PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2020, **22**, 18902

Neural network assisted analysis of bimetallic nanocatalysts using X-ray absorption near edge structure spectroscopy†

Nicholas Marcella, *[a] Yang Liu,[b] Janis Timoshenko, [a] Erjia Guan,[a] Mathilde Luneau,[c] Tanya Shirman, [de] Anna M. Plonka,[a] Jessi E. S. van der Hoeven,[cd] Joanna Aizenberg, [cde] Cynthia M. Friend [cd] and Anatoly I. Frenkel *[af]

X-ray absorption spectroscopy is a common method for probing the local structure of nanocatalysts. One portion of the X-ray absorption spectrum, the X-ray absorption near edge structure (XANES) is a useful alternative to the commonly used extended X-ray absorption fine structure (EXAFS) for probing three-dimensional geometry around each type of atomic species, especially in those cases when the EXAFS data quality is limited by harsh reaction conditions and low metal loading. A methodology for quantitative determination of bimetallic architectures from their XANES spectra is currently lacking. We have developed a method, based on the artificial neural network, trained on *ab initio* site-specific XANES calculations, that enables accurate and rapid reconstruction of the structural descriptors (partial coordination numbers) from the experimental XANES data. We demonstrate the utility of this method on the example of a series of PdAu bimetallic nanoalloys. By validating the neural network-yielded metal–metal coordination numbers based on the XANES analysis by previous EXAFS characterization, we obtained new results for *in situ* restructuring of dilute (2.6 at% Pd in Au) PdAu nanoparticles, driven by their gas and temperature treatments.

Received 18th April 2020,
Accepted 1st May 2020

DOI: 10.1039/d0cp02098b

rsc.li/pccp

Introduction

In bimetallic alloys, different mixing motifs of atomic species (homogeneous or heterogeneous, random or non-random in the case of the former, or various types or segregation in the case of the latter) are known to strongly influence the electronic properties of the surface atoms and hence a material's catalytic, optical, magnetic, and electronic properties.[1–5] Furthermore, just as the size and shape[6,7] and the degree of structural order[8] can change dynamically in reaction conditions, so can the compositions of nanoparticles[9–11] in a complex relationship with other material properties. It is for the latter reason that, in order to accurately measure those dynamic changes, measurements should be taken under *in situ* conditions.[12] Extended X-ray absorption fine structure (EXAFS) has long been a preferred method for studies of bimetallic nanocatalysts due to its ability to monitor structural changes through the measurements of partial coordination numbers (C_{A-A}, C_{A-B}, C_{B-A} and C_{B-B} in an A_xB_{1-x} bimetallic material), bond lengths, and their disorders in a broad range of *in situ* and *operando* conditions.[13] However, the accuracy of EXAFS analysis in metal catalysts has limitations in many cases. For example, at low weight loadings of the X-ray absorbing atoms and in the presence of other factors affecting the data quality (such as low atomic numbers of the absorbing elements, high temperatures, strongly absorbing reactor walls, supports and/or solvents), the low signal to noise ratio of EXAFS spectra may hinder their deciphering by the universally used fitting methods. In addition, particularly in nanoscale systems such as nanocatalysts, the interfacial effects (catalyst-adsorbate, catalyst-support) result in a significant, detectable, asymmetry in bond length distributions which introduces artifacts in the EXAFS fitting procedure.[14–16] Another section of the X-ray absorption coefficient spectrum, the X-ray absorption near edge structure (XANES), has been used, until recently, only qualitatively or semi-quantitatively for structural refinement of catalyst data due to the lack of analytical methods for structural refinement – such as the EXAFS equation that is used for fitting EXAFS spectra. While a true quantitative fitting procedure has not

[a] *Department of Materials Science and Chemical Engineering, Stony Brook University, Stony Brook, New York 11794, USA. E-mail: nicholas.marcella@stonybrook.edu, anatoly.frenkel@stonybrook.edu*
[b] *Department of Chemistry, Stony Brook University, Stony Brook, New York 11794, USA*
[c] *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA*
[d] *John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*
[e] *Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, Massachusetts 02138, USA*
[f] *Chemistry Division, Brookhaven National Laboratory, Upton, New York 11973, USA*
† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0cp02098b

yet been realized, several computational approaches were developed over the last several decades for the quantitative modeling of the XANES spectra, using the atomic coordinates of appropriate model structures.[17–19] This approach, however, suffers from its model-dependency ("given the structure, calculate the spectrum"), which limits the available XANES fitting approaches (*e.g.* MXAN code and FitIt)[20,21] to cases dealing with only a few degrees of freedom.

Recently, we have developed two approaches to help improve the analytical capability of XANES and EXAFS techniques for metal nanoparticle characterization using machine learning tools:[22] Neural Network (NN) assisted EXAFS analysis (NN-EXAFS),[22–24] and Neural Network assisted XANES analysis (NN-XANES).[25] NN-EXAFS can extract the partial radial distribution functions (RDF) for AA, AB, BA and BB pairs in a A_xB_{1-x} bimetallic material, which is not possible by conventional EXAFS analysis.[15,26–29] However, the consequences of harsh reaction conditions are quite severe for EXAFS data quality, especially in dilute alloys containing a few percent of one of the atoms, as mentioned above. They limit the applicability of this method to many types of catalytic studies.[30] NN-XANES, on another hand, can, potentially, be a preferred approach in those cases, due to the presence of intense features in the XANES region that are much less hampered by the thermal disorder and low weight loading, and can be acquired relatively quickly and with better signal-to-noise ratio than EXAFS data. Previously, NN-XANES was developed and validated for use with monometallic Pt,[25,31] Ag,[32,33] Cu,[32] and CuO_x size-selective cluster catalysts.[34] Therefore, NN-XANES was demonstrated to be a viable alternative to EXAFS for catalytic studies, but it has never been used for bimetallic nanomaterials.

In the remainder of this article we present a NN-XANES approach for use with bimetallic systems and demonstrate its utility on the example of the PdAu nanoalloy, an important catalytic system for the selective oxidation of methane, CO oxidation, and selective alkyne hydrogenation reactions.[35–40] We will describe the NN approach, present the details of the neural network training and testing, a demonstration of the utility of our method for *in situ* observation of restructuring in dilute alloy catalysts, followed by the discussion of the results, and, finally, conclusions.

Neural network-based approach to XANES data analysis

Here we use the same approach for the creation of a NN method for the analysis of bimetallic nanocatalysts that was used in our previously reported method for NN-XANES analysis in monometallic systems.[25] The main idea is to train a NN to learn the association between the XANES spectrum and the descriptors of the three-dimensional arrangement of nearest neighbors to the X-ray absorbing species. For bimetallic materials, the X-ray absorption edge of each type of atomic species can be measured, thus providing information on the first nearest neighbor pairs of four types: A–A, A–B, B–B, and B–A, through the partial coordination numbers (CNs) of the first nearest pairs are C_{A-A}, C_{A-B}, C_{B-A} and C_{B-B}. Therefore, our method for obtaining the CNs relies on two independent, "absorber-specific", NNs, each with "pair-specific" outputs, *i.e.*, C_{A-A}, C_{A-B} for absorber A, and C_{B-A} and C_{B-B} for absorber B (Fig. 1).

The first nearest neighbor partial CNs are particularly useful as they are directly related to the compositional motifs of bimetallic nanoparticles as well as the particle sizes and shapes.[13,41–44] For example, they can be used to directly extract the Cowley short range order parameter, for characterizing either mixing or segregation behaviors of components in the nanoalloy from the measured CN values.[13] To discover such a relationship and serve in a predictive capacity, the NN requires training on large sets of labeled data (for which the relationship between the structure – *i.e.*, the CNs – and the spectrum is known). The NN-based method is interpolative in nature, employing large numbers of learnable parameters (weights and biases), whose number (and, hence, complexity of the model) can be increased by increasing the number of NN layers and the number of nodes per layer. In the case of NN-XANES, which utilizes fully connected multilayer perceptron (MLP) layers and convolutional layers, the number of learnable parameters can

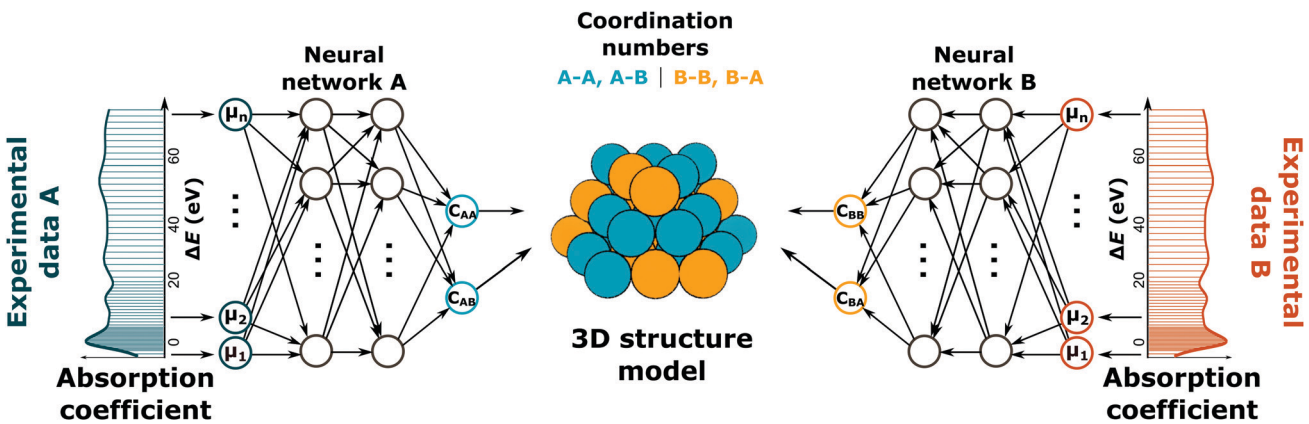

Fig. 1 A schematic that represents the application of NN-XANES to an A_xB_{1-x} bimetallic system. Partial first coordination numbers are extracted from the XANES of A and B absorbing components. The partial coordination numbers (A–A, A–B, B–B, and B–A) are used to deduce the average nanoparticle structure.

quickly reach the order of hundreds of thousands. It is therefore required that the training data set included hundreds of thousands of labeled spectra. Because it is very challenging to obtain a sufficient amount of labeled experimental data for this purpose, hence, we follow a previously developed approach[25] and train the networks on synthetic data (*i.e.*, theoretical XANES spectra calculated *ab initio* using FEFF9[17] code). The use of NNs for local structure predictions from experimental XANES data after training on purely theoretical data has been shown to work in our previous NN-XANES and NN-EXAFS examples, as well as in recent applications of the NN approach to scanning transmission electron microscopy (STEM) data and nuclear magnetic resonance (NMR) spectroscopy.[22,34,45–47] Training, for these case, is the process by which the cost function is minimized by refining the NN weights and biases, where the cost function is defined as the mean squared deviation of the NN outputs from the corresponding "target" values (*i.e.*, true values of CNs that are known for the training data).[22]

To optimize NN hyperparameters, such as number of nodes, layers, learning rate, regularization parameters and number of training iterations, we consider also the cost function for a validation dataset, which is analogous to the cost function for the training dataset, but is calculated for examples that are not directly used in the optimization of NN weights and biases. In all previous works, the validation cost function was calculated using theoretical examples that were excluded from the training data set. For bimetallic NPs, however, we found that approach to be insufficiently accurate, due to systematic differences between theoretically simulated XANES spectra and actual experimental data. In the new approach described here, we use an experimental validation set, which is composed of the data for which we have a good knowledge of the corresponding sample's structure *via* EXAFS. Information regarding the experimental data sets are listed in Table 1. We elaborate on the data sets, and how they were used for validation and testing in the next section.

Table 1 Experimental data for the validation or testing of the absorber-specific NNs

	Sample characteristics		Synchrotron & conditions		Validation (V) or testing (T)	
Dataset	Pd at%	Support or surfactant	Facility	Gas[a]	Pd[b]	Au[c]
Peptides	83	R5-Peptide	APS	Air	V	V
Peptides	67	R5-Peptide	APS	Air	V	V
Peptides	50	R5-Peptide	APS	Air	V	V
Peptides	33	R5-Peptide	APS	Air	V	V
Peptides	25	R5-Peptide	APS	Air	V	V
RCT-1	9	RCT-SiO_2	SSRL	He	T	V
RCT-1	2	RCT-SiO_2	SSRL	He	T	V
RCT-2	9	RCT-SiO_2	NSLSII	He	T	—
RCT-2	4	RCT-SiO_2	NSLSII	He	T	—
RCT-2	25	RCT-SiO_2	NSLSII	He	T	—
RCT-2	25	RCT-SiO_2	NSLSII	H_2	T	—
TiO_2	24	TiO_2	APS	Air	—	T
TiO_2	15	TiO_2	APS	Air	—	T
TiO_2	12	TiO_2	APS	Air	—	T
TiO_2	5	TiO_2	APS	Air	—	T
TiO_2	4	TiO_2	APS	Air	—	T
TiO_2	3	TiO_2	APS	Air	—	T

[a] All data were taken at 25 °C. [b] Pd absorber-specific NN. [c] Au absorber-specific NN.

Because of our reliance on theoretical training data for NN training, the method requires that theoretical spectroscopy codes generating such training data provide good qualitative agreement between theory and experiment for reference materials with known structure, such as bulk standards. In previous NN-XANES works with monometallic (Pt, Ag or Cu) nanoparticles,[25,32] we assumed that FEFF9 simulation parameters, which were optimized to ensure the best possible match between the theoretical and experimental bulk Pt, Ag and Cu spectra, remain optimal also for the simulation of XANES spectra in their respective monometallic nanoparticles. In the present work, we similarly began by evaluating the agreement between FEFF9 simulation and experimental XANES data for bulk Pd and Au. We present the optimal calculated theoretical Pd K-edge and Au L_3-edge XANES spectra of bulk standards and the respective experimental counterparts in Fig. 2. The α, β, γ, and δ symbols mark specific features in the bulk Pd XANES (Fig. 2A), and h, i, j, and k mark features in the Au XANES (Fig. 2B). The observed type of agreement between theory and experiment (in the energy range up to 67 eV from the Pd K-edge and up to 76 eV from the Au L_3-edge) appears to be satisfactory for our method, as we will show in the validation and testing sections.

We also investigate the capability for FEFF calculations, using the optimized parameters, to capture major qualitative trends in alloyed NPs such as the size and concentration dependence of spectral features. To do this, we look for trends in our experimental data sets labeled "peptides" and "RCT-1" in Table 1, which are data used later in NN validation. The XANES from the Pd K-edge and Au L_3-edge is plotted in Fig. 3, labeled "Experiment", and described

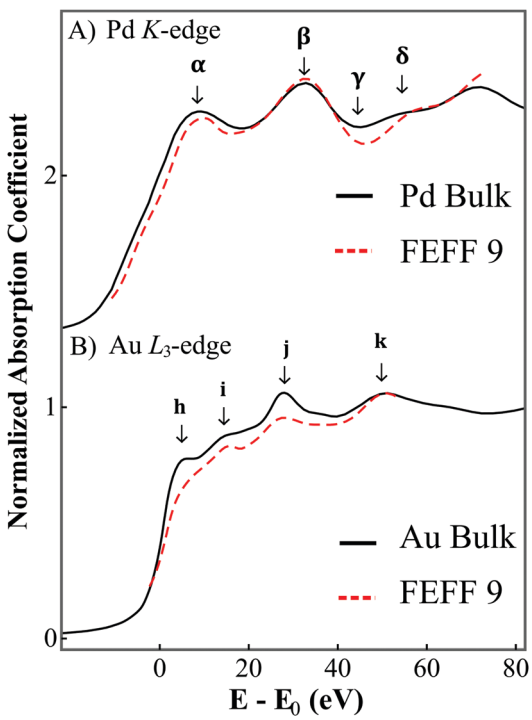

Fig. 2 Experimental and theoretical XANES of bulk Pd and Au at the (A) Pd K-edge and (B) Au L_3-edge. The theoretical spectra are calculated with FEFF9 code.

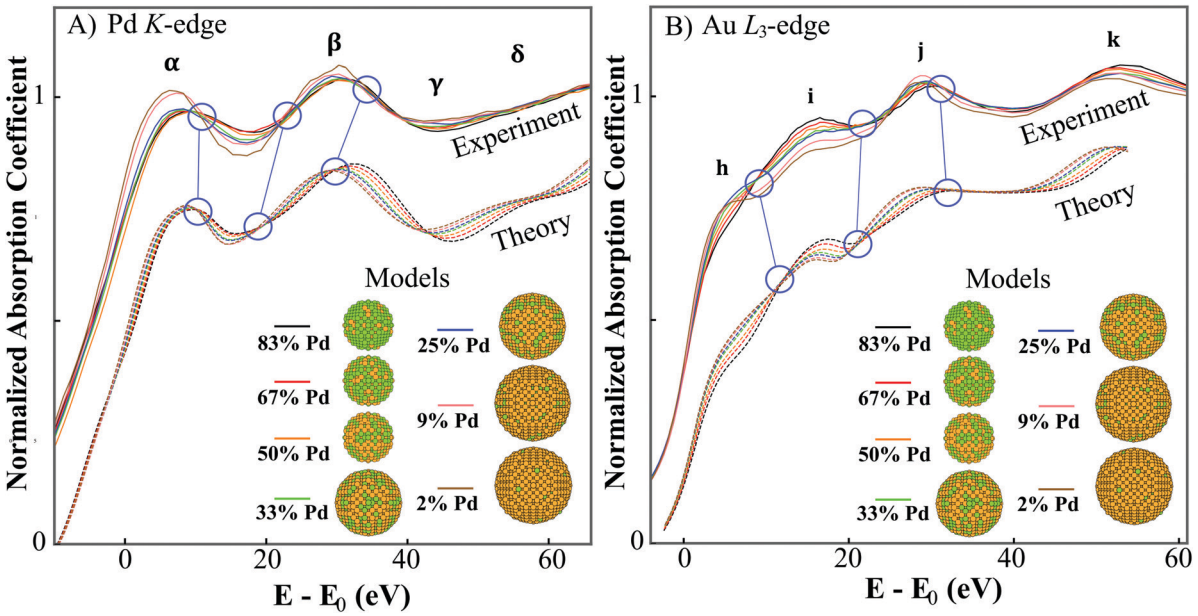

Fig. 3 Experimental and theoretical XANES of peptides and RCT-1 data sets at the (A) Pd K-edge and (B) Au L_3-edge. The theoretical spectra (for 83, 67, 50, 33, 25, 9, and 2 at% Pd) are calculated with FEFF9 code. The theoretical data in A and B were shifted vertically for clarity. Several connected circles are guides to the eye, illustrating similar trends between the theoretical and experimental spectra.

in Table S1, in the ESI,† where information obtained from transmission electron microscopy (TEM) and compositional measurements are listed.[23] Circles mark isosbestic points present due to changes in composition and size between the NPs. Using the experimental data as a reference, we then constructed 7 representative NPs where the theoretical XANES were simulated using basic structure models constructed based on the size and composition of the experimental systems. We approximated each particle as quasi-spherical with an average lattice constant determined by Vegard's law:

$$a_{\text{eff}} = x_{\text{Pd}}a_{\text{Pd}} + (1 - x_{\text{Pd}})a_{\text{Au}}, \quad (1)$$

where a_{eff} is the effective lattice constant, x_{Pd} is the concentration of Pd, a_{Pd} and a_{Au} are the lattice constants of Pd and Au, respectively. Including a varied lattice constant in the training set is important because XANES is sensitive not only to the CNs, but to the nearest neighbor distance as well.[32] For each particle model, the particle-average Pd or Au XANES spectra were calculated by averaging contributions independently calculated for all absorbing atoms of the same species (Pd or Au) in the model. The results of this calculation are shown in Fig. 3, labeled "Theory". While the theory does not reproduce the experimental data, evident by the differences in horizontal placement of the circles in Fig. 3 between the experiment and the theory, we do observe similar isosbestic points in the theory as in our experimental reference systems, which demonstrates that theory reproduces qualitative trends in compositional and size dependences. We will show, in the following sections, by validation and testing of the NN, that such a contrast in XANES produced by FEFF, due to size and composition dependence, is adequate for NN training.

After testing the feasibility of FEFF9 calculations for NN training, we created a set of theoretical training data for which site-specific XANES calculations at the Au L_3 and Pd K edges were made using the optimal parameters, listed in Note S1 in the ESI.† This dataset is herein referred to as the "site-specific training data". To create the site-specific training data, two distinct sets of atomistic models of PdAu NPs were created for calculations at the Pd K-edge (18 538 models) and Au L_3-edge (15 756 models). The total number of atoms, particle composition, and effective lattice constant in these models were variable, and thus randomly generated to ensure maximum training data diversity. The number of atoms in each particle model ranges from 10 to 185, the composition – from 1 to 90% Pd, and effective lattice constant – from 3.89 to 4.08 Å. The geometric templates used to create the models are the same as reported in ref. 25. In addition to the bimetallic NP models, 1200 monometallic Pd NPs and 2663 monometallic Au NPs, also with number of atoms ranging from 10 to 185 and effective lattice constants between 3.89 and 4.08 Å were constructed. Furthermore, to ensure that dilute Pd species are represented, we included the Pd site calculations (2859 dilute Pd sites) made for the 7 PdAu NP models mentioned in the last section and shown in Fig. 3. The first partial CNs for each absorbing atom were extracted from the atomistic coordinates of the 41 016 structure models, resulting in site-specific training data in the form of XANES-CNs associations. The extent of the diversity of the final site-specific training data set is very important for ensuring the NN is able to interpolate well. We examine the diversity in Fig. S1–S6 in the ESI.† We can see that the local compositions (*i.e.* composition of the absorbing site) varies between 0 and 100% Pd and all possible atoms locations are represented. To maximize diversity, and increase the size of our training data set, we trained the NNs on linear combinations of the site-specific training data. This approach, introduced in our previous works,[22,25] mimics the particle-averaging effect in experimental XANES data, and takes advantage of the fact that particle-average XANES $\mu(E)$ and coordination numbers

C_{AA} and C_{AB} are linear combinations of the XANES spectra and coordination numbers, respectively, calculated for each absorbing site j: $\mu(E) = \sum_j \mu_j(E)/N_a$, and $C_{AA} = \sum_j C^j_{AA}/N_a$, $C_{AB} = \sum_j C^j_{AB}/N_a$. Therefore, using a relatively small set of XANES-CNs pairs obtained for individual sites, we can generate large, diverse, sets of labeled examples for the NN training set. To create each example for NN training, we linearly combine three randomly selected XANES-CNs pairs from the pool of site-specific examples. In that way, approximately 1.3×10^{12} possible synthetic training examples can be generated. In the next section, we described how the training data function was used to minimize the validation cost function and how additional testing data were used to test the neural network models.

Neural network training and validation

Here we used an experimental validation set, and subsequent validation cost function, to optimize the Pd and Au absorber-specific neural networks. The experimental data selected for the validation set came from the previously published EXAFS fitting and NN-EXAFS analysis.[23] We examined R5-peptide-templated (peptide) NPs with nominal Pd concentrations of 25, 33, 50, 67, and 83 at% Pd with NP sizes from 3 to 4 nm, as determined by TEM,[48] as well as dilute Pd in Au NPs with Pd concentration of 2 and 9 at% Pd, with sizes between 5–6 nm, synthesized using sequential reduction method, and incorporated into raspberry colloid-templated (RCT) porous SiO_2 using a previously published procedure.[37,49,50] These two data sets are referred to as the "peptide" and "RCT-1" data sets respectively in Table 1, both of which contain Pd K-edge and Au L_3-edge measurements for each sample. XANES spectra for the Pd K-edge and Au L_3-edge of both peptide and RCT-1 data sets are shown in Fig. 3.

Before training, the data were pre-processed (aligned, interpolated, normalized) as described in Note S2 in the ESI.† For both the Pd and Au absorber-specific NNs, we use an early stopping training method based on the cost functions shown in Fig. S7 in the ESI.† This method ensures that the NN model, trained on purely theoretical data, is able to generalize to experimental data. We found that a simple convolutional neural network (CNN) architecture provides the lowest validation loss, where the peptide data was the validation set for the Pd NN, and the peptide and the RCT-1 data were the validation set for the Au NN. To ensure the stability of the models, predictions from 10 independently trained NNs were compared in terms of the median absolute deviation. The results of are presented in Fig. S8, in the ESI.† We see that, for both Pd and Au absorber-specific NNs, the median absolute deviation in CN predictions is very low, characterizing the prediction of the median CNs as very stable. Technical details of training are also included in Note S2 in the ESI.† The layers used in the final Pd K-edge CNN are listed in Table S2 (ESI†) and the layers used in the final Au L_3-edge CNN are listed in Table S3, both located in the ESI.† We also present the entire NN architecture, as implemented in Mathematica 12,[51] in Fig. S9 and S10 in the ESI.†

The resulting absorber-specific NN predictions on the experimental validation sets (described in Table 1) are shown in Fig. 4 with the error bars of the absolute predictions determined by our method described in Note S3 in the ESI.† At the validation cost

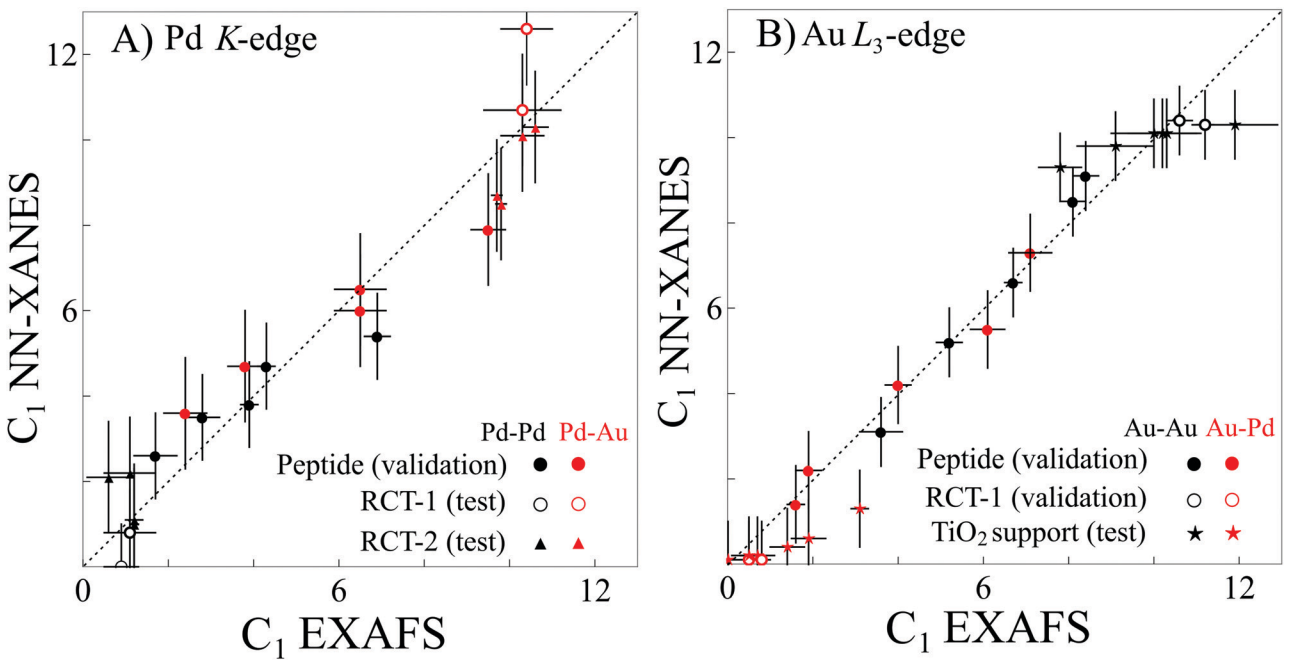

Fig. 4 The XANES-derived *vs.* EXAFS-derived first partial coordination numbers with respect to (A) the Pd K-edge and (B) the Au L_3-edge. Coordination numbers A–A are in black and A–B are in red. In (A) filled in circles are the peptide templated data (validation set), open circles are the RCT-1 data (test set), and triangles are the RCT-2 data (test set). In (B) filled in circles are the peptide templated data (validation set), open circles are the RCT-1 data (validation set), and stars are the TiO_2 supported data (test set).

minima, the coordination number values and relative trends obtained by using NN-XANES agree with those obtained by conventional EXAFS fitting and NN-EXAFS (Fig. 4). While the agreement is good, an additional testing phase was completed, as described in the next section, in which NN predictions were made on data that were not included in the validation cost function.

Neural network testing

Testing the trained NN with data that were not included in the validation cost function is a common way to benchmark the NN before it is applied to unknown systems. For that purpose, we used XANES spectra for several, previously characterized PdAu NPs (4 nm NPs with 4, 15, and 24 at% Pd as well as 6 nm NPs with 3, 5, and 12 at% Pd), all of which were synthesized using a seed-mediated colloidal synthesis method and deposited on TiO_2. This dataset, herein referred to as the "TiO_2-supported" NP dataset, consisted of spectra collected at beamline 12-BM-B, Advanced Photon Source (APS), and their analysis reported in ref. 23. We also use another dataset, referred to as "RCT-2", of Pd K-edge spectra recently measured at beamline ISS (8-ID), National Synchrotron Light Source II (NSLS-II), of 5–6 nm PdAu NPs with Pd concentrations of 4, 9, and 25 at% Pd were synthesized using a sequential reduction method, and incorporated into RCT SiO_2 using a previously published procedure.[37,49,50] The measurements of the 4% and 9% samples were taken *in situ* under He, while the 25% sample was measured under He and H_2. The Pd K-edge data collected for the 25% Pd RCT-2 NPs and their analysis results were reported in ref. 52, while the details of XAS data collection and NN-EXAFS analysis of the rest of the RCT-2 data are included in Note S4, located in the ESI.† The RCT-1 and RCT-2 datasets were used to test the Pd absorber-specific NN, while the TiO_2-supported data were used to test the Au absorber-specific NN. While both the Pd K-edge and Au L_3-edge spectra were collected for the TiO_2-supported dataset, we only use this data to test the Au absorber-specific NN. The Pd K-edge data collected at beamline 12-BM-B appears to have lower energy resolution compared to the rest of the Pd K-edge data that makes NN predictions *via* XANES unstable for this dataset, see Fig. S11 and S12 in the ESI.† Predictions made on the test data sets are plotted *vs.* EXAFS-derived partial coordination numbers in Fig. 4. In all cases, the NN-XANES predictions agree with the CN values, and trends, derived from EXAFS analyses. The results are tabulated in Table S4 in the ESI.†

Application of the NN-XANES for detection of surface restructuring in dilute alloys

Here we demonstrate the utility of our approach for studying gas and temperature treatment effects on component restructuring in dilute (2.6 at% Pd) PdAu/RCT catalyst ($Pd_{2.6}Au_{97.4}$/RCT), for which the conventional EXAFS analysis for the Pd K-edge could not yield conclusive results on Pd–Pd and Pd–Au coordination numbers. As were the other dilute Pd (in Au) catalysts in the RCT-1 and RCT-2 data sets, $Pd_{2.6}Au_{97.4}$/RCT was synthesized using sequential reduction method, and incorporated into raspberry colloid-templated (RCT) porous SiO_2 matrix using a previously published procedure.[37,49,50] In our recently published work on the 4 at% Pd PdAu/RCT catalyst[53] and work on the 25 at% Pd PdAu/RCT catalyst[52] we obtained that Pd species redistribute within Au host in response to high temperature hydrogen treatment. The *in situ* XANES data (Fig. 5) were collected at room temperature under He flow after an initial calcination at 400 °C in 20% O_2/He balance for 1 hour (the O_2 treatment), and then collected at room temperature under He after calcination at 400 °C in 100% H_2 for 30 minutes (the H_2 treatment). More experimental details are in Note S5 in the ESI.† Visual examination of the Pd K-edge XANES data indicates that the local composition around Pd changes between different regimes. For example, energy shifts between the edge positions of the spectra, corresponding to different treatments, and the spectrum of Pd foil reference, reflect the changes in Pd alloying with Au. The larger is the shift towards lower energy, the greater is expected to be the degree of alloying between Pd and Au, consistent with the dissolution of surface Pd into the bulk, expected under H_2 treatment.[52,53]

Quantitative NN-XANES analysis was performed using the Pd and Au absorber-specific NNs. The result of the analysis is presented in Table 2. We see that, after the H_2 treatment, the first partial coordination numbers for Au–Au pairs are similar to those obtained after initial O_2 treatment, as expected for the majority component of the dilute alloy. In contrast, the first coordination numbers for Pd–Pd pairs decrease after the H_2 treatment from 0.17 to 0.10, consistent with partial dissolution of Pd from sites closer to the surface into the bulk as demonstrated in the recent work for larger Pd concentrations.[53] To determine if this change in coordination number is significant, and not due to stochastic differences in NN training, we take the median absolute deviation of predictions made with 10 independently trained NNs. All 10 NN models predict a decrease in the Pd–Pd coordination number after the H_2

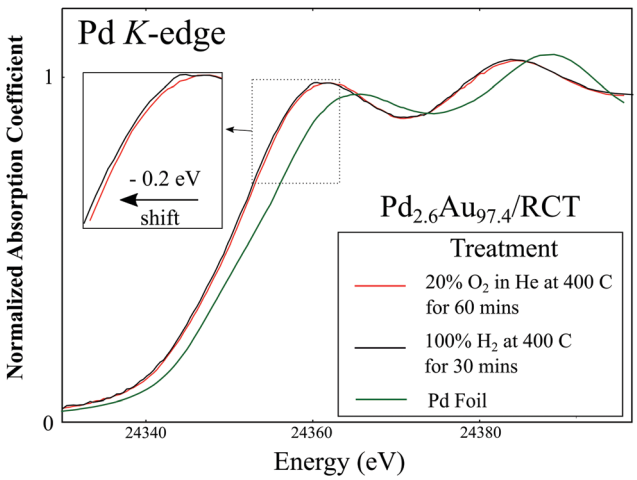

Fig. 5 Pd K-edge XANES spectra of the $Pd_{2.6}Au_{97.4}$/RCT catalyst after O_2 and H_2 treatments and bulk Pd reference foil. The XANES data were collected at room temperature under He.

Table 2 NN-XANES analysis of $Pd_{2.6}Au_{97.4}$/RCT from the Pd K-edge and Au L_3-edge

Sample	Treatment[a]	N[c] Au–Au	N[c] Au–Pd	N Au–M	N[b] Pd–Pd	N[b] Pd–Au	N Pd–M
$Pd_{2.6}Au_{97.4}$	20% O_2/He@400 °C, 1 h	10.13	0.2	10.33	0.17	11.72	11.89
$Pd_{2.6}Au_{97.4}$	100% H_2@400 °C, 30 min	10.16	0.17	10.33	0.10	11.84	11.94

[a] All data were taken at 25 °C under He. [b] Pd absorber-specific NN. [c] Au absorber-specific NN.

treatment, with a median decrease of 0.070 and median absolute deviation of 0.001.

Discussion

In this work, we showed that it is possible to extract the partial coordination numbers from the XANES spectrum of bimetallic alloys using absorber-specific neural networks. This opens the door for various applications. Most significantly, the structural characterization of bimetallic nanoalloys in general, and nanocatalysts in particular, *via* XANES is now possible in materials in which EXAFS analysis is limited by signal quality. For example, structural changes detected in harsh reaction conditions, particularly in dilute Pd in Au catalysts, can now be understood with similar level of detail (*i.e.*, the partial coordination numbers and, hence, restructuring of the catalyst components that may occur under *in situ* conditions, can be extracted). We provided an example of this capability, by showing that *in situ* XANES measurements can detect (and neural network-assisted analysis can correctly recognize) the consequences of Pd restructuring under varying conditions. The neural network-assisted data analysis method we report provides structural parameters with sufficient accuracy for modeling in only seconds, therefore enabling new applications of XANES at the beamline, such as the real-time monitoring of NP deactivation or high throughput sample characterization. The PdAu absorber-specific networks that we constructed can be immediately applied to the analysis of XANES in novel PdAu systems or used to analyze previously collected data for which structural characterization was not originally considered, or not possible by conventional EXAFS, due to, *e.g.*, harsh reaction conditions, but for which XANES data are of sufficiently good quality.

We have demonstrated that during and after NN training, validation, and subsequent testing with experimental data sets from additional systems, which were measured at different beamlines, helped us evaluate and improve robustness of NNs trained on theoretical data. We believe that these insights will be crucial for those using this method to analyze PdAu systems and those who will utilize NN-XANES in general for the analysis of other bimetallic systems. We confirmed that the trained NNs are able to make accurate predictions on metallic PdAu data, at the Pd K and Au L_3 absorption edges, with different compositions, particle sizes, supports, and synthesis methods. Due to the intrinsic structural complexity of bimetallic materials, there was no guarantee that one general NN model (for bulk PdAu alloys), as opposed to a specific NN model (for a given PdAu nanoalloy system, *i.e.*, in some size range and/or composition motif), would be sufficient to capture the spectrum-structure relationship in a range of NP sizes and compositional distributions. Here we see that the general NN model performs well in the case of metallic PdAu. We see this clearly at the Au L_3-edge, where the theoretical-experimental agreement is better, and beamline resolution effects are relatively forgiving, compared to the Pd K edge (*vide infra*). As we create more bimetallic models for different systems for which our method can be applied, we will be able to determine the limits of applicability of our method beyond the several NP systems we tested it on. We also see that our training method results in NNs that make accurate predictions, even in the face of moderate resolution and systematic differences between beamlines and, even without explicitly simulating these effects in the training data. This is an important feature, especially at the Pd K-edge (and other edges at relatively high energy), where the energy resolution is known to suffer at some beamlines. The most significant example is derived from the Pd_9Au_{91} RCT supported sample, which is included in datasets RCT-1 and RCT-2, where RCT-1 was measured at the BL2-2 beamline of Stanford Synchrotron Radiation Lightsource (SSRL) with a Si(220) monochromator, and the RCT-2 samples was measured at the NSLSII with a Si(111) monochromator. The Pd absorber-specific NN predicts Pd–Pd and Pd–Au coordination numbers of 1.0 ± 1.0 and 10.3 ± 1.3 *vs.* 0.8 ± 0.8 and 10.7 ± 0.9 for the data collected at the SSRL and NSLSII, respectively. The predictions are in very good agreement with each other despite some spectral differences that are caused by the energy resolution difference of the beamline monochromators. A comparison of the spectra collected at the two beamlines is shown in Fig. S13 in the ESI.† We do see a limit, though, in the Pd absorber-specific NN's ability to deal with low-resolution Pd K-edge data, as mentioned in the neural network testing section. However, in that case, the resolution was much lower than usual, as quantified by an amplitude reduction factor of 0.70 for Pd foil that was measured at the time of the experiment. At some limit of the resolution, the XANES may simply loose interpretable information. On the other hand, using specialized beamlines and secondary analyzers, it is possible to measure XANES with high energy resolution. The high-energy resolution fluorescence detection (HERFD) mode of XANES measurement provides a significant enhancement to XANES spectra,[54,55] improving its sensitivity to the structure. In principle, a combined HERFD-NN-XANES approach could be used to not only make superior structural predictions, but a HERFD experimental validation set could be used in neural network training, resulting in more accurate predictions on non-HERFD XANES. Together with other recent studies,[56] this work further advances the potentiality of machine learning approaches and XAFS method for characterization of unique structural motifs in bimetallic nanoparticles and their transformations under reaction conditions.

Conclusions

The NN-XANES method has been developed for the extraction of the first partial coordination numbers from the XANES of PdAu nanocatalysts. Experimental validation was performed with a set of well-defined PdAu NPs with sufficient EXAFS quality for providing *a priori* knowledge, in combination with another experimental testing step to benchmark the NN's predictive power over a range of compositions and supports. We have demonstrated that this method provides robust predictions of first partial coordination numbers that agree with those derived by conventional EXAFS fitting and NN-EXAFS methods, and we showed that the method can be used to investigate restructuring in dilute bimetallic catalysts. One can now extend this method to other bimetallic compositions, such as dilute catalysts and size-selective clusters.

Author contributions

NM, YL, JT and AIF developed the method, NM and YL designed and utilized neural network, ML, JESH, TS, JA and CMF synthesized catalysts and developed *in situ* treatment procedures, NM, YL, JT, EG, ML, AMP and AIF conducted synchrotron experiments, NM and AIF wrote the manuscript, all authors discussed the results and manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The work was supported as part of the Integrated Mesoscale Architectures for Sustainable Catalysis (IMASC), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award No. DE-SC0012573. This research used resources of the Center for Functional Nanomaterials, which is a U.S. DOE Office of Science Facility, and the Scientific Data and Computing Center, a component of the Computational Science Initiative, at Brookhaven National Laboratory under Contract No. DE-SC0012704. This research used beamline 8-ID of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. We are grateful to Mr David Verbart, Mr Alexandre Foucher, and Dr Eli Stavitski for help with data collection at the ISS beamline.

Notes and references

1 F. Gao and D. W. Goodman, *Chem. Soc. Rev.*, 2012, **41**, 8009–8020.
2 R. Burch, *Acc. Chem. Res.*, 1982, **15**, 24–31.
3 F. Tao, *Chem. Soc. Rev.*, 2012, **41**, 7977–7979.
4 P. Liu and J. K. Nørskov, *Phys. Chem. Chem. Phys.*, 2001, **3**, 3814–3818.
5 Z. W. Ulissi, M. T. Tang, J. Xiao, X. Liu, D. A. Torelli, M. Karamad, K. Cummins, C. Hahn, N. S. Lewis, T. F. Jaramillo, K. Chan and J. K. Nørskov, *ACS Catal.*, 2017, **7**, 6600–6608.
6 Y. Li, D. Zakharov, S. Zhao, R. Tappero, U. Jung, A. Elsen, P. Baumann, R. G. Nuzzo, E. A. Stach and A. I. Frenkel, *Nat. Commun.*, 2015, **6**, 7583.
7 M. S. Nashner, A. I. Frenkel, D. L. Adler, J. R. Shapley and R. G. Nuzzo, *J. Am. Chem. Soc.*, 1997, **119**, 7760–7771.
8 L. Li, L. L. Wang, D. D. Johnson, Z. F. Zhang, S. I. Sanchez, J. H. Kang, R. G. Nuzzo, Q. Wang, A. I. Frenkel, J. Li, J. Ciston, E. A. Stach and J. C. Yang, *J. Am. Chem. Soc.*, 2013, **135**, 13062–13072.
9 F. Tao, M. E. Grass, Y. W. Zhang, D. R. Butcher, F. Aksoy, S. Aloni, V. Altoe, S. Alayoglu, J. R. Renzas, C. K. Tsung, Z. W. Zhu, Z. Liu, M. Salmeron and G. A. Somorjai, *J. Am. Chem. Soc.*, 2010, **132**, 8697–8703.
10 F. Tao, M. E. Grass, Y. W. Zhang, D. R. Butcher, J. R. Renzas, Z. Liu, J. Y. Chung, B. S. Mun, M. Salmeron and G. A. Somorjai, *Science*, 2008, **322**, 932–934.
11 H. L. L. Xin, S. Alayoglu, R. Z. Tao, A. Genc, C. M. Wang, L. Kovarik, E. A. Stach, L. W. Wang, M. Salmeron, G. A. Somorjai and H. M. Zheng, *Nano Lett.*, 2014, **14**, 3203–3207.
12 A. I. Frenkel, M. W. Cason, A. Elsen, U. Jung, M. W. Small, R. G. Nuzzo, F. D. Vila, J. J. Rehr, E. A. Stach and J. C. Yang, *J. Vac. Sci. Technol., A*, 2014, **32**, 020801.
13 A. I. Frenkel, *Chem. Soc. Rev.*, 2012, **41**, 8163–8178.
14 A. Yevick and A. I. Frenkel, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **81**, 115451.
15 S. T. Chill, R. M. Anderson, D. F. Yancey, A. I. Frenkel, R. M. Crooks and G. Henkelman, *ACS Nano*, 2015, **9**, 4036–4042.
16 J. Timoshenko, Z. Duan, G. Henkelman, R. M. Crooks and A. I. Frenkel, *Annu. Rev. Anal. Chem.*, 2019, **12**, 501–522.
17 J. J. Rehr, J. J. Kas, F. D. Vila, M. P. Prange and K. Jorissen, *Phys. Chem. Chem. Phys.*, 2010, **12**, 5503–5513.
18 O. Bunău and Y. Joly, *J. Phys.: Condens. Matter*, 2009, **21**, 345501.
19 A. Filipponi, A. Di Cicco and C. R. Natoli, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1995, **52**, 15122–15134.
20 M. Benfatto, A. Congiu-Castellano, A. Daniele and S. Della Longa, *J. Synchrotron Radiat.*, 2001, **8**, 267–269.
21 G. Smolentsev and A. V. Soldatov, *Comput. Mater. Sci.*, 2007, **39**, 569–574.
22 J. Timoshenko and A. I. Frenkel, *ACS Catal.*, 2019, **9**, 10192–10211.
23 J. Timoshenko, C. J. Wrasman, M. Luneau, T. Shirman, M. Cargnello, S. R. Bare, J. Aizenberg, C. M. Friend and A. I. Frenkel, *Nano Lett.*, 2019, **19**, 520–529.

24 J. Timoshenko, A. Anspoks, A. Cintins, A. Kuzmin, J. Purans and A. I. Frenkel, *Phys. Rev. Lett.*, 2018, **120**, 225502.
25 J. Timoshenko, D. Lu, Y. Lin and A. I. Frenkel, *J. Phys. Chem. Lett.*, 2017, **8**, 5091–5098.
26 G. Bunker, *Nucl. Instrum. Methods Phys. Res.*, 1983, **207**, 437–444.
27 A. Frenkel, *J. Synchrotron Radiat.*, 1999, **6**, 293–295.
28 J. Timoshenko and A. I. Frenkel, *Catalysis*, 2017, **280**, 274–282.
29 A. Jentys, *Phys. Chem. Chem. Phys.*, 1999, **1**, 4059–4063.
30 J. Matos, L. K. Ono, F. Behafarid, J. R. Croy, S. Mostafa, A. T. DeLaRiva, A. K. Datye, A. I. Frenkel and B. Roldan Cuenya, *Phys. Chem. Chem. Phys.*, 2012, **14**, 11457–11467.
31 M. Ahmadi, J. Timoshenko, F. Behafarid and B. Roldan Cuenya, *J. Phys. Chem. C*, 2019, **123**, 10666–10676.
32 J. Timoshenko, A. Halder, B. Yang, S. Seifert, M. J. Pellin, S. Vajda and A. I. Frenkel, *J. Phys. Chem. C*, 2018, **122**, 21686–21693.
33 S. Roese, A. Kononov, J. Timoshenko, A. I. Frenkel and H. Hövel, *Langmuir*, 2018, **34**, 4811–4819.
34 Y. Liu, N. Marcella, J. Timoshenko, A. Halder, B. Yang, L. Kolipaka, M. J. Pellin, S. Seifert, S. Vajda, P. Liu and A. I. Frenkel, *J. Chem. Phys.*, 2019, **151**, 164201.
35 N. Agarwal, S. J. Freakley, R. U. McVicker, S. M. Althahban, N. Dimitratos, Q. He, D. J. Morgan, R. L. Jenkins, D. J. Willock, S. H. Taylor, C. J. Kiely and G. J. Hutchings, *Science*, 2017, **358**, 223–227.
36 X. H. Peng, Q. M. Pan and G. L. Rempel, *Chem. Soc. Rev.*, 2008, **37**, 1619–1628.
37 M. Luneau, T. Shirman, A. C. Foucher, K. Duanmu, D. M. A. Verbart, P. Sautet, E. A. Stach, J. Aizenberg, R. J. Madix and C. M. Friend, *ACS Catal.*, 2020, **10**, 441–450.
38 J. Liu, J. Shan, F. R. Lucci, S. Cao, E. C. H. Sykes and M. Flytzani-Stephanopoulos, *Catal. Sci. Technol.*, 2017, **7**, 4276–4284.
39 M. Luneau, T. Shirman, A. Filie, J. Timoshenko, W. Chen, A. Trimpalis, M. Flytzani-Stephanopoulos, E. Kaxiras, A. I. Frenkel, J. Aizenberg, C. M. Friend and R. J. Madix, *Chem. Mater.*, 2019, **31**, 5759–5768.
40 T. Ward, L. Delannoy, R. Hahn, S. Kendell, C. J. Pursell, C. Louis and B. D. Chandler, *ACS Catal.*, 2013, **3**, 2644–2653.
41 M. S. Nashner, A. I. Frenkel, D. Somerville, C. W. Hills, J. R. Shapley and R. G. Nuzzo, *J. Am. Chem. Soc.*, 1998, **120**, 8093–8101.
42 M. R. Knecht, M. G. Weir, A. I. Frenkel and R. M. Crooks, *Chem. Mater.*, 2008, **20**, 1019–1028.
43 A. Frenkel, *Z. Kristallogr. - Cryst. Mater.*, 2007, **222**, 605–611.
44 A. I. Frenkel, Q. Wang, S. I. Sanchez, M. W. Small and R. G. Nuzzo, *J. Chem. Phys.*, 2013, **138**, 064202.
45 M. Ziatdinov, O. Dyck, A. Maksov, X. Li, X. Sang, K. Xiao, R. R. Unocic, R. Vasudevan, S. Jesse and S. V. Kalinin, *ACS Nano*, 2017, **11**, 12742–12752.
46 J. Madsen, P. Liu, J. Kling, J. B. Wagner, T. W. Hansen, O. Winther and J. Schiøtz, *Adv. Theory Simul.*, 2018, **1**, 1800037.
47 J. Cuny, Y. Xie, C. J. Pickard and A. A. Hassanali, *J. Chem. Theory Comput.*, 2016, **12**, 765–773.
48 N. A. Merrill, E. M. McKee, K. C. Merino, L. F. Drummy, S. Lee, B. Reinhart, Y. Ren, A. I. Frenkel, R. R. Naik, N. M. Bedford and M. R. Knecht, *ACS Nano*, 2015, **9**, 11968–11979.
49 M. Luneau, T. Shirman, A. Filie, J. Timoshenko, W. Chen, A. Trimpalis, M. Flytzani-Stephanopoulos, E. Kaxiras, A. I. Frenkel, J. Aizenberg, C. M. Friend and R. J. Madix, *Chem. Mater.*, 2019, **31**, 5759–5768.
50 E. Shirman, T. Shirman, A. V. Shneidman, A. Grinthal, K. R. Phillips, H. Whelan, E. Bulger, M. Abramovitch, J. Patil, R. Nevarez and J. Aizenberg, *Adv. Funct. Mater.*, 2018, **28**, 1704559.
51 *Mathematica, version 12.0*, Wolfram Research, Inc., Champaign, Illinois, 2019.
52 E. Guan, A. C. Foucher, N. Marcella, T. Shirman, M. Luneau, A. R. Head, D. M. A. Verbart, J. Aizenberg, C. M. Friend, D. Stacchiola, E. A. Stach and A. I. Frenkel, *ChemCatChem*, 2020, **12**, 717–721.
53 M. Luneau, E. Guan, W. Chen, A. C. Foucher, N. Marcella, T. Shirman, D. M. A. Verbart, J. Aizenberg, M. Aizenberg, E. A. Stach, R. J. Madix, A. I. Frenkel and C. M. Friend, *Commun. Chem.*, 2020, **3**, 46.
54 E. Kleymenov, J. Sa, J. Abu-Dahrieh, D. Rooney, J. A. van Bokhoven, E. Troussard, J. Szlachetko, O. V. Safonova and M. Nachtegaal, *Catal. Sci. Technol.*, 2012, **2**, 373–378.
55 D. K. Pappas, A. Martini, M. Dyballa, K. Kvande, S. Teketel, K. A. Lomachenko, R. Baran, P. Glatzel, B. Arstad, G. Berlier, C. Lamberti, S. Bordiga, U. Olsbye, S. Svelle, P. Beato and E. Borfecchia, *J. Am. Chem. Soc.*, 2018, **140**, 15270–15278.
56 J. Timoshenko, H. S. Jeon, I. Sinev, F. T. Haase, A. Herzog and B. Roldan Cuenya, *Chem. Sci.*, 2020, **11**, 3727–3736.

