Supplementary Information

Revealing the Hidden Third Dimension of Point Defects in Two-Dimensional MXenes

Grace Guinan, Michelle A. Smeaton, Brian C. Wyatt, Steven Goldy, Hilary Egan, Andrew Glaws, Garritt J. Tucker, Babak Anasori, and Steven R. Spurgeon

Date: October 17, 2025

Contents

Supplementary Note 1: Methods

A. Synthesis

To synthesize the MXenes used in this study, we first synthesized a graphite-based Ti_3AlC_2 MAX phase precursor using molar ratios of 3:2:2 of Ti:Al:C, by first ball milling elemental Ti (325 mesh, Alfa Aesar), aluminum (325 mesh, Alfa Aesar), and graphite at a 2:1 ball-to-powder mass ratio using yttria-stabilized zirconia grinding balls in a high-density polyethylene (HDPE) container. After ball milling this graphite-based Ti_3AlC_2 MAX phase, we followed the typical synthesis protocol for $Ti_3C_2T_x$ MXene, as discussed in-depth in a previous step-by-step article from our group[1]. In short, this mixed powder was placed into an alumina crucible inside an alumina tube furnace and sintered at 1400 °C for 4 h under 100 mL/min Ar flow. After sintering, the MAX block was drilled into a fine powder and washed using 3 M hydrochloric acid (HCl) at a volumetric ratio of 30 mL/g MAX powder overnight to remove any intermetallic impurities. After HCl washing, the powder was washed to a neutral pH using repeated centrifugation using deionized (DI) water, filtered, dried in air, and sieved using a 71 μm pore size sieve. After sieving, the powder was then placed into a solution containing 5% HF (3 mL HF/g MAX), 9.1% HF (6 mL/g MAX), or 12.5% HF (9 mL/g MAX) with 18 mL 12 M HCl/g MAX and 9 mL DI water per gram MAX and stirred at 35 °C for 24 h. After 24 h, the solution was centrifuged to neutralize the pH of the solution and placed into an aqueous solution containing 1 g of LiCl per gram starting MAX to 50 mL of DI water and stirred overnight at room temperature to delaminate the MXene into single flakes. After delamination, the solution was washed again using centrifugation to remove the excess Li, vortex mixed for 30 min and finally run at 2380 RCF acceleration for 30 min to yield the single-to-few flake solution of $Ti_3C_2T_x$ MXenes. These MXenes were then used to prepare for STEM imaging. To quickly screen the quality of these MXenes, X-ray diffraction (XRD) was used (Anton Paar XRDynamic 500, monochromatic Cu Kα radiation, 3 – 65 ° 2θ full range, 25 s/step).

B. STEM Imaging

STEM samples were prepared by diluting MXene solutions and drop casting onto ultrathin carbon on lacey carbon support TEM grids. High-angle annular dark-field (STEM-HAADF) images of the samples were acquired using a probe-corrected Thermo Fisher Scientific Spectra 200 STEM operating at 60 kV with a convergence semi-angle of 30 mrad. The beam current was limited to 10-25 pA to minimize sample damage. STEM-HAADF images were acquired as stacks of low dwell time frames, which were subsequently rigidly aligned to obtain high signal-to-noise ratio (SNR) images. Rigid image registration was performed using a method optimized for very low SNR images, like those collected here, to minimize the possibility of unit cell jump errors, which would inhibit vacancy analysis[2]. To minimize the buildup of carbon contamination during high-resolution imaging, the prepared grids were first cleaned with acetone and methanol and then dried under vacuum for at least 12 hours before loading into the STEM. During imaging, beam showering in TEM mode was used to further prevent contamination buildup. All images are shown in Supplementary Figures S8-10.

C. Machine Learning and Statistical Analyses

We used AtomAI's Segmentor class (atomai.models.Segmentor) as our framework, which defines a U-Net neural network for semantic segmentation of microscopy images.[9]. AtomAI, a Python package built on top of PyTorch for deep learning in microscopy, is typically used to locate the positions of atoms in STEM images and has been applied to identify impurity atoms and defect locations in graphene[9]. Due to the beam sensitivity and surface contamination of our MXene samples it was more beneficial to train a network to pick out defects than atoms, therefore, we approached the training of our models differently than previous graphene work. We trained two models (1) Lattice Capture Neural Network and (2) Defect Spotter Neural Network. We used the first model to locate the positions of atoms if our image was pristine, hexagonal lattice and we used the second model to label vacancies. This two-model approach proved more robust against variations in image contrast and local disorder than a single model trained to identify both atoms and defects simultaneously.

A major challenge of training models in materials science is the lack of labeled training data. Specifically for this study, quality STEM images are of MXenes are difficult and time consuming to take. Therefore, we did not have access to a large training dataset, regardless of whether it was labeled or unlabeled. Instead, we turned to a previous paper[3], and used crops from one high resolution, large frame of view image to train our models. We then used 2D Gaussian fitting to label the data (note we were only able to do this because of the sharpness of the large image and 2D Gaussian fitting was unreliable for our own data), giving us binary masks of atomic positions in the cropped images. The crops were square with length randomly chosen between 150 and 300 pixels, allowing the model to be applicable across a range of resolutions. Crops were then resized to 256 × 256 pixels (2^n being an ideal size for a neural network). Data augmentation, including rotation and gaussian noise, was incorporated as part of the neural network's training process. This step was imperative to the success of the model outside of the training dataset. It was necessary for the model to be successful on MXene images taken on a completely different STEM instrument, allowing for a range of lattice directions, resolutions and sharpness.

The Lattice Capture Neural Network was trained as 3 ensembled models each with 1000 training cycles of batch size 15, trained on 1000 crops of our parent image. Gaussian noise in range [40,60] and rotation were incorporated in the data augmentation. The Defect Spotter Neural Network was one model, trained with 350 training cycles and 300 crops of our parent image. Gaussian noise in range [40,100], rotation, and jitter was incorporated in the data augmentation. All training data and model weights are included in the files and code to run the models is included in Jupyter Notebooks in our Github repository.

Once we obtained all atomic/defect locations, we (1) differentiated the middle (M") from outer (M') layers and (2) combined this with a Delaunay triangulation to define 3D defect motifs. For (1), we enforced a hexagonal structure in our dots using their lattice directions (when finding the dots using ML, the structure was not perfectly hexagonal and sometimes morphed a little, future work could investigate the mechanisms causing this). Then, we separated each dot into its relative layer (i.e. we noted that dot 1 and dot 2 must be in the same layer due to the projection pattern of the layers, but we could not yet say which layer that was). For each image we now had three groups of relative layers, and we counted the number of defects in each of these. We consistently saw one layer with significantly less

defects than the other two and were able to label this layer as the middle layer (M"), since previous theoretical studies[4] have shown defects are more likely in outer layers. The other two layers were both classified as outer layers (M'). We could not differentiate between the upper or lower layers, instead, we categorized the one with more defects as M'_{max} and the one with less as M'_{min}. For (2), we took these layer classifications and conducted Delaunay triangulations on each layer, as well as the projection of all three. This allowed us to understand the connectivity between defects within layers as well as between layers, which led to the motifs we defined in Figure 2 of the main text.

D. Modeling

Using the Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS)[5] we conducted hybrid Monte Carlo (MC) Molecular dynamics (MD) calculations. The bond-order potential (BOP) developed by Plummer et al.[6] simulated all interatomic interactions. We started with a large (30nmx30nm) supercell of pristine MXene (Ti_3C_2) to minimize edge effects and establish statistical significance. The nanosheet was seeded with a random distribution of Ti vacancies, C vacancies, and surface terminations (O and F). Cases were populated with 1.5% and 3.5% Ti vacancies to recreate the experimentally observed conditions. Since carbon is difficult to resolve experimentally, C vacancies were seeded with 0, 1, 5, and 10 %. This is in line with experimental suggestion[6-8] and extends to higher concentrations observed in aged MXenes[9]. Surface termination distributions were also randomly seeded. Preliminary calculations showed that the different species (O, F, and vacancy) orient randomly. Additionally, the O and F termination bonds in the interatomic potential are similar enough that there is not a significant energy difference, so we simplified to a single species (O). We chose three surface termination coverages 0, 50, and 80% in order to be in line with literature[9-11] and examine the impacts of varying levels of coverage on Ti vacancy clustering.

The vacancies are represented by non-interacting ghost atoms. These atoms are placeholders in the lattice that have no velocity or forces to contribute to the energy of the system. Placeholders are needed to perform Monte Carlo (MC) swaps to rapidly sample and compare energy configurations. Previous works have used MC methods to examine point

defects[12-15], and some MD codes use ghost atoms, but their use to track vacancy configurations is a nuance of our approach. Each MC calculation included 100,000 attempted swaps. Every step, one candidate atom is randomly selected from each eligible group. The candidate groups are paired and include Ti and ghost-V_{Ti}, C and ghost-V_C, O and ghost-V_T (surface termination). For each grouped pair of atoms, their positions are reversed and the potential energy calculated according to the interatomic potential. Simply, the potential describes the energy between any two atoms at a given distance under the influence of an outside atom. The sum of these energies is the system potential energy. After the swap, if the potential energy is lower, the swap is accepted. If it is not, then there is a chance of acceptance based on the Metropolis criterion at 300K, which is summarized by Equation 1.

$$R < e^{\frac{-\Delta E_{swap}}{k_B T}} \quad (1)$$

Where R is a randomly generated number between 0 and 1, ΔE_{swap} is the change in energy from the swap, k_B is the Boltzmann constant and T is the temperature[12,16]. If it is still not accepted, then the two candidates are returned to their starting location and another pair is tested. After an accepted swap, that is the new starting configuration for the next step. In this way the distribution progression follows a path of swaps governed by the random seed candidate atom selection and entropic acceptance.

Cases included 100,000 attempted swaps because convergence the potential energy was converged at this point for the 30×30 nm samples. Fig. S4 shows this convergence. The number of swaps needed to show convergence is directly proportional to the number of candidate atoms. Additional studies were conducted with fewer atoms and relatively more swaps, these showed similar potential energy behavior but were not as statistically robust.

Supplementary Note 2: Configuration and Validation of ML Models

Here, we provide a visual of our model training and validation process. Absent ground truth data, it is challenging to validate model performance. It is possible to create simulated data; however, making simulated data resemble experimental data is imprecise and can introduce significant bias. We chose to conduct expert comparison of the performance of our model to 2D Gaussian Fitting (the method we used to train the model). As shown in Supplementary Figure 1b, 2D Gaussian Fitting does not provide informative or regular atomic positions (and cannot find defects). In comparison, our NN enforces a regular, hexagonal lattice, can capture the locations of defects (even for a large cluster of defects, as in the 12.5% HF sample), and is more robust to noise in low-dose images.

Supplementary Figure 1: ML Pipeline

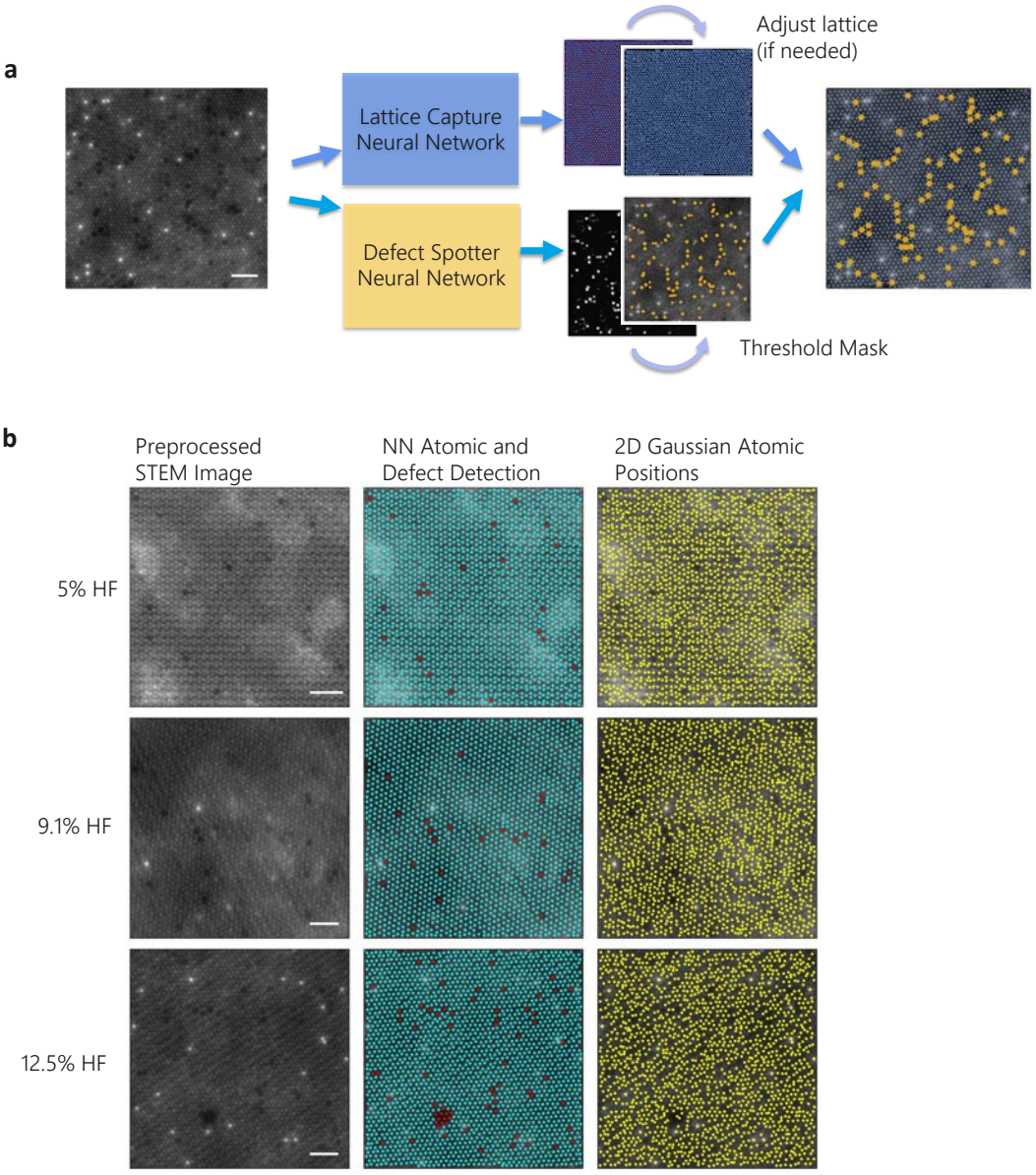

FIG. S1: Model training and performance. (a) Model pipeline. STEM-HAADF images are run through two different neural networks, one to find all atomic positions by enforcing a hexagonal lattice structure (Lattice Capture) and another specifically to detect defect positions (Defect Spotter). Models are combined to find final atomic and defect positions. (b) Performance of our model for three sample images in comparison to 2D Gaussian Fitting. Atomic (blue) and defect (red) positions are visually more regular in comparison to atomic positions found using Atomap's[17] 2D Gaussian Fitting Model (yellow). Scale bars 1 nm.

Supplementary Figure 2: Layer Deconvolution

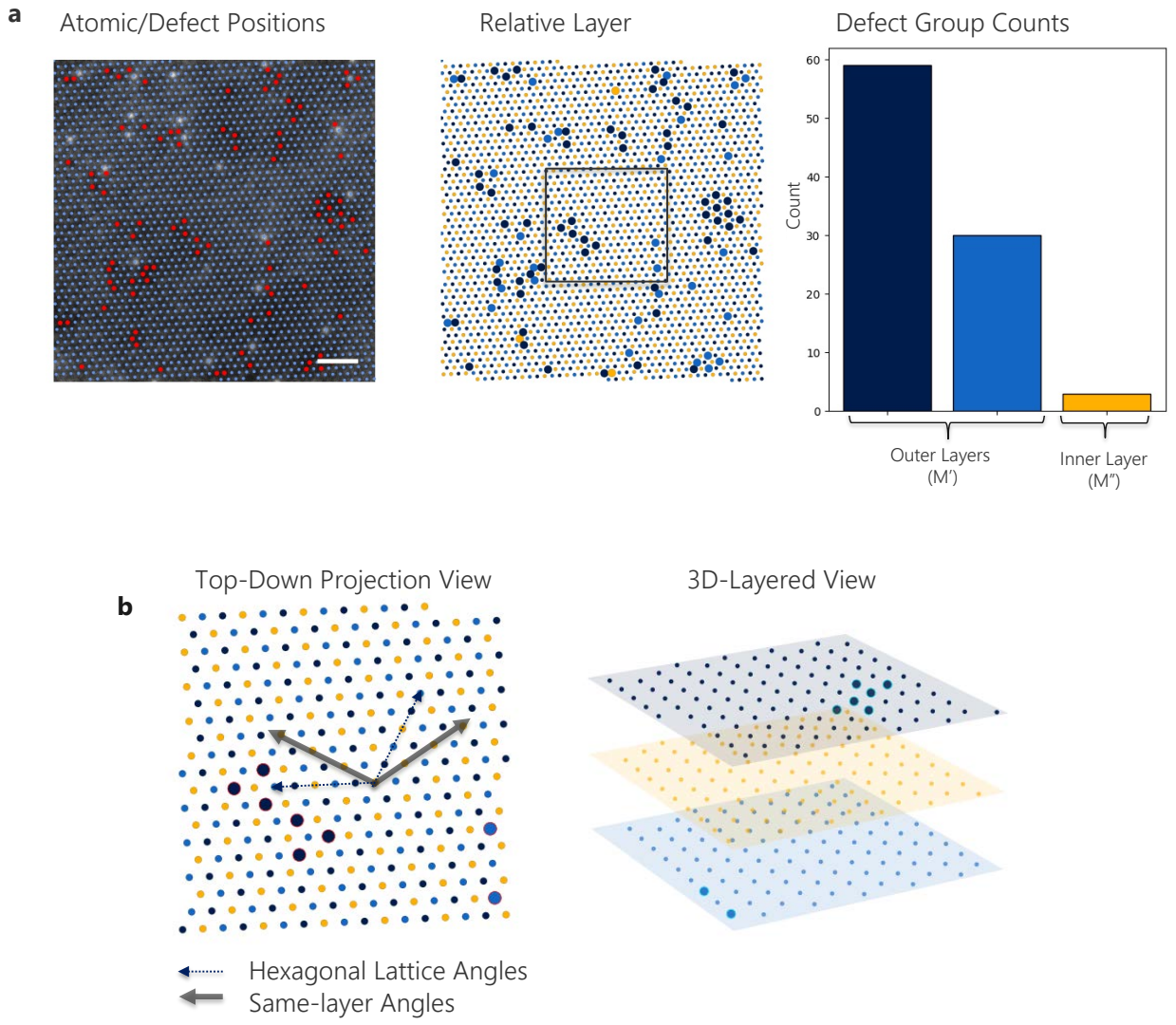

FIG. S2. Layer deconvolution method. (a) Steps to identify outer (M') vs. inner (M") layers, given atomic (blue) and defect (red) positions. Dots are grouped by relative layer and histogram of defect groups reveals outer vs. inner layers. Scale bar 1 nm. (b) Close up of relative layer labeling. When looking at $Ti_3C_2T_x$ MXene in top-down projection view, colored by layer, same-layer angles are 30 degrees off from hexagonal lattice angles. These three layers can be discriminated in the 3D-Layered View with outer layers M' (light/dark blue) and middle layer M" (orange).

Supplementary Figure 3: Delaunay Triangulation

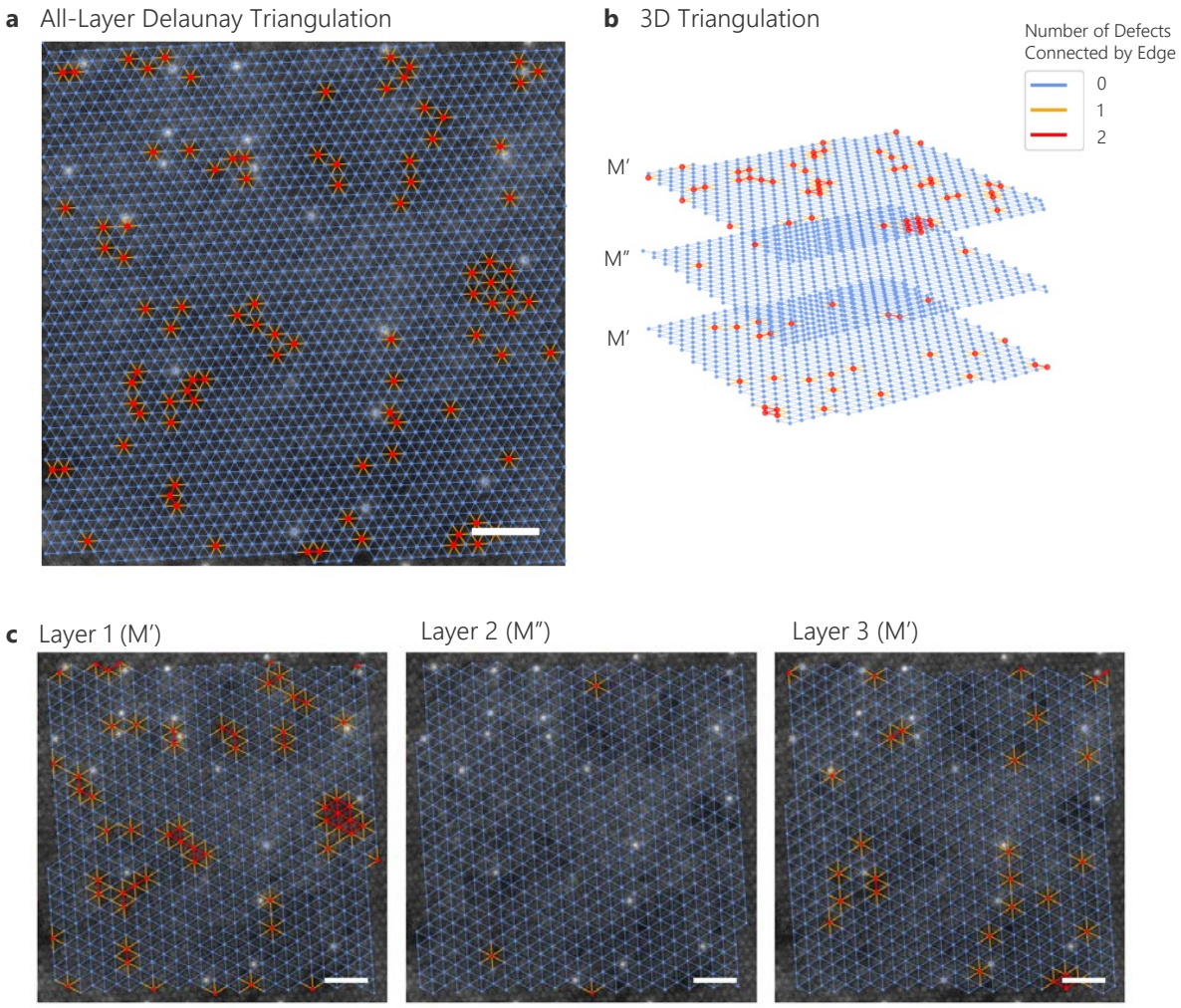

FIG. S3. Delaunay triangulation to identify motifs. (a) Delaunay triangulation on all atomic and defect positions. Red lines indicate adjacent defects (across layers). (b) Visualizing Delaunay triangulation in 3D. (c). Separating out the three layers. Now red lines indicate adjacent defects within layers. By applying the Delaunay triangulation, we can understand how defects form within and between layers. Scale bars 1 nm.

Supplementary Figure 4: Bootstrapping Analysis

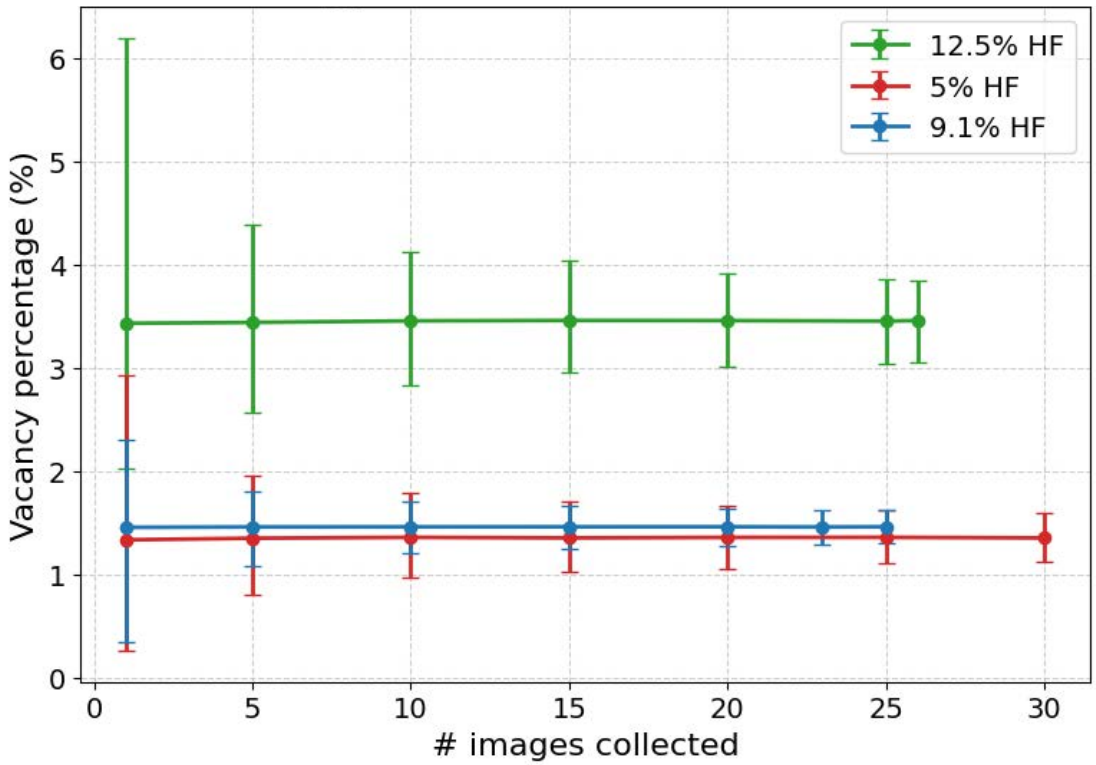

FIG. S4. Bootstrapping across number of images. Bootstrapped vacancy percentage calculations, showing 95% confidence interval across HF concentrations. By adding images, we can tighten 95% confidence bars around all three samples and differentiate the vacancy concentration in the 12.5% HF sample from 5% and 9% HF samples. However, the error bars of the 5% and 9.1% HF samples are still overlapping.

Supplementary Figure 5: XRD Measurements

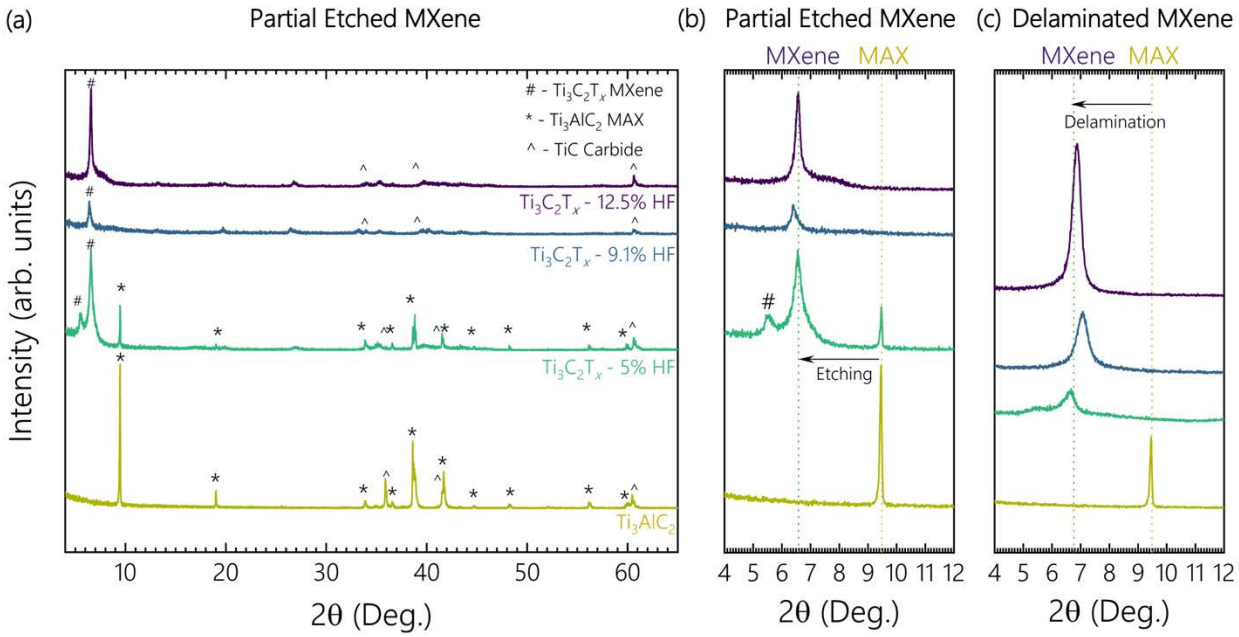

FIG. S5. X-ray diffraction (XRD) of $Ti_3C_2T_x$ MXenes from the Ti_3AlC_2 used in this study. (a-b) The $Ti_3C_2T_x$ MXenes synthesized using 5% HF demonstrate partial etching, as shown by the presence of both MAX and MXene peaks, while the 9.1 and 12.5% HF show only MXene after etching. As the MXenes undergo a delamination process, the (c) final fully etched $Ti_3C_2T_x$ MXenes used in this study for imaging were separated from both remaining MAX and impurity phases.

Supplementary Figures 6-7: Additional Modeling Figures

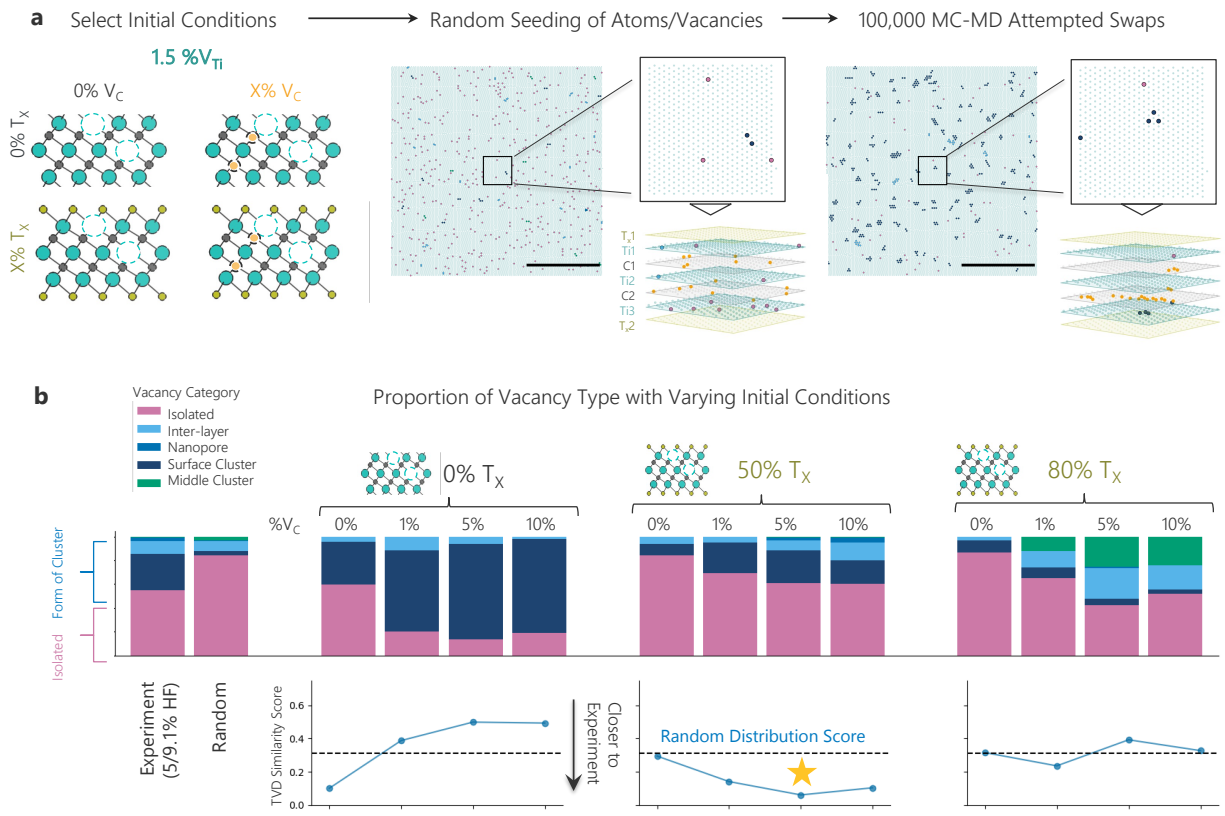

FIG. S6. 1.5 V_{Ti} MCMD results. (a) Workflow for Monte Carlo (MC)–Molecular Dynamics (MD) simulations. A Ti vacancy concentration of 1.5% (from 5 and 9.1% HF experiments) was used. Candidate $Ti_3C_2T_x$ grids were seeded with varying C vacancy (0, 1, 5, 10%) and surface termination (0, 50, 80%) levels, example shown in random and final MC step configurations (scale bars: 10 nm). (b) Vacancy clustering in relaxed configurations compared with experiment and random. Line plots show the Total Variation Distance (TVD) between clustering distributions from the 12.5% HF sample and MC–MD runs, with the random TVD score (dotted line) as reference. Yellow star indicates run with the best TVD score (50% T_x, 5% V_C).

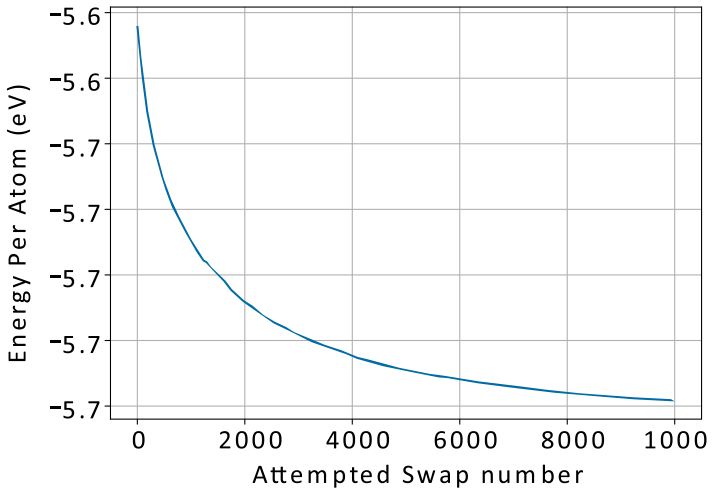

FIG. S7. Energy curve across MC-MD swaps. The MC-MD algorithm attempts to lower the overall energy in the system through random swaps of atoms and "ghost" vacancies. Here, we choose to run the MC-MD until the overall energy in the system flattens out, around 1000 attempted swaps.

Supplementary Figures 8-10: MXene STEM Images

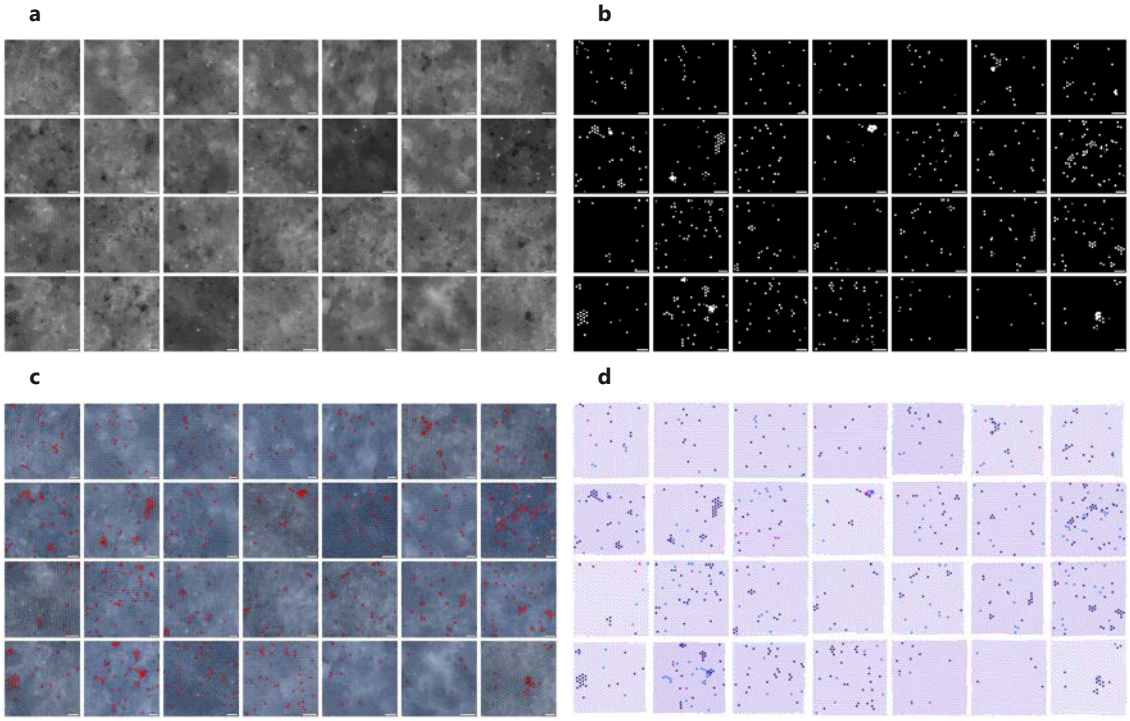

FIG. S8. 5% HF MXene images. (a) Stacked STEM images of MXenes etched with 5% HF. (b). Neural network-outputted masks from the defect finder NN. (c) All located atoms (blue) and defects (red). (d) Atoms and defects colored by layer. Scale bars 1 nm.

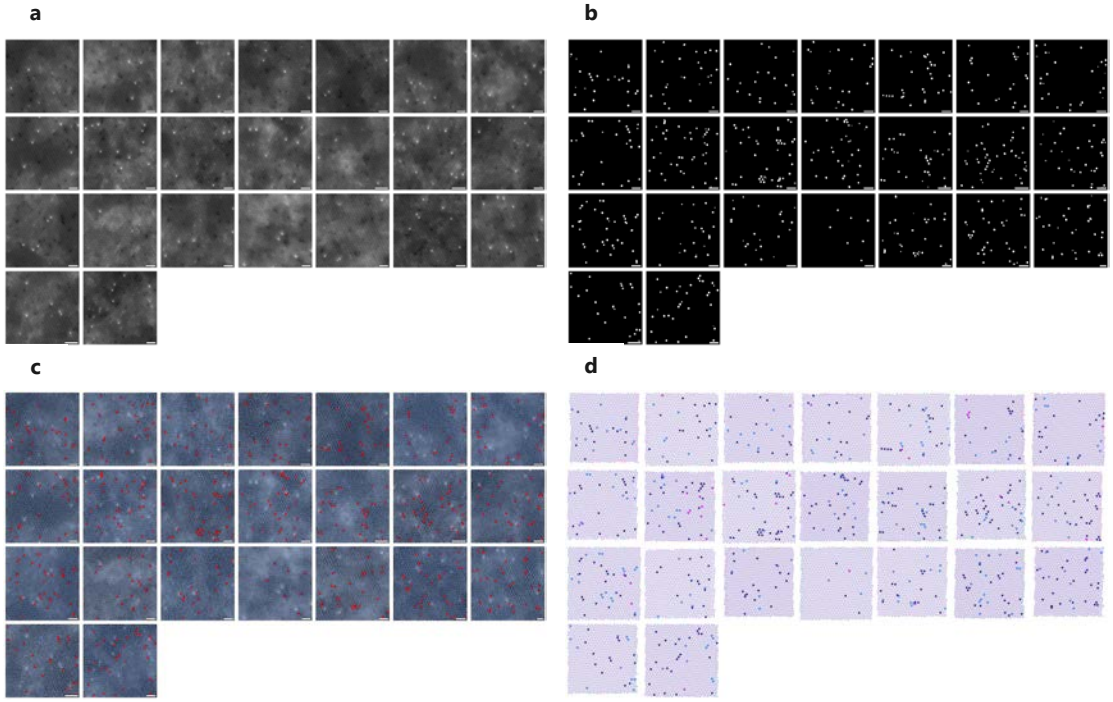

FIG. S9. 9.1% HF MXene images. (a) Stacked STEM images of MXenes etched with 5% HF. (b). Neural network-outputted masks from the defect finder NN. (c) All located atoms (blue) and defects (red). (d) Atoms and defects colored by layer. Scale bars 1 nm.

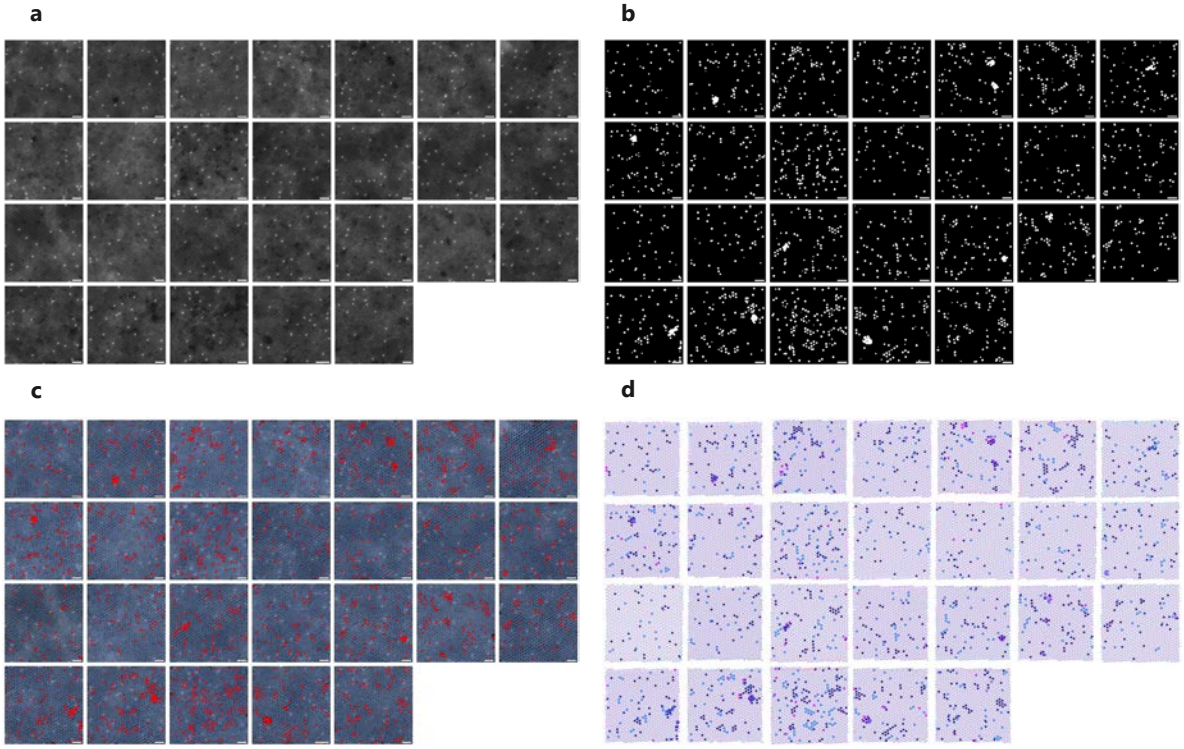

FIG. S10. 12.5% HF MXene images. (a) Stacked STEM images of MXenes etched with 5% HF. (b). Neural network-outputted masks from the defect finder NN. (c) All located atoms (blue) and defects (red). (d) Atoms and defects colored by layer. Scale bars 1 nm.

Supplementary References

1. Thakur, A. *et al.* Step-by-step guide for synthesis and delamination of $Ti_3C_2T_x$ MXene. *Small Methods* **7**, 2300030 (2023). https://doi.org/10.1002/smtd.202300030.
2. Savitzky, B. H. *et al.* Image registration of low signal-to-noise cryo-STEM data. *Ultramicroscopy* **191**, 56–65 (2018). https://doi.org/10.1016/j.ultramic.2018.04.008.
3. Sang, X. *et al.* Atomic defects in monolayer titanium carbide ($Ti_3C_2T_x$) MXene. *ACS Nano* **10**, 9193–9200 (2016). https://doi.org/10.1021/acsnano.6b05240.
4. He, R., Wan, Y., Zhao, P., Guo, P., Jiang, Z. & Zheng, J. First-principles investigation of native point defects in two-dimensional Ti_3C_2. *Computational and Theoretical Chemistry* **1150**, 26–39 (2019). https://doi.org/10.1016/j.comptc.2019.01.006.
5. Thompson, A. P. *et al.* LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **271**, 108171 (2022). https://doi.org/10.1016/j.cpc.2021.108171.
6. Plummer, G. & Tucker, G. J. Bond-order potentials for the Ti_3AlC_2 and Ti_3SiC_2 MAX phases. *Physical Review B* **100**, 214114 (2019). https://doi.org/10.1103/PhysRevB.100.214114.
7. Hu, T., Yang, J. & Wang, X. Carbon vacancies in Ti_2CT_2 MXenes: defects or a new opportunity? *Physical Chemistry Chemical Physics* **19**, 31773–31780 (2017). https://doi.org/10.1039/C7CP06593K.
8. Wyatt, B. C. *et al.* Alkali cation stabilization of defects in 2D MXenes at ambient and elevated temperatures. *Nature Communications* **15**, 6353 (2024). https://doi.org/10.1038/s41467-024-50713-2.
9. Halim, J. *et al.* X-ray photoelectron spectroscopy of select multi-layered transition metal carbides (MXenes). *Applied Surface Science* **362**, 406–417 (2016). https://doi.org/10.1016/j.apsusc.2015.11.089.

10. Hope, M. A. *et al.* NMR reveals the surface functionalisation of Ti_3C_2 MXene. *Physical Chemistry Chemical Physics* **18**, 5099–5102 (2016). https://doi.org/10.1039/C6CP00330C.
11. Li, Y., Huang, S., Wei, C., Wu, C. & Mochalin, V. N. Adhesion of two-dimensional titanium carbides (MXenes) and graphene to silicon. *Nature Communications* **10**, 3014 (2019). https://doi.org/10.1038/s41467-019-10982-8.
12. Tavenner, J. P., Mendelev, M. I. & Lawson, J. W. Molecular dynamics based kinetic Monte Carlo simulation for accelerated diffusion. *Computational Materials Science* **218**, 111929 (2023). https://doi.org/10.1016/j.commatsci.2022.111929.
13. Ibragimova, R., Puska, M. J. & Komsa, H. P. pH-dependent distribution of functional groups on titanium-based MXenes. *ACS Nano* **13**, 9171–9181 (2019). https://doi.org/10.1021/acsnano.9b03511.
14. Fasolino, A., Los, J. H. & Katsnelson, M. I. Intrinsic ripples in graphene. *Nature Materials* **6**, 858–861 (2007). https://doi.org/10.1038/nmat2011.
15. Gienger, J., Severin, N., Rabe, J. P. & Sokolov, I. M. Reconstructing interaction potentials in thin films from real-space images. *Physical Review E* **93**, 043306 (2016). https://doi.org/10.1103/PhysRevE.93.043306.
16. Chib, S. & Greenberg, E. Understanding the Metropolis–Hastings algorithm. *The American Statistician* **49**, 327–335 (1995). https://doi.org/10.1080/00031305.1995.10476177.
17. Nord, M., Vullum, P. E., MacLaren, I., Tybell, P. T. M. & Holmestad, R. Atomap: a new software tool for the automated analysis of atomic-resolution images using two-dimensional Gaussian fitting. *Advanced Structural and Chemical Imaging* **3**, 9 (2017). https://doi.org/10.1186/s40679-017-0042-5.