

Supporting Information for

Managing Expectations and Imbalanced Training Data in Reactive Force Field Development: an Application to Water Adsorption on Alumina

Loïc Dumortier,^{†,‡} Céline Chizallet,[¶] Benoit Creton,[†] Theodorus de Bruin,[†] and
Toon Verstraelen^{*,‡}

[†]*IFP Energies nouvelles, 1 et 4 Avenue de Bois-Préau, 92852 Rueil-Malmaison, France*

[‡]*Center for Molecular Modeling (CMM), Ghent University, Technologiepark-Zwijnaarde 46,
B-9052, Zwijnaarde, Belgium*

[¶]*IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP3, 69360 Solaize, France*

E-mail: toon.verstraelen@ugent.be

Contents

| | |
|--|----|
| S1 Additional Display Items | 3 |
| S2 Comparison of Non-Equilibrium ReaxFF and DFT energies | 29 |
| References | 32 |

S1 Additional Display Items

Table S1: Structures in the training set and contributions to each category of data. Bonds are represented by pairs of chemical elements, angles by triplets. The symbol \cdots denotes a hydrogen bond. Internal coordinates with oxygen not bound to aluminum are discarded. For hydrated alumina surfaces and edges, and for boehmite surfaces, the internal coordinates without hydrogen are not considered.

| Structure | Chem. Form. | O-H | O \cdots H | Al-O | Al-Al | O-Al-O | Al-O-Al | Al-O-H | Al-O \cdots H | H-O \cdots H | H \cdots O \cdots H |
|------------------------|--|-----|--------------|------|-------|--------|---------|--------|-----------------|----------------|-------------------------|
| alpha_bulk | Al ₁₂ O ₁₈ | | | 72 | 84 | 180 | 108 | | | | |
| boehm_bulk | Al ₃₂ O ₆₄ H ₃₂ | 32 | 32 | 192 | 128 | 480 | 224 | 64 | 64 | 32 | |
| boehm_surf-001_00w | Al ₄₈ O ₉₆ H ₄₈ | 48 | 32 | 272 | 160 | 640 | 288 | 96 | 64 | 32 | |
| boehm_surf-001_08w | Al ₄₈ O ₁₁₂ H ₈₀ | 80 | 40 | | 160 | | | 160 | 80 | 40 | |
| boehm_surf-010_00w | Al ₇₂ O ₁₄₄ H ₇₂ | 72 | 48 | 432 | 288 | 1080 | 504 | 144 | 96 | 48 | |
| boehm_surf-100_00w | Al ₄₈ O ₉₆ H ₄₈ | 48 | 48 | 256 | 160 | 576 | 272 | 80 | 80 | 48 | |
| boehm_surf-100_12w | Al ₄₈ O ₁₂₀ H ₉₆ | 96 | 80 | | 160 | | | 128 | 112 | 96 | 24 |
| boehm_surf-100_16w | Al ₄₈ O ₁₂₈ H ₁₁₂ | 112 | 89 | | 160 | | | 144 | 137 | 98 | 24 |
| boehm_surf-101_00w | Al ₄₈ O ₉₆ H ₄₈ | 49 | 33 | 254 | 168 | 560 | 250 | 100 | 68 | 18 | 6 |
| boehm_surf-101_12w | Al ₄₈ O ₁₂₀ H ₉₆ | 96 | 66 | | 156 | | | 156 | 102 | 72 | 18 |
| gamma_bulk | Al ₁₆ O ₂₄ | | | 88 | 88 | 204 | 120 | | | | |
| gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | | | 481 | 424 | 1003 | 598 | | | | |
| gamma_edge-100-110_01w | Al ₉₆ O ₁₄₅ H ₂ | 2 | | | 425 | | | 3 | | | |
| gamma_edge-100-110_02w | Al ₉₆ O ₁₄₆ H ₄ | 4 | 1 | | 426 | | | 6 | 2 | | |
| gamma_edge-100-110_03w | Al ₉₆ O ₁₄₇ H ₆ | 6 | 2 | | 422 | | | 10 | 4 | | |
| gamma_edge-100-110_04w | Al ₉₆ O ₁₄₈ H ₈ | 8 | 3 | | 426 | | | 13 | 6 | | 1 |
| gamma_edge-100-110_05w | Al ₉₆ O ₁₄₉ H ₁₀ | 10 | 4 | | 424 | | | 17 | 8 | | 2 |
| gamma_edge-100-110_07w | Al ₉₆ O ₁₅₁ H ₁₄ | 14 | 7 | | 424 | | | 23 | 13 | 2 | 2 |
| gamma_surf-100_00w | Al ₆₄ O ₉₆ | | | 320 | 280 | 656 | 384 | | | | |
| gamma_surf-100_01w | Al ₆₄ O ₁₀₀ H ₈ | 8 | 4 | | 284 | | | 16 | 4 | 4 | |
| gamma_surf-100_02w | Al ₆₄ O ₁₀₄ H ₁₆ | 16 | 8 | | 282 | | | 24 | 8 | 8 | 4 |
| gamma_surf-100_03w | Al ₆₄ O ₁₀₈ H ₂₄ | 24 | 16 | | 280 | | | 40 | 16 | 16 | 8 |
| gamma_surf-100_04w | Al ₆₄ O ₁₁₂ H ₃₂ | 32 | 20 | | 280 | | | 52 | 28 | 20 | 8 |
| gamma_surf-110_00w | Al ₆₄ O ₉₆ | | | 304 | 240 | 600 | 368 | | | | |
| gamma_surf-110_01w | Al ₆₄ O ₁₀₀ H ₈ | 8 | 4 | | 232 | | | 16 | 8 | | |
| gamma_surf-110_02w | Al ₆₄ O ₁₀₄ H ₁₆ | 16 | 8 | | 240 | | | 24 | 16 | | |
| gamma_surf-110_03w | Al ₆₄ O ₁₀₈ H ₂₄ | 24 | 12 | | 240 | | | 44 | 16 | 8 | 4 |
| gamma_surf-110_04w | Al ₆₄ O ₁₁₂ H ₃₂ | 32 | 24 | | 240 | | | 52 | 44 | 12 | 8 |
| gamma_surf-110_05w | Al ₆₄ O ₁₁₆ H ₄₀ | 40 | 28 | | 240 | | | 60 | 52 | 12 | 12 |
| gamma_surf-110_06w | Al ₆₄ O ₁₂₀ H ₄₈ | 48 | 32 | | 236 | | | 88 | 32 | 32 | 12 |
| gamma_surf-111_00w | Al ₄₀ O ₆₀ | | | 212 | 190 | 480 | 286 | | | | |
| gamma_surf-111_04w | Al ₃₂ O ₅₆ H ₁₆ | 16 | 8 | | 144 | | | 30 | 14 | 6 | |
| gamma_surf-111_05w | Al ₃₂ O ₅₈ H ₂₀ | 20 | 10 | | 146 | | | 36 | 14 | 10 | 2 |
| gamma_surf-111_06w | Al ₃₂ O ₆₀ H ₂₄ | 24 | 12 | | 146 | | | 44 | 16 | 12 | 2 |
| monomer | AlO ₄ H ₅ | 5 | | 4 | | 6 | | 5 | | | |
| water | O H ₂ | | | | | | | | | | |
| total | | 990 | 671 | 2887 | 8383 | 6465 | 3402 | 1675 | 1104 | 626 | 137 |

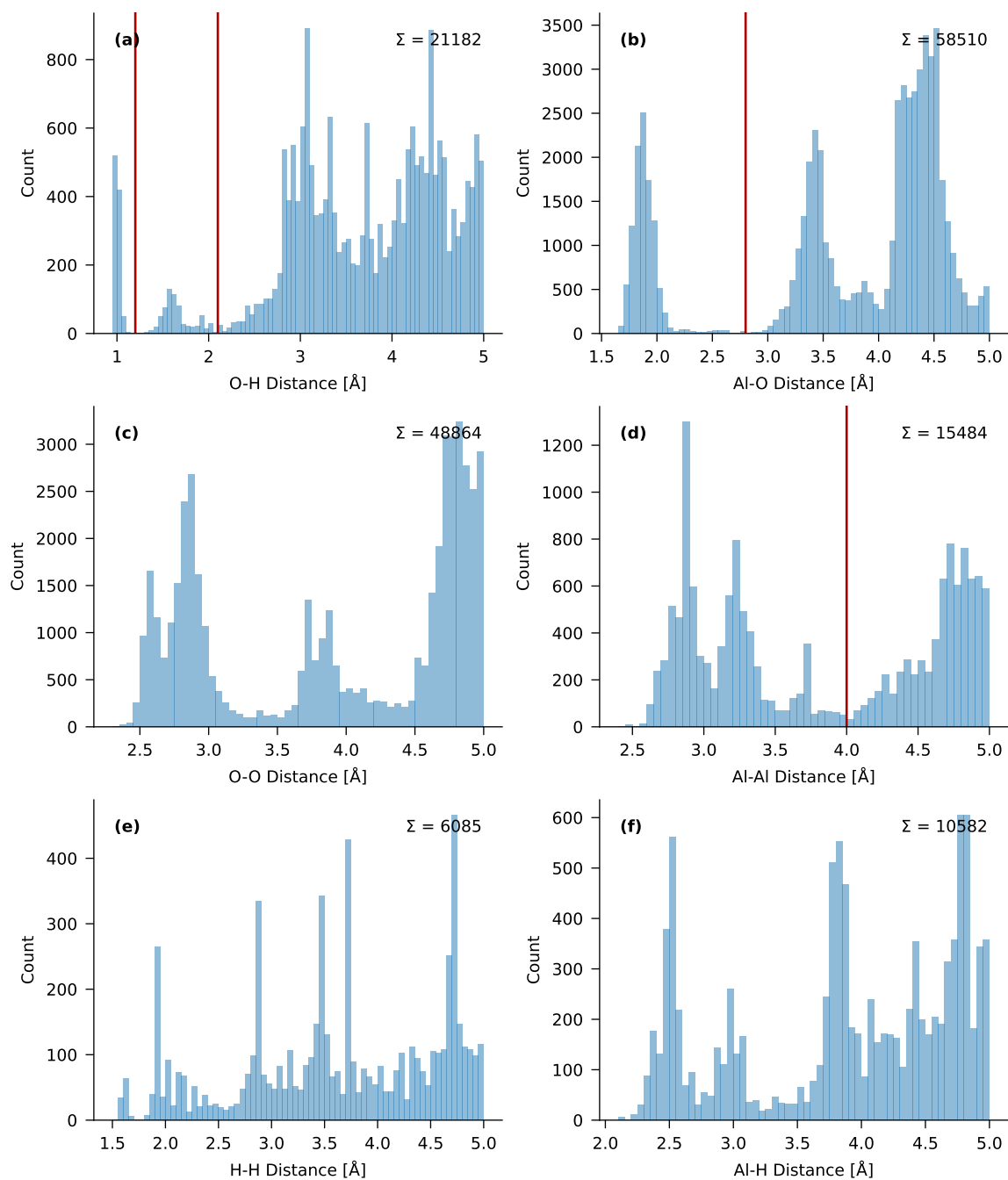


Figure S1: Histograms of all interatomic distances in the training set up to 5 Å, grouped per pair of chemical elements. Cutoffs for OH and AlO pairs depicted as vertical red lines: 1.2 Å for O–H bonds, 2.1 Å for hydrogen bonds and 2.8 Å for Al–O bonds. See main text for a more detailed description.

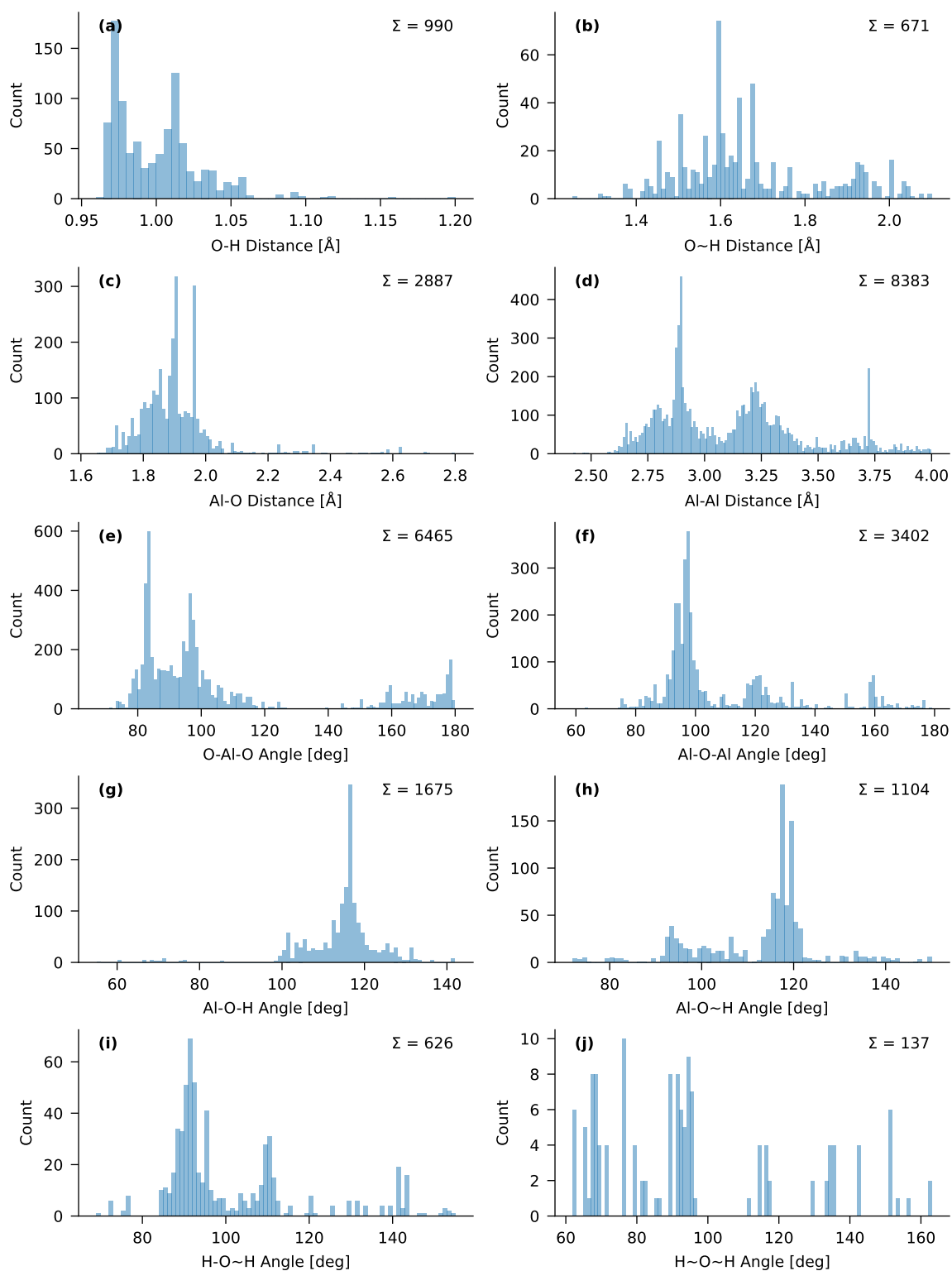


Figure S2: Histograms of internal coordinates in the training set. In the labels on the horizontal axis, a dash represents a regular bond and a tilde represents a hydrogen bond.

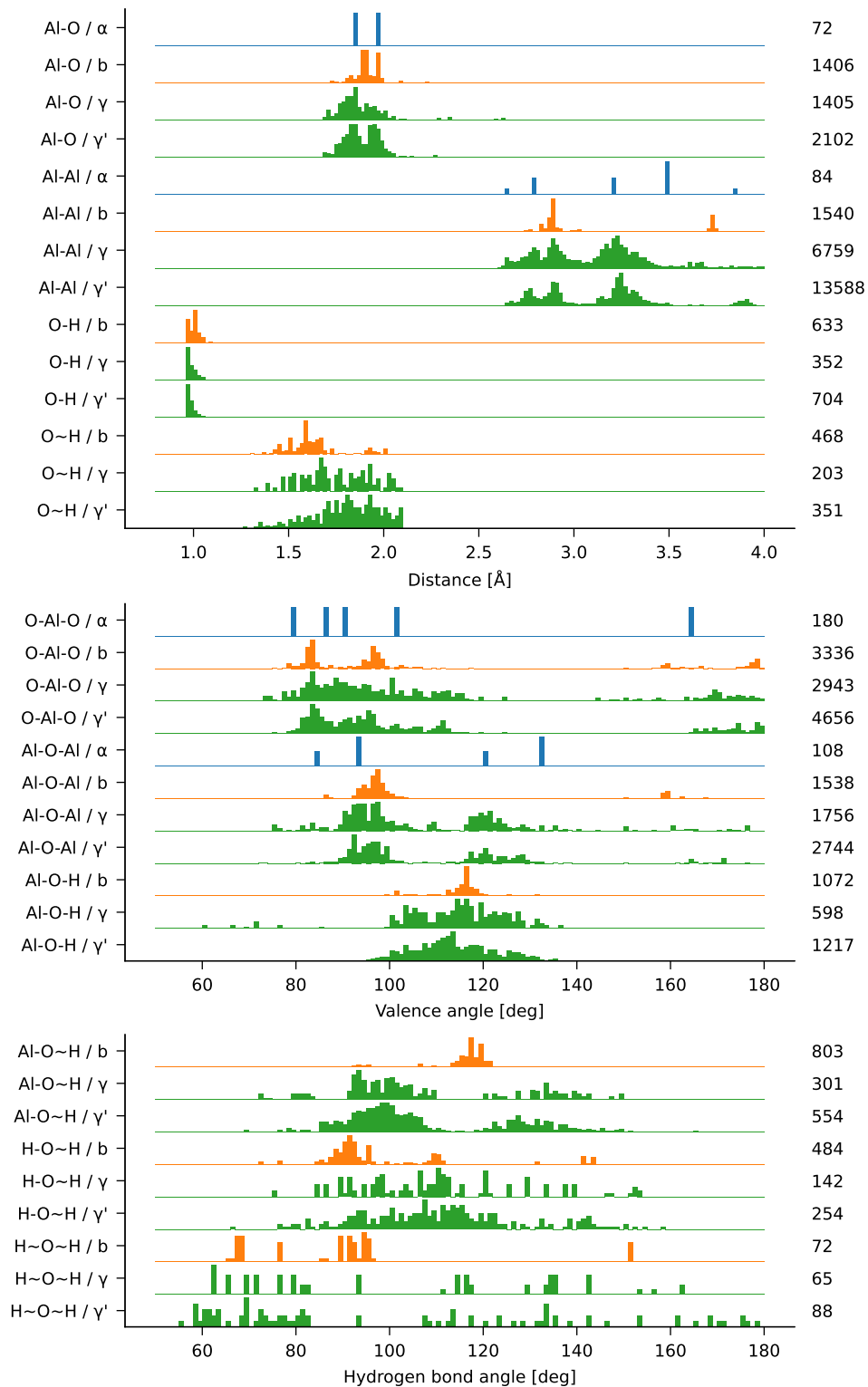


Figure S3: Histograms of internal coordinates, using the notation of Figure S2, per material. Histograms for the training set: $\alpha = \alpha\text{-Al}_2\text{O}_3$ (blue), b = boehmite (orange), $\gamma = \gamma\text{-Al}_2\text{O}_3$ (green). Histograms for the validation set: $\gamma' = \gamma\text{-Al}_2\text{O}_3$ (green). The number of internal coordinates within each class is shown to the right of the corresponding histogram.

Table S2: Overview of all chemical equations in the training set. Reactants are given negative coefficients. For each reaction, three reaction energies are in kcal mol⁻¹: the reference DFT result (R), the prediction with the Joshi force field (J) and the prediction with the new force field in this work (B). Water adsorption energies are normalized to the number of water molecules. All other reaction energies are normalized on the number of Al atoms. The categories are defined in the main text.

| Category | Coeff. | Structure | Chem. Form. | Reaction energy |
|------------|--------|------------------------|--|-----------------|
| BSH | -1/16× | boehm_surf-001_00w | Al ₄₈ O ₉₆ H ₄₈ | R -24.7 |
| | -1× | water | O H ₂ | J 28.0 |
| | 1/16× | boehm_surf-001_08w | Al ₄₈ O ₁₁₂ H ₈₀ | B -21.9 |
| BSH | -1/24× | boehm_surf-100_00w | Al ₄₈ O ₉₆ H ₄₈ | R -46.0 |
| | -1× | water | O H ₂ | J -1.7 |
| | 1/24× | boehm_surf-100_12w | Al ₄₈ O ₁₂₀ H ₉₆ | B -38.0 |
| BSH | -1/32× | boehm_surf-100_00w | Al ₄₈ O ₉₆ H ₄₈ | R -38.0 |
| | -1× | water | O H ₂ | J -4.2 |
| | 1/32× | boehm_surf-100_16w | Al ₄₈ O ₁₂₈ H ₁₁₂ | B -32.2 |
| BSH | -1/8× | boehm_surf-100_12w | Al ₄₈ O ₁₂₀ H ₉₆ | R -14.0 |
| | -1× | water | O H ₂ | J -11.9 |
| | 1/8× | boehm_surf-100_16w | Al ₄₈ O ₁₂₈ H ₁₁₂ | B -15.1 |
| BSH | -1/24× | boehm_surf-101_00w | Al ₄₈ O ₉₆ H ₄₈ | R -41.0 |
| | -1× | water | O H ₂ | J -4.3 |
| | 1/24× | boehm_surf-101_12w | Al ₄₈ O ₁₂₀ H ₉₆ | B -39.3 |
| GEH | -1× | gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | R -104.1 |
| | -1× | water | O H ₂ | J -101.6 |
| | 1× | gamma_edge-100-110_01w | Al ₉₆ O ₁₄₅ H ₂ | B -103.0 |
| GEH | -1/2× | gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | R -81.2 |
| | -1× | water | O H ₂ | J -86.6 |
| | 1/2× | gamma_edge-100-110_02w | Al ₉₆ O ₁₄₆ H ₄ | B -80.3 |
| GEH | -1× | gamma_edge-100-110_01w | Al ₉₆ O ₁₄₅ H ₂ | R -58.3 |
| | -1× | water | O H ₂ | J -71.5 |
| | 1× | gamma_edge-100-110_02w | Al ₉₆ O ₁₄₆ H ₄ | B -57.5 |
| GEH | -1/3× | gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | R -72.1 |
| | -1× | water | O H ₂ | J -18.0 |
| | 1/3× | gamma_edge-100-110_03w | Al ₉₆ O ₁₄₇ H ₆ | B -71.2 |
| GEH | -1/2× | gamma_edge-100-110_01w | Al ₉₆ O ₁₄₅ H ₂ | R -56.1 |
| | -1× | water | O H ₂ | J 23.8 |
| | 1/2× | gamma_edge-100-110_03w | Al ₉₆ O ₁₄₇ H ₆ | B -55.3 |
| GEH | -1× | gamma_edge-100-110_02w | Al ₉₆ O ₁₄₆ H ₄ | R -53.9 |
| | -1× | water | O H ₂ | J 119.1 |
| | 1× | gamma_edge-100-110_03w | Al ₉₆ O ₁₄₇ H ₆ | B -53.0 |
| GEH | -1/4× | gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | R -67.5 |
| | -1× | water | O H ₂ | J -71.0 |

| | | | | | |
|------------|-------|------------------------|---|---|--------|
| | 1/4× | gamma_edge-100-110_04w | Al ₉₆ O ₁₄₈ H ₈ | B | -67.0 |
| GEH | -1/3× | gamma_edge-100-110_01w | Al ₉₆ O ₁₄₅ H ₂ | R | -55.3 |
| | -1× | water | O H ₂ | J | -60.8 |
| | 1/3× | gamma_edge-100-110_04w | Al ₉₆ O ₁₄₈ H ₈ | B | -55.0 |
| GEH | -1/2× | gamma_edge-100-110_02w | Al ₉₆ O ₁₄₆ H ₄ | R | -53.9 |
| | -1× | water | O H ₂ | J | -55.4 |
| | 1/2× | gamma_edge-100-110_04w | Al ₉₆ O ₁₄₈ H ₈ | B | -53.7 |
| GEH | -1× | gamma_edge-100-110_03w | Al ₉₆ O ₁₄₇ H ₆ | R | -53.8 |
| | -1× | water | O H ₂ | J | -230.0 |
| | 1× | gamma_edge-100-110_04w | Al ₉₆ O ₁₄₈ H ₈ | B | -54.5 |
| GEH | -1/5× | gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | R | -63.3 |
| | -1× | water | O H ₂ | J | -22.9 |
| | 1/5× | gamma_edge-100-110_05w | Al ₉₆ O ₁₄₉ H ₁₀ | B | -63.8 |
| GEH | -1/4× | gamma_edge-100-110_01w | Al ₉₆ O ₁₄₅ H ₂ | R | -53.1 |
| | -1× | water | O H ₂ | J | -3.2 |
| | 1/4× | gamma_edge-100-110_05w | Al ₉₆ O ₁₄₉ H ₁₀ | B | -54.0 |
| GEH | -1/3× | gamma_edge-100-110_02w | Al ₉₆ O ₁₄₆ H ₄ | R | -51.4 |
| | -1× | water | O H ₂ | J | 19.5 |
| | 1/3× | gamma_edge-100-110_05w | Al ₉₆ O ₁₄₉ H ₁₀ | B | -52.8 |
| GEH | -1/2× | gamma_edge-100-110_03w | Al ₉₆ O ₁₄₇ H ₆ | R | -50.2 |
| | -1× | water | O H ₂ | J | -30.3 |
| | 1/2× | gamma_edge-100-110_05w | Al ₉₆ O ₁₄₉ H ₁₀ | B | -52.7 |
| GEH | -1× | gamma_edge-100-110_04w | Al ₉₆ O ₁₄₈ H ₈ | R | -46.6 |
| | -1× | water | O H ₂ | J | 169.4 |
| | 1× | gamma_edge-100-110_05w | Al ₉₆ O ₁₄₉ H ₁₀ | B | -51.0 |
| GEH | -1/7× | gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | R | -55.8 |
| | -1× | water | O H ₂ | J | -12.6 |
| | 1/7× | gamma_edge-100-110_07w | Al ₉₆ O ₁₅₁ H ₁₄ | B | -57.1 |
| GEH | -1/6× | gamma_edge-100-110_01w | Al ₉₆ O ₁₄₅ H ₂ | R | -47.8 |
| | -1× | water | O H ₂ | J | 2.2 |
| | 1/6× | gamma_edge-100-110_07w | Al ₉₆ O ₁₅₁ H ₁₄ | B | -49.5 |
| GEH | -1/5× | gamma_edge-100-110_02w | Al ₉₆ O ₁₄₆ H ₄ | R | -45.7 |
| | -1× | water | O H ₂ | J | 16.9 |
| | 1/5× | gamma_edge-100-110_07w | Al ₉₆ O ₁₅₁ H ₁₄ | B | -47.9 |
| GEH | -1/4× | gamma_edge-100-110_03w | Al ₉₆ O ₁₄₇ H ₆ | R | -43.6 |
| | -1× | water | O H ₂ | J | -8.6 |
| | 1/4× | gamma_edge-100-110_07w | Al ₉₆ O ₁₅₁ H ₁₄ | B | -46.6 |
| GEH | -1/3× | gamma_edge-100-110_04w | Al ₉₆ O ₁₄₈ H ₈ | R | -40.2 |
| | -1× | water | O H ₂ | J | 65.1 |
| | 1/3× | gamma_edge-100-110_07w | Al ₉₆ O ₁₅₁ H ₁₄ | B | -43.9 |
| GEH | -1/2× | gamma_edge-100-110_05w | Al ₉₆ O ₁₄₉ H ₁₀ | R | -37.0 |
| | -1× | water | O H ₂ | J | 13.0 |
| | 1/2× | gamma_edge-100-110_07w | Al ₉₆ O ₁₅₁ H ₁₄ | B | -40.4 |
| GSH | -1/4× | gamma_surf-100_00w | Al ₆₄ O ₉₆ | R | -24.0 |

| | | | | | |
|------------|--------|--------------------|---|---|--------|
| | -1× | water | O H ₂ | J | -2.0 |
| | 1/4× | gamma_surf-100_01w | Al ₆₄ O ₁₀₀ H ₈ | B | -23.8 |
| GSH | -1/8× | gamma_surf-100_00w | Al ₆₄ O ₉₆ | R | -25.0 |
| | -1× | water | O H ₂ | J | -10.7 |
| | 1/8× | gamma_surf-100_02w | Al ₆₄ O ₁₀₄ H ₁₆ | B | -26.0 |
| GSH | -1/4× | gamma_surf-100_01w | Al ₆₄ O ₁₀₀ H ₈ | R | -26.0 |
| | -1× | water | O H ₂ | J | -19.5 |
| | 1/4× | gamma_surf-100_02w | Al ₆₄ O ₁₀₄ H ₁₆ | B | -28.2 |
| GSH | -1/12× | gamma_surf-100_00w | Al ₆₄ O ₉₆ | R | -24.3 |
| | -1× | water | O H ₂ | J | -6.8 |
| | 1/12× | gamma_surf-100_03w | Al ₆₄ O ₁₀₈ H ₂₄ | B | -26.8 |
| GSH | -1/8× | gamma_surf-100_01w | Al ₆₄ O ₁₀₀ H ₈ | R | -24.4 |
| | -1× | water | O H ₂ | J | -9.2 |
| | 1/8× | gamma_surf-100_03w | Al ₆₄ O ₁₀₈ H ₂₄ | B | -28.3 |
| GSH | -1/4× | gamma_surf-100_02w | Al ₆₄ O ₁₀₄ H ₁₆ | R | -22.8 |
| | -1× | water | O H ₂ | J | 1.1 |
| | 1/4× | gamma_surf-100_03w | Al ₆₄ O ₁₀₈ H ₂₄ | B | -28.3 |
| GSH | -1/16× | gamma_surf-100_00w | Al ₆₄ O ₉₆ | R | -21.7 |
| | -1× | water | O H ₂ | J | -3.6 |
| | 1/16× | gamma_surf-100_04w | Al ₆₄ O ₁₁₂ H ₃₂ | B | -22.4 |
| GSH | -1/12× | gamma_surf-100_01w | Al ₆₄ O ₁₀₀ H ₈ | R | -21.0 |
| | -1× | water | O H ₂ | J | -4.1 |
| | 1/12× | gamma_surf-100_04w | Al ₆₄ O ₁₁₂ H ₃₂ | B | -22.0 |
| GSH | -1/8× | gamma_surf-100_02w | Al ₆₄ O ₁₀₄ H ₁₆ | R | -18.5 |
| | -1× | water | O H ₂ | J | 3.5 |
| | 1/8× | gamma_surf-100_04w | Al ₆₄ O ₁₁₂ H ₃₂ | B | -18.8 |
| GSH | -1/4× | gamma_surf-100_03w | Al ₆₄ O ₁₀₈ H ₂₄ | R | -14.1 |
| | -1× | water | O H ₂ | J | 6.0 |
| | 1/4× | gamma_surf-100_04w | Al ₆₄ O ₁₁₂ H ₃₂ | B | -9.3 |
| GSH | -1/4× | gamma_surf-110_00w | Al ₆₄ O ₉₆ | R | -92.4 |
| | -1× | water | O H ₂ | J | 50.4 |
| | 1/4× | gamma_surf-110_01w | Al ₆₄ O ₁₀₀ H ₈ | B | -108.6 |
| GSH | -1/8× | gamma_surf-110_00w | Al ₆₄ O ₉₆ | R | -50.2 |
| | -1× | water | O H ₂ | J | 7.3 |
| | 1/8× | gamma_surf-110_02w | Al ₆₄ O ₁₀₄ H ₁₆ | B | -61.8 |
| GSH | -1/4× | gamma_surf-110_01w | Al ₆₄ O ₁₀₀ H ₈ | R | -8.0 |
| | -1× | water | O H ₂ | J | -35.8 |
| | 1/4× | gamma_surf-110_02w | Al ₆₄ O ₁₀₄ H ₁₆ | B | -15.0 |
| GSH | -1/12× | gamma_surf-110_00w | Al ₆₄ O ₉₆ | R | -42.5 |
| | -1× | water | O H ₂ | J | 42.3 |
| | 1/12× | gamma_surf-110_03w | Al ₆₄ O ₁₀₈ H ₂₄ | B | -50.1 |
| GSH | -1/8× | gamma_surf-110_01w | Al ₆₄ O ₁₀₀ H ₈ | R | -17.6 |
| | -1× | water | O H ₂ | J | 38.2 |
| | 1/8× | gamma_surf-110_03w | Al ₆₄ O ₁₀₈ H ₂₄ | B | -20.8 |

| | | | | | |
|-----|--------|--------------------|---|---|-------|
| GSH | -1/4× | gamma_surf-110_02w | Al ₆₄ O ₁₀₄ H ₁₆ | R | -27.2 |
| | -1× | water | O H ₂ | J | 112.3 |
| | 1/4× | gamma_surf-110_03w | Al ₆₄ O ₁₀₈ H ₂₄ | B | -26.7 |
| GSH | -1/16× | gamma_surf-110_00w | Al ₆₄ O ₉₆ | R | -39.6 |
| | -1× | water | O H ₂ | J | 28.9 |
| | 1/16× | gamma_surf-110_04w | Al ₆₄ O ₁₁₂ H ₃₂ | B | -44.9 |
| GSH | -1/12× | gamma_surf-110_01w | Al ₆₄ O ₁₀₀ H ₈ | R | -22.1 |
| | -1× | water | O H ₂ | J | 21.7 |
| | 1/12× | gamma_surf-110_04w | Al ₆₄ O ₁₁₂ H ₃₂ | B | -23.7 |
| GSH | -1/8× | gamma_surf-110_02w | Al ₆₄ O ₁₀₄ H ₁₆ | R | -29.1 |
| | -1× | water | O H ₂ | J | 50.5 |
| | 1/8× | gamma_surf-110_04w | Al ₆₄ O ₁₁₂ H ₃₂ | B | -28.1 |
| GSH | -1/4× | gamma_surf-110_03w | Al ₆₄ O ₁₀₈ H ₂₄ | R | -31.0 |
| | -1× | water | O H ₂ | J | -11.3 |
| | 1/4× | gamma_surf-110_04w | Al ₆₄ O ₁₁₂ H ₃₂ | B | -29.5 |
| GSH | -1/20× | gamma_surf-110_00w | Al ₆₄ O ₉₆ | R | -35.2 |
| | -1× | water | O H ₂ | J | 16.0 |
| | 1/20× | gamma_surf-110_05w | Al ₆₄ O ₁₁₆ H ₄₀ | B | -39.4 |
| GSH | -1/16× | gamma_surf-110_01w | Al ₆₄ O ₁₀₀ H ₈ | R | -20.9 |
| | -1× | water | O H ₂ | J | 7.4 |
| | 1/16× | gamma_surf-110_05w | Al ₆₄ O ₁₁₆ H ₄₀ | B | -22.1 |
| GSH | -1/12× | gamma_surf-110_02w | Al ₆₄ O ₁₀₄ H ₁₆ | R | -25.1 |
| | -1× | water | O H ₂ | J | 21.9 |
| | 1/12× | gamma_surf-110_05w | Al ₆₄ O ₁₁₆ H ₄₀ | B | -24.5 |
| GSH | -1/8× | gamma_surf-110_03w | Al ₆₄ O ₁₀₈ H ₂₄ | R | -24.1 |
| | -1× | water | O H ₂ | J | -23.4 |
| | 1/8× | gamma_surf-110_05w | Al ₆₄ O ₁₁₆ H ₄₀ | B | -23.4 |
| GSH | -1/4× | gamma_surf-110_04w | Al ₆₄ O ₁₁₂ H ₃₂ | R | -17.3 |
| | -1× | water | O H ₂ | J | -35.4 |
| | 1/4× | gamma_surf-110_05w | Al ₆₄ O ₁₁₆ H ₄₀ | B | -17.2 |
| GSH | -1/24× | gamma_surf-110_00w | Al ₆₄ O ₉₆ | R | -33.5 |
| | -1× | water | O H ₂ | J | 21.5 |
| | 1/24× | gamma_surf-110_06w | Al ₆₄ O ₁₂₀ H ₄₈ | B | -37.5 |
| GSH | -1/20× | gamma_surf-110_01w | Al ₆₄ O ₁₀₀ H ₈ | R | -21.7 |
| | -1× | water | O H ₂ | J | 15.8 |
| | 1/20× | gamma_surf-110_06w | Al ₆₄ O ₁₂₀ H ₄₈ | B | -23.3 |
| GSH | -1/16× | gamma_surf-110_02w | Al ₆₄ O ₁₀₄ H ₁₆ | R | -25.1 |
| | -1× | water | O H ₂ | J | 28.7 |
| | 1/16× | gamma_surf-110_06w | Al ₆₄ O ₁₂₀ H ₄₈ | B | -25.3 |
| GSH | -1/12× | gamma_surf-110_03w | Al ₆₄ O ₁₀₈ H ₂₄ | R | -24.4 |
| | -1× | water | O H ₂ | J | 0.8 |
| | 1/12× | gamma_surf-110_06w | Al ₆₄ O ₁₂₀ H ₄₈ | B | -24.9 |
| GSH | -1/8× | gamma_surf-110_04w | Al ₆₄ O ₁₁₂ H ₃₂ | R | -21.1 |
| | -1× | water | O H ₂ | J | 6.8 |

| | | | | | |
|------------|--------|------------------------|---|---|-------|
| | 1/8× | gamma_surf-110_06w | Al ₆₄ O ₁₂₀ H ₄₈ | B | -22.6 |
| GSH | -1/4× | gamma_surf-110_05w | Al ₆₄ O ₁₁₆ H ₄₀ | R | -24.9 |
| | -1× | water | O H ₂ | J | 49.1 |
| | 1/4× | gamma_surf-110_06w | Al ₆₄ O ₁₂₀ H ₄₈ | B | -27.9 |
| GSH | -1/10× | gamma_surf-111_00w | Al ₄₀ O ₆₀ | R | -42.1 |
| | -1× | water | O H ₂ | J | 7.2 |
| | 1/8× | gamma_surf-111_04w | Al ₃₂ O ₅₆ H ₁₆ | B | -31.1 |
| GSH | -2/25× | gamma_surf-111_00w | Al ₄₀ O ₆₀ | R | -45.8 |
| | -1× | water | O H ₂ | J | 5.6 |
| | 1/10× | gamma_surf-111_05w | Al ₃₂ O ₅₈ H ₂₀ | B | -32.8 |
| GSH | -1/2× | gamma_surf-111_04w | Al ₃₂ O ₅₆ H ₁₆ | R | -60.5 |
| | -1× | water | O H ₂ | J | -0.7 |
| | 1/2× | gamma_surf-111_05w | Al ₃₂ O ₅₈ H ₂₀ | B | -39.8 |
| GSH | -1/15× | gamma_surf-111_00w | Al ₄₀ O ₆₀ | R | -43.1 |
| | -1× | water | O H ₂ | J | 13.5 |
| | 1/12× | gamma_surf-111_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -32.5 |
| GSH | -1/4× | gamma_surf-111_04w | Al ₃₂ O ₅₆ H ₁₆ | R | -45.0 |
| | -1× | water | O H ₂ | J | 26.2 |
| | 1/4× | gamma_surf-111_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -35.3 |
| GSH | -1/2× | gamma_surf-111_05w | Al ₃₂ O ₅₈ H ₂₀ | R | -29.4 |
| | -1× | water | O H ₂ | J | 53.2 |
| | 1/2× | gamma_surf-111_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -30.9 |
| SUR | -1/32× | boehm_bulk | Al ₃₂ O ₆₄ H ₃₂ | R | 11.2 |
| | 1/48× | boehm_surf-001_00w | Al ₄₈ O ₉₆ H ₄₈ | J | 5.5 |
| | | | | B | 9.0 |
| SUR | -1/32× | boehm_bulk | Al ₃₂ O ₆₄ H ₃₂ | R | 2.9 |
| | 1/72× | boehm_surf-010_00w | Al ₇₂ O ₁₄₄ H ₇₂ | J | 2.1 |
| | | | | B | 3.1 |
| SUR | -1/32× | boehm_bulk | Al ₃₂ O ₆₄ H ₃₂ | R | 23.0 |
| | 1/48× | boehm_surf-100_00w | Al ₄₈ O ₉₆ H ₄₈ | J | -0.4 |
| | | | | B | 19.1 |
| SUR | -1/32× | boehm_bulk | Al ₃₂ O ₆₄ H ₃₂ | R | 18.6 |
| | 1/48× | boehm_surf-101_00w | Al ₄₈ O ₉₆ H ₄₈ | J | 7.5 |
| | | | | B | 16.4 |
| SUR | -1/16× | gamma_bulk | Al ₁₆ O ₂₄ | R | 13.5 |
| | 1/96× | gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | J | 7.0 |
| | | | | B | 14.0 |
| SUR | -1/16× | gamma_bulk | Al ₁₆ O ₂₄ | R | 10.4 |
| | 1/64× | gamma_surf-100_00w | Al ₆₄ O ₉₆ | J | 6.6 |
| | | | | B | 13.0 |
| SUR | -1/16× | gamma_bulk | Al ₁₆ O ₂₄ | R | 21.2 |
| | 1/64× | gamma_surf-110_00w | Al ₆₄ O ₉₆ | J | 7.4 |
| | | | | B | 26.6 |
| SUR | -1/16× | gamma_bulk | Al ₁₆ O ₂₄ | R | 13.5 |

| | | | | | |
|------------|-------|------------------------|---|---|-------|
| | 1/40× | gamma_surf-111_00w | Al ₄₀ O ₆₀ | J | 5.6 |
| | | | | B | 11.1 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -27.8 |
| | 1/12× | alpha_bulk | Al ₁₂ O ₁₈ | J | -44.9 |
| | 5/2× | water | O H ₂ | B | -31.9 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -38.8 |
| | 1/32× | boehm_bulk | Al ₃₂ O ₆₄ H ₃₂ | J | -38.6 |
| | 2× | water | O H ₂ | B | -36.5 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -27.6 |
| | 1/48× | boehm_surf-001_00w | Al ₄₈ O ₉₆ H ₄₈ | J | -33.1 |
| | 2× | water | O H ₂ | B | -27.4 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -35.9 |
| | 1/72× | boehm_surf-010_00w | Al ₇₂ O ₁₄₄ H ₇₂ | J | -36.5 |
| | 2× | water | O H ₂ | B | -33.4 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -15.8 |
| | 1/48× | boehm_surf-100_00w | Al ₄₈ O ₉₆ H ₄₈ | J | -39.0 |
| | 2× | water | O H ₂ | B | -17.4 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -20.2 |
| | 1/48× | boehm_surf-101_00w | Al ₄₈ O ₉₆ H ₄₈ | J | -31.1 |
| | 2× | water | O H ₂ | B | -20.1 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -25.3 |
| | 1/16× | gamma_bulk | Al ₁₆ O ₂₄ | J | -48.6 |
| | 5/2× | water | O H ₂ | B | -26.4 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -11.7 |
| | 1/96× | gamma_edge-100-110_00w | Al ₉₆ O ₁₄₄ | J | -41.6 |
| | 5/2× | water | O H ₂ | B | -12.4 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -14.8 |
| | 1/64× | gamma_surf-100_00w | Al ₆₄ O ₉₆ | J | -42.0 |
| | 5/2× | water | O H ₂ | B | -13.4 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -4.1 |
| | 1/64× | gamma_surf-110_00w | Al ₆₄ O ₉₆ | J | -41.2 |
| | 5/2× | water | O H ₂ | B | 0.2 |
| FOR | -1× | monomer | AlO ₄ H ₅ | R | -11.7 |
| | 1/40× | gamma_surf-111_00w | Al ₄₀ O ₆₀ | J | -43.0 |
| | 5/2× | water | O H ₂ | B | -15.3 |

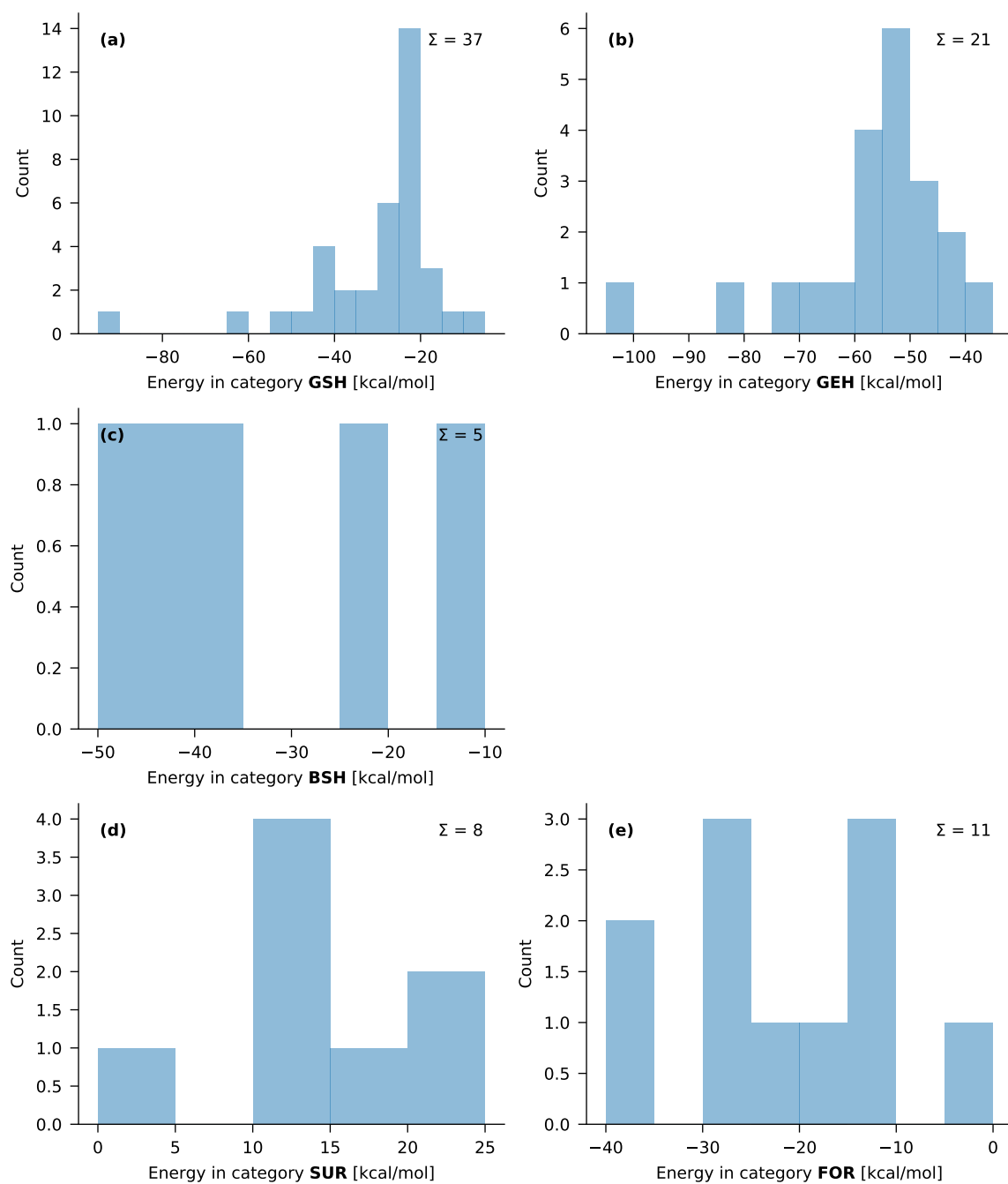


Figure S4: Histograms of reaction energies in the training set.

Table S3: List of activated parameters and their corresponding block, initial values from 1 and bounds.

| Name | Unit | Atoms | Block | Joshi 2014 | Lower bound | Upper bound | This work |
|----------------------|------------------------|---------|-------|------------|-------------|-------------|-----------|
| p_val3 | 1 | Al | ATM | 1.5000 | 1.2000 | 3.0000 | 2.9993 |
| p_val5 | 1 | Al | ATM | 2.5791 | 2.0633 | 3.0949 | 2.4980 |
| D_e ^{sigma} | kcal mol ⁻¹ | Al-H | BND | 92.8579 | 0.8579 | 122.7844 | 77.1779 |
| p_be1 | 1 | Al-H | BND | -0.6528 | -0.7834 | 1.0000 | -0.7823 |
| p_ovun1 | 1 | Al-H | BND | 0.1551 | 0.0100 | 0.5000 | 0.3089 |
| p_be2 | 1 | Al-H | BND | 10.0663 | 0.2281 | 13.0000 | 2.0985 |
| p_bo1 | 1 | Al-H | BND | -0.0842 | -0.3320 | -0.0674 | -0.1398 |
| p_bo2 | 1 | Al-H | BND | 7.1758 | 5.0015 | 15.0000 | 11.7930 |
| D_e ^{sigma} | kcal mol ⁻¹ | Al-O | BND | 182.0654 | 118.9203 | 232.7313 | 166.4329 |
| p_be1 | 1 | Al-O | BND | -0.0920 | -1.0000 | -0.0736 | -0.0963 |
| p_ovun1 | 1 | Al-O | BND | 0.1688 | 0.0100 | 0.4562 | 0.1457 |
| p_be2 | 1 | Al-O | BND | 0.0010 | 0.0008 | 1.5477 | 0.7025 |
| p_bo1 | 1 | Al-O | BND | -0.1959 | -0.2351 | -0.0740 | -0.2053 |
| p_bo2 | 1 | Al-O | BND | 6.0894 | 4.6533 | 7.3073 | 7.1635 |
| D_e ^{sigma} | kcal mol ⁻¹ | Al-Al | BND | 34.0777 | 27.2622 | 65.7742 | 31.9434 |
| r_0 ^{sigma} | Å | Al-H | OFD | 1.7276 | 1.3821 | 2.0731 | 1.4893 |
| r_0 ^{sigma} | Å | Al-O | OFD | 1.5646 | 1.2517 | 1.8775 | 1.6172 |
| p_val1 | 1 | Al-H-O | ANG | 4.2750 | 3.4200 | 20.0000 | 19.7283 |
| p_val2 | 1 | Al-H-O | ANG | 1.0250 | 0.8200 | 4.8339 | 3.9208 |
| p_val4 | 1 | Al-H-O | ANG | 1.4750 | 1.0100 | 1.7700 | 1.0931 |
| Theta_0,0 | deg | Al-O-H | ANG | 88.6163 | 64.6197 | 106.3396 | 66.0975 |
| p_val1 | 1 | Al-O-H | ANG | 10.1310 | 4.2037 | 19.7491 | 16.7556 |
| p_val2 | 1 | Al-O-H | ANG | 1.6896 | 1.3517 | 10.0000 | 2.4332 |
| p_val4 | 1 | Al-O-H | ANG | 1.0000 | 0.8000 | 3.0000 | 1.4778 |
| Theta_0,0 | deg | Al-O-Al | ANG | 13.8580 | 5.2474 | 64.5513 | 64.5486 |
| p_val1 | 1 | Al-O-Al | ANG | 12.3669 | 9.8935 | 40.0000 | 12.8201 |
| p_val2 | 1 | Al-O-Al | ANG | 4.4355 | 0.5527 | 9.9945 | 7.0734 |
| p_val4 | 1 | Al-O-Al | ANG | 1.1908 | 0.9526 | 3.0000 | 2.9786 |
| Theta_0,0 | deg | H-Al-O | ANG | 41.8108 | 0.0000 | 64.8437 | 9.7233 |
| p_val1 | 1 | H-Al-O | ANG | 17.3800 | 5.4547 | 30.9495 | 23.4162 |
| p_val2 | 1 | H-Al-O | ANG | 2.6618 | 0.9702 | 3.1942 | 3.0390 |
| p_val4 | 1 | H-Al-O | ANG | 1.0100 | 0.8080 | 3.0000 | 1.0218 |
| Theta_0,0 | deg | O-Al-O | ANG | 55.4358 | 43.7395 | 84.7469 | 61.5592 |
| p_val1 | 1 | O-Al-O | ANG | 22.1089 | 7.3926 | 40.0000 | 24.2337 |
| p_val2 | 1 | O-Al-O | ANG | 3.7402 | 1.2450 | 4.4882 | 1.2622 |
| p_val4 | 1 | O-Al-O | ANG | 2.2064 | 1.0123 | 3.0000 | 2.9975 |

Table S4: Structures in the validation set and contributions to each category of data. Bonds are represented by pairs of chemical elements, angles by triplets. The symbol \dots denotes a hydrogen bond. Internal coordinates with oxygen not bound to aluminum are discarded. For hydrated alumina surfaces, the internal coordinates without hydrogen are not considered.

| Structure | Chem. Form. | O-H | O...H | Al-O | Al-Al | O-Al-O | Al-O-Al | Al-O-H | Al-O...H | H-O...H | H...O...H |
|--------------------------------|---|-----|-------|------|-------|--------|---------|--------|----------|---------|-----------|
| gamma_bulk | Al ₃₂ O ₄₈ | | | 176 | 176 | 408 | 240 | | | | |
| gamma_surf-001_00w | Al ₆₄ O ₉₆ | | | 336 | 316 | 736 | 432 | | | | |
| gamma_surf-001_01w | Al ₆₄ O ₉₈ H ₄ | 4 | 2 | | 316 | | | 8 | 2 | 2 | |
| gamma_surf-001_02w | Al ₆₄ O ₁₀₀ H ₈ | 8 | 4 | | 318 | | | 12 | 4 | 4 | 2 |
| gamma_surf-001_03w | Al ₆₄ O ₁₀₂ H ₁₂ | 12 | 8 | | 316 | | | 20 | 8 | 8 | 4 |
| gamma_surf-001_04w | Al ₆₄ O ₁₀₄ H ₁₆ | 16 | 10 | | 316 | | | 26 | 14 | 10 | 4 |
| gamma_surf-110b_from_boehm_00w | Al ₈₀ O ₁₂₀ | | | 414 | 416 | 902 | 532 | | | | |
| gamma_surf-110b_from_boehm_02w | Al ₈₀ O ₁₂₄ H ₈ | 8 | | | 436 | | | 16 | | | |
| gamma_surf-110b_from_boehm_03w | Al ₈₀ O ₁₂₆ H ₁₂ | 12 | | | 436 | | | 22 | | | |
| gamma_surf-110b_from_boehm_04w | Al ₈₀ O ₁₂₈ H ₁₆ | 16 | 8 | | 436 | | | 24 | 8 | 8 | |
| gamma_surf-110b_from_bulk_00w | Al ₆₄ O ₉₆ | | | 336 | 324 | 744 | 440 | | | | |
| gamma_surf-110b_from_bulk_02w | Al ₆₄ O ₉₈ H ₄ | 4 | 2 | | 324 | | | 6 | 4 | | |
| gamma_surf-110b_from_bulk_04w | Al ₆₄ O ₁₀₀ H ₈ | 8 | 4 | | 324 | | | 12 | 4 | 4 | |
| gamma_surf-110b_from_bulk_06w | Al ₆₄ O ₁₀₂ H ₁₂ | 12 | 6 | | 324 | | | 16 | 10 | 4 | |
| gamma_surf-110b_from_bulk_08w | Al ₆₄ O ₁₀₄ H ₁₆ | 16 | 10 | | 324 | | | 20 | 18 | 6 | 2 |
| gamma_surf-1101_A1_00w | Al ₄₈ O ₇₂ | | | 250 | 238 | 552 | 324 | | | | |
| gamma_surf-1101_A1_01w | Al ₄₈ O ₇₄ H ₄ | 4 | 2 | | 234 | | | 8 | 4 | | |
| gamma_surf-1101_A1_02w | Al ₄₈ O ₇₆ H ₈ | 8 | 4 | | 240 | | | 14 | 6 | 2 | |
| gamma_surf-1101_A1_03w | Al ₄₈ O ₇₈ H ₁₂ | 12 | 4 | | 240 | | | 22 | 6 | 4 | |
| gamma_surf-1101_A1_04w | Al ₄₈ O ₈₀ H ₁₆ | 16 | 12 | | 236 | | | 26 | 22 | 6 | 4 |
| gamma_surf-1101_A1_05w | Al ₄₈ O ₈₂ H ₂₀ | 20 | 14 | | 236 | | | 30 | 26 | 8 | 4 |
| gamma_surf-1101_A1_06w | Al ₄₈ O ₈₄ H ₂₄ | 24 | 17 | | 236 | | | 30 | 37 | 4 | 9 |
| gamma_surf-1101_A2_00w | Al ₅₆ O ₈₄ | | | 296 | 276 | 660 | 392 | | | | |
| gamma_surf-1101_A2_01w | Al ₅₆ O ₈₆ H ₄ | 4 | 2 | | 274 | | | 8 | 4 | | |
| gamma_surf-1101_A2_02w | Al ₅₆ O ₈₈ H ₈ | 8 | 4 | | 282 | | | 14 | 8 | | 2 |
| gamma_surf-1101_A2_03w | Al ₅₆ O ₉₀ H ₁₂ | 12 | 7 | | 284 | | | 20 | 11 | 5 | 1 |
| gamma_surf-1101_A2_04w | Al ₅₆ O ₉₂ H ₁₆ | 16 | 10 | | 284 | | | 24 | 18 | 8 | |
| gamma_surf-1101_A2_05w | Al ₅₆ O ₉₄ H ₂₀ | 20 | 14 | | 282 | | | 26 | 30 | 4 | 6 |
| gamma_surf-1101_A2_06w | Al ₅₆ O ₉₆ H ₂₄ | 24 | 18 | | 282 | | | 30 | 38 | 6 | 8 |
| gamma_surf-1101_L1_04w | Al ₄₈ O ₈₀ H ₁₆ | 16 | 6 | | 240 | | | 32 | 12 | 4 | |
| gamma_surf-1101_L2_00w | Al ₅₆ O ₈₄ | | | 294 | 286 | 654 | 384 | | | | |
| gamma_surf-1101_L2_01w | Al ₅₆ O ₈₆ H ₄ | 4 | 2 | | 276 | | | 8 | 4 | | |
| gamma_surf-1101_L2_02w | Al ₅₆ O ₈₈ H ₈ | 8 | 4 | | 282 | | | 14 | 8 | | 2 |
| gamma_surf-1101_L2_03w | Al ₅₆ O ₉₀ H ₁₂ | 12 | 4 | | 282 | | | 22 | 10 | 2 | |
| gamma_surf-1101_L2_04w | Al ₅₆ O ₉₂ H ₁₆ | 16 | 6 | | 282 | | | 32 | 16 | 2 | |
| gamma_surf-111_D1_03w | Al ₃₂ O ₅₄ H ₁₂ | 12 | 5 | | 144 | | | 22 | 6 | 4 | 2 |
| gamma_surf-111_D1_04w | Al ₃₂ O ₅₆ H ₁₆ | 16 | 5 | | 146 | | | 30 | 6 | 5 | 2 |
| gamma_surf-111_D1_05w | Al ₃₂ O ₅₈ H ₂₀ | 20 | 10 | | 146 | | | 36 | 14 | 10 | 2 |
| gamma_surf-111_D1_06w | Al ₃₂ O ₆₀ H ₂₄ | 24 | 14 | | 146 | | | 44 | 17 | 14 | 5 |
| gamma_surf-111_D2_03w | Al ₃₂ O ₅₄ H ₁₂ | 12 | 8 | | 144 | | | 22 | 12 | 4 | 2 |
| gamma_surf-111_D2_04w | Al ₃₂ O ₅₆ H ₁₆ | 16 | 9 | | 146 | | | 30 | 14 | 7 | 2 |

| | | | | | | | | | | | |
|-------------------------|--|-----|-----|------|-------|------|------|------|-----|-----|----|
| gamma_surf-111_D2_05w | Al ₃₂ O ₅₈ H ₂₀ | 20 | 12 | | 146 | | | 36 | 15 | 12 | 6 |
| gamma_surf-111_D2_06w | Al ₃₂ O ₆₀ H ₂₄ | 24 | 15 | | 146 | | | 44 | 18 | 15 | 6 |
| gamma_surf-111_P1_1_05w | Al ₄₀ O ₇₀ H ₂₀ | 20 | 3 | | 184 | | | 43 | 6 | 3 | |
| gamma_surf-111_P1_1_06w | Al ₄₀ O ₇₂ H ₂₄ | 24 | 4 | | 178 | | | 54 | 8 | 4 | |
| gamma_surf-111_P1_2_03w | Al ₄₀ O ₆₆ H ₁₂ | 12 | 2 | | 178 | | | 24 | 4 | 2 | |
| gamma_surf-111_P1_2_04w | Al ₄₀ O ₆₈ H ₁₆ | 16 | 4 | | 178 | | | 34 | 8 | 4 | |
| gamma_surf-111_P1_2_05w | Al ₄₀ O ₇₀ H ₂₀ | 20 | 5 | | 178 | | | 44 | 10 | 5 | |
| gamma_surf-111_P1_2_06w | Al ₄₀ O ₇₂ H ₂₄ | 24 | 2 | | 182 | | | 54 | 4 | 2 | |
| gamma_surf-111_P2_1_05w | Al ₄₀ O ₇₀ H ₂₀ | 20 | 14 | | 196 | | | 27 | 15 | 13 | 4 |
| gamma_surf-111_P2_1_06w | Al ₄₀ O ₇₂ H ₂₄ | 24 | 16 | | 200 | | | 38 | 16 | 16 | 4 |
| gamma_surf-111_P2_2_04w | Al ₄₀ O ₆₈ H ₁₆ | 16 | 10 | | 194 | | | 26 | 16 | 6 | |
| gamma_surf-111_P2_2_05w | Al ₄₀ O ₇₀ H ₂₀ | 20 | 14 | | 190 | | | 32 | 18 | 12 | 3 |
| gamma_surf-111_P2_2_06w | Al ₄₀ O ₇₂ H ₂₄ | 24 | 15 | | 192 | | | 35 | 15 | 15 | 2 |
| water | O H ₂ | | | | | | | | | | |
| total | | 704 | 351 | 2102 | 13588 | 4656 | 2744 | 1217 | 554 | 254 | 88 |

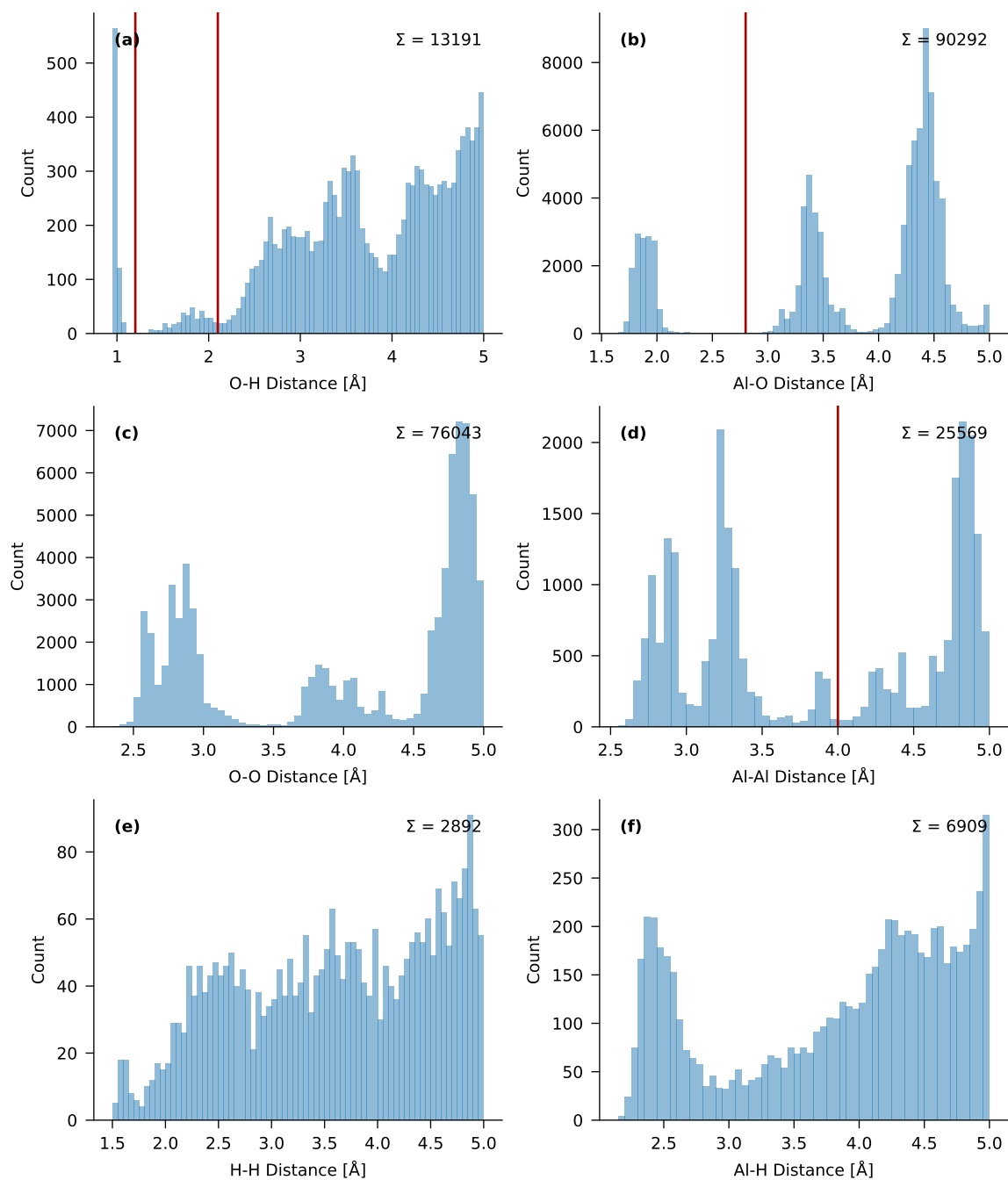


Figure S5: Histograms of all interatomic distances in the validation set up to 5 Å, grouped per pair of chemical elements. Cutoffs for OH and AlO pairs depicted as vertical red lines: 1.2 Å for O–H bonds, 2.1 Å for hydrogen bonds and 2.8 Å for Al–O bonds. See main text for a more detailed description.

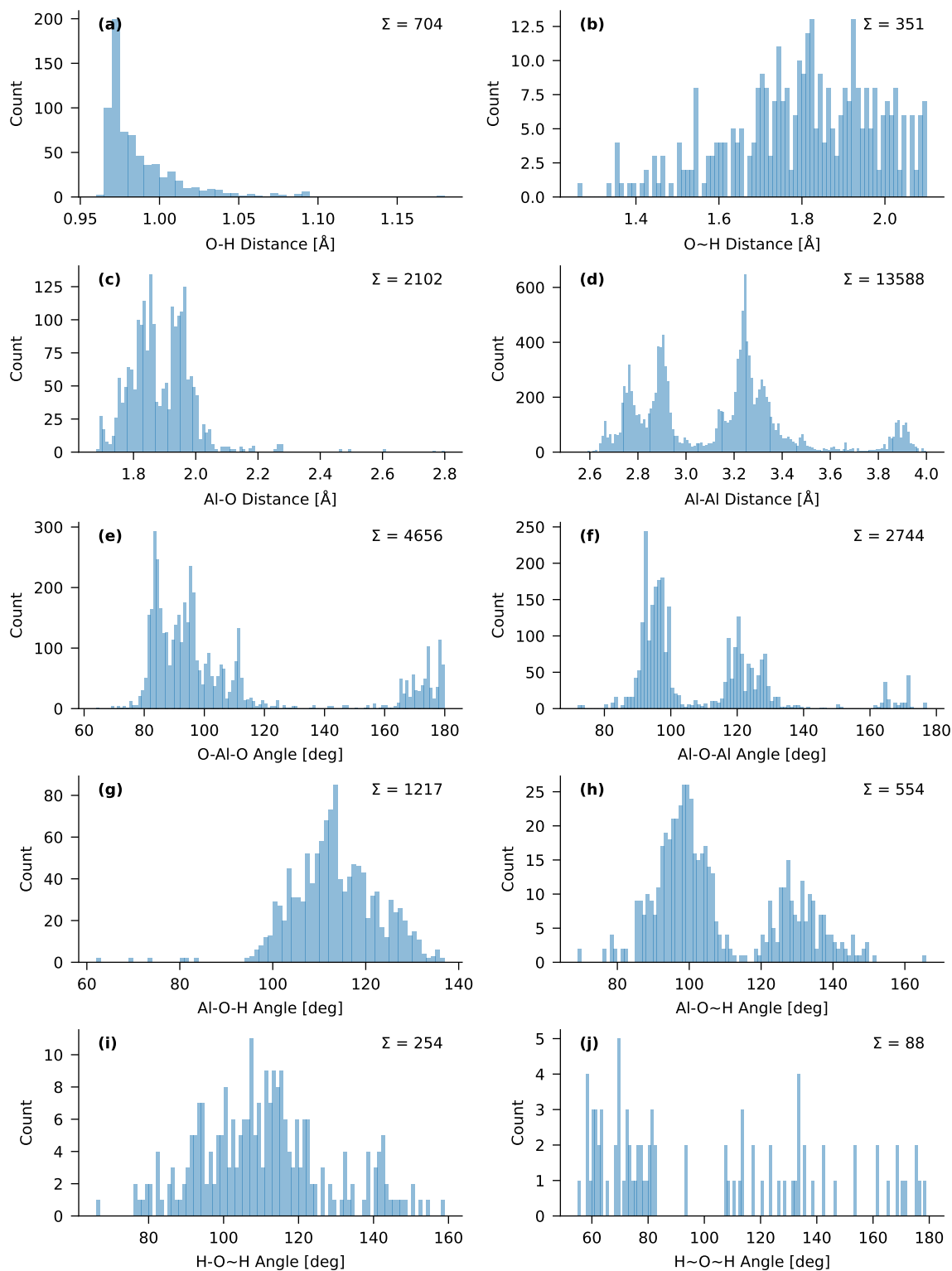


Figure S6: Histograms of internal coordinates in the validation set. In the labels on the horizontal axis, a dash represents a regular bond and a tilde represents a hydrogen bond.

Table S5: Overview of all chemical equations in the validation set. Reactants are given negative coefficients. For each reaction, three reaction energies are in kcal mol⁻¹: the reference DFT result (R), the prediction with the Joshi force field (J) and the prediction with the new force field in this work (B). Water adsorption energies are normalized to the number of water molecules. All other reaction energies are normalized on the number of Al atoms. The categories are defined in the main text.

| Category | Coeff. | Structure | Chem. Form. | Reaction energy |
|----------|--------|--------------------------------|---|-----------------|
| GSH | -1/2× | gamma_surf-001_00w | Al ₆₄ O ₉₆ | R -22.5 |
| | -1× | water | O H ₂ | J -14.3 |
| | 1/2× | gamma_surf-001_01w | Al ₆₄ O ₉₈ H ₄ | B -31.6 |
| GSH | -1/4× | gamma_surf-001_00w | Al ₆₄ O ₉₆ | R -24.4 |
| | -1× | water | O H ₂ | J -5.0 |
| | 1/4× | gamma_surf-001_02w | Al ₆₄ O ₁₀₀ H ₈ | B -27.5 |
| GSH | -1/2× | gamma_surf-001_01w | Al ₆₄ O ₉₈ H ₄ | R -26.3 |
| | -1× | water | O H ₂ | J 4.2 |
| | 1/2× | gamma_surf-001_02w | Al ₆₄ O ₁₀₀ H ₈ | B -23.4 |
| GSH | -1/6× | gamma_surf-001_00w | Al ₆₄ O ₉₆ | R -23.5 |
| | -1× | water | O H ₂ | J 0.2 |
| | 1/6× | gamma_surf-001_03w | Al ₆₄ O ₁₀₂ H ₁₂ | B -28.4 |
| GSH | -1/4× | gamma_surf-001_01w | Al ₆₄ O ₉₈ H ₄ | R -24.0 |
| | -1× | water | O H ₂ | J 7.4 |
| | 1/4× | gamma_surf-001_03w | Al ₆₄ O ₁₀₂ H ₁₂ | B -26.8 |
| GSH | -1/2× | gamma_surf-001_02w | Al ₆₄ O ₁₀₀ H ₈ | R -21.6 |
| | -1× | water | O H ₂ | J 10.6 |
| | 1/2× | gamma_surf-001_03w | Al ₆₄ O ₁₀₂ H ₁₂ | B -30.1 |
| GSH | -1/8× | gamma_surf-001_00w | Al ₆₄ O ₉₆ | R -20.8 |
| | -1× | water | O H ₂ | J -12.3 |
| | 1/8× | gamma_surf-001_04w | Al ₆₄ O ₁₀₄ H ₁₆ | B -24.1 |
| GSH | -1/6× | gamma_surf-001_01w | Al ₆₄ O ₉₈ H ₄ | R -20.2 |
| | -1× | water | O H ₂ | J -11.6 |
| | 1/6× | gamma_surf-001_04w | Al ₆₄ O ₁₀₄ H ₁₆ | B -21.6 |
| GSH | -1/4× | gamma_surf-001_02w | Al ₆₄ O ₁₀₀ H ₈ | R -17.1 |
| | -1× | water | O H ₂ | J -19.5 |
| | 1/4× | gamma_surf-001_04w | Al ₆₄ O ₁₀₄ H ₁₆ | B -20.7 |
| GSH | -1/2× | gamma_surf-001_03w | Al ₆₄ O ₁₀₂ H ₁₂ | R -12.6 |
| | -1× | water | O H ₂ | J -49.7 |
| | 1/2× | gamma_surf-001_04w | Al ₆₄ O ₁₀₄ H ₁₆ | B -11.3 |
| GSH | -1/4× | gamma_surf-110b_from_boehm_00w | Al ₈₀ O ₁₂₀ | R -60.7 |
| | -1× | water | O H ₂ | J 5.5 |
| | 1/4× | gamma_surf-110b_from_boehm_02w | Al ₈₀ O ₁₂₄ H ₈ | B -50.1 |
| GSH | -1/6× | gamma_surf-110b_from_boehm_00w | Al ₈₀ O ₁₂₀ | R -58.9 |
| | -1× | water | O H ₂ | J 10.3 |

| | | | | | |
|-----|-------|--------------------------------|---|---|-------|
| | 1/6× | gamma_surf-110b_from_boehm_03w | Al ₈₀ O ₁₂₆ H ₁₂ | B | -56.7 |
| GSH | -1/2× | gamma_surf-110b_from_boehm_02w | Al ₈₀ O ₁₂₄ H ₈ | R | -55.1 |
| | -1× | water | O H ₂ | J | 19.7 |
| | 1/2× | gamma_surf-110b_from_boehm_03w | Al ₈₀ O ₁₂₆ H ₁₂ | B | -70.0 |
| GSH | -1/8× | gamma_surf-110b_from_boehm_00w | Al ₈₀ O ₁₂₀ | R | -58.2 |
| | -1× | water | O H ₂ | J | -7.7 |
| | 1/8× | gamma_surf-110b_from_boehm_04w | Al ₈₀ O ₁₂₈ H ₁₆ | B | -53.6 |
| GSH | -1/4× | gamma_surf-110b_from_boehm_02w | Al ₈₀ O ₁₂₄ H ₈ | R | -55.7 |
| | -1× | water | O H ₂ | J | -21.0 |
| | 1/4× | gamma_surf-110b_from_boehm_04w | Al ₈₀ O ₁₂₈ H ₁₆ | B | -57.2 |
| GSH | -1/2× | gamma_surf-110b_from_boehm_03w | Al ₈₀ O ₁₂₆ H ₁₂ | R | -56.3 |
| | -1× | water | O H ₂ | J | -61.7 |
| | 1/2× | gamma_surf-110b_from_boehm_04w | Al ₈₀ O ₁₂₈ H ₁₆ | B | -44.4 |
| GSH | -1/2× | gamma_surf-110b_from_bulk_00w | Al ₆₄ O ₉₆ | R | -43.6 |
| | -1× | water | O H ₂ | J | -26.1 |
| | 1/2× | gamma_surf-110b_from_bulk_02w | Al ₆₄ O ₉₈ H ₄ | B | -49.6 |
| GSH | -1/4× | gamma_surf-110b_from_bulk_00w | Al ₆₄ O ₉₆ | R | -45.8 |
| | -1× | water | O H ₂ | J | -43.6 |
| | 1/4× | gamma_surf-110b_from_bulk_04w | Al ₆₄ O ₁₀₀ H ₈ | B | -50.9 |
| GSH | -1/2× | gamma_surf-110b_from_bulk_02w | Al ₆₄ O ₉₈ H ₄ | R | -48.0 |
| | -1× | water | O H ₂ | J | -61.0 |
| | 1/2× | gamma_surf-110b_from_bulk_04w | Al ₆₄ O ₁₀₀ H ₈ | B | -52.1 |
| GSH | -1/6× | gamma_surf-110b_from_bulk_00w | Al ₆₄ O ₉₆ | R | -39.1 |
| | -1× | water | O H ₂ | J | -32.6 |
| | 1/6× | gamma_surf-110b_from_bulk_06w | Al ₆₄ O ₁₀₂ H ₁₂ | B | -41.3 |
| GSH | -1/4× | gamma_surf-110b_from_bulk_02w | Al ₆₄ O ₉₈ H ₄ | R | -36.8 |
| | -1× | water | O H ₂ | J | -35.9 |
| | 1/4× | gamma_surf-110b_from_bulk_06w | Al ₆₄ O ₁₀₂ H ₁₂ | B | -37.2 |
| GSH | -1/2× | gamma_surf-110b_from_bulk_04w | Al ₆₄ O ₁₀₀ H ₈ | R | -25.6 |
| | -1× | water | O H ₂ | J | -10.7 |
| | 1/2× | gamma_surf-110b_from_bulk_06w | Al ₆₄ O ₁₀₂ H ₁₂ | B | -22.2 |
| GSH | -1/8× | gamma_surf-110b_from_bulk_00w | Al ₆₄ O ₉₆ | R | -36.4 |
| | -1× | water | O H ₂ | J | -26.9 |
| | 1/8× | gamma_surf-110b_from_bulk_08w | Al ₆₄ O ₁₀₄ H ₁₆ | B | -38.1 |
| GSH | -1/6× | gamma_surf-110b_from_bulk_02w | Al ₆₄ O ₉₈ H ₄ | R | -34.0 |
| | -1× | water | O H ₂ | J | -27.1 |
| | 1/6× | gamma_surf-110b_from_bulk_08w | Al ₆₄ O ₁₀₄ H ₁₆ | B | -34.3 |
| GSH | -1/4× | gamma_surf-110b_from_bulk_04w | Al ₆₄ O ₁₀₀ H ₈ | R | -26.9 |
| | -1× | water | O H ₂ | J | -10.2 |
| | 1/4× | gamma_surf-110b_from_bulk_08w | Al ₆₄ O ₁₀₄ H ₁₆ | B | -25.4 |
| GSH | -1/2× | gamma_surf-110b_from_bulk_06w | Al ₆₄ O ₁₀₂ H ₁₂ | R | -28.2 |
| | -1× | water | O H ₂ | J | -9.7 |
| | 1/2× | gamma_surf-110b_from_bulk_08w | Al ₆₄ O ₁₀₄ H ₁₆ | B | -28.6 |
| GSH | -1/2× | gamma_surf-110l_A1_00w | Al ₄₈ O ₇₂ | R | -86.5 |

| | | | | | |
|------------|--------|------------------------|--|---|-------|
| | -1× | water | O H ₂ | J | 61.0 |
| | 1/2× | gamma_surf-1101_A1_01w | Al ₄₈ O ₇₄ H ₄ | B | -83.3 |
| GSH | -1/4× | gamma_surf-1101_A1_00w | Al ₄₈ O ₇₂ | R | -58.5 |
| | -1× | water | O H ₂ | J | 4.1 |
| | 1/4× | gamma_surf-1101_A1_02w | Al ₄₈ O ₇₆ H ₈ | B | -65.9 |
| GSH | -1/2× | gamma_surf-1101_A1_01w | Al ₄₈ O ₇₄ H ₄ | R | -30.5 |
| | -1× | water | O H ₂ | J | -52.8 |
| | 1/2× | gamma_surf-1101_A1_02w | Al ₄₈ O ₇₆ H ₈ | B | -48.5 |
| GSH | -1/6× | gamma_surf-1101_A1_00w | Al ₄₈ O ₇₂ | R | -65.3 |
| | -1× | water | O H ₂ | J | 21.2 |
| | 1/6× | gamma_surf-1101_A1_03w | Al ₄₈ O ₇₈ H ₁₂ | B | -67.2 |
| GSH | -1/4× | gamma_surf-1101_A1_01w | Al ₄₈ O ₇₄ H ₄ | R | -54.6 |
| | -1× | water | O H ₂ | J | 1.4 |
| | 1/4× | gamma_surf-1101_A1_03w | Al ₄₈ O ₇₈ H ₁₂ | B | -59.2 |
| GSH | -1/2× | gamma_surf-1101_A1_02w | Al ₄₈ O ₇₆ H ₈ | R | -78.8 |
| | -1× | water | O H ₂ | J | 55.5 |
| | 1/2× | gamma_surf-1101_A1_03w | Al ₄₈ O ₇₈ H ₁₂ | B | -69.8 |
| GSH | -1/8× | gamma_surf-1101_A1_00w | Al ₄₈ O ₇₂ | R | -57.0 |
| | -1× | water | O H ₂ | J | 15.1 |
| | 1/8× | gamma_surf-1101_A1_04w | Al ₄₈ O ₈₀ H ₁₆ | B | -55.8 |
| GSH | -1/6× | gamma_surf-1101_A1_01w | Al ₄₈ O ₇₄ H ₄ | R | -47.1 |
| | -1× | water | O H ₂ | J | -0.2 |
| | 1/6× | gamma_surf-1101_A1_04w | Al ₄₈ O ₈₀ H ₁₆ | B | -46.6 |
| GSH | -1/4× | gamma_surf-1101_A1_02w | Al ₄₈ O ₇₆ H ₈ | R | -55.4 |
| | -1× | water | O H ₂ | J | 26.1 |
| | 1/4× | gamma_surf-1101_A1_04w | Al ₄₈ O ₈₀ H ₁₆ | B | -45.6 |
| GSH | -1/2× | gamma_surf-1101_A1_03w | Al ₄₈ O ₇₈ H ₁₂ | R | -32.0 |
| | -1× | water | O H ₂ | J | -3.4 |
| | 1/2× | gamma_surf-1101_A1_04w | Al ₄₈ O ₈₀ H ₁₆ | B | -21.4 |
| GSH | -1/10× | gamma_surf-1101_A1_00w | Al ₄₈ O ₇₂ | R | -49.8 |
| | -1× | water | O H ₂ | J | 5.5 |
| | 1/10× | gamma_surf-1101_A1_05w | Al ₄₈ O ₈₂ H ₂₀ | B | -48.5 |
| GSH | -1/8× | gamma_surf-1101_A1_01w | Al ₄₈ O ₇₄ H ₄ | R | -40.7 |
| | -1× | water | O H ₂ | J | -8.4 |
| | 1/8× | gamma_surf-1101_A1_05w | Al ₄₈ O ₈₂ H ₂₀ | B | -39.8 |
| GSH | -1/6× | gamma_surf-1101_A1_02w | Al ₄₈ O ₇₆ H ₈ | R | -44.1 |
| | -1× | water | O H ₂ | J | 6.4 |
| | 1/6× | gamma_surf-1101_A1_05w | Al ₄₈ O ₈₂ H ₂₀ | B | -36.9 |
| GSH | -1/4× | gamma_surf-1101_A1_03w | Al ₄₈ O ₇₈ H ₁₂ | R | -26.7 |
| | -1× | water | O H ₂ | J | -18.1 |
| | 1/4× | gamma_surf-1101_A1_05w | Al ₄₈ O ₈₂ H ₂₀ | B | -20.4 |
| GSH | -1/2× | gamma_surf-1101_A1_04w | Al ₄₈ O ₈₀ H ₁₆ | R | -21.4 |
| | -1× | water | O H ₂ | J | -32.8 |
| | 1/2× | gamma_surf-1101_A1_05w | Al ₄₈ O ₈₂ H ₂₀ | B | -19.4 |

| | | | | | |
|-----|--------|------------------------|--|---|-------|
| GSH | -1/12× | gamma_surf-1101_A1_00w | Al ₄₈ O ₇₂ | R | -43.6 |
| | -1× | water | O H ₂ | J | -1.2 |
| | 1/12× | gamma_surf-1101_A1_06w | Al ₄₈ O ₈₄ H ₂₄ | B | -42.5 |
| GSH | -1/10× | gamma_surf-1101_A1_01w | Al ₄₈ O ₇₄ H ₄ | R | -35.0 |
| | -1× | water | O H ₂ | J | -13.6 |
| | 1/10× | gamma_surf-1101_A1_06w | Al ₄₈ O ₈₄ H ₂₄ | B | -34.4 |
| GSH | -1/8× | gamma_surf-1101_A1_02w | Al ₄₈ O ₇₆ H ₈ | R | -36.2 |
| | -1× | water | O H ₂ | J | -3.8 |
| | 1/8× | gamma_surf-1101_A1_06w | Al ₄₈ O ₈₄ H ₂₄ | B | -30.8 |
| GSH | -1/6× | gamma_surf-1101_A1_03w | Al ₄₈ O ₇₈ H ₁₂ | R | -22.0 |
| | -1× | water | O H ₂ | J | -23.6 |
| | 1/6× | gamma_surf-1101_A1_06w | Al ₄₈ O ₈₄ H ₂₄ | B | -17.8 |
| GSH | -1/4× | gamma_surf-1101_A1_04w | Al ₄₈ O ₈₀ H ₁₆ | R | -17.0 |
| | -1× | water | O H ₂ | J | -33.7 |
| | 1/4× | gamma_surf-1101_A1_06w | Al ₄₈ O ₈₄ H ₂₄ | B | -16.0 |
| GSH | -1/2× | gamma_surf-1101_A1_05w | Al ₄₈ O ₈₂ H ₂₀ | R | -12.5 |
| | -1× | water | O H ₂ | J | -34.6 |
| | 1/2× | gamma_surf-1101_A1_06w | Al ₄₈ O ₈₄ H ₂₄ | B | -12.7 |
| GSH | -1/2× | gamma_surf-1101_A2_00w | Al ₅₆ O ₈₄ | R | -82.6 |
| | -1× | water | O H ₂ | J | -40.3 |
| | 1/2× | gamma_surf-1101_A2_01w | Al ₅₆ O ₈₆ H ₄ | B | -89.7 |
| GSH | -1/4× | gamma_surf-1101_A2_00w | Al ₅₆ O ₈₄ | R | -66.4 |
| | -1× | water | O H ₂ | J | -41.8 |
| | 1/4× | gamma_surf-1101_A2_02w | Al ₅₆ O ₈₈ H ₈ | B | -75.6 |
| GSH | -1/2× | gamma_surf-1101_A2_01w | Al ₅₆ O ₈₆ H ₄ | R | -50.2 |
| | -1× | water | O H ₂ | J | -43.2 |
| | 1/2× | gamma_surf-1101_A2_02w | Al ₅₆ O ₈₈ H ₈ | B | -61.5 |
| GSH | -1/6× | gamma_surf-1101_A2_00w | Al ₅₆ O ₈₄ | R | -60.3 |
| | -1× | water | O H ₂ | J | -35.4 |
| | 1/6× | gamma_surf-1101_A2_03w | Al ₅₆ O ₉₀ H ₁₂ | B | -65.2 |
| GSH | -1/4× | gamma_surf-1101_A2_01w | Al ₅₆ O ₈₆ H ₄ | R | -49.1 |
| | -1× | water | O H ₂ | J | -32.9 |
| | 1/4× | gamma_surf-1101_A2_03w | Al ₅₆ O ₉₀ H ₁₂ | B | -52.9 |
| GSH | -1/2× | gamma_surf-1101_A2_02w | Al ₅₆ O ₈₈ H ₈ | R | -48.1 |
| | -1× | water | O H ₂ | J | -22.5 |
| | 1/2× | gamma_surf-1101_A2_03w | Al ₅₆ O ₉₀ H ₁₂ | B | -44.3 |
| GSH | -1/8× | gamma_surf-1101_A2_00w | Al ₅₆ O ₈₄ | R | -52.7 |
| | -1× | water | O H ₂ | J | -10.4 |
| | 1/8× | gamma_surf-1101_A2_04w | Al ₅₆ O ₉₂ H ₁₆ | B | -56.2 |
| GSH | -1/6× | gamma_surf-1101_A2_01w | Al ₅₆ O ₈₆ H ₄ | R | -42.7 |
| | -1× | water | O H ₂ | J | -0.4 |
| | 1/6× | gamma_surf-1101_A2_04w | Al ₅₆ O ₉₂ H ₁₆ | B | -45.1 |
| GSH | -1/4× | gamma_surf-1101_A2_02w | Al ₅₆ O ₈₈ H ₈ | R | -38.9 |
| | -1× | water | O H ₂ | J | 21.0 |

| | | | | | |
|-----|--------|------------------------|--|---|-------|
| | 1/4× | gamma_surf-1101_A2_04w | Al ₅₆ O ₉₂ H ₁₆ | B | -36.8 |
| GSH | -1/2× | gamma_surf-1101_A2_03w | Al ₅₆ O ₉₀ H ₁₂ | R | -29.8 |
| | -1× | water | O H ₂ | J | 64.5 |
| | 1/2× | gamma_surf-1101_A2_04w | Al ₅₆ O ₉₂ H ₁₆ | B | -29.3 |
| GSH | -1/10× | gamma_surf-1101_A2_00w | Al ₅₆ O ₈₄ | R | -46.1 |
| | -1× | water | O H ₂ | J | -5.3 |
| | 1/10× | gamma_surf-1101_A2_05w | Al ₅₆ O ₉₄ H ₂₀ | B | -50.1 |
| GSH | -1/8× | gamma_surf-1101_A2_01w | Al ₅₆ O ₈₆ H ₄ | R | -37.0 |
| | -1× | water | O H ₂ | J | 3.5 |
| | 1/8× | gamma_surf-1101_A2_05w | Al ₅₆ O ₉₄ H ₂₀ | B | -40.2 |
| GSH | -1/6× | gamma_surf-1101_A2_02w | Al ₅₆ O ₈₈ H ₈ | R | -32.6 |
| | -1× | water | O H ₂ | J | 19.0 |
| | 1/6× | gamma_surf-1101_A2_05w | Al ₅₆ O ₉₄ H ₂₀ | B | -33.0 |
| GSH | -1/4× | gamma_surf-1101_A2_03w | Al ₅₆ O ₉₀ H ₁₂ | R | -24.9 |
| | -1× | water | O H ₂ | J | 39.8 |
| | 1/4× | gamma_surf-1101_A2_05w | Al ₅₆ O ₉₄ H ₂₀ | B | -27.4 |
| GSH | -1/2× | gamma_surf-1101_A2_04w | Al ₅₆ O ₉₂ H ₁₆ | R | -20.0 |
| | -1× | water | O H ₂ | J | 15.2 |
| | 1/2× | gamma_surf-1101_A2_05w | Al ₅₆ O ₉₄ H ₂₀ | B | -25.5 |
| GSH | -1/12× | gamma_surf-1101_A2_00w | Al ₅₆ O ₈₄ | R | -42.0 |
| | -1× | water | O H ₂ | J | -17.4 |
| | 1/12× | gamma_surf-1101_A2_06w | Al ₅₆ O ₉₆ H ₂₄ | B | -47.1 |
| GSH | -1/10× | gamma_surf-1101_A2_01w | Al ₅₆ O ₈₆ H ₄ | R | -33.9 |
| | -1× | water | O H ₂ | J | -12.8 |
| | 1/10× | gamma_surf-1101_A2_06w | Al ₅₆ O ₉₆ H ₂₄ | B | -38.5 |
| GSH | -1/8× | gamma_surf-1101_A2_02w | Al ₅₆ O ₈₈ H ₈ | R | -29.8 |
| | -1× | water | O H ₂ | J | -5.1 |
| | 1/8× | gamma_surf-1101_A2_06w | Al ₅₆ O ₉₆ H ₂₄ | B | -32.8 |
| GSH | -1/6× | gamma_surf-1101_A2_03w | Al ₅₆ O ₉₀ H ₁₂ | R | -23.7 |
| | -1× | water | O H ₂ | J | 0.6 |
| | 1/6× | gamma_surf-1101_A2_06w | Al ₅₆ O ₉₆ H ₂₄ | B | -29.0 |
| GSH | -1/4× | gamma_surf-1101_A2_04w | Al ₅₆ O ₉₂ H ₁₆ | R | -20.7 |
| | -1× | water | O H ₂ | J | -31.3 |
| | 1/4× | gamma_surf-1101_A2_06w | Al ₅₆ O ₉₆ H ₂₄ | B | -28.8 |
| GSH | -1/2× | gamma_surf-1101_A2_05w | Al ₅₆ O ₉₄ H ₂₀ | R | -21.4 |
| | -1× | water | O H ₂ | J | -77.7 |
| | 1/2× | gamma_surf-1101_A2_06w | Al ₅₆ O ₉₆ H ₂₄ | B | -32.1 |
| GSH | -1/2× | gamma_surf-1101_L2_00w | Al ₅₆ O ₈₄ | R | -60.8 |
| | -1× | water | O H ₂ | J | -15.8 |
| | 1/2× | gamma_surf-1101_L2_01w | Al ₅₆ O ₈₆ H ₄ | B | -58.3 |
| GSH | -1/4× | gamma_surf-1101_L2_00w | Al ₅₆ O ₈₄ | R | -44.6 |
| | -1× | water | O H ₂ | J | 10.5 |
| | 1/4× | gamma_surf-1101_L2_02w | Al ₅₆ O ₈₈ H ₈ | B | -51.9 |
| GSH | -1/2× | gamma_surf-1101_L2_01w | Al ₅₆ O ₈₆ H ₄ | R | -28.4 |

| | | | | | |
|------------|-------|------------------------|--|---|-------|
| | -1× | water | O H ₂ | J | 36.8 |
| | 1/2× | gamma_surf-1101_L2_02w | Al ₅₆ O ₈₈ H ₈ | B | -45.5 |
| GSH | -1/6× | gamma_surf-1101_L2_00w | Al ₅₆ O ₈₄ | R | -50.5 |
| | -1× | water | O H ₂ | J | 25.8 |
| | 1/6× | gamma_surf-1101_L2_03w | Al ₅₆ O ₉₀ H ₁₂ | B | -53.2 |
| GSH | -1/4× | gamma_surf-1101_L2_01w | Al ₅₆ O ₈₆ H ₄ | R | -45.3 |
| | -1× | water | O H ₂ | J | 46.6 |
| | 1/4× | gamma_surf-1101_L2_03w | Al ₅₆ O ₉₀ H ₁₂ | B | -50.6 |
| GSH | -1/2× | gamma_surf-1101_L2_02w | Al ₅₆ O ₈₈ H ₈ | R | -62.3 |
| | -1× | water | O H ₂ | J | 56.5 |
| | 1/2× | gamma_surf-1101_L2_03w | Al ₅₆ O ₉₀ H ₁₂ | B | -55.7 |
| GSH | -1/8× | gamma_surf-1101_L2_00w | Al ₅₆ O ₈₄ | R | -53.0 |
| | -1× | water | O H ₂ | J | 24.9 |
| | 1/8× | gamma_surf-1101_L2_04w | Al ₅₆ O ₉₂ H ₁₆ | B | -49.3 |
| GSH | -1/6× | gamma_surf-1101_L2_01w | Al ₅₆ O ₈₆ H ₄ | R | -50.4 |
| | -1× | water | O H ₂ | J | 38.5 |
| | 1/6× | gamma_surf-1101_L2_04w | Al ₅₆ O ₉₂ H ₁₆ | B | -46.2 |
| GSH | -1/4× | gamma_surf-1101_L2_02w | Al ₅₆ O ₈₈ H ₈ | R | -61.5 |
| | -1× | water | O H ₂ | J | 39.3 |
| | 1/4× | gamma_surf-1101_L2_04w | Al ₅₆ O ₉₂ H ₁₆ | B | -46.6 |
| GSH | -1/2× | gamma_surf-1101_L2_03w | Al ₅₆ O ₉₀ H ₁₂ | R | -60.7 |
| | -1× | water | O H ₂ | J | 22.1 |
| | 1/2× | gamma_surf-1101_L2_04w | Al ₅₆ O ₉₂ H ₁₆ | B | -37.5 |
| GSH | -1/2× | gamma_surf-111_D1_03w | Al ₃₂ O ₅₄ H ₁₂ | R | -60.7 |
| | -1× | water | O H ₂ | J | -5.4 |
| | 1/2× | gamma_surf-111_D1_04w | Al ₃₂ O ₅₆ H ₁₆ | B | -49.0 |
| GSH | -1/4× | gamma_surf-111_D1_03w | Al ₃₂ O ₅₄ H ₁₂ | R | -49.5 |
| | -1× | water | O H ₂ | J | 1.5 |
| | 1/4× | gamma_surf-111_D1_05w | Al ₃₂ O ₅₈ H ₂₀ | B | -38.6 |
| GSH | -1/2× | gamma_surf-111_D1_04w | Al ₃₂ O ₅₆ H ₁₆ | R | -38.4 |
| | -1× | water | O H ₂ | J | 8.5 |
| | 1/2× | gamma_surf-111_D1_05w | Al ₃₂ O ₅₈ H ₂₀ | B | -28.3 |
| GSH | -1/6× | gamma_surf-111_D1_03w | Al ₃₂ O ₅₄ H ₁₂ | R | -42.9 |
| | -1× | water | O H ₂ | J | 15.4 |
| | 1/6× | gamma_surf-111_D1_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -34.8 |
| GSH | -1/4× | gamma_surf-111_D1_04w | Al ₃₂ O ₅₆ H ₁₆ | R | -34.0 |
| | -1× | water | O H ₂ | J | 25.8 |
| | 1/4× | gamma_surf-111_D1_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -27.7 |
| GSH | -1/2× | gamma_surf-111_D1_05w | Al ₃₂ O ₅₈ H ₂₀ | R | -29.6 |
| | -1× | water | O H ₂ | J | 43.2 |
| | 1/2× | gamma_surf-111_D1_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -27.1 |
| GSH | -1/2× | gamma_surf-111_D2_03w | Al ₃₂ O ₅₄ H ₁₂ | R | -50.3 |
| | -1× | water | O H ₂ | J | 46.1 |
| | 1/2× | gamma_surf-111_D2_04w | Al ₃₂ O ₅₆ H ₁₆ | B | -27.4 |

| | | | | | |
|------------|-------|-------------------------|--|---|-------|
| GSH | -1/4× | gamma_surf-111_D2_03w | Al ₃₂ O ₅₄ H ₁₂ | R | -44.4 |
| | -1× | water | O H ₂ | J | 39.7 |
| | 1/4× | gamma_surf-111_D2_05w | Al ₃₂ O ₅₈ H ₂₀ | B | -34.8 |
| GSH | -1/2× | gamma_surf-111_D2_04w | Al ₃₂ O ₅₆ H ₁₆ | R | -38.5 |
| | -1× | water | O H ₂ | J | 33.3 |
| | 1/2× | gamma_surf-111_D2_05w | Al ₃₂ O ₅₈ H ₂₀ | B | -42.1 |
| GSH | -1/6× | gamma_surf-111_D2_03w | Al ₃₂ O ₅₄ H ₁₂ | R | -39.4 |
| | -1× | water | O H ₂ | J | 24.6 |
| | 1/6× | gamma_surf-111_D2_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -34.0 |
| GSH | -1/4× | gamma_surf-111_D2_04w | Al ₃₂ O ₅₆ H ₁₆ | R | -33.9 |
| | -1× | water | O H ₂ | J | 13.8 |
| | 1/4× | gamma_surf-111_D2_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -37.3 |
| GSH | -1/2× | gamma_surf-111_D2_05w | Al ₃₂ O ₅₈ H ₂₀ | R | -29.4 |
| | -1× | water | O H ₂ | J | -5.6 |
| | 1/2× | gamma_surf-111_D2_06w | Al ₃₂ O ₆₀ H ₂₄ | B | -32.4 |
| GSH | -1/2× | gamma_surf-111_P1_1_05w | Al ₄₀ O ₇₀ H ₂₀ | R | -21.2 |
| | -1× | water | O H ₂ | J | 1.7 |
| | 1/2× | gamma_surf-111_P1_1_06w | Al ₄₀ O ₇₂ H ₂₄ | B | -22.9 |
| GSH | -1/2× | gamma_surf-111_P1_2_03w | Al ₄₀ O ₆₆ H ₁₂ | R | -51.2 |
| | -1× | water | O H ₂ | J | 47.5 |
| | 1/2× | gamma_surf-111_P1_2_04w | Al ₄₀ O ₆₈ H ₁₆ | B | -36.4 |
| GSH | -1/4× | gamma_surf-111_P1_2_03w | Al ₄₀ O ₆₆ H ₁₂ | R | -41.3 |
| | -1× | water | O H ₂ | J | 34.9 |
| | 1/4× | gamma_surf-111_P1_2_05w | Al ₄₀ O ₇₀ H ₂₀ | B | -35.9 |
| GSH | -1/2× | gamma_surf-111_P1_2_04w | Al ₄₀ O ₆₈ H ₁₆ | R | -31.4 |
| | -1× | water | O H ₂ | J | 22.3 |
| | 1/2× | gamma_surf-111_P1_2_05w | Al ₄₀ O ₇₀ H ₂₀ | B | -35.4 |
| GSH | -1/6× | gamma_surf-111_P1_2_03w | Al ₄₀ O ₆₆ H ₁₂ | R | -35.0 |
| | -1× | water | O H ₂ | J | 24.3 |
| | 1/6× | gamma_surf-111_P1_2_06w | Al ₄₀ O ₇₂ H ₂₄ | B | -33.6 |
| GSH | -1/4× | gamma_surf-111_P1_2_04w | Al ₄₀ O ₆₈ H ₁₆ | R | -26.9 |
| | -1× | water | O H ₂ | J | 12.7 |
| | 1/4× | gamma_surf-111_P1_2_06w | Al ₄₀ O ₇₂ H ₂₄ | B | -32.2 |
| GSH | -1/2× | gamma_surf-111_P1_2_05w | Al ₄₀ O ₇₀ H ₂₀ | R | -22.3 |
| | -1× | water | O H ₂ | J | 3.1 |
| | 1/2× | gamma_surf-111_P1_2_06w | Al ₄₀ O ₇₂ H ₂₄ | B | -29.1 |
| GSH | -1/2× | gamma_surf-111_P2_1_05w | Al ₄₀ O ₇₀ H ₂₀ | R | -23.0 |
| | -1× | water | O H ₂ | J | 60.2 |
| | 1/2× | gamma_surf-111_P2_1_06w | Al ₄₀ O ₇₂ H ₂₄ | B | -15.2 |
| GSH | -1/2× | gamma_surf-111_P2_2_04w | Al ₄₀ O ₆₈ H ₁₆ | R | -46.0 |
| | -1× | water | O H ₂ | J | 29.6 |
| | 1/2× | gamma_surf-111_P2_2_05w | Al ₄₀ O ₇₀ H ₂₀ | B | -20.8 |
| GSH | -1/4× | gamma_surf-111_P2_2_04w | Al ₄₀ O ₆₈ H ₁₆ | R | -17.4 |
| | -1× | water | O H ₂ | J | 3.1 |

| | | | | | |
|------------|--------|--------------------------------|--|---|-------|
| | 1/4× | gamma_surf-111_P2_2_06w | Al ₄₀ O ₇₂ H ₂₄ | B | -21.7 |
| GSH | -1/2× | gamma_surf-111_P2_2_05w | Al ₄₀ O ₇₀ H ₂₀ | R | 11.2 |
| | -1× | water | O H ₂ | J | -23.5 |
| | 1/2× | gamma_surf-111_P2_2_06w | Al ₄₀ O ₇₂ H ₂₄ | B | -22.6 |
| SUR | -1/32× | gamma_bulk | Al ₃₂ O ₄₈ | R | 5.1 |
| | 1/64× | gamma_surf-001_00w | Al ₆₄ O ₉₆ | J | 2.5 |
| | | | | B | 4.9 |
| SUR | -1/32× | gamma_bulk | Al ₃₂ O ₄₈ | R | 5.7 |
| | 1/80× | gamma_surf-110b_from_boehm_00w | Al ₈₀ O ₁₂₀ | J | 2.0 |
| | | | | B | 3.6 |
| SUR | -1/32× | gamma_bulk | Al ₃₂ O ₄₈ | R | 7.7 |
| | 1/64× | gamma_surf-110b_from_bulk_00w | Al ₆₄ O ₉₆ | J | 5.7 |
| | | | | B | 7.3 |
| SUR | -1/32× | gamma_bulk | Al ₃₂ O ₄₈ | R | 9.3 |
| | 1/48× | gamma_surf-1101_A1_00w | Al ₄₈ O ₇₂ | J | 1.4 |
| | | | | B | 7.9 |
| SUR | -1/32× | gamma_bulk | Al ₃₂ O ₄₈ | R | 7.4 |
| | 1/56× | gamma_surf-1101_A2_00w | Al ₅₆ O ₈₄ | J | 4.3 |
| | | | | B | 6.6 |
| SUR | -1/32× | gamma_bulk | Al ₃₂ O ₄₈ | R | 7.2 |
| | 1/56× | gamma_surf-1101_L2_00w | Al ₅₆ O ₈₄ | J | 1.1 |
| | | | | B | 5.2 |

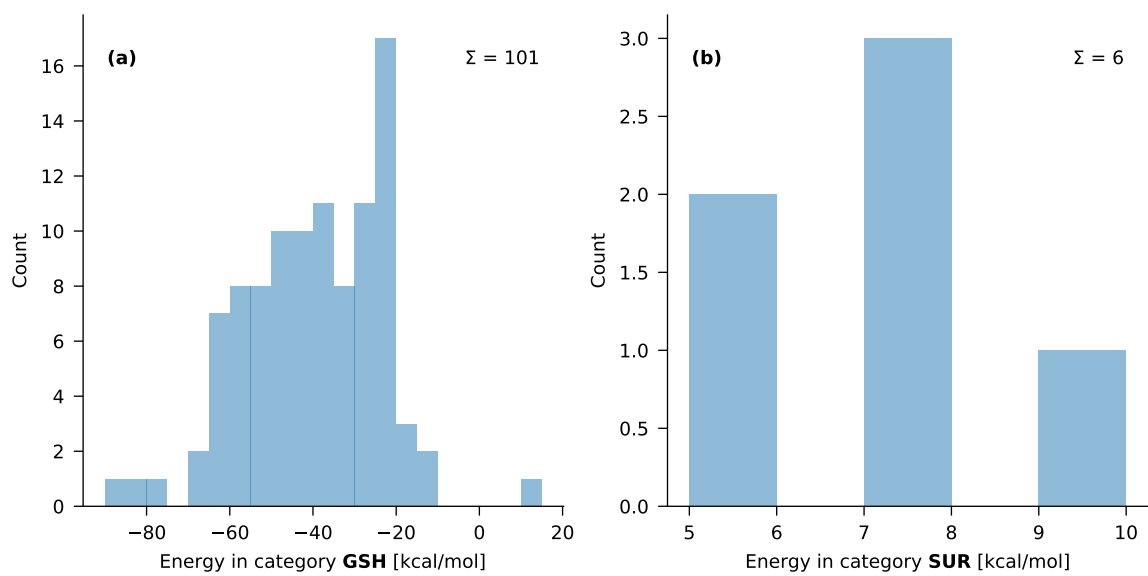


Figure S7: Histograms of reaction energies in the validation set.

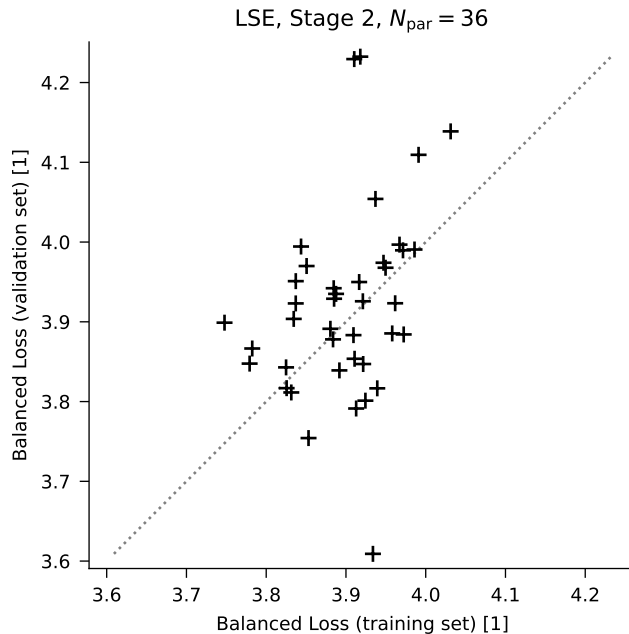


Figure S8: Parity plot comparing the value of the Balanced Loss function, for the training and validation sets, for the best solution from all 40 CMA runs (LSE form, optimization stage 2, see main text for details). Despite that the 40 optimized parameter vectors different significantly, their performance for the training and validation sets is similar.

Table S6: Comparison of root-mean-square-errors (RMSEs) of the initial force field by Joshi *et al.*¹ and the force field optimized in this work (BL). The RMSEs are computed for categories of training data and structures for which there is no counterpart in the validation set, i.e. not related to γ -Al₂O₃ surfaces. For reference, the standard deviation (SD) on the reference data is included.

| Category | Unit | Ref. SD | Joshi RMSE | BL RMSE | # |
|------------|------------------------|------------|---------------|------------|------|
| Al-O | Å | 0.08 | 0.35 | 0.07 | 1482 |
| Al-Al | Å | 0.37 | 0.33 | 0.07 | 1624 |
| O-H | Å | 0.03 | 0.26 | 0.04 | 638 |
| O...H | Å | 0.15 | 0.42 | 0.12 | 468 |
| Al-O-Al | deg | 22.0 | 9.3 | 3.3 | 1646 |
| Al-O-H | deg | 7.6 | 14.2 | 5.1 | 1077 |
| O-Al-O | deg | 32.2 | 11.0 | 3.2 | 3522 |
| Al-O...H | deg | 8.6 | 12.8 | 4.9 | 803 |
| H-O...H | deg | 15.8 | 18.2 | 8.1 | 484 |
| H...O...H | deg | 21.5 | 15.1 | 7.1 | 72 |
| BSH | kcal mol ⁻¹ | 11.8 | 38.1 | 4.7 | 5 |
| GEH | kcal mol ⁻¹ | 14.6 | 84.5 | 1.9 | 21 |
| FOR | kcal mol ⁻¹ | 10.3 | 22.4 | 2.4 | 11 |

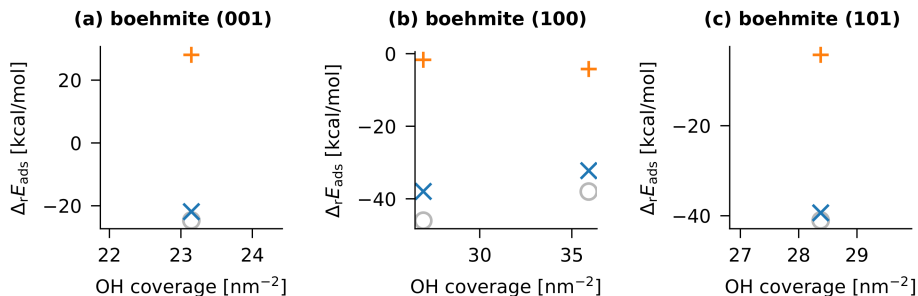


Figure S9: Adsorption energies, as defined in Eq. (3) in the main text, on boehmite surfaces as a function of the OH coverage, computed with different models: DFT (gray circle), Joshi (orange plus), this work (blue cross). All structures are derived from the training set.

S2 Comparison of Non-Equilibrium ReaxFF and DFT energies

A hydrated alumina slab (structure `gamma_surf-1101_A1_06w` from the validation set) was used as a starting point for a constant-temperature molecular dynamics (MD) run with VASP using the same level of theory as the training set. A Nosé-Hoover thermostat with a temperature of 1000 K and a relaxation time of 40 fs was used to stimulate the desorption of water from the surface. The simulation ran for 500 steps of 1 fs, and snapshots were taken every 10 steps for further analysis. Two desorption events occurs during the first 200 fs, after which no additional water molecules desorb.

The total single-point energies, E_{total} , of the selected snapshots are compared in Fig. S10(a). The DFT energies are shown in black and ReaxFF results with the new parameters in blue. ReaxFF energies obtained with the parameters of Joshi *et al.*¹ are depicted in red. The average is subtracted from both time series because these energies are only comparable up to a constant. While the fluctuations in DFT and ReaxFF energy are correlated, they also show significant deviations. This is expected, since the majority of the atoms are in the slab, not the water, and our training set emphasizes hydration reactions, not the vibrations in the alumina slab.

To show that our ReaxFF parameters can reproduce the relevant contribution to the DFT reference energy, additional single-point energies were calculated on the same snapshots from which some atoms were removed:

- E_{slab} is the energy of the slab and the water molecules that remain adsorbed, but without the atoms of the two water molecules that desorb during the simulation.
- $E_{2\text{H}_2\text{O}}$ is the energy computed for only the atoms of the two desorbing water molecules.

The energies of these two complementary subsystems are shown in Fig. S10(b) and Fig. S10(c), respectively. (Since these are also absolute energies, the average is again subtracted in both plots.) Finally, using these data, also the energy difference $(E_{\text{total}} - E_{\text{slab}} - E_{2\text{H}_2\text{O}})/2$ was computed and is shown in Fig. S10(d). In this case, no average was subtracted because the difference in energy has a chemically meaningful reference.

Fig. S10(b) shows deviations between ReaxFF and DFT energies that are very similar to those in Fig. S10(a), confirming that these deviations are due to the internal energy of the alumina slab. The results in Fig. S10(c) and Fig. S10(d) show a fair agreement between the DFT and ReaxFF energies obtained with our new parameters. Mainly Fig. S10(d) is of interest, because it is closely related to the hydration energies in the training set: This energy difference is calculated similarly to energy training data (Eq. (3) in the main text) but is now evaluated using non-equilibrium snapshots instead of optimized geometries. Because of this similarity, it is reasonable to expect a correspondence of the energies in Fig. S10(d). The average of the relative error over the first 200 fs is 25 % for the Balanced Loss parameters, which is comparable to relative errors on adsorption energies in the training set. For Fig. S10(a) and Fig. S10(b), however, a similar agreement would have been coincidental, since no related data were used for training. Also note that the new ReaxFF parameters show a clear improvement in Fig. S10(d) with respect to the parameters of Joshi *et al.*,¹ for which the relative error is 79 %. In Fig. S10(c) both ReaxFF parameterizations yield the same results because the parameters for water were not refitted.

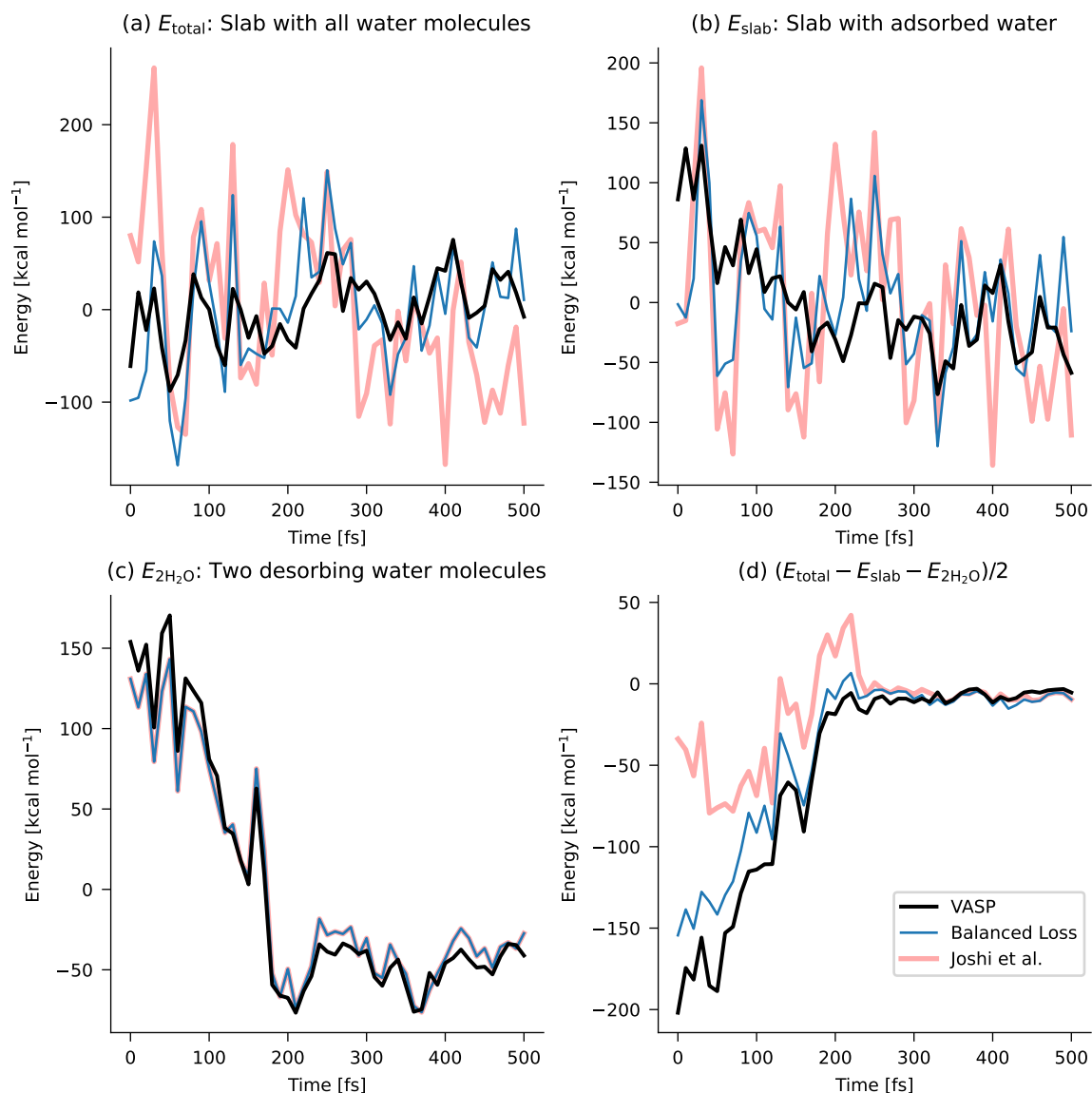


Figure S10: Comparison of DFT and ReaxFF energies for snapshots taken from a 1000 K DFT molecular dynamics simulation. (black: DFT, blue: ReaxFF with Balanced Loss parameters, red: ReaxFF with Joshi *et al.* parameters¹). Panel (a) contains the total energy of the system, panel (b) the energy of the slab and the water molecules that remain adsorbed, and panel (c) the energy of only the desorbing water molecules. Panel (d) depicts the instantaneous hydration energy as computed with Eq. (3) in the main text.

References

- (1) Joshi, K. L.; Psfogiannakis, G.; Duin, A. C. T. v.; Raman, S. Reactive molecular simulations of protonation of water clusters and depletion of acidity in H-ZSM-5 zeolite. *Phys. Chem. Chem. Phys.* **2014**, *16*, 18433–18441.