajectory. Each step in SSW, also termed as an MC step, comprises three independent parts, namely, (i) the climbing, (ii) the relaxation and (iii) the Metropolis MC. Such an MC step utilizes the climbing module to move uphill and the relaxation module to locate minimum. Once a minimum is reached, the Metropolis MC is used to judge whether the structure will be accepted or refused. The climbing procedure lies at the heart of the SSW method, involving repeated bias-potential-driven structure extrapolation and local geometry optimization to minima R_t^n on the biased PES, which drags gradually R_i^0 to a high energy configuration R_i^H, where "*i*" is the index of the current MC step. Starting from the current minimum R_i^0, SSW first generates a random direction N_i^0, a normalized vector defining the direction to change the current geometry. The random direction is then softened by the so-called biased rotation towards one eigenvector of Hessian matrix with small eigenvalues (not necessarily the lowest one). A series of bias Gaussian potentials v_n (*n* is the index of the bias potential, *n*=1,2$\cdots$*H*) is added one by one consecutively along the softened direction N_i^n and thus creates a series of local minima R_t^n along the moving trajectory on the modified PES. For each configuration, we parallelly run 3-10 different trajectories. The SSW search is regarded as converged when all the trajectories find the same GM, which is not updated within 1000 SSW steps.

1.3 DFT Calculations

The G-NN potential training relies on DFT calculated energy, forces, and stresses of structures.[1] To achieve high accuracy and data consistency, our DFT calculations in VASP utilized the following setups as well as utilized for generating all G-NN potentials in LASP G-NN library: the DFT functional being at the optB88-vdW functional level,[12,13] which shown good accuracy in computing the geometry and energies of several solid phase of alumina[14]; the kinetic energy cut off being 450 eV; the projector augmented wave (PAW) pseudopotential[15] utilized to describe ionic core electrons; the fully automatic Monkhorst–Pack K-mesh with 25 times the reciprocal lattice vectors[16] for the first Brillouin zone k-point sampling. To optimize the structure, we minimized the total energy until the total forces on each atom were smaller than 0.01 eV/Å. It should be mentioned that we have confirmed all important minima by DFT calculations and thus, if otherwise specifically mentioned, all energetics data reported in this work are from DFT calculations.

INCAR file:
NELMIN = 5
NELMDL= -5
EDIFF = 1e-5
EDIFFG = -1e-2
GGA = PE

ISMEAR = 0
NCORE=8
ENCUT = 450
LWAVE = .FALSE.
LCHARG = .FALSE.
LUSE_VDW = .TRUE.
PARAM1 = 0.1833333333
PARAM2 = 0.22
AGGAC = 0.0
GGA = BO
LASPH = .TRUE.
IBRION = 2
ISYM = 0
NSW = 99
ISIF = 3

1.4 Free-Energy Calculations

The relative energy (ΔE) of different hydrated alumina structures as computed based on Eq. (S1) are computed for evaluating the stability of $(Al_2O_3)_xH_2O$ at different ratio (Fig. S1). The free energy of the GM of AlOOH, HAl_5O_8, HAl_9O_{14} and Al_2O_3 are computed to construct the phase diagram (Fig. S1). For solid states, the free energy considering the zero-point energy (ZPE), vibrational entropy $S_{vib}$and *PV* terms as Eq. S2. For water molecules (Eq. S3), the free energy G_{H2O} includes DFT energy (E_{H2O}), zero-point energy (ZPE), and the standard gas-phase thermodynamic correction terms at varied temperatures and pressures.[17]

$\Delta E((Al_2O_3)_xH_2O) = [E((Al_2O_3)_xH_2O) - 2\cdot E(AlOOH) - (x-1)\cdot E(Al_2O_3)] \times \frac{1}{x}$. (S1)

$G_{solid} = E_{solid} + ZPE - TS_{vib} + PV$ (S2)

$G_{H_2O} = E_{H_2O} + H_T + ZPE - \Delta TS_{H_2O} + RTlnP$ (S3)

1.5 RDF Calculations

To compute the radial distribution functions (RDF) of model γ-AD and Ouyang-2009, we first build 2×2×2 supercell for model γ-AD (320 atoms per cell) and 2×2×1 for model Ouyang-2009 (640 atoms per cell). Molecular dynamics (MD) simulations are then performed using G-NN under 300 K for 10 ps in the isothermal-isobaric (NPT) ensemble, capturing snapshots every 10 fs to generate the corresponding (RDF) of two models.

1.6 Calculation detail of the pH_{PZC} of γ-AD

We build a slab model of the γ-AD p(1×1) (512) surface (14.67 Å × 19.55 Å), corresponding to the (110) surface of the fcc oxygen sublattice, to simulate the the surface point-of-zero charge, i.e. pH_{PZC} value of γ-AD. The model consists of six layer (240 atoms) with the bottom four layer fixed at the bulk-truncated position and a 15 Å vacuum above the surface. The vacuum is filled with 136 H_2O molecules, which are confined to the surface by a Lennard-Jones wall potential positioned in the vacuum to maintain the water density at ~1.0 g·cm^{-3}. To equilibrate the water-Al_2O_3 interface, we begin with the most stable structure obtained from 10, 000 SSW-NN steps, followed by a 100 ps MD simulation at 298 K in the NVT ensemble. The final Al_2O_3 surface is covered by 9 H_2O molecule and 13 dissociated OH^- adsorbed on Al atoms and 13 H^+ adsorbed on O atoms.

We then calculate the free energy profiles for protonation and deprotonation of the equilibrated surface through enhanced MD simulations from three independent runs, which were performed in the canonical ensemble (NVT) with Nóse-Hoover thermostat at 298 K, using a small time-step of 0.5fs. To obtain the free energy profile, we apply the umbrella sampling (US) method, a biased MD technique that calculates the mean force along the collective variable. The transition from protonated to deprotonated states was divided into 71 windows based on a collective variable,

with a simulation time of 10 ps for each window. The US method imposes a harmonic bias potential (Eq. S4) on a designated reaction coordinate to constrain the system within each window i. The value of κ was set to 10 eV in this work. The biased probability for each window $P_i^b(\xi)$ is collected from MD trajectories (Eq. S5), enabling an accurate calculation of free energy changes. The derivative of free energy with respect to the reaction coordinated ξ, $(\partial G_i)/\partial\xi$, was derived as shown in Eq. S6, where $\overline{\xi_i^b}$ represents the mean value of ξ, and σ_i^b represented the variance. The obtained free energy change of protonation (ΔF_p) and deprotonation (ΔF_d) process are then used to calculate the point of zero charge (pH_{PZC}) by Eq. S7.

$$\omega_i = \frac{\kappa}{2}\left(\xi - \xi_i^{ref}\right)^2 \quad (S4)$$

$$P_i^b(\xi) = \frac{\int e^{-\beta[E+\omega_i(\xi')]}\delta(\xi'-\xi)d^N r}{\int e^{-\beta[E+\omega_i(\xi')]}d^N r}, \qquad \beta = (k_B T)^{-1} \quad (S5)$$

$$\frac{\partial G_i}{\partial \xi} = \frac{\xi - \overline{\xi_i^b}}{\beta\left(\sigma_i^b\right)^2} - \kappa\left(\xi - \xi_i^{ref}\right) \quad (S6)$$

$$pH_{PZC} = \frac{\mathrm{p}K_w - log_{10}(e^{\frac{\Delta F_p}{RT}}) + log_{10}(e^{\frac{\Delta F_d}{RT}})}{2} \quad (S7)$$

1.7 Definition of the collective variable in the US method

The collective variable (ξ) used in the US method is the coordination number of surface oxygen atoms with nearby hydrogen atoms, calculated using Eq. S8-9, where n_H represents the total number of H atoms in the system, n_{s-O} denotes the number of surface oxygen atoms within a cutoff distance (2.0 Å) from the ith H atom. The coordination number CN_{ji} of each oxygen atom O_j, contributed by each hydrogen atom H_i, is computed based on the distance d_{ij} between O_j and H_i, with d_0 (-0.16 Å) and r_0 (1.37 Å) as predefined parameters. The equilibrium state has a ξ value of 42.26 $\pm$ 0.02, corresponding to 44 H atoms on the surface.

$$\xi = \sum_i^{n_H}\sum_j^{n_{s-O}} \frac{{CN_{ji}}^2}{\sum_j^{n_{s-O}} CN_{ji}} \quad (S8)$$

$$CN_{ji} = \sum_j^{n_{ns-O}} \frac{1}{1+\left(\frac{d_{ij}-d_0}{r_0}\right)^{20}} \quad (S9)$$

2. Training dataset of the AlOH G-NN potential

Table S1 | Structure information in the first principles global dataset. Listed data are the number of the structures in the global dataset, as distinguished by the chemical formula (Species), the number of atoms per cell (N_{atm}), the type of structures, being layer (N_{lay}) and bulk (N_{bul}). Total number of structures (N_{tot}) are also summarized.

Species	N_{atm}	N_{lay}	N_{bul}	N_{tot}
O6-Al4	10	5	117	122
O12-Al8	20	1086	257	1343
O30-Al20	50	49	30	79
O33-Al22	55	51	25	76
O36-Al24	60	42	24	66
H1-O35-Al23	59	54	19	73
H2-O28-Al18	48	0	3	3

H2-O31-Al20	53	0	223	223
H2-O34-Al22	58	60	33	93
H3-O18-Al11	32	156	42	198
H3-O33-Al21	57	62	27	89
H3-O36-Al23	62	0	125	125
H4-O20-Al12	36	329	277	606
H4-O32-Al20	56	95	304	399
H4-O62-Al40	106	0	2	2
H5-O31-Al19	55	67	37	104
H6-O18-Al10	34	202	221	423
H6-O21-Al12	39	2	80	82
H6-O33-Al20	59	7	171	178
H6-O36-Al22	64	75	155	230
H8-O16-Al8	32	0	2	2
H8-O19-Al10	37	189	27	216
H8-O22-Al12	42	689	522	1211
H8-O34-Al20	62	4	280	284
H8-O40-Al24	72	169	3	172
H8-O64-Al40	112	0	1	1
H9-O18-Al9	36	154	44	198
H9-O36-Al21	66	0	160	160
H10-O20-Al10	40	170	36	206
H10-O23-Al12	45	534	516	1050
H10-O35-Al20	65	1	150	151
H12-O24-Al12	48	0	185	185
H12-O36-Al20	68	109	200	309
H12-O66-Al40	118	0	2	2
H14-O7	21	1	77	78
H15-O36-Al19	70	8	215	223
H16-O8	24	1	405	406
H16-O38-Al20	74	131	0	131
H16-O44-Al24	84	379	3	382
H16-O68-Al40	124	0	7	7
H18-O36-Al18	72	90	4	94
H20-O40-Al20	80	87	0	87
H20-O46-Al24	90	299	7	306
H20-O70-Al40	130	0	3	3
H30-O15	45	1	11	16
H34-O41-Al16	91	80	49	129
H36-O42-Al16	94	132	97	229
H37-O41-Al15	93	71	53	124
H38-O43-Al16	97	291	196	487
H39-O42-Al15	96	105	123	228
H40-O41-Al14	95	134	100	234

H40-O44-Al16	100	189	75	264
H41-O43-Al15	99	67	49	116
H68-O82-Al32	182	102	9	111
H72-O84-Al32	188	199	9	208
H74-O82-Al30	186	110	4	114
H76-O86-Al32	194	409	21	430
H78-O84-Al30	192	228	5	233
H80-O82-Al28	190	188	14	202
H82-O86-Al30	198	101	6	107
total	--	7764	5842	13610

3. Benchmark of G-NN potential against DFT calculations

Table S2 | Benchmark of NN calculations for AlOH systems as compared with DFT results. Listed data includes the compositions, structure, DFT energy, NN energy and energy differences between DFT energy and NN energy (E_{diff}, meV/atom)

No.	Species	N_{atom}	E_{DFT}/eV	E_{nn}/eV	$\|\Delta(E_{nn}-E_{DFT})\|$/(meV/atom)
1	Al_2O_3	50	0.000	0.034	0.679
			-0.747	-0.071	0.083
			0.143	0.060	1.662
2	$(Al_2O_3)_9H_2O$	48	0.000	-0.042	0.873
			-0.030	-0.014	0.345
			-0.009	0.006	0.305
3	$(Al_2O_3)_5H_2O$	56	0.000	-0.032	0.563
			0.298	0.296	0.027
			0.354	0.365	0.197
4	$Al_{196}O_{294}$	480	0.000	1.845	3.844
			1.320	2.878	3.246
			4.686	5.993	2.722
			-0.921	0.411	2.775
			-1.029	0.500	3.184
			1.109	2.756	3.431
			5.178	6.774	3.326
			2.465	4.015	3.229
			1.699	3.341	3.421
			5.041	6.875	3.820

4. Thermodynamics of bulk $Al_2O_3(H_2O)_x$

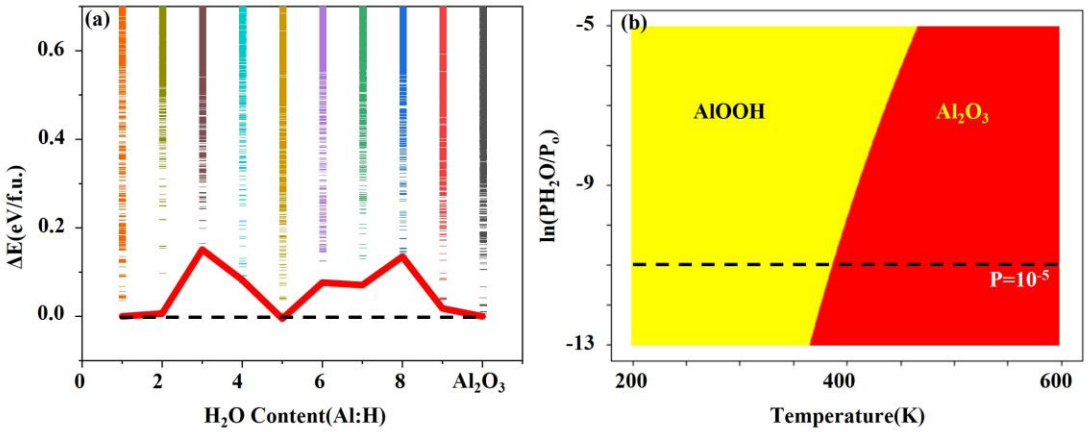


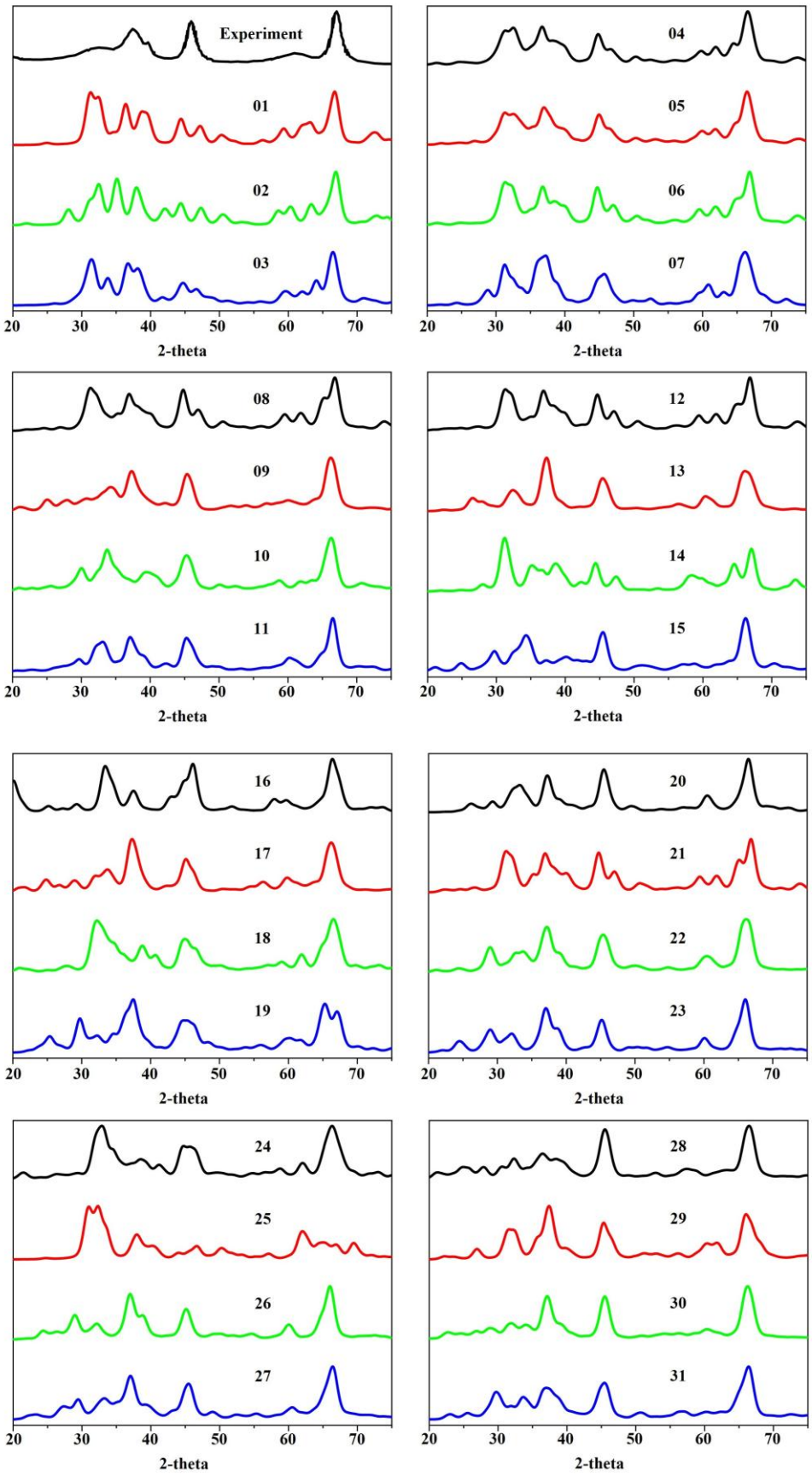
Figure S1| Thermodynamics of bulk $Al_2O_3(H_2O)_x$ at different water content. (a) Thermodynamics convex hull diagram for bulk $Al_2O_3(H_2O)_x$ with AlOOH and Al_2O_3 as the energy zero points. The relative energies of different structures are computed using Eq. S1. The color bars represent the energy of local minima. (b) The phase diagram of AlOOH (yellow) and Al_2O_3 (red).The black line denote the typical experimental water vapor pressure.

5. The summary of structural details for 62 structures.

Table S3 | The summary of structural details for 62 structures. Listed data includes space group, the DFT energy differences between the structure and θ phase (ΔE(θ=0), meV/atom), and atom/unit cell

Index	Space group	ΔE(θ=0) (meV/atom)	Atom /Unit cell	Index	Space group	ΔE(θ=0) (meV/atom)	Atom /Unit cell
1	12	0	20(10)	32	1	22.84	50
2	62	2.39	20	33	2	22.86	50
3	2	3.78	50	34	1	22.93	40
4	1	5.23	50	35	2	23.97	40
5	1	6.5	40	36	12	23.98	50(25)
6	2	7.14	50	37	8	24.05	80(40)
7	15	7.81	40(20)	38	1	24.41	25
8	2	8.95	40	39	2	25.09	40
9	2	11.26	40	40	2	25.19	50
10	2	11.27	50	41	2	25.24	40
11	2	12.99	50	42	2	25.48	40
12	2	13.91	50	43	2	26.01	50
13	2	14.07	40	44	2	26.19	40
14	63	14.19	80(40)	45	10	26.26	40
15	2	14.22	40	46	2	26.3	40
16	14	14.22	40	47	1	26.37	40
17	2	15.42	40	48	1	26.62	50
18	2	15.5	50	49	1	26.69	50
19	15	16.01	80(40)	50	2	27.22	50
20	2	16.82	40	51	1	27.24	50
21	2	16.91	40	52	13	27.26	40
22	1	18.08	40	53	1	27.54	40
23	8	19.32	80(40)	54	2	28.25	20
24	2	19.63	40	55	2	28.4	40
25	2	19.75	10	56	2	28.53	50
26	1	21.24	40	57	1	28.6	40
27	2	21.34	50	58	1	28.71	40
28	2	21.35	50	59	2	28.75	40
29	5	22.19	40(20)	60	2	29.25	40
30	1	22.45	50	61	11	29.31	40
31	2	22.78	50	62	1	29.59	50

6. The theoretical simulated XRD patterns for 62 Al_2O_3 minima. The original peaks are expanded in gaussian type with FWHM parameters (u=0.01, v=-0.001, w=2.000).



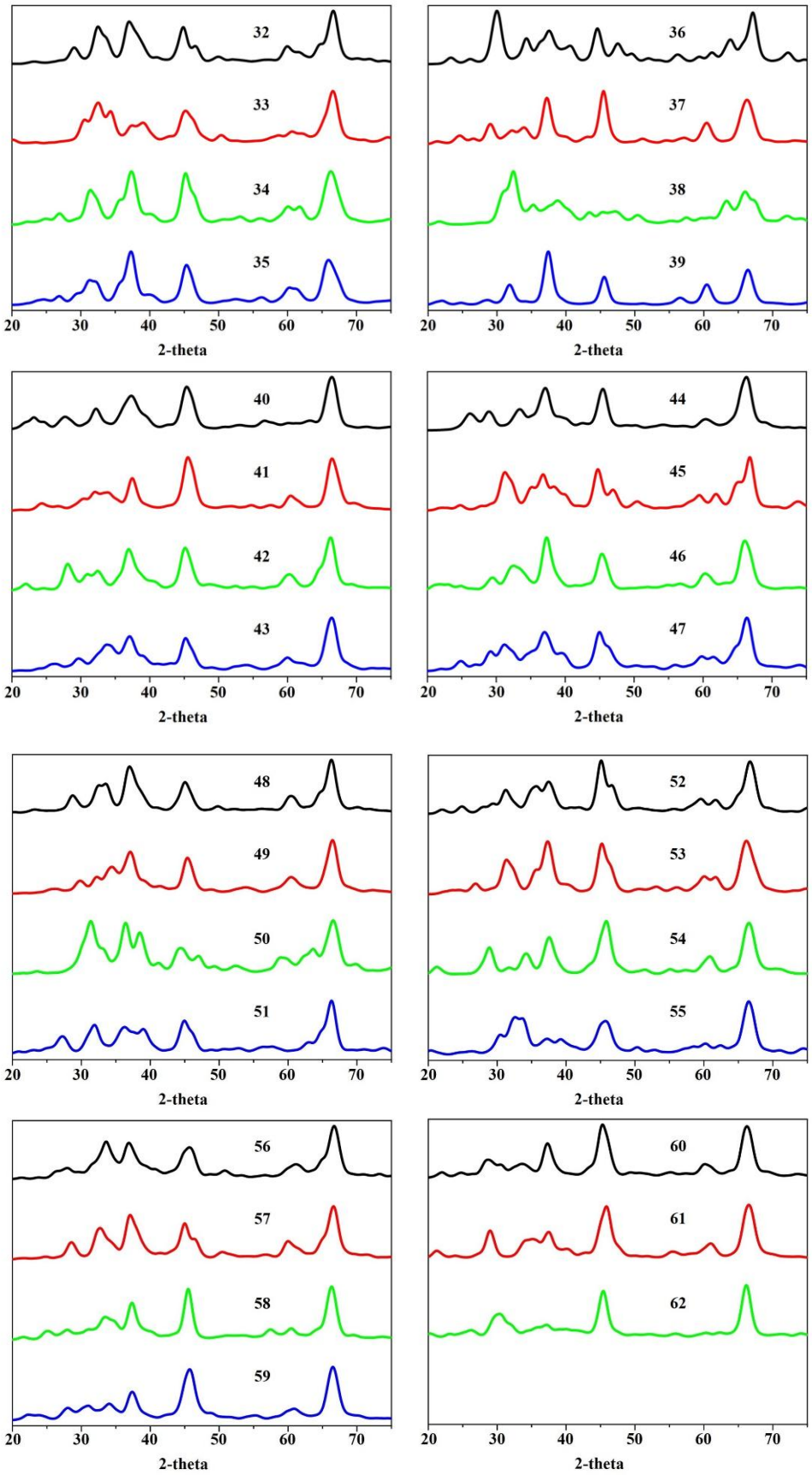
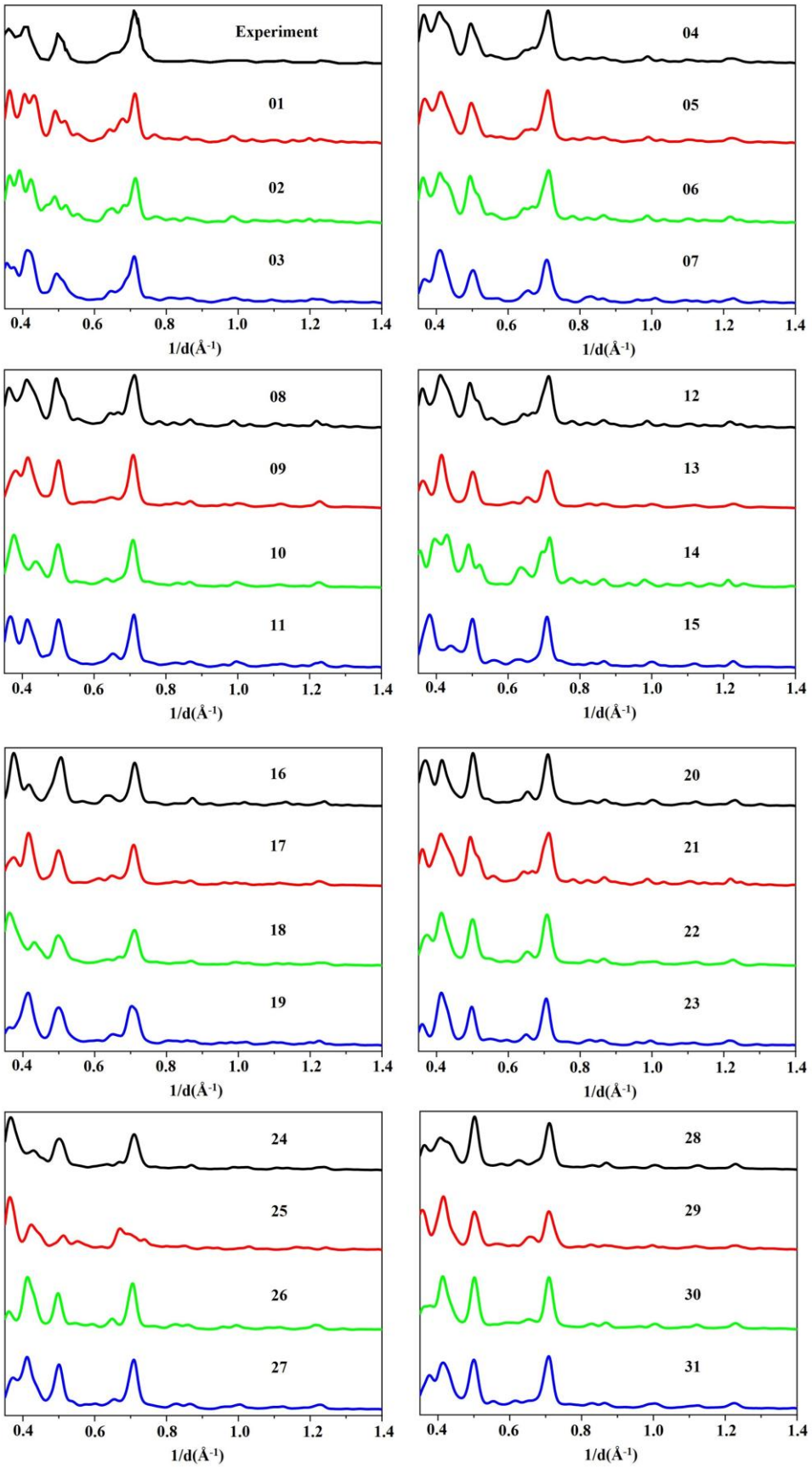


Figure S2| The theoretical XRD patterns for 62 structures. (The sequence corresponds to Table S3)

7. The theoretical simulated ED patterns for 62 Al_2O_3 minima. The original peaks are expanded in gaussian type with FWHM parameters (u=0.01, v=-0.001, w=2.000).



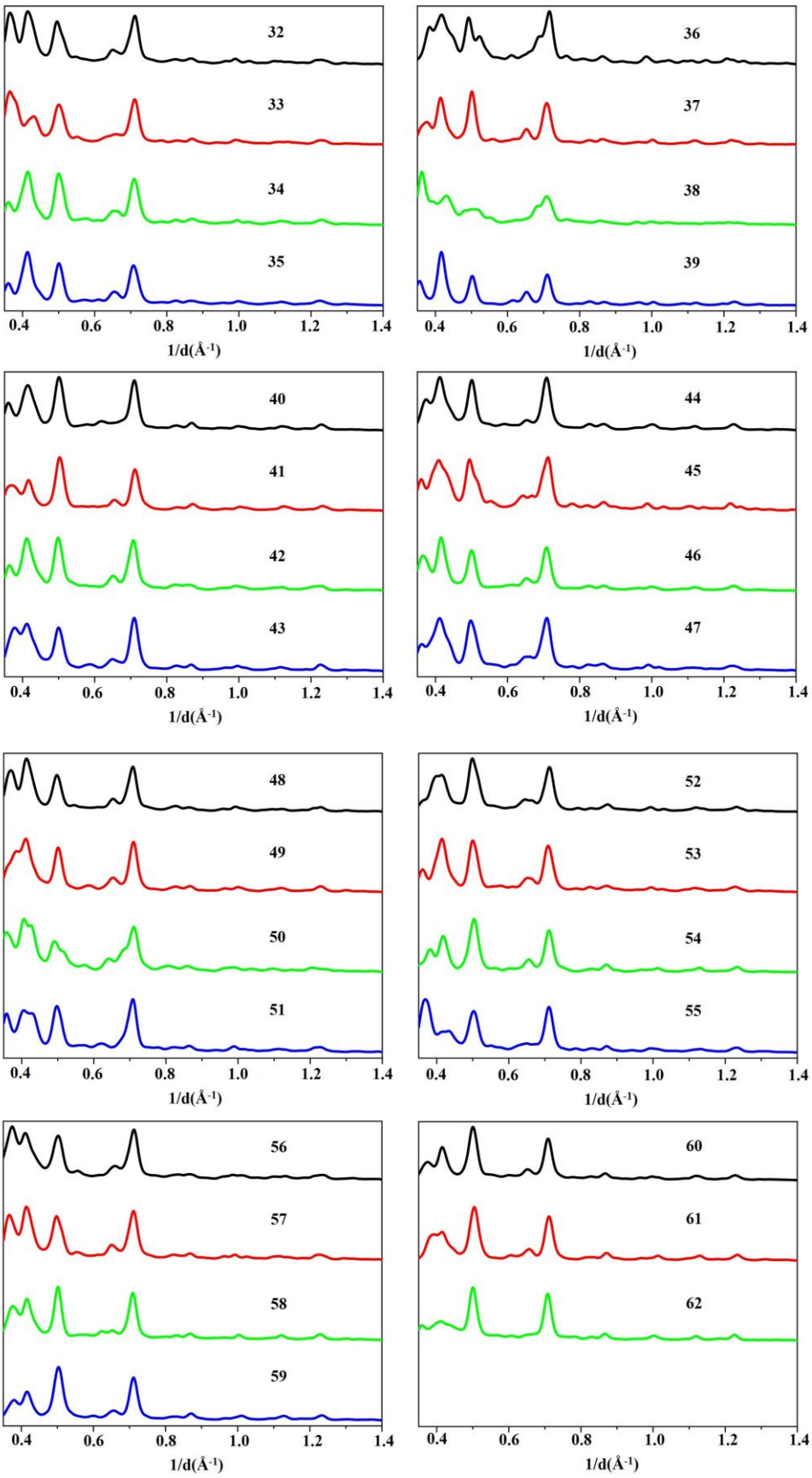


Figure S3| The theoretical XRD patterns for 62 structures. (The sequence corresponds to Table S3)

8.Dynamic behavior of Al cation in MD-trajectory under 1800K

We further performed 1000 molecular dynamics (MD) simulations using G-NN in parallel for **γ-AD** model (under 1800 K, 0.1 ns in the isothermal-isobaric (NPT) ensemble of 480-atom simulation cell), from which we obtained 16883 distinct local minima and calculated the P_{occ}. The results are shown in Fig. S5. We found that (i) all O anions only vibrate at the equilibrium lattice sites as the same under 1000K. (ii) The L1-L3 sites are the same as under 1000K. (iii) Under 1800K, we identified minor dynamic behaviors, which are the L4 sites (P_{occ} <1%) include 8 octahedral and 8 tetragonal sites. Overall, 32 octahedral and tetrahedral sites per cell can be occupied from our simulation.

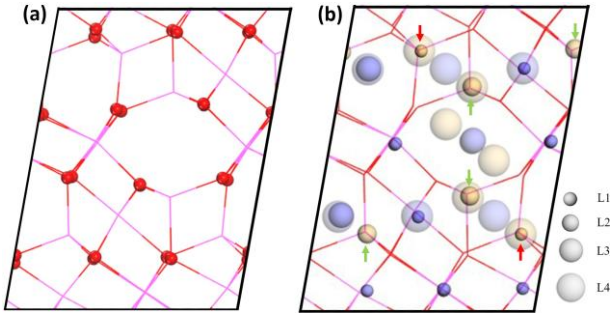


Figure S4| The dynamic behavior of γ-AD model statistically averaged from 1000 MD trajectories. (a) showing O anion behavior; (b) showing Al cation behavior. The overall occupancy probability P_{occ} of each site from high to low (L1 to L4) is represented by different sized transparent balls from small to large. Al atoms in L1-L3 are the same as that in Fig. 2 in the manuscript, where L4 sites only appear under 1800K.

9. XRD patterns of nanoparticle and the comparison between our model and previous models in the literature.

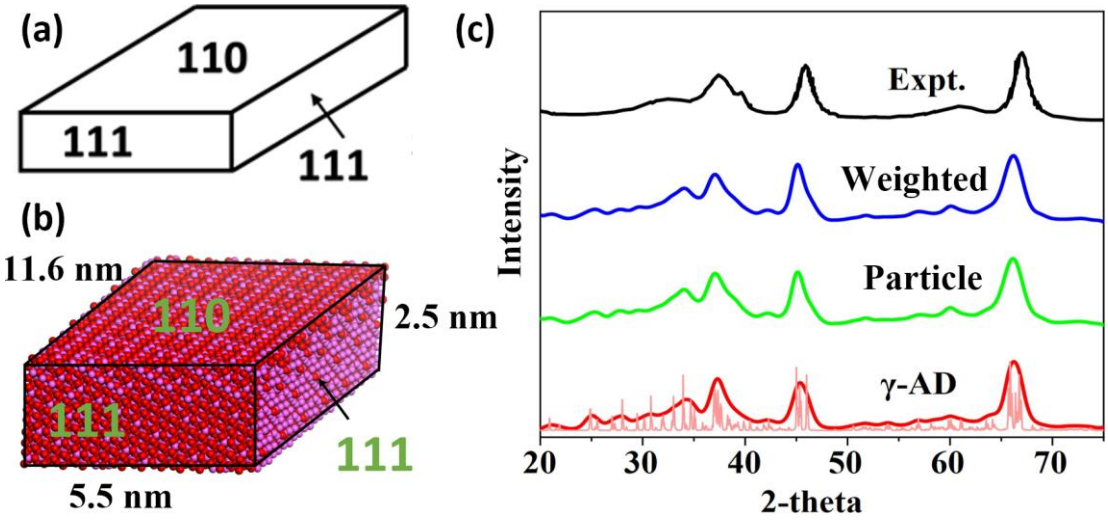


Figure S5| (a) The proposed rhombic motif γ-Al_2O_3 nanoparticle in literature. (b) The actual nanoparticle used to simulate the XRD patterns. (c) The simulated XRD patterns of nanoparticle (green), and a model that considers the dynamic behavior of Al (blue). The XRD pattern of experiment (black) and of 40-atom bulky γ-AD with gaussian expansion (red) are also provided for comparison.

10. The relevant peaks within the 28°-33° range in the XRD of γ-AD.

Table S4 The relevant peaks with 2θ falling within the 28°-33° range. The (hkl) value corresponds to the Miller indices. The intensity of each peak is relative to the strongest $(\bar{1}\bar{3}2)$ peak of **γ-AD** located at 2θ=45.05°.

(hkl) of γ-AD	2θ (°)	(hkl) of fcc sublattice	Relative intensity (%)
$(1\bar{2}1)$	28.04	$(7,7,1)$	46.58
$(\bar{2}10)$	28.41	$(1,\overline{15},\bar{1})$	2.71
(130)	29.47	$(11,3,\overline{11})$	10.76
$(2\bar{1}1)$	30.15	$(11,27,13)$	24.64
(221)	30.83	$(7,\bar{9},\overline{31})$	43.76
$(\bar{1}21)$	31.30	$(5,\overline{27},19)$	5.64
$(\bar{2}\bar{1}1)$	31.89	$(27,29,17)$	15.58
$(\bar{2}01)$	32.03	$(\overline{13},15,\bar{3})$	12.49
(002)	32.05	$(13,\bar{3},11)$	9.78

11. Determination of the pH_{PZC} of γ-AD

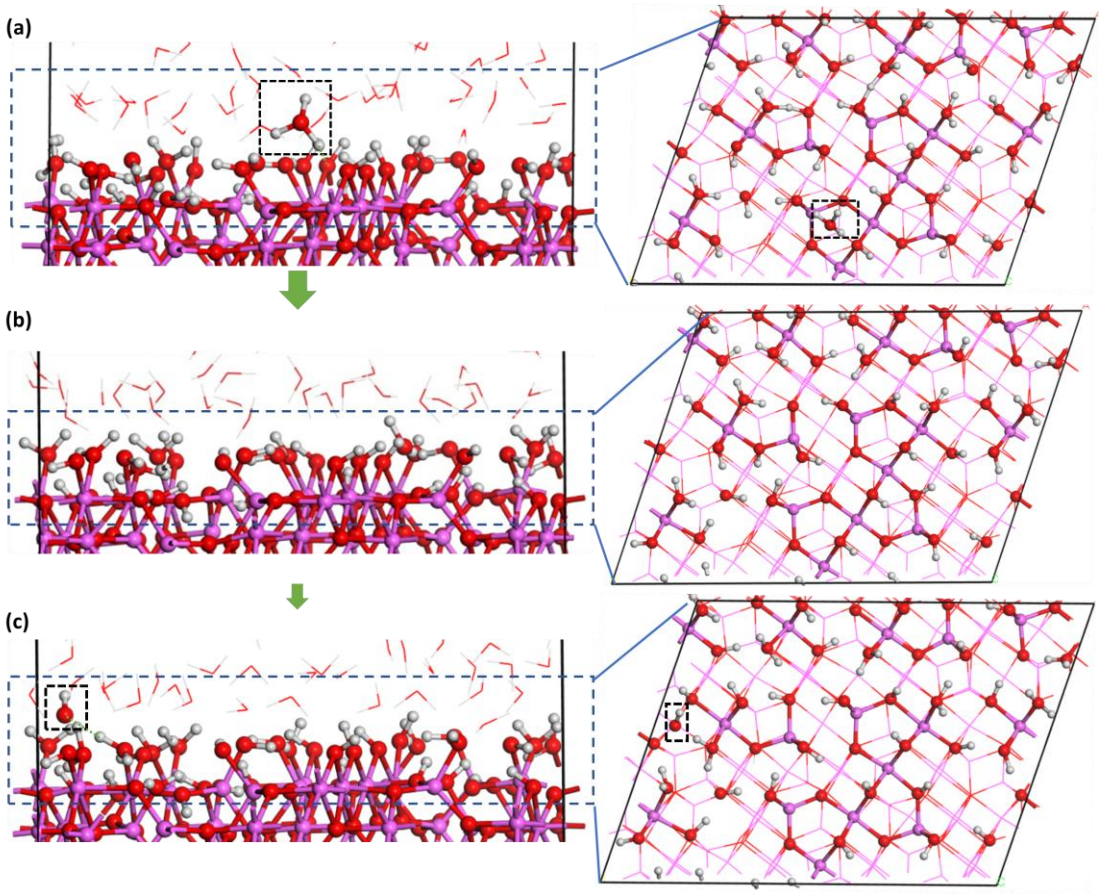


Figure S6| The surface OH distribution for (a) the deprotonation, (b) the equilibrium and (c) the protonation state. The H_3O^+ and the OH^- in solvent that generated from the deprotonation and protonation process of the surface are highlighted by black dashed box.

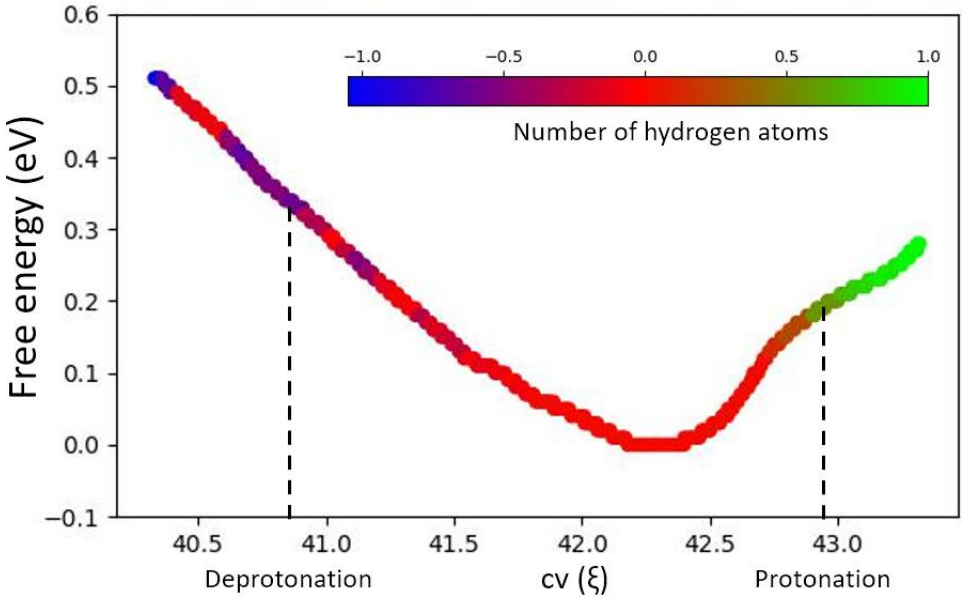


Figure S7| The free energy profile of one US trajectory for **γ-AD** (512) surface deprotonation and protonation, where the energy of the equilibrium state of the water-Al_2O_3 interface is set to be zero. X axis represents the value of collect variable (ξ) that defined by Eq. S7. The color of the profile curve indicates changes in the number of hydrogen atoms on the surface, which is averaged over the value of ξ along the MD trajectory. The number of hydrogen atom is set to be zero for equilibrium state. The deprotonated and protonated states are defined as the averaged hydrogen atom first pass -0.5 and 0.5, respectively. The ξ values for the deprotonated, equilibrium, and protonated states of the surface are 40.83, 42.24, and 42.92, respectively.

12.The arc file of γ-Ad. Please find the cif files for the above 62 Al_2O_3 configuration online.

!BIOSYM archive 2
PBC=ON

		Energy	0	0.0083	-289.620539		P-1	
!DATE								
PBC	7.43668054	9.35288890	5.68087596	90.13046397	79.80948764	79.31320604		
O	1.994600885	1.045411036	5.432860820	CORE	1 O	O	0.0000	1
O	6.700709564	8.012368422	4.769230581	CORE	2 O	O	0.0000	2
O	2.916378825	3.382399772	2.587959877	CORE	3 O	O	0.0000	3
O	1.805388771	5.645919255	1.310892188	CORE	4 O	O	0.0000	4
O	3.685484925	8.036643801	1.049467136	CORE	5 O	O	0.0000	5
O	5.097133629	3.374339514	0.567504233	CORE	6 O	O	0.0000	6
O	5.633347942	3.510733135	3.363126076	CORE	7 O	O	0.0000	7
O	4.038841628	7.982073631	3.851531636	CORE	8 O	O	0.0000	8
O	1.501675452	1.129638131	2.635383974	CORE	9 O	O	0.0000	9
O	6.334809268	8.165725816	1.957562851	CORE	10 O	O	0.0000	10
O	2.437134452	5.479446217	4.088771098	CORE	11 O	O	0.0000	11
O	3.427872349	3.326823701	5.186563735	CORE	12 O	O	0.0000	12
Al	5.146691603	0.087276688	0.437321376	CORE	13 Al	Al	0.0000	13
Al	7.754032256	1.590805646	1.433065305	CORE	14 Al	Al	0.0000	14
Al	6.016084705	4.441484868	1.796615051	CORE	15 Al	Al	0.0000	15
Al	6.664229118	2.690461602	4.526894256	CORE	16 Al	Al	0.0000	16
Al	3.461614455	2.757018375	0.950832054	CORE	17 Al	Al	0.0000	17
Al	2.423942399	2.021079477	3.937501130	CORE	18 Al	Al	0.0000	18
Al	4.570852006	8.971438756	2.354478399	CORE	19 Al	Al	0.0000	19
O	8.181551180	7.942423331	0.154718540	CORE	20 O	O	0.0000	20
O	3.475442501	0.975465945	0.818348778	CORE	21 O	O	0.0000	21

```
O     7.259773240  5.605434594  2.999619483 CORE  22 O   O   0.0000  22
O     8.370763294  3.341915112  4.276687172 CORE  23 O   O   0.0000  23
O     6.490667140  0.951190566  4.538112223 CORE  24 O   O   0.0000  24
O     5.079018436  5.613494853  5.020075127 CORE  25 O   O   0.0000  25
O     4.542804123  5.477101232  2.224453284 CORE  26 O   O   0.0000  26
O     6.137310437  1.005760735  1.736047723 CORE  27 O   O   0.0000  27
O     8.674476613  7.858196236  2.952195386 CORE  28 O   O   0.0000  28
O     3.841342797  0.822108551  3.630016509 CORE  29 O   O   0.0000  29
O     7.739017613  3.508388150  1.498808261 CORE  30 O   O   0.0000  30
O     6.748279716  5.661010666  0.401015625 CORE  31 O   O   0.0000  31
Al    5.029460462  8.900557679  5.150257984 CORE  32 Al  Al  0.0000  32
Al    2.422119809  7.397028720  4.154514055 CORE  33 Al  Al  0.0000  33
Al    4.160067360  4.546349499  3.790964308 CORE  34 Al  Al  0.0000  34
Al    3.511922947  6.297372764  1.060685104 CORE  35 Al  Al  0.0000  35
Al    6.714537609  6.230815992  4.636747306 CORE  36 Al  Al  0.0000  36
Al    7.752209666  6.966754890  1.650078230 CORE  37 Al  Al  0.0000  37
Al    5.605300059  0.016395611  3.233100960 CORE  38 Al  Al  0.0000  38
Al    0.867200462  4.595334159  0.000000000 CORE  39 Al  Al  0.0000  39
Al    1.369735761  4.493917183  2.793789680 CORE  40 Al  Al  0.0000  40
end
end
```


(1) Huang, S.-D.; Shang, C.; Kang, P.-L.; Zhang, X.-J.; Liu, Z.-P. LASP: Fast global potential energy surface exploration. *WIREs Computational Molecular Science* **2019**, *9*, e1415

(2) Shang, C.; Zhang, X.-J.; Liu, Z.-P. Stochastic surface walking method for crystal structure and phase transition pathway prediction. *Phys. Chem. Chem. Phys.* **2014**, *16*, 17845-17856

(3) Shang, C.; Liu, Z.-P. Stochastic Surface Walking Method for Structure Prediction and Pathway Searching. *J. Chem. Theory Comput.* **2013**, *9*, 1838-1845

(4) Huang, S.-D.; Shang, C.; Zhang, X.-J.; Liu, Z.-P. Material discovery by combining stochastic surface walking global optimization with a neural network. *Chem. Sci.* **2017**, *8*, 6327-6337

(5) 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

(6) Shang, C.; Liu, Z. P. AlOH dataset download. *http://www.lasphub.com/supportings/Trainfile_AlOH.tar.gz*.

(7) Kang, P.-L.; Yang, Z.-X.; Shang, C.; Liu, Z.-P. Global Neural Network Potential with Explicit Many-Body Functions for Improved Descriptions of Complex Potential Energy Surface. *J. Chem. Theory Comput.* **2023**, *19*, 7972-7981

(8) Ma, S.; Shang, C.; Liu, Z.-P. Heterogeneous catalysis from structure to activity via SSW-NN method. *Journal of Chemical Physics* **2019**, *151*, 050901

(9) Shang, C.; Liu, Z. P. AlOH potential download. *http://www.lasphub.com/supportings/AlOH.pot*.

(10) Iannuzzi, M.; Laio, A.; Parrinello, M. Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics. *Phys. Rev. Lett.* **2003**, *90*, 238302

(11) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087-1092

(12) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. Van der Waals Density Functional for General Geometries. *Phys. Rev. Lett.* **2004**, *92*, 246401

(13) Klimeš, J.; Bowler, D. R.; Michaelides, A. Chemical accuracy for the van der Waals density functional. *J. Phys.: Condens. Matter* **2010**, *22*, 022201

(14) van Gog, H. First-principles study of dehydration interfaces between diaspore and corundum, gibbsite and boehmite, and boehmite and γ-Al2O3: Energetic stability, interface charge effects, and dehydration defects. *Appl. Surf. Sci.* **2021**, *541*, 148501

(15) Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method.

Phys. Rev. B **1999**, *59*, 1758-1775
(16) Monkhorst, H. J.; Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **1976**, *13*, 5188-5192
(17) Chase, M. W., Jr. NIST-JANAF Themochemical Tables, Fourth Edition. *J. Phys. Chem. Ref. Data* **1998**, *Monograph 9*, 1-1951