PCCP

Iterative Reverse Monte Carlo and Molecular Statics for Improved Atomic Structure Modeling: A Case Study of Zinc Oxide Grown by Atomic Layer Deposition

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-08-2021-003742.R1
Article Type:	Paper
Date Submitted by the Author:	29-Oct-2021
Complete List of Authors:	Gettler, Ryan; University of Missouri, Department of Biomedical, Biological, and Chemical Engineering Koenig, Henry; University of Missouri, Department of Biomedical, Biological, and Chemical Engineering Young, Matthias; University of Missouri, Department of Biomedical, Biological, and Chemical Engineering; University of Missouri, Department of Chemistry

Iterative Reverse Monte Carlo and Molecular Statics for Improved Atomic Structure Modeling: A Case Study of Zinc Oxide Grown by Atomic Layer Deposition

Ryan C. Gettler[1], Henry D. Koenig[1], Matthias J. Young[1,2,*]

[1]Department of Biomedical, Biological and Chemical Engineering, University of Missouri, Columbia, MO, USA

[2]Department of Chemistry, University of Missouri, Columbia, MO, USA

*email: matthias.young@missouri.edu

1. Abstract

Reverse Monte Carlo (RMC) modeling is a common method to derive atomic structure models of materials from experimental diffraction data. However, conventional RMC modeling does not impose energetic constraints and can produce non-physical local structures within the simulation volume. Although previous strategies have introduced energetic constraints during RMC modeling, these approaches have limitations in computational cost and physical accuracy. In this work, we periodically introduce molecular statics (MS) energy minimizations during RMC modeling in an iterative RMC-MS approach. We test this iterative RMC-MS approach using diffraction data collected by *in-operando* high energy X-ray diffraction during atomic layer deposition of ZnO as a sample case. For MS relaxations we employ ReaxFF pair potentials previously established for ZnO. We find that RMC-MS and RMC provide equivalent agreement with experimental data, but RMC-MS structures are on average 0.6 eV/atom lower in energy and are more consistent with known ZnO atomic structure features. The iterative RMC-MS approach we report can accommodate large systems with minimal additional computational burden beyond a standard RMC simulation and can leverage established pair potentials for immediate application to study a wide range of materials.

2. Introduction

In recent years, significant advancements in material synthesis and processing techniques have allowed for the controlled fabrication of increasingly complex, finely-tuned materials. Knowledge of the atomic configurations of these systems is paramount to understanding and optimizing the physical properties for the desired application. One example of this can be found in semiconductor devices, where the number and arrangement of defects in a semiconductor lattice is crucial to device performance, and determination of the material structure is necessary to explain and predict the physical properties.[1,2] Characterization techniques like X-ray diffraction have matured in parallel with materials synthesis techniques, allowing for the acquisition of more detailed compositional and structural data.[3–5] In general, experimental characterization provides only salient structural features, and the local atom environments must be elucidated using simulation techniques.[6,7] Atomic-scale modeling of the experimental data has accordingly become increasingly challenging, requiring ever more detail and computational effort to produce satisfactory results. It is thus necessary to implement new modeling procedures that accurately capture the large-scale experimental observations, are consistent with thermodynamics, and require little computational resources.

Reverse Monte Carlo (RMC) simulation is one common method for deriving the structure of a material from corresponding X-ray, neutron, and electron diffraction data. Based solely on simple input structures and constraints, RMC can give valuable insight into large scale systems with thousands of atoms at low computational cost. The ease of use has made it a popular choice for modeling a wide range of systems and material phases, including amorphous and crystalline solids as well as liquids.[8–11] The accuracy of a RMC model is assumed to be purely statistical, meaning that any configuration of a given system is considered valid provided it fits the experimental data.[12] As a result, simulations often result in non-unique and non-physical atom arrangements.[8,10,12] In this sense, the simplicity of the RMC approach is also its biggest weakness.

To address the weaknesses of RMC models, one may include additional constraints during RMC simulation steps to improve the physical accuracy of the resulting model. These constraints can be energetic,[13] geometric,[14] or others, for example using three-body correlations,[15] or distance window constraints.[16] Among these options for physical constraints, the use of classical molecular dynamics (MD) is appealing because it is low-computational cost and can describe the energetics of atomic interactions for longer-range and complex coordination environments. On its own, MD provides structures that are physically accurate, however MD structures are not expected to conform to a specific structure measured experimentally unless the experimental data is used to guide the MD model. This has motivated researchers to include aspects of both MD and RMC modeling to establish structural models from experimental diffraction data. In its simplest implementation, MD simulations can be used to generate an input structure closely resembles the XRD data, followed by RMC manipulation to improve the fit. This type of RMC and MD approach yields more energetically favorable configurations than RMC alone, but thermodynamic control of the system is limited to the initial structure, and the error between simulated and measured structures is often higher.[17,18]

Several more sophisticated implementations of combined stochastic/MD methods have been developed to resolve material structures with higher accuracy.[10,12,15,19–21] Of these, two prominent examples are the empirical potential structure refinement (EPSR) and RMC_POT methods.[10,12,19] The EPSR technique involves iterative Monte Carlo steps with adjustment of an interatomic pair potential until the model agrees with experiment and has been particularly successfully in fluid simulations.[10,22,23] Unfortunately, this requires 'biasing' of the physics to fit the experiment, meaning the pair potential is not necessarily consistent with known physics, and the solutions are not unique.[12,24] Similarly, the RMC_POT method introduces potential energy factors as constraints and has shown improvements in physical accuracy, providing lower system energies compared to standard RMC simulations.[19,22] The RMC_POT algorithm is limited to small systems (< 100 atoms) with simple Lennard-Jones type interactions, however, and thus has not been used in a wide range of applications.[22]

In addition to EPSR and RMC_POT, a select few studies have employed other techniques that integrate aspects of MD and RMC in unique ways. One study employed a combination of *ab initio* MD modeling and Monte Carlo (MC) modeling to generate a starting structure that was subsequently refined using iterative steps of RMC and MC modeling.[20] While this iterative MD-MC-RMC approach provided impressive results, the system dimensions were restricted due to the

high computational cost of *ab initio* MD and MC steps, similar to the EPSR and RMC_POT methods described above. Another report introduced the force enhanced atomic refinement (FEAR) algorithm, in which traditional weighting factors and rejection parameters employed in stochastic modeling techniques are replaced with a force constraint based on electrostatic forces between atoms.[25] However, this study was restricted to 1536 atoms due to the high computational cost of the force functions, and the software used in this study is not publicly available. Another report applied statistical sampling of different MD quench temperatures to arrive at a starting structure that most closely resembled their experimental data, and then employed hybrid RMC with force constraints to optimize the starting structure determined from MD.[26] Again, this study was limited to a system size of ~600 atoms, and required many MD simulations to find a starting structure that closely matched the experimental system. In general, prior reports combining MD and RMC steps limit the number of atoms that can be used in a model system and employ specialized software that is not readily available to the broader community.

In this work, we report on a new iterative approach combining RMC and molecular statics (MS) relaxations using off-the-shelf open-source software packages, and demonstrate this iterative approach for a system size of >20,000 atoms. Here we employ MS energy minimizations between a set number of RMC translations. By periodically relaxing the structure, we effectively impose an additional constraint throughout the simulation that can accommodate systems and potentials of high complexity at low computational cost. The described combined RMC-MS technique should thus simultaneously produce physical structures that correlate well with experimental measurements with a wide range of application.

As a test material for benchmarking the RMC-MS approach, we employ zinc oxide (ZnO) grown by atomic layer deposition (ALD). ALD has been studied for a wide range of applications including photovoltaics,[27–29] catalysis,[30,31] energy storage,[32,33] and semiconductor devices,[34] but deriving the atomic assembly of ALD-grown materials is challenging because of their nanoscale and disordered structures. Recent work from our group and others has established pair distribution function (PDF) characterization and RMC modeling of ALD-grown materials to reveal new physical insights into the atomic structure of these films.[11,35–38] However, as this work expands and is used to guide subsequent studies,[39] it is critical to establish confidence in the physical accuracy of these atomic structure models. In this context, we seek to examine (and improve) the physical accuracy of atomic structures derived from RMC modeling. ALD-grown ZnO is particularly well-suited for this study because bulk ALD ZnO exhibits long-range crystallinity with a known structure, providing a valuable benchmark for atomic structure.[40] Furthermore, the implementation of *in operando* measurement of atomic structure during ALD growth reported previously allows for the rapid generation of a large data set of ZnO diffraction data for analysis and modeling.[11]

3. Methods

3.1 In operando ALD/HE-XRD

High energy X-ray diffraction (HE-XRD) was performed *in operando* according to previously established procedures.[11] Briefly, a custom atomic layer deposition (ALD) reactor stage designed

for powder HE-XRD was connected to a gas-handling system[41] for delivery of ALD precursors. The stage consisted of a polyimide capillary filled with hydroxyl-terminated CNT powder and fixed within the X-ray beam path. The capillary was positioned within a small ALD reactor chamber sealed with polyimide windows to facilitate transmission and allow for data acquisition at diffraction angles up to 45°. HE-XRD measurements were done at 11-ID-C at APS using a wavelength of 0.1173 Å.

The stage chamber was kept at 150 °C and ~1 torr using nitrogen carrier gas during deposition and diffraction measurement. ZnO was grown by sequential dosing of diethyl zinc (DEZ) and water[42–44] separated by a nitrogen purge between each precursor using the following cycle timing: 10 s DEZ dose, 120 s purge, 10 s water dose, 120 s purge. HE-XRD patterns were obtained by 72 repeat measurements at 5 s exposure time after each reported ALD cycle, which were averaged before analyzing. We caution that DEZ is pyrophoric and should be handled with care.

3.2 Pair Distribution Function

The pair distribution function *G(r)* is defined as

$$G(r) = 4\pi r\rho_0\left[\sum_{i,j}[w_{ij}g_{ij(r)}] - 1\right] \qquad (1)$$

where ρ_0 is the atomic number density, $g_{ij}(r)$ is the partial pair distribution for each *ij* pair, and w_{ij} is weight factor related to the component mole fractions and X-ray scattering factors.[7,8] *G(r)* can be obtained from the Fourier transform of the structure factor *S(Q)*, resulting in Eqn. 2,

$$G(r) = \frac{2}{\pi}\int Q[S(Q) - 1]\sin(Qr)dQ \qquad (2)$$

where *S(Q)* is determined from 2-D diffraction data and *Q* is the photon momentum.[45]

2-D HE-XRD diffraction data was calibrated and integrated into 1-D diffraction scans using GSASII.[46] Pair distribution functions (PDF's) were generated from HE-XRD diffraction data after every 5 ALD cycles using the pdfgetx3 package.[47] The diffraction pattern of the uncoated CNT substrate was used as a background. *S(Q)* was integrated from Q_{min} = 0.8 Å^{-1} to Q_{max} = 22.2 Å^{-1}.

3.3 RMC Modeling

RMC calculations were performed using the fullrmc package.[48] Starting atomic coordinates were generated using repeating units of hexagonal ZnO (HM space group $P6_3mc$186) positioned in a cube of 64.4 Å edge length. Simulations were run for 10 million attempted structural perturbations, i.e. perturbations that reduce the mean square error (MSE) between the experimental and model PDF. For each ALD growth cycle, the starting simulations box contained 21,952 total atoms (10,976 Zn and 10,976 O), however, 100 atom removal steps were implemented every 1 million translations using the *AtomsRemovesGenerator* implemented in fullrmc, in which 100 Zn and/or O atom eliminations were attempted, the PDF was updated, and the MSE recalculated. If the atom

removal improved the MSE, the change was accepted, and translations continued as described. Over 10 million attempted translation steps, a total of 1,000 atom removal steps were attempted, which corresponds to 9% of each of the Zn and O atoms present in the starting structure. RMC modeling requires setting a minimum approach distance (*r*) for each set of pairs in the box, which are presented in Table 1. Additionally, the coordination number (CN) for Zn-O pairs was constrained to a value between 3 and 5 (inclusive) for pairs separated by 1.5 to 2.5 Å.

Table 1: Minimum Pair Distance Constraints

Pair	**Zn-Zn**	**O-O**	**Zn-O**
r (Å)	2.5	1.2	1.7

3.4 Molecular Statics Simulations

Molecular Statics (MS) relaxations were done with LAMMPS software using a reactive force field (ReaxFF) potential developed by Tainter et al.[49,50] Briefly, ReaxFF defines the covalent interactions of the system in terms of a bond order *BO'*, which is a function of interatomic separation r_{ij} and is separated into single (σ), double (π), and triple ($\pi\pi$) bond terms (Eqn. 3).[49,51,52]

$$BO'_{ij} = BO'^{\sigma}_{ij} + BO'^{\pi}_{ij} + BO'^{\pi\pi}_{ij} \qquad (3)$$

The bond energy has the form of a Morse potential, with fitting parameters p_i determined by DFT calculations (Eqn. 4). Defining E_{bond} solely as a function of the bond order allows bonds to form or break during the simulation and ensures that all covalent terms disappear as r_{ij} increases beyond the energy minimum.[51] A more detailed explanation of ReaxFF potentials can be found in references 43 and 45.[51,53]

$$E_{bond} = -D_e^{\sigma} BO_{ij}^{\sigma} exp\left[p_1\left(1-\left(BO_{ij}^{\sigma}\right)^{p_2}\right)\right] - D_e^{\pi} BO_{ij}^{\pi} - D_e^{\pi\pi} BO_{ij}^{\pi\pi} \qquad (4)$$

Force minimization was accomplished via the conjugate gradient method at constant N and T. Following adjustment of atomic coordinates, an external pressure was applied to the simulation box, allowing the system to contract or expand during the relaxation before updating the energy. Simulations were run sequentially in this fashion until the energy converged to a minimum precision of 10^{-9} eV for the total system energy between successive calculations.

Post-simulation processing was done using Interactive Structural Analysis of Amorphous and Crystalline Solids (ISAACS), Visual Molecular Dynamics (VMD), and Visualization for Electronic and Structural Analysis (VESTA) software.[54–56] Bond distributions, CNs, and PDF's were calculated using ISAACS for all reported model structures. VESTA was used to generate the atomic structure representations. LAMMPS, fullrmc, and the post-processing software all require different data file formats (lammpstrj, pdb, and xyz, respectively). To accommodate the various formatting requirements, scripts were written using the VMD module in python to convert files to the necessary type before the corresponding analysis.

4. Results and Discussion

A plot of *G(r)* vs. *r* gives the radial distribution of atom pairs separated by distance *r* relative to the average total number of pairs in the system.[11] Figure 1a, 1b, and 1c show sample pair distribution functions (PDFs) obtained from XRD overlayed with the simulated PDF's for 5, 25, and 45 ALD cycles, respectively. A peak in the PDF corresponds to a pair-density greater than (positive) or less than (negative) the average pair-density at a given radius from the origin.[8,11] Well-defined, periodic peaks in the PDF are thus indicative of repeating units in the structure and are present at long range ($r \geq 5$ Å) order for 30 growth cycles and above. In Figure 1d, we plot a heatmap representation of the PDF data for every ALD growth cycle from 1 ALD growth cycle up to 50 ALD growth cycles. This corresponds to ZnO thicknesses of ~2 to ~100 Å based on a steady-state growth rate of ~2.0Å/cycle for ZnO grown using DEZ and H_2O at 150 °C.[44,57] We observe from this PDF data that the long-range order in ZnO increases with the number of ALD growth cycles. We highlight this observation with a dashed grey line in Figure 1d. (Note that no periodic peaks in the x-direction are evident to the bottom right of this grey line, and pronounced periodic peaks are evident in the x-direction above this grey line.) The observation of long-range order at a higher number of ALD growth cycles is consistent with previous reports for ALD growth of ZnO showing the formation of crystalline ZnO with long-range order.[40]

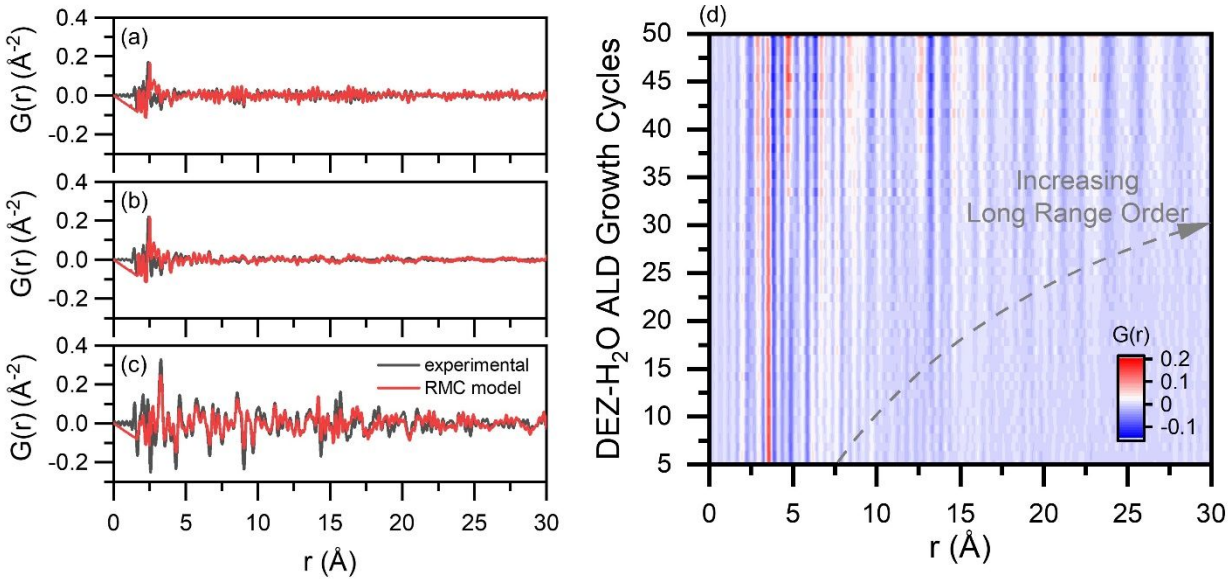

Figure 1: Sample total PDF for ZnO from HE-XRD data (black) and corresponding PDFs derived from RMC simulations (red) following (a) 5, (b) 25, and (c) 45 ALD growth cycles. Panel (d) shows the change in PDF peak intensity with the number of ALD cycles, where red indicates $G(r) > 0$ and blue indicates $G(r) < 0$.

RMC simulations produced structures with PDFs that agreed with the experimental PDFs within an MSE of 4.5 for all sampled ALD growth cycles. The average molecular formula across all simulations was $Zn_{0.99}O$, indicating that the atom removal steps did not significantly affect the Zn

to O atom ratio. While the RMC models produced the correct pair distribution function, the predicted average bond length (Figure 2) is longer than the expected value of 1.8-2.0 Å typical of ZnO.[49,58,59] This discrepancy was observed for the full range of ALD growth cycles studied here (1-50) and is especially pronounced $\leq$ 25 ALD growth cycles. This could be explained by interfacial effects. One might expect that at a low number of ALD cycles, surface/interface chemistry dominates bond character, and the structure may differ from bulk ZnO. As the number of growth cycles increases beyond 25, the average bond length approaches the ideal value of 2.0 Å.[58] We observe a reverse trend for CN in Figure 2, where the CN deviates from the expected 4-coordinate system of tetrahedral ZnO for longer growth times. These low values of CN are neither consistent with known ZnO crystal structures nor with the expected coordination environment of Zn from crystal field theory.[58,60,61] We hypothesize that these seemingly incorrect CN values may be due to non-physical artifacts arising from the rudimentary constraints employed during RMC modeling (*vide infra*).[8]

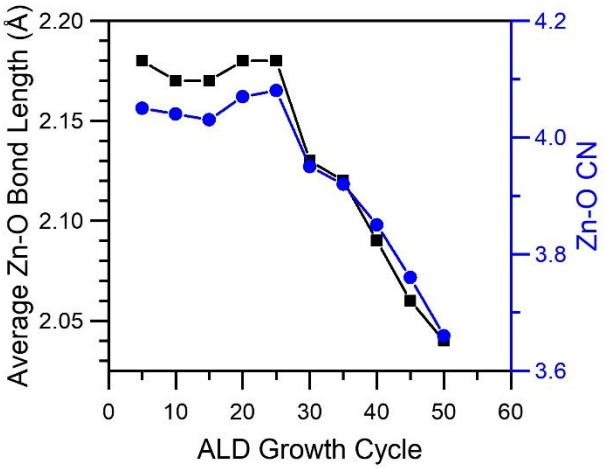

Figure 2: Average ZnO bond length and coordination number as predicted by RMC simulation versus number of ALD ZnO growth cycles.

Because we expected that non-physical artifacts were driving the Zn-O CNs <4 as depicted in Figure 2, we set out to remove non-physical artifacts using MS relaxation steps at various intervals during RMC modeling. To evaluate whether these MS relaxation steps improved the outcome, we compared RMC modeling of PDF data without MS relaxations against an alternating sequence of RMC and MS relaxation (RMC-MS). Specifically, we introduced an MS energy minimization step following each block of two million RMC translations until ten million total RMC translations were achieved. We compared the resulting structural metrics against structures derived from ten million total RMC translations in two million step intervals without MS relaxations. By introducing MS relaxation steps, we expected to remove any non-physical local structure artifacts, and arrive at a more physically realistic and accurate representation of the atomic structure of ALD-grown ZnO. In Figure 3 we compare structural metrics derived from RMC and RMC-MS modeling of diffraction data at 45 ALD ZnO growth cycles. We compare the MSE (Figure 3a),

bond lengths (Figure 3b), CN (Figure 3c), and cohesive energy (E_c, Figure 3d), for these two modeling approaches, following each of the five RMC cycles consisting of two million translation steps. The cohesive energy E_c is the difference in energy between an ideal gas of free atoms and the solid state of the material.

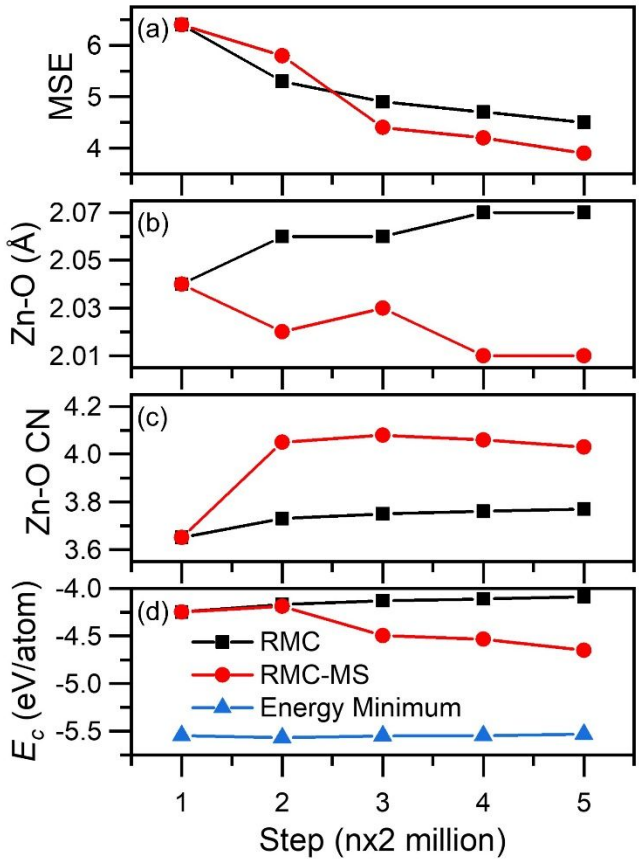

Figure 3: Comparison of structural metrics derived from RMC and RMC-MS with MS relaxation every two million RMC iterations for 45 ALD growth cycles including (a) MSE between experimental and simulated PDF, (b) average Zn-O bond length, (c) average Zn-O CN, and (d) cohesive energy per atom. All reported values were obtained immediately following the RMC steps except for the "energy minimum", which was obtained immediately following the MS step and prior to subsequent RMC-MS.

The results presented in Figure 3 reveal several interesting outcomes. First, somewhat surprisingly, relaxing the ZnO structure to its lowest energy state using MS relaxations between RMC cycles does not substantially increase the MSE in Figure 3a, but in fact in this instance reduces the MSE relative to RMC-only calculations for ≥3 RMC-MS loops. This suggests that the use of RMC modeling alone may produce nonphysical local structure environments that become trapped within the structure over the course of the RMC modeling and prevent improvement in the

model structure overall. Incorporating MS allows the structure to break out of these non-physical local atomic structure arrangements and access atomic structures that are more consistent with the experimental data (lower MSE). Second, incorporating MS between RMC translation cycles improves agreement of the calculated bond length (Figure 3b) and CN (Figure 3c) with reported literature.[58] Crystalline ZnO is expected to have a Zn-O first-coordination-sphere pair distance of ~2.0 Å, and exhibit tetrahedral Zn-O coordination (CN=4).[40,59] RMC-only modeling yields an average Zn-O bond length of 2.07 Å and an average Zn-O CN of 3.70. These values are 4% larger, and 8% smaller than the maximum expected values, respectively. RMC-MS modeling yields an average Zn-O bond length of 2.01 Å and an average Zn-O CN of 4.03. These values within 0.5% and 0.8% of the expected values, respectively. Lastly, and perhaps most interesting, is the observation that including MS relaxations reduces the cohesive energy of the resulting model structures as depicted in Figure 3d. After 3 RMC-MS loops, the cohesive energy is -4.50 eV/atom vs. -4.13 eV/atom using RMC only, corresponding to 8% lower energy. Following 5 RMC-MS loops, the cohesive energy is 0.56 eV/atom lower than the value obtained using the same number of RMC translation steps without MS relaxation. This highlights that RMC-only modeling introduces non-physical local structures that increase the overall energy of the system, and that inserting MS relaxations into the RMC modeling process removes these non-physical local structures and produces more energetically favorable and physically-relevant structures. It should be emphasized that the choice of pair potential function is crucial to obtaining physically accurate structures. For comparison, we implemented the scheme in Figure 3 using a Tersoff[62] potential (not shown) in place of the ReaxFF potential and obtained a final bond length of 2.08 Å, a CN of 2.9, and E_c of -2.5 eV/atom. In terms of these structural metrics, RMC-only modeling more closely captures the expected structural metrics than the use of a simplistic (Tersoff) potential within the iterative RMC-MS approach.

We note that if the RMC-MS iteration process ends on an MS step, the system relaxes into more thermodynamically preferred state, but one that is not consistent with experimental observations. The blue trace in Figure 3d reflects the results from ending in a MS step. Although this leads to a 0.46 eV/atom lower energy, the MSE for the structure ending on an MS step after five RMC-MS loops is 24.5% -- a factor of ~6 larger than the MSE of 4.2% from ending on an RMC step. Because of this, each RMC-MS simulation ends in an RMC step. We also note that the decrease in cohesive energy using RMC-MS relative to RMC-only arises in part from changes in the total cell volume during the MS relaxation, which is not implemented in the RMC modeling. However, this is not the only factor driving a decrease in cohesive energy. In Figure 3d, the cohesive energy for the first two RMC-MS loops is equivalent to the cohesive energy for the first two RMC-only loops. If the volume change from MS relaxation alone significantly impacted the cohesive energy, the cohesive energy with no volume change (RMC-only) would be expected to be substantially different from the cohesive energy with a volume change (RMC-MS)—even on the first and second modeling loops. Based on this, we expect that the difference in cohesive energy between RMC and RMC-MS modeling arises because of non-physical local structure artifacts in the RMC model structures, which are removed upon MS relaxation.

To confirm this understanding, we examine differences in local structure and local energy between RMC and RMC-MS models. In Figure 4, we depict an example of the differences in local structure

observed between RMC (Figure 4a) and RMC-MS (Figure 4b) modeling within a cube of 8 Å side length (from the origin) following the 5 modeling loops represented in Figure 3. These sample structures are qualitatively different. The sample simulation volume from RMC-only modeling (Figure 4a) includes over-coordinated Zn atoms (Figure 4a-1), clustering of O atoms (Figure 4a-2), and under-coordinated O (Figure 4a-3). In contrast, the structure derived from RMC-MS modeling (Figure 4b) depicts a structure that is closer to the expected crystal structure for ZnO, with predominantly tetrahedrally coordinated Zn centers in repeating units with a relatively low amount of disorder. We emphasize that the more ordered structure depicted in Figure 4b corresponds to a lower MSE (Figure 3a), and has a (~1 eV/atom) lower cohesive energy for the atoms shown than Figure 4a, indicating that this structure is both (a) more consistent with the experimental PDF data and (b) more consistent with physical expectations. As such, it is a more probable representation of the actual atomic structure in the physical sampled measured by HE-XRD.

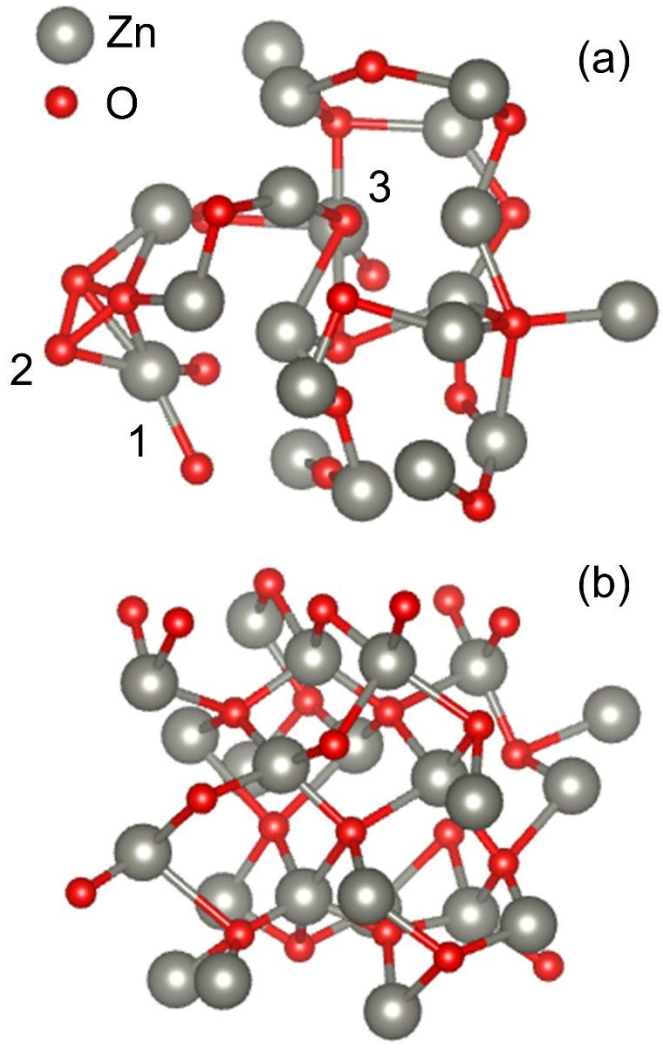

Figure 4: A portion of the simulated ZnO atomic structure for 45 ALD growth cycles of ZnO. Panel (a) shows the model generated by RMC; panel (b) shows the model using RMC-MS. Labels 1, 2, and 3 in panel (a) mark over-coordinated Zn, clustering of O atoms, and under-coordinated O, respectively.

In Figure 5, we expand on the small region of local structure presented in Figure 4, and depict 3-D representations showing high-energy atoms throughout the entire simulation volume for both RMC (Figure 5a) and RMC-MS (Figure 5b) modeling approaches as well as ring size distributions for each (Figure 5c-d). In Figure 5a-b, we isolate and depict only atoms having cohesive energy values ≥-2 eV/atom and ≤ 0 eV/atom, where the atom color represents the specific energy of each of these high-energy atoms. We note that cohesive energies > 0 eV/atom were omitted for clarity in this figure due to the small fraction of atoms having energy greater than this value (f=0.006 for RMC and f=0.003 for RMC-MS). This allows us to visualize how prevalent high-energy local structure regions are in the atomic models derived from each approach, and how these atoms are distributed throughout the structure. We note that the RMC-MS relaxations allow for changes in the total cell volume as well as atom removal and result in an extension of the dimensions of the simulations box for this modeling. The number of atoms in the RMC structure with an energy ≥-2 eV/atom is 8720, and the number of atoms in the RMC-MS structure with an energy ≥-2 eV/atom is 6892. Dividing these numbers by the total number of atoms in each simulation, the fraction (f) of atoms having an energy E ≥-2 eV is 0.4 for the RMC model and 0.3 for the RMC-MS model. This corresponds to a 33% increase in the number of high energy atoms using RMC-only modeling relative to RMC-MS. These atoms with E ≥-2 eV contribute 0.1 eV/atom of the overall ~0.6 eV/atom higher cohesive energy for the RMC structure. We also note that high energy atoms in Figure 5a (RMC) appear to be randomly distributed throughout the simulation volume, whereas high energy atoms in Figure 5b (RMC-MS) seem to appear as localized clusters. This clustering of high-energy atoms is consistent with grain boundaries expected to form between nanocrystalline domains in ZnO.[40] The differences in high-energy atom fractions are reflected in the ring-size distributions (Figure 5c-d). Using Guttman's ring definition,[63] we expect ZnO grown by ALD to be in the wurtzite phase and contain predominately 6-membered rings.[64,65] Indeed, we find that the RMC-MS approach (Figure 5d) increases the fraction of 6-membered rings by 8% relative to RMC (Figure 5c).

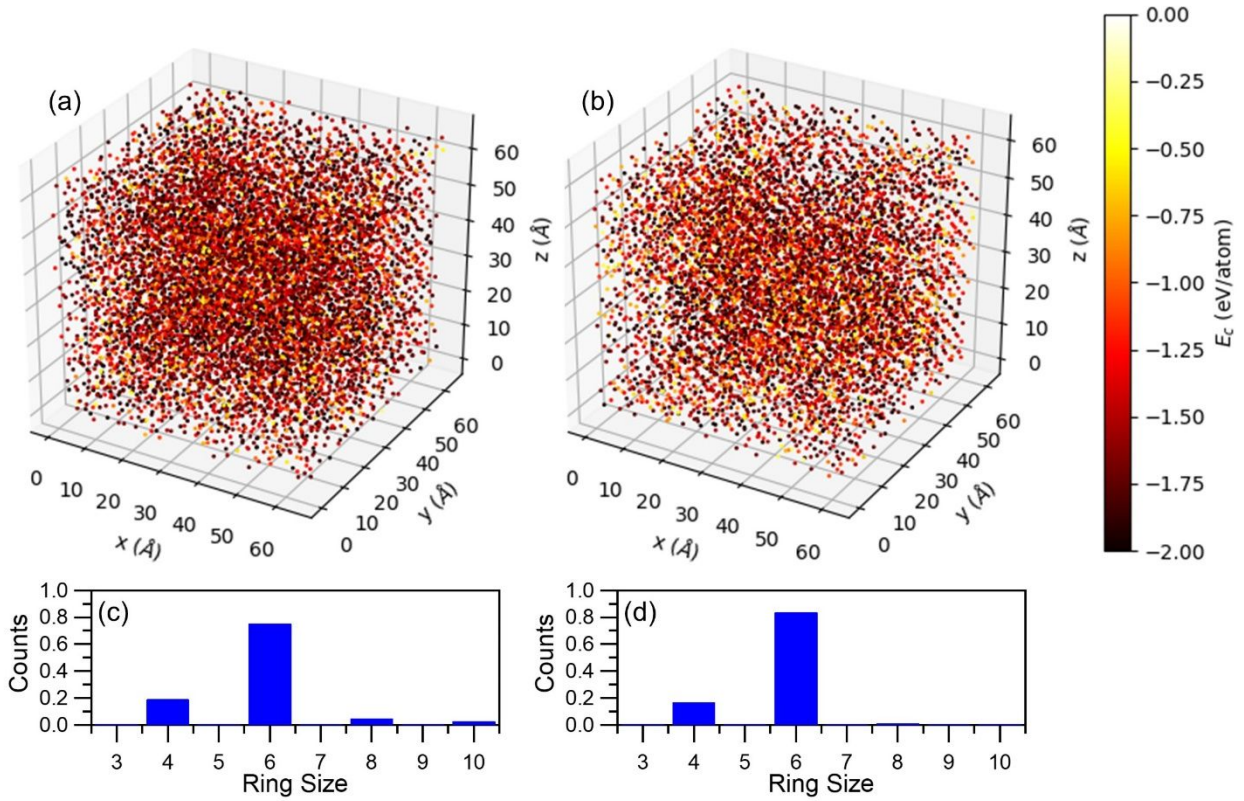

Figure 5: Locations of high-energy atoms within the simulation domain for (a) RMC modeling and (b) RMC-MS modeling and ring-size distributions for (c) RMC modeling and (d) RMC-MS modeling of ZnO at 45 ALD cycles.

From the above analysis, we conclude that the use of alternating RMC-MS modeling steps in sequence provides substantial improvement in the physical accuracy of model structures without decreasing the MSE of the fit to experimental data. As such, we expect that the atomic structures derived from ZnO using RMC-MS modeling are more accurate representations of the actual atomic structure of the experimental samples, and are more relevant for subsequent analysis to understand material structure and properties. Based on the trends in the results presented in Figure 3, we expect that at least three RMC-MS loops are necessary to provide such an improvement over RMC-only modeling, with diminishing, but still significant, returns in model results up to five RMC-MS loops. Based on this analysis, we remodeled a portion of the ZnO structures in Figure 2 from 5 to 45 ALD cycles in 10 ALD cycle increments using five RMC-MS loops—equivalent to the modeling procedures described above for 45 ALD growth cycles. We depict the results of this structural modeling in Figure 6.

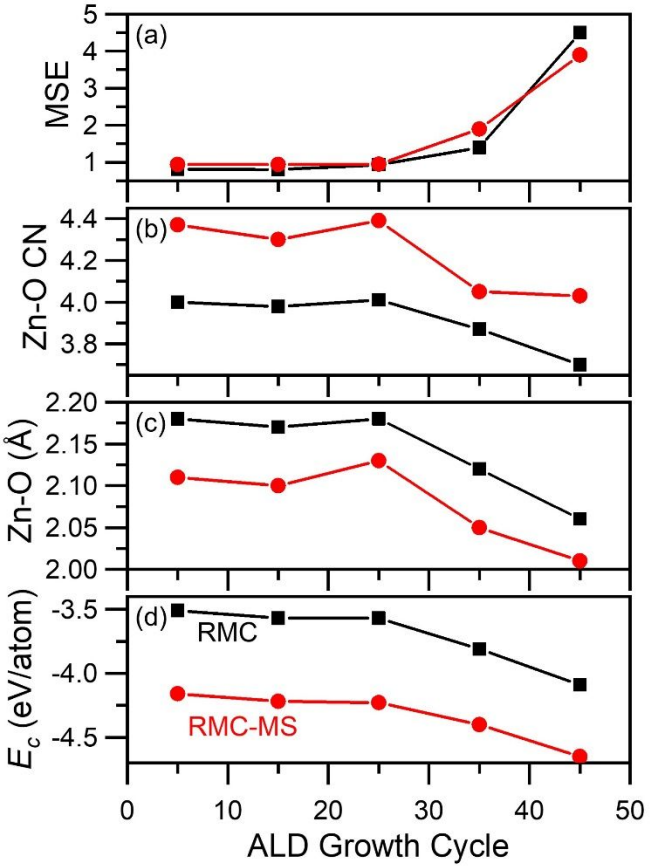

Figure 6: Comparison of the simulated structural properties with and without energy minimization across multiple growth cycles. Panel (a) shows the MSE between the experimental and model PDF, (b) and (c) show changes in average CN and bond length, respectively, and (d) gives the cohesive energy difference by ALD cycle.

Comparing the RMC and RMC-MS modeling in Figure 6, we find that the RMC-MS provides qualitatively similar results to RMC-only modeling but that RMC-MS results are more consistent with physical expectations. In Figure 6a, the MSE is equivalent between RMC and RMC-MS modeling for all numbers of ZnO ALD growth cycles, indicating that both RMC and RMC-MS results are equally consistent with the experimental data. The qualitative trends in Zn-O bond length (Figure 6b) and Zn-O CN (Figure 6c) are also consistent between RMC and RMC-MD modeling. Figure 6b shows a decrease in Zn-O bond length and Figure 6c shows a decrease in Zn-O CN with an increasing number of ALD growth cycles, regardless of the modeling approach. However, the exact values of Zn-O bond length and Zn-O CN are substantially different between RMC and RMC-MS. On average across all ALD growth cycles, Zn-O bond lengths derived from RMC modeling are 3.0(±0.5)% larger than Zn-O bond lengths derived from RMC-MS modeling. Likewise, Zn-O CN values derived from RMC modeling are 7.4(±2.0)% smaller than Zn-O CN values derived from RMC-MS modeling. These differences in local atomic structure metrics are reflected in the total cohesive energy, where RMC-MS simulations produce structures which are 0.62(±0.04) eV/atom lower in energy than RMC-only structures. The lower cohesive energy using combined RMC-MS indicates a more physically accurate structural model.

Above, we discussed that RMC-only modeling yielded incorrect Zn-O CN values <4 for thick ZnO films (>30 ALD ZnO growth cycles). Here, we find that RMC-MS produces Zn-O CN values more in line with our expectations. As the number of ALD growth cycles increases, we would expect the structure to be more consistent with bulk ZnO. Indeed, the value of Zn-O CN derived from RMC-MS modeling is 4.05 at 35 ALD ZnO growth cycles and 4.03 at 45 ALD ZnO growth cycles, consistent with literature value of 4.0 expected for bulk ZnO.[59] This suggests that the RMC-MS modeling accurately captures the bulk ZnO structure, and therefore gives us more confidence in the atomic structure metrics derived from the RMC-MS models and allows us to interpret how the atomic structure evolves during early-stage nucleation at a lower number of ALD cycles before a bulk structure is established.

Examining the RMC-MS results versus the number of ALD ZnO growth cycles in Figure 6, we can understand how the atomic structure of ZnO evolves with increasing ALD cycles. We expect that the ZnO structure at the film-substrate and film-atmosphere interfaces will be different from the bulk ZnO structure. At thinner ZnO film thicknesses, we expect the measured structural metrics in Figure 6 to be representative of ZnO at the film-substrate and film-atmosphere interfaces. At thicker ZnO film thicknesses (higher number of ALD cycles), we expect these interfaces to contribute less, and that the structural metrics will reflect bulk ZnO. The evolution of the structural metrics with an increasing number of ALD growth cycles can then be analyzed to elucidate the structure present at the interfaces versus the bulk ZnO structure. In Figure 6b, we see that as the number of ALD ZnO growth cycles increases, the Zn-O bond length decreases from 2.12 Å at 5 ALD growth cycles to 2.01 Å at 45 growth cycles. Likewise in Figure 6c, we see that as the number of ALD ZnO growth cycles increases, the Zn-O CN decreases from 4.35 at 5 ALD growth cycles to 4.03 at 45 ALD growth cycles. In Figure 7, we depict a schematic of the qualitative interpretation of these trends. Here, the material structure comprises two phases of material which evolve in fractional composition with increasing ALD growth cycles. Specifically, we suggest that at the interfaces of ZnO with vacuum and/or CNT substrate surface, a ZnO phase exists that contains over-coordinated Zn-O with longer Zn-O bond lengths relative to the bulk ZnO phase. During early-stage nucleation, the structural features of the *interfacial* phase contribute a larger amount to the ensemble average structure measurements by HE-XRD. As the number of ALD growth cycles increases, the *bulk* ZnO phase contributes a larger fraction to the ensemble average structure measurement by HE-XRD, driving the Zn-O bond length to the bulk value of 2.0 Å and the CN to the bulk value of 4.0. This general two-phase interpretation of ALD growth behavior is consistent with our prior observations on ALD growth of AlO_x using trimethylaluminum and water and suggests that such two-phase behavior may be a more widespread general phenomenon during ALD growth.[11]

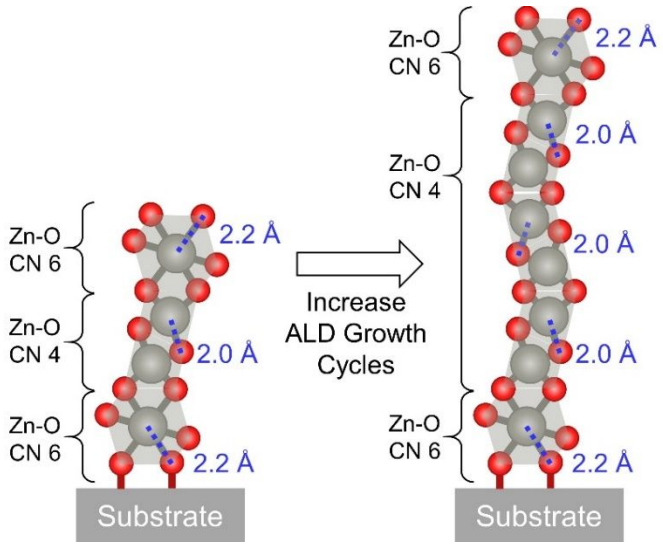

Figure 7: Schematic depiction of the evolution in local structural metrics of ZnO with increasing ALD growth cycles.

We emphasize that the iterative RMC-MS modeling improved the accuracy of the structural metrics derived from experimental pair distribution function data and allowed us to correctly interpret the evolution of ZnO atomic structure during ALD growth. Although RMC-only modeling produced the same general trends in the structural metrics with increasing number of ALD growth cycles, we were at an impasse as to the physical interpretation of this data because bulk values of CN from RMC modeling were significantly lower than the expected value. Without the use of RMC-MS, we may have concluded, for example, that bulk oxygen vacancies were present in significant concentration to yield under-coordinated Zn-O centers. However, this interpretation is not supported by physical understanding. Oxygen vacancy formation energies are predicted to be > 3.8 eV/atom in ZnO, and are also found to increase in formation energy with increasing crystallite size.[66–68] In this sense, our present study suggests that *quantitative* interpretation of RMC-only modeling results (that do not employ careful physical constraints) may result in erroneous physical interpretation of material structures.

We note that due to the lack of features in the PDF beyond ~5 Å for a low number of ALD cycles, a greater number of configurations are expected to satisfy the model constraints. Based on this, we considered that five RMC-MD iterations may not be sufficient to reduce the structure to its most reasonable state for a small number of ALD growth cycles. To examine this, we examined the change in cohesive energy versus the number of RMC-MS looped iterations for 5, 15, and 25 ALD ZnO growth cycles, as depicted in Figure 8. This data indicates that the cohesive energy converges to < 0.1 eV/atom within five RMC-MS loops, just as we observed for 45 ALD growth cycles presented above. This suggests that five RMC-MS loops are sufficient to provide substantial improvement in the atomic structure models for each condition, and supports the above two-phase interpretation of the evolution of the ZnO atomic structure depicted in Figure 7.

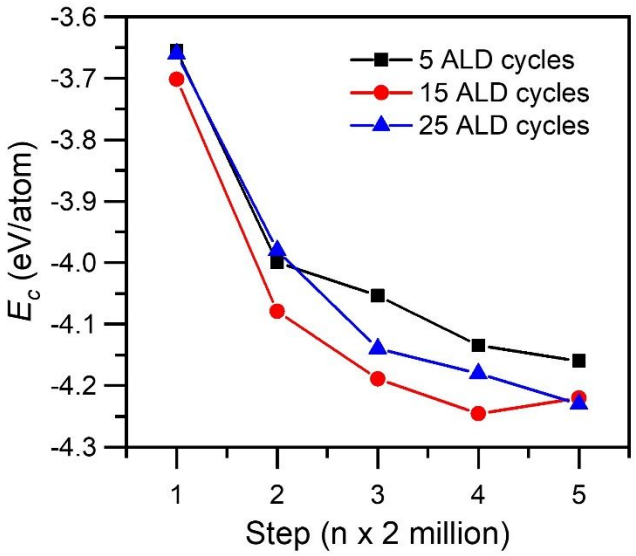

Figure 8: Cohesive energy per 2 million RMC translations for 5, 15, and 25 ALD cycles during alternating RMC-MS modeling.

Finally, we note that the iterative RMC-MS method we report here employs a total of 10^7 RMC steps with an additional ~500 MS calculations to achieve the reported convergence criteria and structural metrics. Our simple approach dramatically improves the physical accuracy of the modeled structures over RMC-only modeling with minimal additional computational cost. By incorporating a sophisticated pair potential in MS steps, we capture the benefits of using energy constraints while avoiding both the computational burden of hybrid RMC-MD techniques and non-physical structural metrics introduced by using an inadequate (e.g., Tersoff) potential function. Similarly, by using MS instead of MD minimization, we remove all dynamic variables (temperature, velocities, quenching rates, etc.) from the simulation and further simplify the modeling process.

5. Conclusions

The inclusion of MS relaxations in RMC modeling in iterative RMC-MS steps was found to enhance the physical accuracy of model structures. By periodically imposing MS conditions after a number of RMC steps, we find that the cohesive energy of model structures of ALD-grown ZnO are reduced by 0.62 eV relative to RMC-only model structures, while maintaining equivalent agreement with experimentally determined pair distribution function data. This approach maximizes the aforementioned benefits of RMC and MS without sacrificing physical accuracy or significantly increasing computational expense, and provides more accurate atomic structure models from diffraction data.

Although further work is needed to explore the wide range of possible implementations combining iterative RMC and MS steps, the initial study presented in this work suggests implementation of an MS energy minimization every two million RMC perturbations for a minimum of three complete iterations. This procedure provided sufficient atom translations to match the simulated PDF to the experimental PDF without significantly distorting the model structure. Likewise, our

initial studies indicate that three iterations of RMC and MS loops produces a significant improvement in cohesive energy, while five or more iterations are necessary for the model structure to converge to less than 0.1 eV/atom with respect to cohesive energy.

Using combined RMC-MS modeling, we analyze *in operando* evolution of ZnO atomic structure during ALD growth using synchrotron HE-XRD. We reveal a general two-phase growth behavior consistent with observations for ALD growth of Al_2O_3 using trimethylaluminum and water.[11] The iterative RMC-MS method established in this work is easily applicable to the expanding field of ALD research and can provide valuable information on topics such as growth behavior, crystal structure, surface interactions, and material stability.[69–71]

6. Acknowledgements

The authors acknowledge partial support from the United States Geological Survey, Missouri Water Resource Research Center through Award Number G16AP00066, and Great Rivers Cooperative Ecosystems Studies Unit through Award Number G21AC10041. The use of the 11-ID-C beamline of the Advanced Photon Source (APS) is supported under Contract DE-AC02-06CH11357. HDK acknowledges support from the University of Missouri Discovery Fellows Program. We would like to thank Tommy Sewell and Karl Hammond at the University of Missouri for advice on MD modeling and Predrag Lazic at the University of Missouri for high performance computing support. We would also like to thank Uta Ruett, Olaf Borkiewicz, Yang Ren, Richard Spence, Kevin Beyer, Guy Jennings, Andrey Yakovenko, Charles Kurtz, Angel Yanguas-Gil, and Nicholas M. Bedford for assistance with *in operando* measurements at 11-ID-C.

7. References

[1] S.-H. Wei, Comput. Mater. Sci. **30**, 337 (2004).

[2] H.J. Queisser, Science (80-.). **281**, 945 (1998).

[3] P. Thompson, D.E. Cox, and J.B. Hastings, J. Appl. Crystallogr. **20**, 79 (1987).

[4] J.E. Post and D.L. Bish, in *Mod. Powder Diffr.*, edited by D.L. Bish and J.E. Post (De Gruyter, Berlin, Boston, 1989), pp. 277–308.

[5] S.J.L. Billinge and M.F. Thorpe, editors , *From Semiconductors to Proteins: Beyond the Average Structure* (Springer US, Boston, MA, 2002).

[6] S. Lee and H. Xu, Minerals **10**, 124 (2020).

[7] T. Proffen, S.J.L. Billinge, T. Egami, and D. Louca, Zeitschrift Für Krist. - Cryst. Mater. **218**, 132 (2003).

[8] C.R. Müller, V. Kathriarachchi, M. Schuch, P. Maass, and V.G. Petkov, Phys. Chem. Chem. Phys. **12**, 10444 (2010).

[9] G. Opletal, D.W. Drumm, R.P. Wang, and S.P. Russo, J. Phys. Chem. A **118**, 4790 (2014).

[10] I. Pethes and L. Pusztai, J. Chem. Phys. **146**, 064506 (2017).

[11] M.J. Young, N.M. Bedford, A. Yanguas-Gil, S. Letourneau, M. Coile, D.J. Mandia, B. Aoun, A.S. Cavanagh, S.M. George, and J.W. Elam, ACS Appl. Mater. Interfaces **12**, 22804 (2020).

[12] R.L. McGreevy, J. Phys. Condens. Matter **13**, R877 (2001).

[13] S.K. Jain, R.J.-M. Pellenq, J.P. Pikunic, and K.E. Gubbins, Langmuir **22**, 9942 (2006).

[14] A. Vrhovšek, O. Gereben, A. Jamnik, and L. Pusztai, J. Phys. Chem. B **115**, 13473 (2011).

[15] C. Russ, M. Brunner, C. Bechinger, and H.H. von Grünberg, Europhys. Lett. **69**, 468 (2005).

[16] M.G. Tucker, D.A. Keen, M.T. Dove, A.L. Goodwin, and Q. Hui, J. Phys. Condens. Matter **19**, 335218 (2007).

[17] L. Pusztai, I. Harsányi, H. Dominguez, and O. Pizio, Chem. Phys. Lett. **457**, 96 (2008).

[18] V. Mile, L. Pusztai, H. Dominguez, and O. Pizio, J. Phys. Chem. B **113**, 10760 (2009).

[19] O. Gereben and L. Pusztai, J. Comput. Chem. **33**, 2285 (2012).

[20] N.A. Katcho, P. Zetterström, E. Lomba, J.F. Marco, E. Urones-Garrote, D. Ávila-Brande, A. Gómez-Herrero, L.C. Otero-Díaz, and A.R. Landa-Cánovas, Phys. Rev. B **77**, 195402 (2008).

[21] T.X. Nguyen, N. Cohaut, J.-S. Bae, and S.K. Bhatia, Langmuir **24**, 7912 (2008).

[22] O. Gereben and L. Pusztai, J. Non. Cryst. Solids **407**, 213 (2015).

[23] A.K. Soper, Chem. Phys. **202**, 295 (1996).

[24] A.K. SOPER, Mol. Phys. **99**, 1503 (2001).

[25] A. Pandey, P. Biswas, and D.A. Drabold, Phys. Rev. B **92**, 155205 (2015).

[26] C. Bousige, A. Boţan, F.-J. Ulm, R.J.-M. Pellenq, and B. Coasne, J. Chem. Phys. **142**, 114112 (2015).

[27] A.F. Palmstrom, P.K. Santra, and S.F. Bent, Nanoscale **7**, 12266 (2015).

[28] H. Frankenstein, C.Z. Leng, M.D. Losego, and G.L. Frey, Org. Electron. **64**, 37 (2019).

[29] C. Prasittichai and J.T. Hupp, J. Phys. Chem. Lett. **1**, 1611 (2010).

[30] B.J. O'Neill, D.H.K. Jackson, J. Lee, C. Canlas, P.C. Stair, C.L. Marshall, J.W. Elam, T.F. Kuech, J.A. Dumesic, and G.W. Huber, ACS Catal. **5**, 1804 (2015).

[31] S. Xu, Y. Kim, J. Park, D. Higgins, S.-J. Shen, P. Schindler, D. Thian, J. Provine, J. Torgersen, T. Graf, T.D. Schladt, M. Orazov, B.H. Liu, T.F. Jaramillo, and F.B. Prinz, Nat. Catal. **1**, 624 (2018).

[32] Y.S. Jung, A.S. Cavanagh, L. a. Riley, S.-H.H. Kang, A.C. Dillon, M.D. Groner, S.M. George, and S.-H.H. Lee, Adv. Mater. **22**, 2172 (2010).

[33] X. Meng, X.-Q. Yang, and X. Sun, Adv. Mater. **24**, 3589 (2012).

[34] A. Shearrow, G. Koolstra, S.J. Whiteley, N. Earnest, P.S. Barry, F.J. Heremans, D.D. Awschalom, E. Shirokoff, and D.I. Schuster, Appl. Phys. Lett. **113**, 1 (2018).

[35] M.J. Young, D. Choudhury, S. Letourneau, A. Mane, A. Yanguas-Gil, and J.W. Elam, Chem. Mater. **32**, 992 (2020).

[36] X. He, R.Z. Waldman, D.J. Mandia, N. Jeon, N.J. Zaluzec, O.J. Borkiewicz, U. Ruett, S.B. Darling, A.B.F. Martinson, and D.M. Tiede, ACS Nano **14**, 14846 (2020).

[37] S. Letourneau, M.J. Young, N.M. Bedford, Y. Ren, A. Yanguas-Gil, A.U. Mane, J.W. Elam, and E. Graugnard, ACS Appl. Nano Mater. **1**, 4028 (2018).

[38] A.M. Jasim, X. He, Y. Xing, T.A. White, and M.J. Young, ACS Omega **6**, 8986 (2021).

[39] V. Sharma, S.D. Elliott, T. Blomberg, S. Haukka, M.E. Givens, M. Tuominen, and M. Ritala, Chem. Mater. **33**, 2883 (2021).

[40] R. Boichot, L. Tian, M.-I. Richard, A. Crisci, A. Chaker, V. Cantelli, S. Coindeau, S. Lay, T. Ouled, C. Guichet, M.H. Chu, N. Aubert, G. Ciatto, E. Blanquet, O. Thomas, J.-L. Deschanvres, D.D. Fong, and H. Renevier, Chem. Mater. **28**, 592 (2016).

[41] J.A. Klug, M.S. Weimer, J.D. Emery, A. Yanguas-Gil, S. Seifert, C.M. Schlepütz, A.B.F. Martinson, J.W. Elam, A.S. Hock, and T. Proslier, Rev. Sci. Instrum. **86**, 113901 (2015).

[42] A.W. Ott and R.P.H. Chang, Mater. Chem. Phys. **58**, 132 (1999).

[43] A. Yamada, B. Sang, and M. Konagai, Appl. Surf. Sci. **112**, 216 (1997).

[44] V. Lujala, J. Skarp, M. Tammenmaa, and T. Suntola, Appl. Surf. Sci. **82**–**83**, 34 (1994).

[45] D.A. Keen, J. Appl. Crystallogr. **34**, 172 (2001).

[46] B.H. Toby and R.B. Von Dreele, J. Appl. Crystallogr. **46**, 544 (2013).

[47] P. Juhás, T. Davis, C.L. Farrow, and S.J.L. Billinge, J. Appl. Crystallogr. **46**, 560 (2013).

[48] B. Aoun, J. Comput. Chem. **37**, 1102 (2016).

[49] C.J. Tainter and G.C. Schatz, J. Phys. Chem. C **120**, 2950 (2016).

[50] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).

[51] A.C.T. van Duin, A. Strachan, S. Stewman, Q. Zhang, X. Xu, and W.A. Goddard, J. Phys. Chem. A **107**, 3803 (2003).

[52] K. Chenoweth, A.C.T. van Duin, and W.A. Goddard, J. Phys. Chem. A **112**, 1040 (2008).

[53] D. Raymand, A.C.T. van Duin, M. Baudin, and K. Hermansson, Surf. Sci. **602**, 1020 (2008).

[54] S. Le Roux and V. Petkov, J. Appl. Crystallogr. **43**, 181 (2010).

[55] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graph. **14**, 33 (1996).

[56] K. Momma and F. Izumi, J. Appl. Crystallogr. **44**, 1272 (2011).

[57] J.W. Elam and S.M. George, Chem. Mater. **15**, 1020 (2003).

[58] A. Wander and N.. Harrison, Surf. Sci. **457**, L342 (2000).

[59] M. Zobel, A. Windmüller, E.M. Schmidt, K. Götz, T. Milek, D. Zahn, S.A.J. Kimber, J.M. Hudspeth, and R.B. Neder, CrystEngComm **18**, 2163 (2016).

[60] G. Weirum, G. Barcaro, A. Fortunelli, F. Weber, R. Schennach, S. Surnev, and F.P. Netzer, J. Phys. Chem. C **114**, 15432 (2010).

[61] P. Erhart, A. Klein, and K. Albe, Phys. Rev. B **72**, 085213 (2005).

[62] P. Erhart, N. Juslin, O. Goy, K. Nordlund, R. Müller, and K. Albe, J. Phys. Condens. Matter **18**, 6585 (2006).

[63] L. Guttman, J. Non. Cryst. Solids **116**, 145 (1990).

[64] E. Guziewicz, M. Godlewski, T. Krajewski, Ł. Wachnicki, A. Szczepanik, K. Kopalko, A. Wójcik-Głodowska, E. Przeździecka, W. Paszkowicz, E. Łusakowska, P. Kruszewski, N. Huby, G. Tallarida, and S. Ferrari, J. Appl. Phys. **105**, 122413 (2009).

[65] V. Miikkulainen, M. Leskelä, M. Ritala, and R.L. Puurunen, J. Appl. Phys. **113**, 021301 (2013).

[66] A. Janotti and C.G. Van de Walle, Appl. Phys. Lett. **87**, 122102 (2005).

[67] A. Janotti and C.G. Van de Walle, Reports Prog. Phys. **72**, 126501 (2009).

[68] J. Liu, F. Gao, L. Wu, H. Zhang, W. Hong, G. Jin, Z. Zhai, and C. Fu, Appl. Phys. A **126**, 454 (2020).

[69] G. Dogan, U.T. Sanli, K. Hahn, L. Müller, H. Gruhn, C. Silber, G. Schütz, C. Grévent, and K. Keskinbora, ACS Appl. Mater. Interfaces **12**, 33377 (2020).

[70] S.M. Geyer, R. Methaapanon, R.W. Johnson, W.-H. Kim, D.G. Van Campen, A. Metha, and S.F. Bent, Rev. Sci. Instrum. **85**, 055116 (2014).

[71] R. Methaapanon, S.M. Geyer, S. Brennan, and S.F. Bent, Chem. Mater. **25**, 3458 (2013).

Footnote

Electronic supporting information (ESI) is available: full atomic structure models of ALD-grown ZnO as a function of growth cycles are provided for 5, 15, 25, 35, and 45 ALD cycles in xyz file format.