

---

This is an electronic reprint of the original article.  
This reprint may differ from the original in pagination and typographic detail.

Kurki, Lauri; Oinonen, Niko; Foster, Adam

## Automated Structure Discovery for Scanning Tunneling Microscopy

*Published in:*  
ACS Nano

*DOI:*  
[10.1021/acsnano.3c12654](https://doi.org/10.1021/acsnano.3c12654)

Published: 30/04/2024

*Document Version*  
Publisher's PDF, also known as Version of record

*Published under the following license:*  
CC BY

*Please cite the original version:*  
Kurki, L., Oinonen, N., & Foster, A. (2024). Automated Structure Discovery for Scanning Tunneling Microscopy. *ACS Nano*, 18(17), 11130-11138. <https://doi.org/10.1021/acsnano.3c12654>

---

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Automated Structure Discovery for Scanning Tunneling Microscopy

Lauri Kurki, Niko Oinonen, and Adam S. Foster\*



Cite This: *ACS Nano* 2024, 18, 11130–11138



Read Online

ACCESS |



Metrics & More



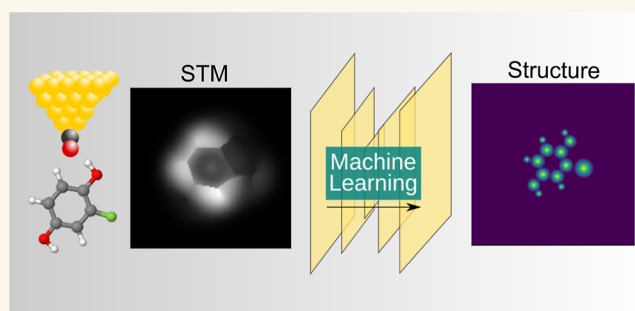
Article Recommendations



Supporting Information

**ABSTRACT:** Scanning tunneling microscopy (STM) with a functionalized tip apex reveals the geometric and electronic structures of a sample within the same experiment. However, the complex nature of the signal makes images difficult to interpret and has so far limited most research to planar samples with a known chemical composition. Here, we present automated structure discovery for STM (ASD-STM), a machine learning tool for predicting the atomic structure directly from an STM image, by building upon successful methods for structure discovery in noncontact atomic force microscopy (nc-AFM). We apply the method on various organic molecules and achieve good accuracy on structure predictions and chemical identification on a qualitative level while highlighting future development requirements for ASD-STM. This method is directly applicable to experimental STM images of organic molecules, making structure discovery available for a wider scanning probe microscopy audience outside of nc-AFM. This work also allows more advanced machine learning methods to be developed for STM structure discovery.

**KEYWORDS:** scanning probe microscopy, scanning tunneling microscopy, tip functionalization, machine learning, convolutional neural network, structure discovery



Scanning probe microscopy (SPM) methods are powerful tools for studying nanoscale systems with atomic resolution. As the basis of SPM methods, scanning tunneling microscopy (STM)<sup>1</sup> and atomic force microscopy (AFM)<sup>2</sup> have been widely utilized in the characterization of various systems, such as biological samples, hybrid inorganic–organic interfaces, and individual steps of on-surface reactions.<sup>3–9</sup> To enhance the spatial accuracy in the characterization, the probe tip can be functionalized using a chemically inert, flexible apex (often a CO molecule) to allow scanning at a very close tip-sample distance where Pauli repulsion is the dominating interaction.<sup>10,11</sup> With respect to the detailed characterization of atomic structures, there exist many demonstrations of the improved spatial resolution of AFM scanning with functionalized tips,<sup>9,12,13</sup> but STM is now also being increasingly used in this bond-resolved mode.<sup>8,13–18</sup> STM in particular benefits significantly from tip functionalization as sharp submolecular features appear in the image, revealing both the molecular skeleton and the electronic structure in high detail within the same experiment—this is impossible using a bare metal tip<sup>19</sup> or by high-resolution AFM. In addition to atomic structural characterization, STM with a functionalized tip offers interesting approaches for electronic structure characterization via Frontier orbitals of the tip apex<sup>20</sup> by improving the visibility

of molecular orbitals<sup>21</sup> and by distinguishing nearby molecular states.<sup>17</sup>

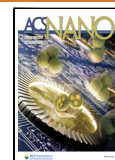
In terms of structure characterization, the main limitations of STM are its inability to see beyond the closest atom to the tip and the insurmountable problem of chemical identification of the atoms. In fact, research using bond-resolved STM has so far been mostly limited to planar molecules consisting of only a priori known atomic species, and the recognition of an unknown sample can require an extensive search through all possible molecules and configurations. The same problems have been experienced in the field of noncontact AFM where recent advances in machine learning image analysis have been proven effective in structure discovery and chemical identification of single molecules and ice structures.<sup>22–27</sup> Beyond machine learning, some of the challenges with respect to chemical identification can be overcome by combining tip-enhanced Raman spectroscopy with STM and AFM to achieve

**Received:** December 15, 2023

**Revised:** March 25, 2024

**Accepted:** April 5, 2024

**Published:** April 22, 2024



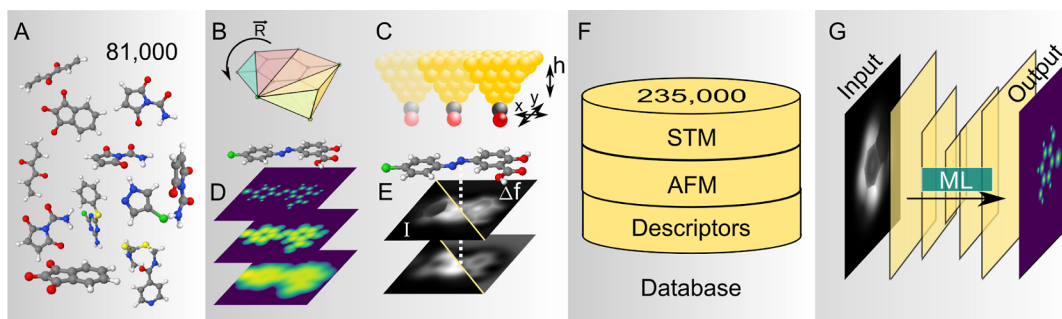


Figure 1. Scheme of ASD-STM. (A) Starting from a large set of organic molecules, (B–E) we simulate STM and AFM for different rotations of the molecules and calculate corresponding physical descriptors. (F) The images are stored in a database and (G) used as training material for a machine learning model for STM structure discovery.

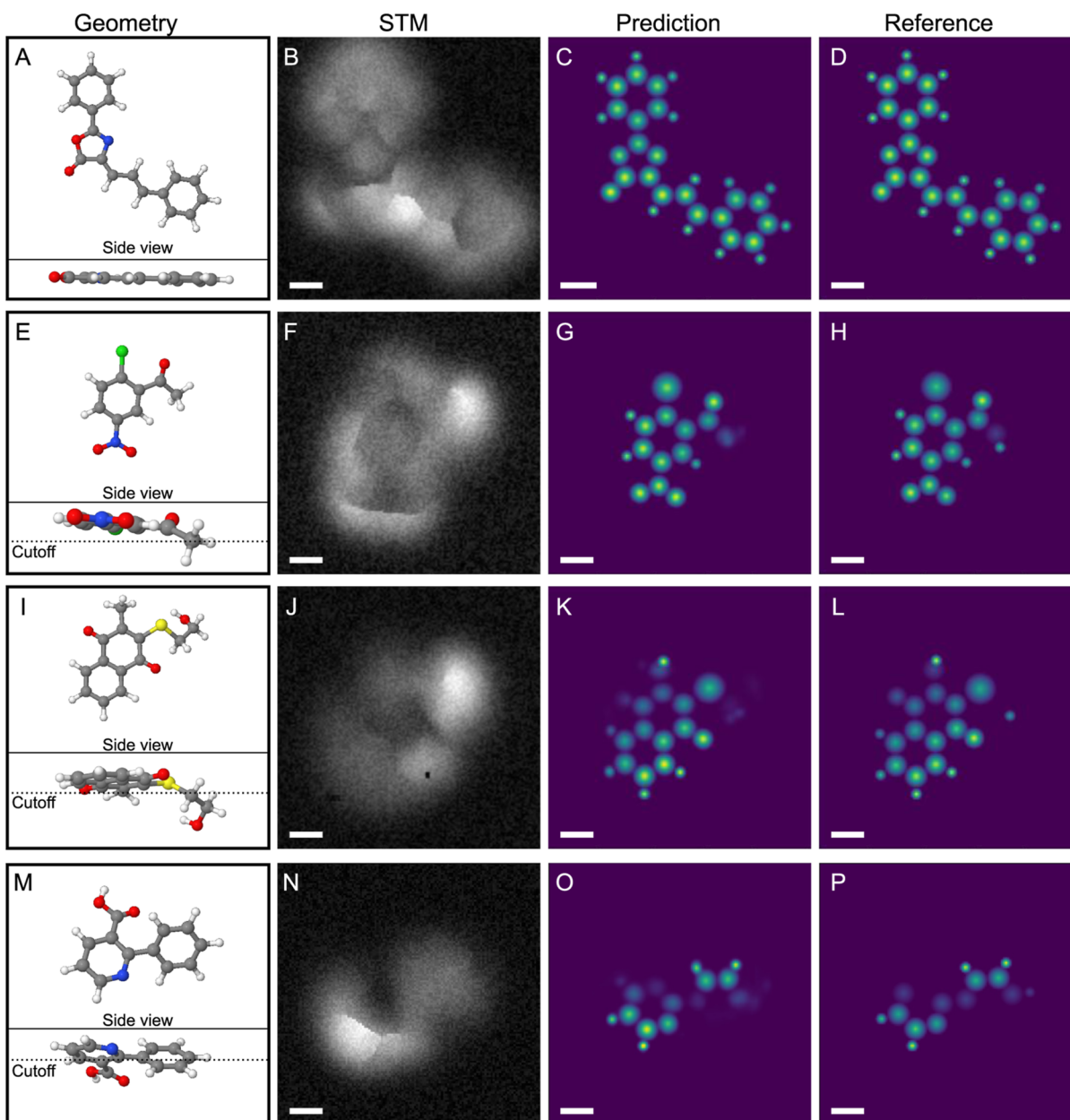
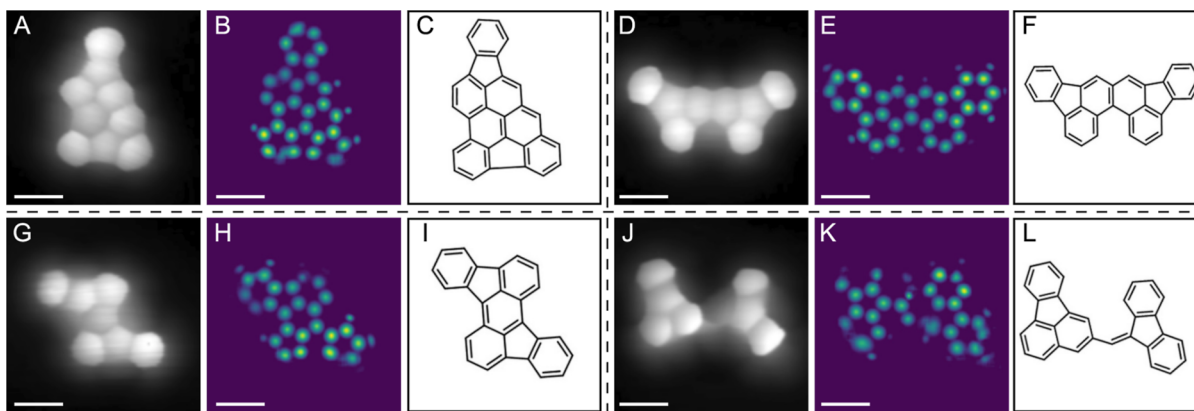


Figure 2. Example predictions of simulated STM images of (A–D)  $C_{18}H_{13}NO_2$ , (E–H)  $C_8H_6ClNO_3$ , (I–L)  $C_{13}H_{12}O_3S$ , and (M–P)  $C_{12}H_9NO_2S$  molecules from the testing set. Each molecule has a column for the geometry (descriptor depth cutoff shown in the side profile), STM image, predicted structure, and reference. Scale bar is 2 Å.



**Figure 3.** Predictions from four experimental STM images. (A,D,G,J) The first column of each example is the STM image used as input for ASD-STM. (B,E,H,K) The predictions are shown in the second column and (C,F,I,L) the third column contains the structural formula. Scale bar is 4 Å. STM images and structural formulas adapted with permission from ref 18 ©2020 ACS.

quantitative chemical sensitivity.<sup>28</sup> For STM, advanced image analysis methods have so far not been developed for bond-resolved imaging but instead the focus has been on, e.g., defect detection,<sup>29</sup> molecule keypoint detection,<sup>30</sup> surface characterization,<sup>31</sup> and atom manipulation<sup>32</sup> as well as autonomous experiments.<sup>33</sup> Machine learning methods have also been used for automating the tip conditioning<sup>34</sup> and tip functionalization<sup>35</sup> in STM. It is clear that the focus in general has been on larger-scale systems and processes, and there is a need for automated characterization methods for bond-resolved imaging with a CO tip.

In this work, we present automated structure discovery for STM (ASD-STM), a machine learning approach for structure characterization directly from experimentally bond-resolved STM images. With ASD-STM, we offer a solution to the problems in STM structure characterization while bridging the gap between image analysis methods for high-resolution AFM and STM. As STM with a functionalized tip is a more easily accessible characterization method compared to noncontact AFM,<sup>36</sup> this work also makes ASD available for a wider SPM audience and allows more sophisticated methods to be developed for sample recognition in STM.

## RESULTS

The proposed method starting from a data set of molecule geometries and ending in a machine learning model that can predict atomic structure directly from experimental images is outlined in Figure 1.

**Predictions on Simulated Images.** To benchmark our method and assess the improvement in prediction accuracy, we use the trained model to predict atomic structures from simulated images not included in the training data. The obvious benefit from analyzing simulated predictions is that the atomic coordinates of the reference, and therefore the image descriptor, are known exactly. Figure 2 shows three examples of predictions from simulated images. The three examples have been taken at varying scanning heights and the sample molecules are different in size and structure, and they contain different chemical species and functional groups to demonstrate the versatility of the method.

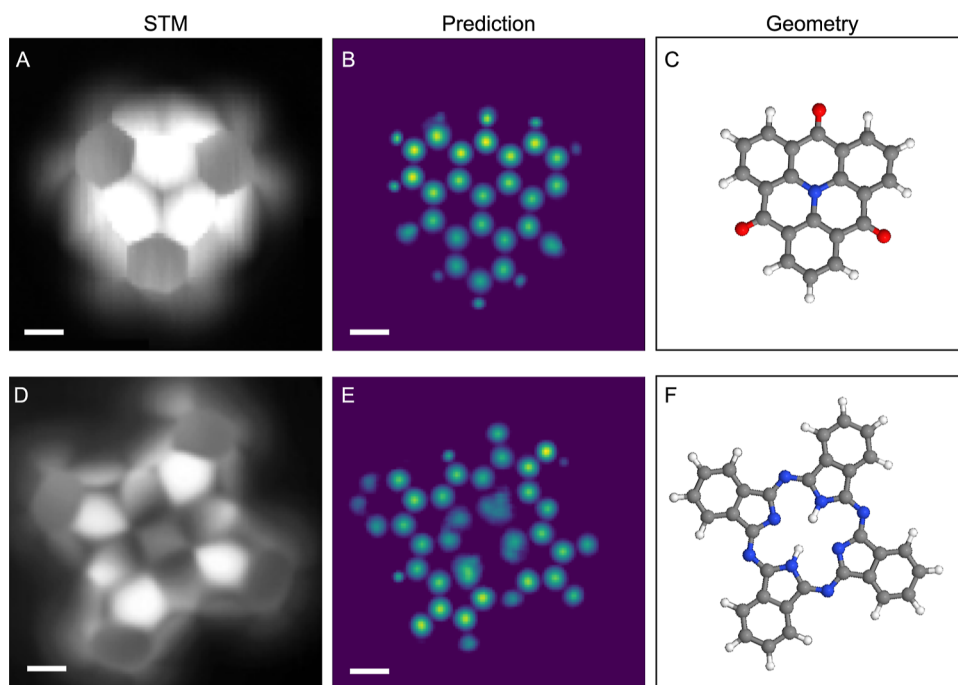
Molecule 1 (Figure 2A–D) is the largest of the three and it has two benzene groups and a 5-ring lactone group with an additional nitrogen heteroatom. The molecule is planar, and all atoms are visible to the tip resulting in the carbon backbone

skeleton being revealed with the exception of the benzene rings which are not prominent in the STM image (Figure 2B). Even though the skeleton is not completely visible, the prediction (Figure 2C) is very close to that of the reference (Figure 2D). All atoms, including hydrogens which are not visible to the eye, are correctly located and the associated chemical identification distinguishes hydrogen atoms from carbon, nitrogen, and oxygen atoms. It is also promising that the model correctly identifies the oxygen bonded to the 5-ring as not hydrogen and that it does not blindly add hydrogen atoms to nitrogen and oxygen atoms of the 5-ring. Overall, the mean absolute error relative to the range of values in the reference is 1.1%.

The second example shows a smaller molecule containing a nitro group and a chlorine atom (Figure 2E–H). The molecule is slightly tilted, and it includes a nonplanar part in which the branch containing a carbonyl group is bending downward. This is projected in the reference descriptor by a barely visible carbon and two of the bonded hydrogen atoms are too deep to be considered (Figure 2H). Again, the prediction resembles the reference very closely (Figure 2G). The chlorine atom is distinguished from hydrogen atoms, and the relative height of the oxygen in the carbonyl group is correctly identified. The location of the rightmost carbon atom is predicted correctly in the scanning plane, but the disk is too bright, meaning that the vertical position prediction is wrong. Since the oxygen atom is approximately 1.0 Å higher than the deepest carbon atom, the disk representing it is very bright in comparison and it appears to shadow the nearby deep carbon atom, meaning that it is not surprising that the model struggles most in this region. The relative error is 0.4%.

The third simulated example contains a sulfur atom, a hydroxy group, and two carbon 6-rings, one of which has two carbonyl groups (Figure 2I–L). In the STM image, the sulfur atom appears as the brightest area (Figure 2J). This molecule is particularly tricky for chemical identification due to the carbonyl groups which could theoretically be hydrogen atoms of a benzene ring, and the sulfur atom bridging the two carbon structures could be, e.g., an oxygen atom forming an ether group. Regardless of the apparent challenges, the model predicts all of these properties correctly (Figure 2K). The disks representing oxygen atoms have a larger radius than the hydrogen atoms of the lower ring, and the sulfur atom has an even bigger radius distinguishing it from oxygen and carbon atoms. The carbon chain originating from the sulfur is not





**Figure 4.** Prediction on the experimental image of the (A–C) TOAT (1,5,9-trioxo-13-azatriangulene) and (D–F) phthalocyanine (2H-Pc) molecules. For both molecules, the columns contain, from left to right, the STM image, structure prediction, and geometry of the molecule. Location of the pyrrolic hydrogen atoms in (F) as proposed in ref 21. Scale bar is 2 Å. STM images in (A) reprinted with permission from ref 38 ©2016 ACS and in (D) from ref 21 ©2023 ACS.

considered in the descriptor due to the depth threshold. Finally, all hydrogen atoms included in the reference are predicted and the tilting of the molecule is captured in the prediction. The relative error for this example is 0.6%. The fourth example consists of two 6-rings twisted with respect to each other, one containing a nitrogen heteroatom and a carboxyl group (Figure 2M–P). Here, the carbon backbone is predicted accurately despite the twisting, and the relative error is 0.7%. To illustrate the accuracy of three-dimensional (3D) molecules, we show more predictions in Supporting Information (see Figure S2). Overall, the prediction accuracy is very good on simulated STM images and the trained model predicts atomic structures of different sizes, distinguishes various chemical species, and even appears to predict seemingly hidden atoms. However, the final validation of the model has to be done on experimental data, and to this end, we first consider four benchmark molecules. Also, even though ASD-STM was trained on constant-height images, extending it to constant-current images could possibly enable easier 3D structure recognition. As a preliminary test of feasibility, we applied the present model to a constant-current image of benzene, and Figure S3 shows that the prediction is quite reasonable. Developing this further will certainly be a future target for the method.

**Experimental Validation.** The first experimental images we apply ASD-STM on are STM images obtained by Song et al.<sup>18</sup> that reveal the molecular skeletons of four hydrocarbons in high detail making them excellent for benchmarking purposes (Figure 3A,D,G,J). The molecules consist of exclusively carbon and hydrogen atoms constructing five- and six-membered rings, barring the fourth example in which a carbon atom links two ring structures (Figure 3L). In all STM images, the rings appear quite distorted, and in many cases, it is not immediately clear how many carbon atoms constitute the

particular ring, and therefore along with general structure discovery, we consider our objective to be to distinguish these rings. The first example consists of nine carbon rings of which seven are six-membered. Most of the rings, seven out of nine, are correctly identified in the prediction (Figure 3B), characterization of one ring is unclear, and characterization of one ring is misclassified. The model also seems to have learned how rings of different sizes connect with each other. Additionally, the prediction suggests that the bottom and top parts of the molecule are slightly closer to the tip than the central parts, but these predictions are difficult to validate.

The second example is an STM image of a symmetric molecule with eight rings (Figure 3D–F). The five rings are fairly large in the image, and the rings at all ends of the molecule appear brighter than in the center. The prediction is very good in terms of the carbon structure with all atoms predicted correctly (Figure 3E). In this case, even many hydrogen atoms are included in the prediction. The third example we consider is smaller than the previous molecules (Figure 3G–I), and again, the model performs well in predicting the atomic positions of carbon atoms. In this case, the bottom row of rings proposes a challenge for the model as the six rings in the corners are misclassified as five rings. A possible explanation for the wrong prediction is that the rings are almost circular and lack characteristic corners. This is supported by the first example which exhibited similar features in the STM image and where similar problems were encountered. It is also worth noting that the STM image contains considerable noise in the form of horizontal lines, but the model succeeds in the prediction regardless. We attribute this to a successful augmentation process during the training phase, where noise and cut-outs were applied to the initially pristine simulated images.

The final example is different from the others in that it includes two discrete ring structures that are connected by carbon atoms (Figure 3J–L). It is clear that the model struggles with the molecule and does not predict the correct structure in this case. The rightmost part is mostly correct, although the model does not seem to fully recognize the rings, as some of the predicted atoms are blurry and slightly deformed. The leftmost part of the molecule has more errors, and only the central five ring is correctly predicted (Figure 3K). A general problem with many machine learning models is that their explainability is poor and it is often difficult to explain the reasons behind a certain prediction.<sup>37</sup> One possible reason would be the particular structure of the molecule; molecules containing two separate but connected carbon ring structures are rare in the data set, meaning that the model has not had proper experience and thus fails in the prediction. Also, it is vital to train the model to withstand noise and other artifacts present in experimental images, which is illustrated by comparing to predictions of these structures from pristine simulated images (see Figure S4). Overall, the predictions correspond well to the molecular structures, and in total, 106 out of 116 (91.4%) carbon atoms and 23 out of 31 (74.2%) carbon rings are predicted correctly.

Predictions on the hydrocarbon examples are promising, but the molecular structures were rather simple, and we want to test the limits of the model on a more challenging example. To this end, we use STM images of the TOAT molecule (1,5,9-trioxo-13-azatriangulene) (Figure 4A)<sup>38</sup> and phthalocyanine (2H-Pc) (Figure 4D).<sup>21</sup> Predicting the atomic structure of TOAT is particularly interesting since the simulation software we used to generate the training data is known for not being able to properly reproduce experimental STM images of TOAT.<sup>39</sup> Regardless, the prediction is in very good agreement with the true molecular structure (Figure 4B). The triangulene backbone is correctly identified, and all oxygen atoms are distinguished from hydrogen atoms, indicated by larger disks in the prediction. In this example, even most hydrogen atoms are located with the only exception being the top right ring missing one hydrogen. The disk representing the center nitrogen seems to have the same radius as the neighboring carbon atoms suggesting an incorrect prediction, but this classification is inconclusive. To assist in the chemical identification, we performed further analysis and trained another model which separates the disks into different classes by atomic type, which allowed us to distinguish the nitrogen heteroatom (Figure S5).<sup>22</sup> Definitive chemical identification has been explored previously for AFM,<sup>23,24</sup> but for STM, it remains a future challenge.

The most challenging example we test here is an STM image of 2H-Pc (Figure 4D), and this prediction (Figure 4E) reveals the limitations of our approach. 2H-Pc is a cyclic molecule consisting of four isoindole units connected by nitrogen atoms (Figure 4F). In this example, a perfect prediction would reveal the configuration of the hydrogen atoms in the central moiety. In the STM image, the inner five rings are bright and clearly exhibit the pentagon-like motif, whereas the outer six rings are less pronounced. The nitrogen atoms connecting the units appear as cones decaying toward the center. In general, the molecular skeleton is not very prominent in the image, and the signal originating from the electronic structure intersects significantly with the mechanical bending of the tip apex. It is apparent that these factors severely affect the ability of ASD-STM to make accurate predictions of the structure. In the

prediction, the five rings apart from the top left ring are correctly identified, and all nitrogen atoms connecting the rings are located. The biggest errors are the missing halves of the six rings and also the central hydrogens that are not included in the prediction. It is worth noting that a similar phenomenon was observed in the case of the fourth hydrocarbon (Figure 3K) where the top half of a ring was missing in the prediction.

## DISCUSSION

While predictions on most experimental examples were good, it is clear that there is room for improvement. Since the quality of data included in the training process is vital to the accuracy of predictions, rare structures affecting the accuracy should be discussed, as it raises the important question of how to select the molecules in the data set. In the scope of SPM structure discovery, there are two main approaches to data set generation. The first option, which this work highlights, is to create a large, diverse, and descriptive molecular data set of various chemical species and structures and to train a versatile model with the aim of predicting the structure of almost any small organic molecule. This approach has also been chosen in most previous sample characterization efforts (e.g.,<sup>22,40</sup>), but recently, another method has been used in ice structure discovery,<sup>26,41</sup> where instead of a diverse data set, a tailored data set is utilized and perfected to make very accurate predictions possible in a constrained problem domain. That is, if the goal was to predict only the geometries of different hydrocarbons or triangulene-based molecules, the model would benefit from a tailored data set with a heavy emphasis on such structures. As methods for quick generation of molecular structures to create vast molecular data sets are becoming more readily available,<sup>42–44</sup> the balance and cost-effectiveness of tailored versus diverse data sets should be the focus of future research.

Second, ASD-STM has been trained on synthetic STM images of molecules in isolation without the substrate due to much cheaper computational cost, and as such, the generated image data set should not be inferred as a source for accurate electronic structures of adsorbed organic molecules. Moreover, as we know that PPSTM cannot accurately reproduce STM images of the TOAT molecule, it was expected that ASD-STM would struggle with the experimental STM image of TOAT. However, the accuracy in the prediction was excellent, which suggests that even though images in the training set are not entirely representative of true images in terms of electronic structure, they capture the characteristic sharp lines coming from CO tip bending accurately, and this seems to be critical for structure discovery. Still, we do note that for systems experiencing significant charge transfer or orbital hybridization ASD-STM is most likely not a sufficient tool, and a more comprehensive approach is needed.<sup>41</sup> Also, synthetic AFM images of isolated molecules in general correspond well to experimental AFM images where the substrate is naturally present, and since for some samples it is possible to simultaneously gather STM and AFM signals, the possibility of incorporating AFM data into the training process should be explored.

Third, we note that the appearance of the molecule in STM is sensitive to Frontier orbitals of the CO tip and consequently to the tip height as at closer tip-sample distances the tip exhibits a strong p-wave character which diminishes to a more s-wave character at longer distances.<sup>21</sup> The s-wave character is also increased at higher absolute bias voltages.<sup>45</sup> In this work,

ASD-STM has been trained using a constant ratio of s and p-wave contributions, but the different orbital ratios should be accounted for if ASD-STM is used for a larger range of tip-sample distances or bias voltages. Finally, although it is inevitable that the accuracy of predictions on experimental images is worse than on simulated images, we will investigate this further in the future with a particular focus on improving the robustness of the model against experimental noise and artifacts.

## CONCLUSIONS

This work presents ASD-STM, a method for predicting the atomic structure of a sample molecule directly from a bond-resolved STM image. We present the workflow and the software required to generate an STM image data set and to train the ML model. Images included in the training set were synthesized by considering the molecules in isolation, and even though the electronic structure can be affected by the substrate, we showed that the approximation of molecules in isolation is reasonable for structure discovery on bond-resolved STM. The model was validated by applying it to experimental images of six different organic molecules, and on most samples, the accuracy was good in terms of atomic structure; additionally, we achieved a qualitative level of chemical identification by distinguishing between hydrogen, carbon, and oxygen atoms. While the chemical domain of the model is quite large within the class of organic molecules, many types of molecules such as organometallic compounds are not within the scope and would require expanding the training set. On the other hand, the example of 2H-Pc clearly demonstrated some of the challenges. The main limitation of the method is the reliance on the sharp submolecular features in the images, which restricts the method to high-quality images and to a range of short tip-sample distances. While the simulated images are augmented with noise and cut-outs during training, we anticipate that these augmentations are not the only difference between simulated and experimental images, and advanced augmentation strategies should be explored to make the model more robust against experimental conditions. Also, we discussed possible further improvements to ASD-STM by varying the orbital contributions in the tunneling calculation, by tailoring the composition of the initial data set of molecules to the problem at hand, and by including simultaneously gathered AFM data in the training process.

Despite the challenges, ASD-STM is readily applicable to STM images of various small organic molecules, and with this work, we allow more sophisticated methods to be developed for STM structure discovery. Finally, as bond-resolved STM images exhibit similar submolecular features as noncontact AFM images at a significantly reduced acquisition time, ASD-STM demonstrates a promising start for accelerating molecular structure discovery in SPM in general.

## METHODS

**Simulated STM Images.** The initial data set contains the optimized geometries and atomic point charges of approximately 81,000 small organic molecules with chemical species lighter than bromine ( $Z \leq 35$ ) with most emphasis on hydrogen, carbon, nitrogen, and oxygen atoms (Figure 1A).<sup>25</sup> Simulating STM images requires that we first obtain the densities of states of all molecules in the data set, and since the molecular structures are already optimized, a single-point calculation is performed for all molecules. All *ab initio* calculations are performed using the FHI-aims code<sup>46,47</sup> with the

PBE functional<sup>48</sup> considering only the  $\Gamma$   $k$ -point. The unit cell is constructed by padding the molecule with 7 Å of empty space on each side to ensure that the periodic images do not interact with and affect the electronic structure.

The STM simulations for all molecules are done using the PPSTM code.<sup>39</sup> PPSTM is based on the Bardeen tunneling theory,<sup>49</sup> and it simulates constant height raster scanning by calculating the tunneling current at each point of a 3D grid of tip positions allowing for a quick simulation of multiple constant height scans at different tip-sample distances (Figure 1C,E). In PPSTM, the tunneling current is composed of contributions from the individual states of the tip and the sample. The sample states are obtained from density functional theory calculations as mentioned, and they are represented by Lorentzian distributions with a 0.1 eV broadening. The tip orbitals are considered as s and p orbitals located at the tip apex, and they provide independent tunneling channels. The tunneling channels are given a weight between 0 and 1, and the correct weights for a certain tip apex are found by matching with experiments (e.g., approximately 10% s and 90% p orbitals for CO tip<sup>8,50</sup>). Parallel to the STM simulations, we run the PPAFM<sup>51,52</sup> code to calculate the relaxation of the CO molecule near the sample, and the relaxed positions are, in turn, used in the STM simulation to achieve a good resemblance to experimental images revealing the molecular skeleton. PPAFM calculates the CO relaxation by considering the CO molecule as a probe particle (PP) and using a mechanical model taking into account the Lennard-Jones and electrostatic potentials between the PP and the sample. The position of the PP approaching the sample is relaxed by minimizing the total force acting upon the PP, and while calculating the relaxed tip positions, we simultaneously obtain AFM images of the molecules with little additional cost by integrating the forces affecting the PP over its path.<sup>53</sup> The AFM images are included in the data set of synthetic images, and they could be used as additional signal in a method developed for simultaneous AFM/STM imaging.

To maximize the amount of information in the SPM images, we emphasized flat regions normal to the scanning direction and also ensured a more even distribution of elements in the resulting images. To this end, we subjected the molecules to a set of rotations calculated separately for each molecule (Figure 1B). The rotations were obtained by computing the convex hull of the atomic coordinates, which defines planar segments of the molecule, and the rotations were weighted to include an even distribution of elements within 0.7 Å of the plane. After applying the rotations, our data set consists of synthetic SPM images of approximately 48,000 unique molecules and 235,000 unique images (Figure 1F). This procedure and the full composition of the data set are described in full detail in a previous study by Oinonen et al.<sup>25</sup>

One example item from the generated data set is shown in Figure 1D,E. All items in the data set include a stack of 10 constant height STM and AFM images taken at 0.1 Å separation over a range of tip-sample distances starting from a close distance scan where the molecular skeleton is revealed to a distance where only the electronic structure at the Fermi level is seen. A stack of images is created as opposed to just one slice because the real tip-sample distance in the experiment is not known, and we want to avoid overfitting to an exact scanning height by letting the model see images taken at a range of different scanning heights. This is achieved by randomly selecting one two-dimensional (2D) slice from the 3D stack of images during training. Consequently, our model does not require a full stack of images at prediction time, making structure discovery easily accessible using only one image. The data set also includes three image descriptors for each molecule—atomic disks, van der Waals spheres, and height map<sup>22</sup> (Figure 1D)—which we have chosen as visual and 2D representations of the molecules that we can calculate analytically from the atomic coordinates. Image descriptors are a convenient tool here, as they are not only visually intuitive for people but also suitable for many machine learning model designs. In all results, we focus on the atomic disk descriptor as it provides the most discrete predictions for the atomic positions and encodes the chemical information in the size of each disk, proportional to the covalent radius of the atom. The relative height of each atom is represented as the brightness of the



disk (background is zero), and atoms deeper than 1.2 Å relative to the top atom are not considered.

**Machine Learning.** In this work, the structure discovery task is formulated as an image-to-image problem, for which we have developed and trained a machine learning model which translates the STM image into a descriptor using an Attention U-Net-type model which utilizes an encoder-decoder architecture and an attention-gating mechanism.<sup>54</sup> On a conceptual level, the encoder compresses the information on the input image (STM image) into a latent space vector, which in turn is translated into a spatially larger representation (image descriptor) by the decoder. The encoder consists of 4 convolutional blocks containing two 2D convolutional layers with a LeakyReLU<sub>0.1</sub><sup>55</sup> activation function in between, followed by batch normalization<sup>56</sup> and a second LeakyReLU activation. Each of the convolutional blocks is preceded by a downscaling block, which downsamples the image using strided convolutions to increase the receptive field. The same activation and batch normalization policies are used in the downscaling. The decoder section consists of four blocks that include an upscaling block, an attention gate, and a convolutional block identical to the encoder. Upsampling is achieved using a transposed 2D convolutional layer, and in the block, the same activation and batch normalization policies are used as in the encoder. The attention gate generates highlighted regions in the input for the model to focus on using a type of skip connection where an additional query signal comes from the upsampled feature map. The attention map is concatenated with the upsampled feature map and is used as input for the convolutional block. The output layer of the model is a 1 × 1 convolutional layer followed by a ReLU activation (see full details in Figure S1).

The ML model is implemented in PyTorch,<sup>57</sup> and it is trained in a supervised setting for which the images were divided into training (180,000 images), validation (20,000), and testing (35,554) sets. The model was trained for 50 epochs, and a batch size of 30 was used. The parameters of the model were optimized using the Adam optimizer<sup>58</sup> to minimize the mean squared error. During training, the simulated STM images are normalized to 0 mean and unit variance, and we also add white noise and cut-outs to the images, representing electronic noise and tip artifacts to make the model not rely on pristine simulated images but instead force it to make predictions on imperfect images. After training, inference on the model can be done using true experimental images to obtain a prediction of the structure with very little computational cost.

## ASSOCIATED CONTENT

### Data Availability Statement

Lauri Kurki; Niko Oinonen; Adam S. Foster. Automated Structure Discovery for Scanning Tunneling Microscopy. 2023, 2312.08854. arXiv. <https://arxiv.org/abs/2312.08854> (accessed March 13, 2024).

### Data Availability Statement

The source code and training data for the machine learning models is available at <https://github.com/SINGROUP/ASD-STM>.

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsnano.3c12654>.

Details of the machine learning model, constant current prediction of benzene, predictions of 3D molecules from simulated images, predictions from simulated images of the four hydrocarbons, and chemical identification (PDF)

## AUTHOR INFORMATION

### Corresponding Author

Adam S. Foster — Department of Applied Physics, Aalto University, Espoo 00076, Finland; WPI Nano Life Science

Institute (WPI-NanoLSI), Kanazawa University, Kanazawa 920-1192, Japan; [orcid.org/0000-0001-5371-5905](https://orcid.org/0000-0001-5371-5905); Email: [adam.foster@aalto.fi](mailto:adam.foster@aalto.fi)

## Authors

Lauri Kurki — Department of Applied Physics, Aalto University, Espoo 00076, Finland; [orcid.org/0000-0002-5027-7847](https://orcid.org/0000-0002-5027-7847)

Niko Oinonen — Department of Applied Physics, Aalto University, Espoo 00076, Finland; Nanolayers Research Computing Ltd., London N12 0HL, U.K.

Complete contact information is available at: <https://pubs.acs.org/10.1021/acsnano.3c12654>

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors acknowledge funding from the Academy of Finland (project no. 346824). A.S.F. was supported by the World Premier International Research Center Initiative (WPI), MEXT, Japan. L.K. acknowledges funding from the Finnish Cultural Foundation. The authors acknowledge the computational resources provided by the Aalto Science-IT project and CSC, Helsinki.

## REFERENCES

- (1) Binnig, G.; Rohrer, H.; Gerber, C.; Weibel, E. Tunneling through a controllable vacuum gap. *Appl. Phys. Lett.* **1982**, *40*, 178–180.
- (2) Binnig, G.; Quate, C. F.; Gerber, C. Atomic force microscope. *Phys. Rev. Lett.* **1986**, *56*, 930–933.
- (3) Driscoll, R. J.; Youngquist, M. G.; Baldeschwieler, J. D. Atomic-scale imaging of DNA using scanning tunnelling microscopy. *Nature* **1990**, *346*, 294–296.
- (4) Pawlak, R.; Vilhena, J. G.; Hinaut, A.; Meier, T.; Glatzel, T.; Barattoff, A.; Gnecco, E.; Pérez, R.; Meyer, E. Conformations and cryo-force spectroscopy of spray-deposited single-strand DNA on gold. *Nat. Commun.* **2019**, *10*, 685.
- (5) Dufrêne, Y. F.; Ando, T.; Garcia, R.; Alsteens, D.; Martinez-Martin, D.; Engel, A.; Gerber, C.; Müller, D. J. Imaging modes of atomic force microscopy for application in molecular and cell biology. *Nat. Nanotechnol.* **2017**, *12*, 295–307.
- (6) Müller, D. J.; Dufrêne, Y. F. Atomic force microscopy as a multifunctional molecular toolbox in nanobiotechnology. *Nat. Nanotechnol.* **2008**, *3*, 261–269.
- (7) Gobbi, M.; Orgiu, E.; Samorì, P. When 2D Materials Meet Molecules: Opportunities and Challenges of Hybrid Organic/Inorganic van der Waals Heterostructures. *Adv. Mater.* **2018**, *30*, 1706103.
- (8) Sun, K.; Silveira, O. J.; Ma, Y.; Hasegawa, Y.; Matsumoto, M.; Kera, S.; Krejčí, O.; Foster, A. S.; Kawai, S. On-surface synthesis of disilabenzene-bridged covalent organic frameworks. *Nat. Chem.* **2023**, *15*, 136–142.
- (9) Gross, L.; Schuler, B.; Pavliček, N.; Fatayer, S.; Majzik, Z.; Moll, N.; Peña, D.; Meyer, G. Atomic Force Microscopy for Molecular Structure Elucidation. *Angew. Chem., Int. Ed.* **2018**, *57*, 3888–3908.
- (10) Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G. The chemical structure of a molecule resolved by atomic force microscopy. *Science* **2009**, *325*, 1110–1114.
- (11) Bartels, L.; Meyer, G.; Rieder, K. H. Controlled vertical manipulation of single CO molecules with the scanning tunneling microscope: A route to chemical contrast. *Appl. Phys. Lett.* **1997**, *71*, 213–215.



- (12) Zhong, Q.; Li, X.; Zhang, H.; Chi, L. Noncontact atomic force microscopy: Bond imaging and beyond. *Surf. Sci. Rep.* **2020**, *75*, 100509.
- (13) Jelinek, P. High resolution SPM imaging of organic molecules with functionalized tips. *J. Phys.: Condens. Matter* **2017**, *29*, 343002.
- (14) Pavliček, N.; Mistry, A.; Majzik, Z.; Moll, N.; Meyer, G.; Fox, D. J.; Gross, L. Synthesis and characterization of triangulene. *Nat. Nanotechnol.* **2017**, *12*, 308–311.
- (15) Seibel, J.; Fittolani, G.; Mirhosseini, H.; Wu, X.; Rauschenbach, S.; Anggara, K.; Seeberger, P. H.; Delbianco, M.; Kühne, T. D.; Schlickum, U.; Kern, K. Visualizing Chiral Interactions in Carbohydrates Adsorbed on Au(111) by High-Resolution STM Imaging. *Angew. Chem., Int. Ed.* **2023**, *62*, No. e202305733.
- (16) Vilas-Varela, M.; Romero-Lara, F.; Vegliante, A.; Calupitan, J. P.; Martínez, A.; Meyer, L.; Uriarte-Amiano, U.; Friedrich, N.; Wang, D.; Schulz, F.; Koval, N. E.; Sandoval-Salinas, M. E.; Casanova, D.; Corso, M.; Artacho, E.; Peña, D.; Pascual, J. I. On-Surface Synthesis and Characterization of a High-Spin Aza-[5]-Triangulene. *Angew. Chem., Int. Ed.* **2023**, *62*, No. e202307884.
- (17) Martinez-Castro, J.; Bolat, R.; Fan, Q.; Werner, S.; Arefi, H. H.; Esat, T.; Sundermeyer, J.; Wagner, C.; Michael Gottfried, J.; Temirov, R.; Ternes, M.; Stefan Tautz, F. Disentangling the electronic structure of an adsorbed graphene nanoring by scanning tunneling microscopy. *Commun. Mater.* **2022**, *3*, 57.
- (18) Song, S.; Guo, N.; Li, X.; Li, G.; Haketa, Y.; Telychko, M.; Su, J.; Lyu, P.; Qiu, Z.; Fang, H.; Peng, X.; Li, J.; Wu, X.; Li, Y.; Su, C.; Koh, M. J.; Wu, J.; Maeda, H.; Zhang, C.; Lu, J. Real-Space Imaging of a Single-Molecule Monoradical Reaction. *J. Am. Chem. Soc.* **2020**, *142*, 13550–13557.
- (19) Weiss, C.; Wagner, C.; Temirov, R.; Tautz, F. S. Direct imaging of intermolecular bonds in scanning tunneling microscopy. *J. Am. Chem. Soc.* **2010**, *132*, 11864–11865.
- (20) Gross, L.; Moll, N.; Mohn, F.; Curioni, A.; Meyer, G.; Hanke, F.; Persson, M. High-Resolution Molecular Orbital Imaging Using a p-Wave STM Tip. *Phys. Rev. Lett.* **2011**, *107*, 086101.
- (21) Néel, N.; Kröger, J. Orbital and Skeletal Structure of a Single Molecule on a Metal Surface Unveiled by Scanning Tunneling Microscopy. *J. Phys. Chem. Lett.* **2023**, *14*, 3946–3952.
- (22) Alldritt, B.; Hapala, P.; Oinonen, N.; Urtev, F.; Krejci, O.; Federici Canova, F.; Kannala, J.; Schulz, F.; Liljeroth, P.; Foster, A. S. Automated structure discovery in atomic force microscopy. *Sci. Adv.* **2020**, *6*, No. eaay6913.
- (23) Carracedo-Cosme, J.; Romero-Muñiz, C.; Pérez, R. A Deep Learning Approach for Molecular Classification Based on AFM Images. *Nanomaterials* **2021**, *11*, 1658.
- (24) Oinonen, N.; Kurki, L.; Ilin, A.; Foster, A. S. Molecule graph reconstruction from atomic force microscope images with machine learning. *MRS Bull.* **2022**, *47*, 895–905.
- (25) Oinonen, N.; Xu, C.; Alldritt, B.; Canova, F. F.; Urtev, F.; Cai, S.; Krejčí, O.; Kannala, J.; Liljeroth, P.; Foster, A. S. Electrostatic Discovery Atomic Force Microscopy. *ACS Nano* **2022**, *16*, 89–97.
- (26) Tang, B.; Song, Y.; Qin, M.; Tian, Y.; Wu, Z. W.; Jiang, Y.; Cao, D.; Xu, L. Machine learning aided atomic structure identification of interfacial ionic hydrates from AFM images. *Natl. Sci. Rev.* **2022**, *10*, nwac282.
- (27) Carracedo-Cosme, J.; Romero-Muñiz, C.; Pou, P.; Pérez, R. Molecular Identification from AFM Images Using the IUPAC Nomenclature and Attribute Multimodal Recurrent Neural Networks. *ACS Appl. Mater. Interfaces* **2023**, *15*, 22692–22704.
- (28) Xu, J.; Zhu, X.; Tan, S.; Zhang, Y.; Li, B.; Tian, Y.; Shan, H.; Cui, X.; Zhao, A.; Dong, Z.; Yang, J.; Luo, Y.; Wang, B.; Hou, J. G. Determining structural and chemical heterogeneities of surface species at the single-bond limit. *Science* **2021**, *371*, 818–822.
- (29) Wang, C.; Li, H.; Hao, Z.; Li, X.; Zou, C.; Cai, P.; Wang, Y.; You, Y. Z.; Zhai, H. Machine learning identification of impurities in the STM images. *Chin. Phys. B* **2020**, *29*, 116805.
- (30) Yuan, S.; Zhu, Z.; Lu, J.; Zheng, F.; Jiang, H.; Sun, Q. Applying a Deep-Learning-Based Keypoint Detection in Analyzing Surface Nanostructures. *Molecules* **2023**, *28*, 5387.
- (31) Choudhary, K.; Garrity, K. F.; Camp, C.; Kalinin, S. V.; Vasudevan, R.; Ziatdinov, M.; Tavazza, F. Computational scanning tunneling microscope image database. *Sci. Data* **2021**, *8*, 57.
- (32) Chen, I. J.; Aapro, M.; Kipnis, A.; Ilin, A.; Liljeroth, P.; Foster, A. S. Precise atom manipulation through deep reinforcement learning. *Nat. Commun.* **2022**, *13*, 7499.
- (33) Rashidi, M.; Wolkow, R. A. Autonomous Scanning Probe Microscopy in Situ Tip Conditioning through Machine Learning. *ACS Nano* **2018**, *12*, 5185–5189.
- (34) Wang, S.; Zhu, J.; Blackwell, R.; Fischer, F. R. Automated tip conditioning for scanning tunneling spectroscopy. *J. Phys. Chem. A* **2021**, *125*, 1384–1390.
- (35) Alldritt, B.; Urtev, F.; Oinonen, N.; Aapro, M.; Kannala, J.; Liljeroth, P.; Foster, A. S. Automated tip functionalization via machine learning in scanning probe microscopy. *Comput. Phys. Commun.* **2022**, *273*, 108258.
- (36) Gross, L. Recent advances in submolecular resolution with scanning probe microscopy. *Nat. Chem.* **2011**, *3*, 273–278.
- (37) Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy* **2020**, *23*, 18.
- (38) van der Heijden, N. J.; Hapala, P.; Rombouts, J. A.; van der Lit, J.; Smith, D.; Mutombo, P.; Švec, M.; Jelinek, P.; Swart, I. Characteristic Contrast in  $\Delta_{\text{min}}$  Maps of Organic Molecules Using Atomic Force Microscopy. *ACS Nano* **2016**, *10*, 8517–8525.
- (39) Krejčí, O.; Hapala, P.; Ondráček, M.; Jelinek, P. Principles and simulations of high-resolution STM imaging with a flexible tip apex. *Phys. Rev. B* **2017**, *95*, 045407.
- (40) Carracedo-Cosme, J.; Romero-Muñiz, C.; Pou, P.; Pérez, R. QUAM-AFM: A Free Database for Molecular Identification by Atomic Force Microscopy. *J. Chem. Inf. Model.* **2022**, *62*, 1214–1223.
- (41) Priante, F.; Oinonen, N.; Tian, Y.; Guan, D.; Xu, C.; Cai, S.; Liljeroth, P.; Jiang, Y.; Foster, A. S. Structure discovery in Atomic Force Microscopy imaging of ice. *ACS Nano* **2024**, *18*, 5546–5555.
- (42) Bilodeau, C.; Jin, W.; Jaakkola, T.; Barzilay, R.; Jensen, K. F. Generative models for molecular discovery: Recent advances and challenges. *WIREs Comput. Mol. Sci.* **2022**, *12*, No. e1608.
- (43) Hoogeboom, E.; Satorras, V. G.; Vignac, C.; Welling, M. Equivariant Diffusion for Molecule Generation in 3D. *arXiv* **2022**, arXiv:2203.17003 <https://arxiv.org/abs/2203.17003>.
- (44) Daigavane, A.; Kim, S.; Geiger, M.; Smidt, T. Symphony: Symmetry-Equivariant Point-Centered Spherical Harmonics for Molecule Generation. *arXiv* **2023**, arXiv:2311.16199 <https://arxiv.org/abs/2311.16199>.
- (45) Pavliček, N.; Swart, I.; Niedenführ, J.; Meyer, G.; Repp, J. Symmetry dependence of vibration-assisted tunneling. *Phys. Rev. Lett.* **2013**, *110*, 136101.
- (46) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.
- (47) Havu, V.; Blum, V.; Havu, P.; Scheffler, M. Efficient integration for all-electron electronic structure calculation using numeric basis functions. *J. Comput. Phys.* **2009**, *228*, 8367–8379.
- (48) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (49) Bardeen, J. Tunneling from a Many-Particle Point of View. *Phys. Rev. Lett.* **1961**, *6*, 57–59.
- (50) de la Torre, B.; Švec, M.; Foti, G.; Krejčí, O.; Hapala, P.; Garcia-Lekue, A.; Frederiksen, T.; Zbořil, R.; Arnau, A.; Vázquez, H.; Jelinek, P. Submolecular Resolution by Variation of the Inelastic Electron Tunneling Spectroscopy Amplitude and its Relation to the AFM/STM Signal. *Phys. Rev. Lett.* **2017**, *119*, 166001.
- (51) Hapala, P.; Kichin, G.; Wagner, C.; Tautz, F. S.; Temirov, R.; Jelinek, P. Mechanism of high-resolution STM/AFM imaging with functionalized tips. *Phys. Rev. B* **2014**, *90*, 085421.
- (52) Hapala, P.; Temirov, R.; Tautz, F. S.; Jelinek, P. Origin of high-resolution IETS-STM images of organic molecules with functionalized tips. *Phys. Rev. Lett.* **2014**, *113*, 226101.

(53) Giessibl, F. J. A direct method to calculate tip-sample forces from frequency shifts in frequency-modulation atomic force microscopy. *Appl. Phys. Lett.* **2001**, 78, 123–125.

(54) Oktay, O.; Schlemper, J.; Folgoc, L. L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N. Y.; Kainz, B.; Glocker, B.; Rueckert, D. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv* **2018**, arXiv:1804.03999 <https://arxiv.org/abs/1804.03999>.

(55) Xu, B.; Wang, N.; Kong, H.; Chen, T.; Li, M. Empirical Evaluation of Rectified Activations in Convolutional Network. *arXiv* **2015**, arXiv:1505.00853v2 <https://arxiv.org/abs/1505.00853>.

(56) Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In *Proceedings of the 32nd International Conference on Machine Learning*; JMLR: W&CP: Lille, France, 2015; Vol. 31.

(57) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Köpf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*; Curran Associates Inc.: Red Hook, NY, 2019; pp 8026–8037. 32 .

(58) Kingma, D. P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2015**, arXiv:1412.6980v9 <https://arxiv.org/abs/1412.6980>.