Unraveling the Nature of Extraframework Catalytic Ensembles in Zeolites: Flexibility and Dynamics of the Copper-Oxo Trimers in Mordenite

Elena V. Khramenkova¹, Michael G. Medvedev², Guanna Li³,⁴, Evgeny A. Pidko¹*

¹Inorganic Systems Engineering (ISE), Department of Chemical Engineering, Delft University of Technology, 2629 HZ Delft (The Netherlands)
²Zelinsky Institute of Organic Chemistry RAS, Leninsky Prospect, 47, Moscow (Russia)
³Biobased Chemistry & Technology, Wageningen University & Research, 6708 PB Wageningen (The Netherlands)
⁴Organic Chemistry, Wageningen University & Research, 6708 PB Wageningen (The Netherlands)

Corresponding author: Evgeny A. Pidko (e.a.pidko@tudelft.nl)
Contents

1. Computational details ........................................................................................................................ S3
   *Ab Initio* Molecular Dynamics (aiMD) ............................................................................................. S3
   Clustering .............................................................................................................................................. S3
   *Ab initio* Thermodynamics Analysis (aiTD) .................................................................................. S4

2. K-medoids clustering: silhouette score, clusters’ visualization ........................................................ S6
   Figure S1. The silhouette score for the number of clusters varied between 20 and 100 obtained for a different set of trajectories. ................................................................................................................... S6
   Figure S2a. Visualizing 44 clusters in LMMMD@923 K trajectory using dimensionality reduction methods such as a. t-SNE and b. PCA. .............................................................................................................. S7
   Figure S2b. Visualizing 36 clusters in LMMMD@723 K trajectory using dimensionality reduction methods such as a. t-SNE and b. PCA. .............................................................................................................. S7
   Figure S2c. Visualizing 66 clusters in MD@923 K trajectory using dimensionality reduction methods such as a. t-SNE and b. PCA. ............................................................................................................... S8
   Figure S2d. Visualizing 125 clusters of combined sets of trajectories using dimensionality reduction methods such as a. t-SNE and b. PCA. ........................................................................................................ S8

3. DBSCAN clustering .......................................................................................................................... S9
   Figure S3. Visualizing 47 clusters using dimensionality reduction methods: a. t-SNE and b. PCA. .......................................................................................................................... S9
   Table S1. Second-gen clusters ............................................................................................................. S10
   Figure S4. The correlation between active sites optimization at PBE level of theory and their subsequent single point calculation at PBE0 level of theory. ......................................................... S11
   Figure S5. The correlation of the individual Hirshfeld spin on O (a-c) and Cu (d-f) atoms in Cu$_3$O$_{32}^{2+}$ configurations calculated at the PBE-D3(BJ) level of theory with their respective energy calculated at the PBE0-D3(BJ) level of theory. ......................................................................... S12
   Figure S6. The correlation of the individual Hirshfeld charges on O (a-c) and Cu (d-f) atoms in Cu$_3$O$_{32}^{2+}$ configurations calculated at the PBE-D3(BJ) level of theory with their respective energy calculated at the PBE0-D3(BJ) level of theory................................................................................................................ S13
   Figure S7. a. RMSD of the Cu$_3$O$_2^{2+}$ in MOR. b. The active sites configurations corresponding to the highest RMSD deviation. ...................................................................................................... S13

4. Methane activation .......................................................................................................................... S14
   Figure S8. The reaction pathway of the methane activation over one of the lowest lying active sites from family I. ...................................................................................................................................... S14
   Figure S9. The reaction pathways representing the rearrangements from one type of configuration (family I) to other type of configuration (family IV). ............................................................................. S14

5. Supplementary references: .............................................................................................................. S15
1. Computational details

*Ab Initio* Molecular Dynamics (aiMD)

AiMD was used to sample the configurational space of the extraframework clusters confined inside the zeolite micropores. All calculations were carried out using a fully periodic zeolite structure represented by a supercell. A periodic supercell model of MOR was constructed by expanding the elementary cell by one translation along the c-axis and optimized giving the lattice parameters: \(a = b = 13.648\), \(c = 15.015\) Å. The two lattice Si ions in the side-pocket 8-MR channel were substituted with Al to give the Si/Al ratio of 23. The resulting framework negative charge is compensated by the \([\text{Cu}_3(\mu-\text{O})_3]^{2+}\) cluster, for which we presumed the triplet spin state \((s = 3/2)\), which was found to be lower in energy than the low spin state \((s = 1/2)\). Previously it was shown that the energy difference between high and low spin states for this structure is small.1

aiMD simulations were carried out using CP2K software package with the Quickstep module and orbital transformation for faster convergence.2–6 All calculations were carried out in the framework of density functional theory (DFT) using the gradient-corrected PBE-D3(BJ) functional7 with the empirical dispersion correction by Grimme.8–10 The Goedecker-Teter-Hutter pseudopotentials with a combination of periodic Poisson solver were used to calculate the electron repulsion integrals.11–13 A Gaussian basis set DZVP-MOLOPT-SR-GTH was used for Si, O, Al atoms while the TZV2P-MOLOPT-SR-GTH basis set was used for Cu.14 The projected augmented wave (PAW) method with a cutoff energy of 450 Ry was used.15 The relative cutoff was set to 30 Ry. The SCF convergence criterion was set to 10E-5 a.u. Each aiMD trajectory was integrated for 4.25 ps within the canonical ensemble by solving the classical equation of motion with a time step of 0.5 fs, giving 6 MD runs of each type and 18 in total.16–18 aiMD trajectories were simulated at the temperatures of 723 K (**LMMD@723 K**) and 923 K (**LMMD@923 K, MD@923 K**). To improve the configurational sampling, the velocity softening low mode-guided aiMD simulations were implemented as described here.19,20 These trajectories were denoted as LMMD (**LMMD@923 K, LMMD@723 K**). In these simulations, the velocities were directed along the low curvatures for the first 500 steps and then the standard aiMD procedure was activated.

Clustering

A clustering algorithm was employed to determine the structurally distinct configurations in the aiMD trajectories. The clustering process was performed using an unsupervised machine learning algorithm – k-medoids – which is widely applied for clustering MD trajectories.21,22 The algorithm is based on the procedure of minimizing a sum of pairwise dissimilarities between structures and their medoids (the centers of the cluster) as implemented in scikit-learn.23 When k-medoids are generated, k-medoids
algorithm refines them by computing the distance between the centers of clustering and the clusters belonging to them. Every 50th step of each trajectory was included in the clustering procedure. The silhouette index was chosen as a validation measure to establish the object similarity to the other members of the cluster. The distance metric for both clustering method and validation measures was set to “Euclidean” distance which measures the square distance between vectors. For visualization purposes, the clusters were projected into a bidimensional Euclidean space using t-SNE algorithm.\textsuperscript{24–26} We have applied clustering in the subsets of trajectories separately, but the clustering of all the trajectories together was performed but not considered further as it revealed a low silhouette score (see Figure S1d).

Next, the geometries of the cluster’s centers were optimized to local minima using the Quasi-Newton (Broyden-Fletcher-Goldfarb-Shanno) algorithm as implemented in CP2K using the same parameters for the electronic structure calculations as described above. The hybrid functional PBE0 was used to calculate the single point energy of configurations optimized at the PBE level of theory. The auxiliary density matrix method (ADMM) was utilized to reduce the costs for approximate calculations of non-local exchange energy. Vibrational analysis on the optimized structures was carried out using the finite difference method (0.02 Å atomic displacement).

To facilitate further analysis, the centers of the optimized clusters were further grouped using the density-based spatial clustering of applications with noise (DBSCAN)\textsuperscript{27,28} as implemented in scikit-learn.\textsuperscript{23} DBSCAN method is able to detect the core samples of high density and expand clusters from them. The optimal value of $\text{eps}$ parameter, which specifies the proximity of the points to become a part of the cluster, was set to 0.4. The value was chosen based on the k-nearest neighbor of each point. The minimum number of points in a cluster equals 1.

**Ab initio Thermodynamics Analysis (aiTD)**

aiTD was used to assess the stabilities of the different active site configurations under the model conditions of the catalyst activation and catalytic reaction. aiTD enables computing the 2 – or 3 – dimensional dependencies of the Gibbs free energies on the state of the gaseous components in the catalyst’s environment in terms of the T- and P-dependencies.\textsuperscript{29,30} The stability of the extraframework Cu trimers was assessed with the reference to the Cu$_3$(µ$_3$-O)$_2$$^{2+}$ in the side pocket of mordenite. Such a reference state represents the regeneration of the Cu$_3$(µ-O)$_3$$^{2+}$ active species under O$_2$ activation. The following equilibrium was considered

$$\text{Cu}_3\text{O}^{3+}/\text{MOR} + \text{O}_2 \leftrightarrow \text{Cu}_3\text{O}_3^{2+}/\text{MOR}$$

(S1)
where $\text{Cu}_3\text{O}_3^{2+}$ is one of the configurations obtained from aiMD and $\text{Cu}_3\text{O}^{2+}$ is a reduced isomer. $O_2$ is the molecular oxygen. The reaction Gibbs free energy $\Delta G_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}}$ for equilibrium (S1) is defined as follows:

$$
\Delta G_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}} = G_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}}^S - G_{\text{Cu}_3\text{O}^{3+}/\text{MOR}}^S - 2\mu^g_O + \Delta ZPE_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}} \quad (S2)
$$

Gibbs free energies of solids are approximated as their respective electronic DFT-energies calculated at the PBE0-D3(BJ) level of theory, with the vibrational contribution $\Delta ZPE_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}}$ calculated at PBE level. The T- and P-condition dependencies are expressed via the chemical potential of oxygen $\mu^g_O$ which is defined as:

$$
\mu^g_O(T, p) = \frac{1}{2} E_{O_2} + \Delta \mu^g_O(T, p) \quad (S3)
$$

The $E_{O_2}$ is the electronic energy of molecular oxygen. The change in chemical potential $\Delta \mu^g_O(T, p)$ can be expressed as:

$$
\Delta \mu^g_O(T, p) = \Delta \mu_O(T, p^0) + \frac{1}{2} RT \left( \frac{p_{O_2}}{p_{O_2}^0} \right) = \\
\frac{1}{2} \left[ H(T, p^0, O_2) - H(0K, p^0, O_2) - T(S(T, p^0, O_2) - S(0K, p^0, O_2)) + RT \ln \left( \frac{p_{O_2}}{p_{O_2}^0} \right) \right] \quad (S4)
$$

The zero-point energy was defined standardly. We have only considered the contributions of the extraframework Cu-oxo atoms.

Combining (S2),(S3) with the ZPE we arrive at the following:

$$
\Delta G_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}} = E_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}} - E_{\text{Cu}_3\text{O}^{3+}/\text{MOR}} - E_{O_2} - 2\Delta \mu^g_O(T, p) + \Delta ZPE_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}} \quad (S5)
$$

where $E_{\text{Cu}_3\text{O}_3^{2+}/\text{MOR}}, E_{\text{Cu}_3\text{O}^{3+}/\text{MOR}}$ are electronic energies of the $\text{Cu}_3\text{O}_3^{2+}$ structural isomer and $\text{Cu}_3\text{O}^{2+}$ reduced isomer respectively.
2. **K-medoids clustering: silhouette score, clusters’ visualization**

To determine the optimal number of clusters for each set of trajectories, we have used the silhouette score. The number of tested clusters varied between 20 and 100. The value of the silhouette score for each number of clusters is shown in **Figure S1**, where LMMD@723 K, LMMD@923 K, MD@923 K set of trajectories are clustered separately and together and have their own number of desirable k-medoids.

![Figure S1](image)

**Figure S1.** The silhouette score for the number of clusters varied between 20 and 100 obtained for a different set of trajectories.

a. The silhouette score for **LMMD@923 K**. The maximum value of the silhouette score is 0.41 and corresponds to 44 clusters.  
b. The silhouette score for **MD@923 K** set of trajectories. The maximum value of the silhouette score is 0.31 and corresponds to 66 clusters.  
c. The silhouette score for **LMMD@723 K** set of trajectories. The maximum value of the silhouette score is 0.34 and corresponds to 36 clusters.  
d. The silhouette score for all sets of trajectories clustered together. The maximum value of the silhouette score is 0.29 and corresponds to 125 clusters.

Clustering was carried out using k-medoids algorithm. It starts with selecting the number of centers using the value of the highest silhouette score. As the optimal clustering for the defined number of clusters is performed (**Figure S1**), dimensionality reduction algorithms are applied for visualization purposes. We have performed both PCA (Principal component analysis) and t-SNE (t-distributed stochastic neighbor embedding) dimensionality reduction methods. However, for all sets of trajectories, t-SNE distinguishes the data in a better way compared to PCA.
**Figure S2a.** Visualizing 44 clusters in LMMD@923 K trajectory using dimensionality reduction methods such as a. t-SNE and b. PCA.

**Figure S2b.** Visualizing 36 clusters in LMMD@723 K trajectory using dimensionality reduction methods such as a. t-SNE and b. PCA.
Figure S2c. Visualizing 66 clusters in MD@923 K trajectory using dimensionality reduction methods such as a. t-SNE and b. PCA.

Figure S2d. Visualizing 125 clusters of combined sets of trajectories using dimensionality reduction methods such as a. t-SNE and b. PCA.
3. DBSCAN clustering

After optimization as the PBE-D3(BJ) level of theory, the centres of the clusters (first-gen clusters) were clustered using the DBSCAN method. The change in the clustering algorithm was motivated by the low silhouette score value for the k-medoids method and better cluster separation in the DBSCAN algorithm.

Figure S3. Visualizing 47 clusters using dimensionality reduction methods: a. t-SNE and b. PCA.
Table S1. Second-gen clusters.

Their numbers, geometrical configurations of the lowest-lying structure optimized at the PBE level of theory, energies range (relative to the initial structure) obtained at the PBE0 level of theory, and the number of frames from each type of MD which were attributed to each cluster.

| № | Clusters | Energy conf.; Energy range (kJ/mol) | Conf. from LMMD@923K: | Conf. from LMMD@723K: | Conf. from MD@923K: | № | Clusters | Energy conf.; Energy range (kJ/mol) | Conf. from LMMD@923K: | Conf. from LMMD@723K: | Conf. from MD@923K: |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 17 | [−137; −78] 324:0:4 | 17 | [−18] 0:0:55 | 33 | [−59; 38] 0:760:731 |
| 2 | 18 | −124 0:0:5 | 18 | [−7] 85:0:0 | 34 | [−40] 0:7:0 |
| 3 | 19 | [−124; −123] 13:0:0 | 19 | 0 0:0:73 | 35 | [0; 30] 0:0:5 |
| 4 | 20 | −101 6:0:0 | 20 | 5 0:0:3 | 36 | 2 0:15:0 |
| 5 | 21 | [−89; 3] 134:0:0 | 21 | 14 36:0:0 | 37 | [14; 73] 0:36:70 |
| 6 | 22 | −64 1:0:0 | 22 | 17 66:0:0 | 38 | 15 0:0:4 |
| 7 | 23 | [−48; −19] 16:0:0 | 23 | 18 0:0:5 | 39 | 20 0:0:3 |
| 8 | 24 | [−29; −12] 74:0:0 | 24 | 18 1:0:0 | 40 | 31 0:0:27 |
| 9 | 25 | −8 0:0:8 | 25 | 22 0:0:2 | 41 | 43 0:9:0 |
To verify the results obtained at the PBE level of theory, the final configurations were computed with the PBE0 hybrid functional. The recent implementation of the Auxiliary Density Matrix Method (ADMM) in CP2K allows to use the auxiliary basis sets to accelerate the HFX calculations. A primary basis set DZVP-MOLOPT-SR-GTH was used for the PBE calculations.

**Figure S4.** The correlation between active sites optimization at PBE level of theory and their subsequent single point calculation at PBE0 level of theory.
To identify the difference in the electronic structure properties of generated isomers, we have calculated the Hirshfeld spin states and charges on oxygens and copper atoms for every Cu$_3$O$_5^{2+}$ configuration obtained from the second-gen clusters.

**Figure S5.** The correlation of the individual Hirshfeld spin on O (a-c) and Cu (d-f) atoms in Cu$_3$O$_5^{2+}$ configurations calculated at the PBE-D3(BJ) level of theory with their respective energy calculated at the PBE0-D3(BJ) level of theory.

The energies are calculated with respect to the energy of the starting configuration. The colors correspond to the families of the active sites described in the paper.
**Figure S6.** The correlation of the individual Hirshfeld charges on O (a-c) and Cu (d-f) atoms in Cu₃O₃²⁺ configurations calculated at the PBE-D3(BJ) level of theory with their respective energy calculated at the PBE0-D3(BJ) level of theory.

The energies are calculated with respect to the energy of the starting configuration. The colors correspond to the families of the active sites described in the paper.

**Figure S7.** a. RMSD of the Cu₃O²⁺ in MOR. b. The active sites configurations corresponding to the highest RMSD deviation.
4. Methane activation

To investigate the reactivity of the active sites with a peroxo-ligand in the process of methane activation, we have calculated the reaction pathways for the activation of a single methane molecule over the lowest-lying site with a peroxo-ligand (Figure S8). It consists of the rearrangement of the active site and includes the cleavage of the O-O-bond and the activation of the C-H-bond in methane. In Figure S8, the presence of methane leads to the reorganization of the structure. It starts with the barrierless cleavage of the O-O-bond, forming of the adduct that is 129 kJ/mol less stable than the initial configuration (FS). The subsequent activation of methane over structure (TS) proceeds with a low barrier of 27 kJ/mol and directly forms methanol coordinated to one of the Cu atoms.

**Figure S8.** The reaction pathway of the methane activation over one of the lowest lying active sites from family I.

**Figure S9.** The reaction pathways representing the rearrangements from one type of configuration (family I) to other type of configuration (family IV).
5. Supplementary references:

(1) Li, G.; Vassilev, P.; Sanchez-Sanchez, M.; Lercher, J. A.; Hensen, E. J. M.; Pidko, E. A. Stability and Reactivity of Copper Oxo-Clusters in ZSM-5 Zeolite for Selective Methane Oxidation to Methanol. *J. Catal.* **2016**, *338*, 305–312, DOI: 10.1016/j.jcat.2016.03.014.

(2) Schütt, O.; Messmer, P.; Hutter, J.; VandeVondele, J. Gpu-Accelerated Sparse Matrix-Matrix Multiplication for Linear Scaling Density Functional Theory. *Electron. Struct. Calc. Graph. Process. Units* **2016**, *173–190*, DOI: 10.1002/9781118670712.ch8.

(3) Borštnik, U.; VandeVondele, J.; Weber, V.; Hutter, J. Sparse Matrix Multiplication: The Distributed Block-Compressed Sparse Row Library. *Parallel Comput.* **2014**, *40* (5–6), 47–58, DOI: 10.1016/j.parco.2014.03.012.

(4) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. Cp2k: Atomistic Simulations of Condensed Matter Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (1), 15–25, DOI: 10.1002/wcms.1159.

(5) Frigo, M.; Johnson, S. G. The Design and Implementation of FFTW3. *Proc. IEEE* **2005**, *93* (2), 216–231, DOI: 10.1109/JPROC.2004.840301.

(6) VandeVondele, J.; Hutter, J. An Efficient Orbital Transformation Method for Electronic Structure Calculations. *J. Chem. Phys.* **2003**, *118* (10), 4365–4369, DOI: 10.1063/1.1543154.

(7) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865, DOI: 10.1103/PhysRevLett.77.3865.

(8) 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* (15), DOI: 10.1063/1.3382344.

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

(10) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Commun.* **2005**, *167* (2), 103–128, DOI: 10.1016/j.cpc.2004.12.014.

(11) Krack, M. Pseudopotentials for H to Kr Optimized for Gradient-Corrected Exchange-Correlation Functionals. *Theor. Chem. Acc.* **2005**, *114* (1–3), 145–152, DOI: 10.1007/s00214-005-0655-y.

(12) Hartwigs, C.; Goedecker, S.; Hutter, J. Relativistic Separable Dual-Space Gaussian Pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58* (7), 3641, DOI: 10.1103/PhysRevB.58.3641.

(13) Goedecker, S.; Teter, M. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B - Condens. Matter Mater. Phys.* **1996**, *54* (3), 1703–1710, DOI: 10.1103/PhysRevB.54.1703.

(14) VandeVondele, J.; Hutter, J. J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *J. Chem. Phys.* **2007**, *127* (11), 114105, DOI: 10.1063/1.2770708.

(15) Lippert, B. G.; PARRINELLO, J. H. and M. A Hybrid Gaussian and Plane Wave Density Functional Scheme. *Mol. Phys.* **1997**, *92* (3), 477–488, DOI: 10.1080/002689797170220.

(16) Kolafa, J. Time-reversible Always Stable Predictor–Corrector Method for Molecular Dynamics of
Polarizable Molecules. *J. Comput. Chem.* **2004**, *25* (3), 335–342, DOI: 10.1002/jcc.10385.

(17) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81* (1), 511–519, DOI: 10.1063/1.447334.

(18) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52* (2), 255–268, DOI: 10.1080/00268978400101201.

(19) Medvedev, M. G.; Zeifman, A. A.; Novikov, F. N.; Bushmarinov, I. S.; Stroganov, O. V.; Titov, I. Y.; Chilov, G. G.; Svitanko, I. V. Quantifying Possible Routes for SpnF-Catalyzed Formal Diels-Alder Cycloaddition. *J. Am. Chem. Soc.* **2017**, *139* (11), 3942–3945, DOI: 10.1021/jacs.6b13243.

(20) Medvedev, M. G.; Panova, M. V; Chilov, G. G.; Bushmarinov, I. S.; Novikov, F. N.; Stroganov, O. V; Zeifman, A. A.; Svitanko, I. V. Exhaustive Conformational Search for Transition States: The Case of Catechol O-Methyltransferase Active Site. *Mendeleev Commun.* **2017**, *27* (3), 224–227, DOI: 10.1016/j.mencom.2017.05.002.

(21) Park, H.-S.; Jun, C.-H. A Simple and Fast Algorithm for K-Medoids Clustering. *Expert Syst. Appl.* **2009**, *36* (2), 3336–3341, DOI: 10.1016/j.eswa.2008.01.039.

(22) Maranzana, F. E. On the Location of Supply Points to Minimize Transportation Costs. *IBM Syst. J.* **1963**, *2* (2), 129–135, DOI: 10.1057/jors.1964.47.

(23) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

(24) Van Der Maaten, L.; Hinton, G. Visualizing Data Using T-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2625.

(25) Krijthe, J. H.; Van der Maaten, L. Rtsne: T-Distributed Stochastic Neighbor Embedding Using Barnes-Hut Implementation. *R Packag. version 0.13*, https://github.com/jkrijthe/Rtsne (accessed Oct.10, 2015).

(26) Van Der Maaten, L. Accelerating T-SNE Using Tree-Based Algorithms. *J. Mach. Learn. Res.* **2014**, *15* (1), 3221–3245.

(27) Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Kdd*; 1996; Vol. 96, pp 226–231.

(28) Schubert, E.; Sander, J.; Ester, M.; Kriegel, H. P.; Xu, X. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Trans. Database Syst.* **2017**, *42* (3), 1–21, DOI: 10.1145/3068335.

(29) Grajciar, L.; Heard, C. J.; Bondarenko, A. A.; Polynski, M. V.; Meeprasert, J.; Pidko, E. A.; Nachtgall, P. Towards Operando Computational Modeling in Heterogeneous Catalysis. *Chem. Soc. Rev.* **2018**, *47* (22), 8307–8348, DOI: 10.1039/c8cs00398j.

(30) Reuter, K.; Stampf, C.; Scheffler, M. AB Initio Atomistic Thermodynamics and Statistical Mechanics of Surface Properties and Functions. In *Handbook of Materials Modeling*; Yip, S., Ed.; Springer Netherlands: Dordrecht, 2005; pp 149–194, DOI: 10.1007/978-1-4020-3286-8_10.