What makes a good descriptor for heterogeneous ice nucleation on OH-patterned surfaces

Philipp Pedevilla

Thomas Young Centre and London Centre for Nanotechnology, 17-19 Gordon Street, London WC1H 0AH, United Kingdom
and Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, United Kingdom

Martin Fitzner and Angelos Michaelides[*]

Thomas Young Centre and London Centre for Nanotechnology, 17-19 Gordon Street, London WC1H 0AH, United Kingdom
and Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom

Freezing of water is arguably one of the most common phase transitions on Earth and almost always happens heterogeneously. Despite its importance, we lack a fundamental understanding of what makes substrates efficient ice nucleators. Here we address this by computing the ice nucleation (IN) ability of numerous model hydroxylated substrates with diverse surface hydroxyl (OH) group arrangements. Overall, for the substrates considered, we find that neither the symmetry of the OH patterns nor the similarity between a substrate and ice correlate well with the IN ability. Instead, we find that the OH density and the substrate-water interaction strength are useful descriptors of a material's IN ability. This insight allows the rationalization of ice nucleation ability across a wide range of materials and can aid the search and design of novel potent ice nucleators in the future.

I. INTRODUCTION

Nucleation is a process that plays a pivotal role in numerous fields. Self-assembly during biomineralization [1] or nanostructure formation [2], epitaxial growth of semiconductor heterostructures [3] or the controlled formation of quantum dots through heteroepitaxy [4] are just some examples. One of the most common nucleation processes on Earth is the freezing of water, and despite its ubiquity, pure water is surprisingly difficult to freeze. Although frost on car windshields and ice accumulation in the freezer compartment of a fridge are common annoyances that suggest otherwise, the ease of water freezing on its own can hardly be blamed for such events. The thermodynamic freezing point of water is $0\,^\circ$C; pure water, however, can remain in its liquid state until $-43\,^\circ$C [5]. Indeed, most ice on Earth does not freeze by itself (homogeneously) but instead with the help of a large variety of different substrates (heterogeneously). These substrates can be mineral dust, soot particles, organic and even biological materials [6]. It will not come as a surprise, then, that the presence of such particles in, for example, clouds plays a crucial role in determining the amount of ice in them, which in turn has implications for Earth's climate [6–9].

A considerable body of experimental work has been carried out with a view to understanding ice formation. Surface science measurements have provided an atomistic understanding of the initial stages of water clustering and ice formation on well-defined atomically flat surfaces [10,11], but these experiments are currently not applicable under atmospherically relevant conditions. Scanning electron microscopy (SEM) measurements provide insight into nucleation at a micron level, thus shedding light on the dependence of ice nucleation ability on surface topology. Through such experiments it was, for example, revealed that edges and cracks at mineral surfaces can play a crucial role in nucleating ice [12,13]. Droplet

freezing experiments have made the systematic screening of numerous materials' ice nucleating ability possible [14]. This has revealed insight that, e.g., oxidizing graphene flakes improves their nucleating ability [15]. However, neither SEM nor droplet freezing measurements provide direct molecular-level information about *how* and ultimately *why* a substrate is able to aid the formation of ice. What is currently missing, simply stated, is an understanding of what it is that makes materials good or bad ice nucleators. Having this knowledge would have wide-ranging implications in areas such as cryotherapy, aviation, the oil industry, and the atmospheric sciences.

One potential way to tackle this issue is with computer simulations, and studying nucleation *in silico* is indeed a thriving field (see, e.g., [16–30]). For ice nucleation, in particular, it became feasible only recently to use all-atom force field models [20,21,31–33]. However, looking at a wide variety of different substrates in order to extract general trends with such force fields is still out of reach. Advances in force field representations of water, such as the mW potential [34], have made it feasible over the past few years to study a variety of model surfaces and their impact on ice nucleation (IN) ability [23,25–29,35–37]. Despite giving valuable insight into heterogeneous ice nucleation, a clear picture of what makes a good ice nucleator remains elusive.

Here we focus on this issue by providing an interpretation which is able to explain and therefore predict ice nucleation ability on hydroxylated model surfaces with high accuracy. The surfaces of many good ice nucleators, ranging from inorganic to organic and even biological material, are hydroxylated. Therefore understanding the connection between OH group arrangement and ice nucleation efficiency is of paramount importance. Whereas some ice nucleators, such as kaolinite, have OH groups arranged in a structure that resembles ice [38], other potent ice nucleators such as feldspar [14] and cholesterol [39] have no apparent lattice match with ice. Furthermore, well-defined surface science measurements of ice growth on Cu have revealed that free OH groups indeed act as nucleation sites for water adsorption and subsequent ice

---

[*]angelos.michaelides@ucl.ac.uk

growth [11]. With this in mind we computed the nucleation rate constants of a large variety of structurally diverse model hydroxylated substrates, all with distinct OH group patterns. We find that, simply put, if adsorption of small water clusters is weak, water molecules can rearrange in the contact layer, and ice can form, even if the substrate does not resemble an icelike structure. Conversely, if adsorption of such clusters is strong, this rearrangement cannot take place, and ice formation is not promoted. Calculating the adsorption properties of small water clusters with, for example, all-atom force fields or *ab initio* calculations can be readily done nowadays; thus this work opens up the possibility of fast and efficient predictions of the ice nucleating potential of substrates in general.

The rest of this paper is structured as follows: In Sec. II we describe how we constructed the model substrates and the computational approach used to obtain nucleation rates. This is followed in Sec. III with analysis and discussion on descriptors used to characterize IN ability. We then follow this in Sec. IV with a brief discussion of how these descriptors can be applied to some realistic materials and surfaces along with a summary of our results.

II. COMPUTATIONAL APPROACH

We considered eight different tilings of the plane based on triangles, squares, hexagons, octagons, and combinations thereof, see Fig. 1. These particular polygons were chosen because hydroxyl groups and water molecules in realistic systems are most likely to form structures based on these shapes. On kaolinite, for example, hydroxyl groups are arranged in a triangular pattern [38]. Squarelike arrangements of water molecules have also been suggested [40,41]. On some cholesterol faces, hydroxyl groups are also arranged in a nearly squarelike fashion [42]. Hexagonal arrangements of water molecules not only make up ice I, but can also be found on, for example, microcline (001) [43]. Also, unconventional water structures consisting of a mix of octagons and squares are known [44] and hence have been considered in this study

as well. For each tiling, we varied the nearest-neighbor OH distance (d_{OH-OH}) between 2 Å (high-density OH patterns) and 6 Å (relatively low-density OH patterns). This range of OH separations covers the range of OH separations for existing materials, and with future experiments using controlled hydroxylation of inorganic substrates in mind [45–47], extends beyond this range in both the low- and high-density regimes. It is not, however, designed to encompass substrates that are only very sparsely decorated in OH groups, such as the predominantly graphitic surfaces examined in [22].

We used the mW model [34] to represent water-water interactions. The mW water model has proven very successful in studying the behavior of water, and in particular, ice nucleation [23,25–29,35–37]. The mW potential, for example, describes the density of both water and ice better than commonly used all-atom force fields, and also describes the nucleation of stacking-disordered ice well. The interested reader is referred to the Supplemental Material (SM [48]), where we discuss in more detail the suitability of this potential for studying aspects of ice nucleation ability. Hydroxyl groups were modeled as frozen mW molecules, as was done before [22]. Note that although the positions of the OH groups are frozen, the hydrogen bond directions of the OH groups are not. This is because the hydrogen bond network arrangement is simulated by means of a three-body potential which can act in any direction. Therefore, the hydrogen bond network that the OH groups induce is fully flexible. For each OH pattern, two model substrate structures were obtained by adding a Lennard-Jones (9-3) wall potential in the z direction at the bottom of the simulation cell (with interaction strength of either $\varepsilon = 0.05$ eV or $\varepsilon = 0.20$ eV). The resulting adsorption energies cover a physically reasonable range, from -0.11 eV (weak adsorption as found on hydrophobic surfaces) to -0.88 eV (strong adsorption as found on hydrophilic surfaces) per water molecule. Note that we did not tune the adsorption energies to fall within this range, but instead they emerged naturally from the large structural space of OH arrangements. We therefore believe that the substrate space considered in
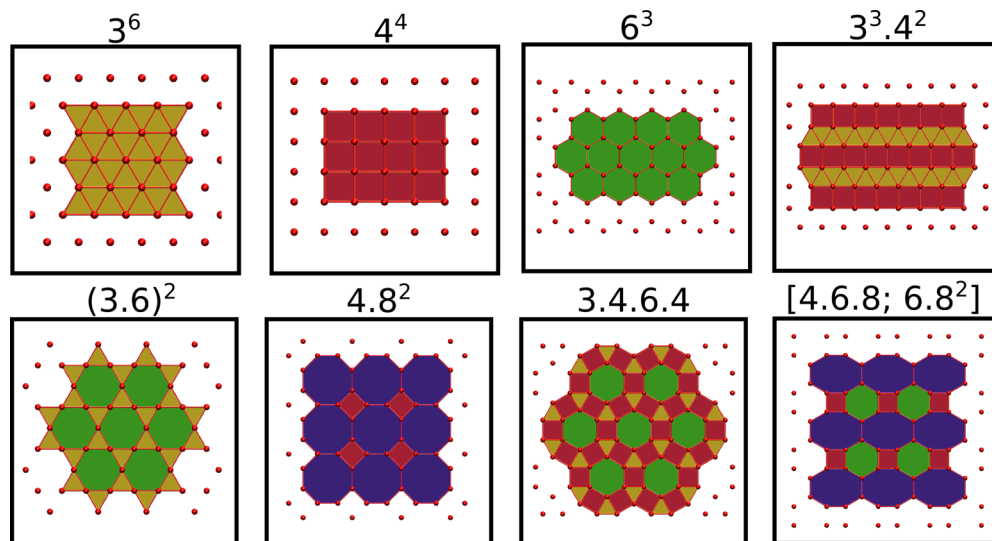

FIG. 1. Range of OH tiling patterns considered in this study with standard vertex notation on top of each structure. All OH patterns consist of triangles (yellow), squares (red), hexagons (green), octagons (blue), and various combinations thereof.

this work mimics a broad range of realistic systems, without representing any particular system explicitly. 4000 water molecules were placed on top of each surface, which is sufficient to restore bulk densities of liquid water [29]. The lateral dimensions of the simulation boxes were chosen to be as close as possible to 50×50 Å, and never less than 40×40 Å. A total of 164 different substrates were generated in this way.

Nucleation rate constants were computed by means of brute force molecular dynamics simulations at 205 K with an established protocol [27–29]. In brief, the procedure involved quenching 15 simulations to the target temperature and measuring the induction times for nucleation, from which we estimated the nucleation rate. The induction times were identified by the drop in potential energy, which coincides with the nucleation of sizable ice clusters (see, e.g., [29]). Further details can be found in the SM [48]. The average freezing temperature obtained from progressively decreasing the temperature of five simulations for each substrate by -1 K/ns starting at 270 K was also calculated; more details about this procedure can be found in Refs. [22–24].

III. RESULTS

Figure 2(a) shows the nucleation rate relative to the homogeneous one as a function of the underlying substrate pattern. The combination of supercooling and simulation time (hundreds of nanoseconds for the longest) in our study leads to nucleation rates within a few orders of magnitude. A nucleation rate that is 2 orders of magnitude larger than the homogeneous rate is the largest rate achieved, similar to previous work [29], and corresponds to immediate freezing within the first few nanoseconds. We therefore focus on the more qualitative distinction between good and bad nucleators, define good ice nucleators to be at least 10 times more efficient than homogeneous nucleation, and label everything that nucleates ice at a slower rate than that a bad nucleator. Note that the distinction between good and bad ice nucleators is a binary classification aimed at finding the key differences between

substrates that are able to accelerate ice nucleation and ones which do not. Figure 2(b) shows the correlation between nucleation rate constants and freezing temperatures. Apart from the strong correlation between the two methods, the classification of good and bad ice nucleators also agrees well between the two approaches. In the case of cooling ramps, we define a good ice nucleator to nucleate above the homogeneous freezing temperature of mW water at a cooling rate of -1 K/ns. The red lines in Fig. 2(b) show the boundaries between good and bad ice nucleators, and the areas in green highlight the regimes in which both protocols lead to the same classification of IN ability. Clearly, the vast majority of points is assigned in the same fashion, irrespective of the computational protocol used to calculate IN ability. Furthermore, the sparsity of the nucleation rates around the boundary between good and bad ice nucleators combined with similar results obtained via two different approaches means that our results are not dependent on the details of the choice of this classification boundary (see SM [48]).

From the data reported in Fig. 2(a), it is clear that for every OH pattern good and bad nucleating efficiencies can be found. Hexagonal surfaces (6^3), i.e., surfaces that resemble the symmetry of ice, are by no means better than, for example, structures that do not seem to resemble ice at all. In scenarios where the substrate structure matches the structure of ice very well (3^6 and 6^3, $d_{OH-OH} \approx 2.7$ Å), a hexagonal contact layer akin to ice forms which leads to fast nucleation, see Fig. 3. This behavior of a substrate templating the structure of ice and subsequently leading to high IN ability is not surprising. However, many other examples of fast nucleation exist where no such templating effect is present. In these cases, the contact layer does not resemble the structure of ice but rearranges during nucleation to form an icelike structure (Fig. 3). Independent of nucleation via templating or contact layer rearrangement, we observe stacking-disordered ice to form, which is expected at strong supercooling [49–52]. We note in passing that in Fig. 2(a) nucleation rate constants below the homogeneous one are reported. These do not mean
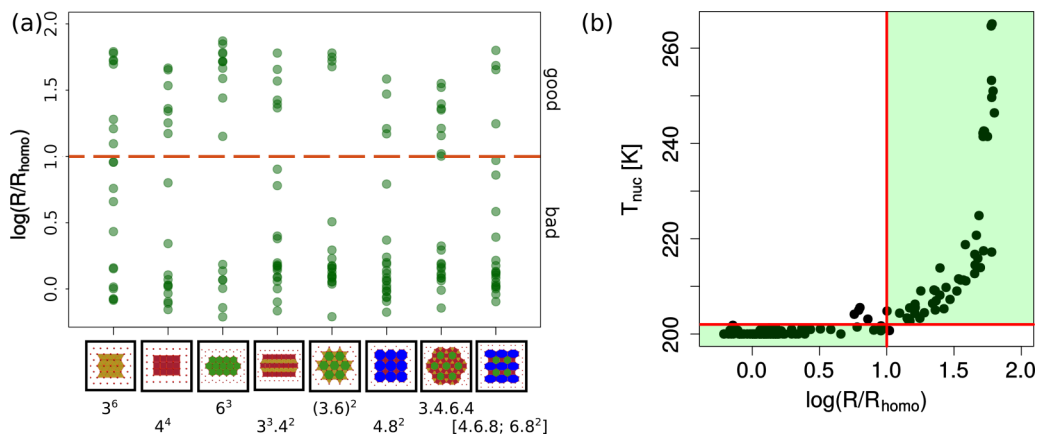

FIG. 2. (a) Ice nucleation rate constants relative to the homogeneous nucleation rate constants for the range of OH tiling patterns considered. Independent of the symmetry of the substrate, both good and bad ice nucleation efficiencies exist. The color code used to label polygons is the same as in Fig. 1. Good and bad ice nucleators are separated by the dashed line, with good IN being at least 10 times more efficient compared to homogeneous nucleation. (b) Classification of IN ability depending on the computational approach to induce nucleation. Nucleation rate constants are plotted on the x axis and freezing temperatures on the y axis. The red lines show the boundary between good and bad ice nucleators for both approaches.
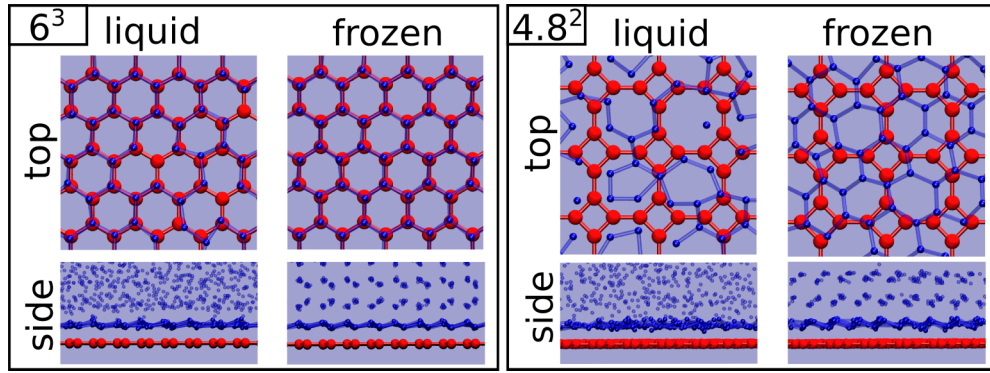
FIG. 3. Contact layer rearrangement upon freezing. The figure shows the water interface before and after ice nucleation for a substrate that matches the structure of ice (6^3) and one that does not (4.8^2). OH groups are shown in red, and water molecules in blue. In the top-view images, only the water molecules in the contact layer are displayed.

that a substrate is truly inhibiting ice nucleation. The finite simulation box size rather means that a contact layer on such a substrate is deactivated and will therefore not contribute to the homogeneous ice nucleation rate constant.

The result that the symmetry of the OH group arrangement of the substrate does not play a crucial role in determining the ice nucleation efficiency is a key finding of this work. Given the widely held view that hexagonal substrates make the best templates for ice nucleation, it comes somewhat as a surprise. In the rest of this paper, we will focus on trying to explain this behavior by identifying features that are able to discriminate between good and bad ice nucleating agents. Various descriptors were considered, and we now discuss how some of the most interesting structural and energetic descriptors perform in discriminating between good and bad IN ability.

Lattice match is frequently used to explain IN ability [53]. The basic idea behind this is simple: ice can form more readily on substrates that themselves look icelike. Recently, while studying the IN ability of alcohol films, Qiu *et al.* [24] showed that the area match between ice and a substrate as well as anisotropy can predict IN ability. Area match is defined as $[(a_{OH} \times b_{OH})/(a_{ice} \times b_{ice}) - 1] \times 100$ and anisotropy as $b_{OH}/(\sqrt{3}a_{OH})$ [24], where a_{OH} and b_{OH} are the rectangular lattice parameters of the OH structure and a_{ice} and b_{ice} are the lattice parameters of ice. Figure 4(a) shows how area mismatch correlates with ice nucleating efficiency for a certain range (similar to the one considered by Qiu *et al.*). We find that, in line with Qiu *et al.*, particular substrates are most efficient with an mismatch close to zero. This is the case for hexagonal (6^3) and triangular (3^6) arrangements shown in red/blue, respectively. However, square arrangements, such as, e.g., (4^4), follow a different trend, shown in green. In addition, the predictive quality of this descriptor diminishes if we consider a larger range of area mismatch, shown in Fig. 4(b). There we projected all area mismatches in the range [0,100]. The largest match considered in our study was 400%. We do the same for the other substrates in [Fig. 4(c)], and it becomes clear that the area match is not able to discriminate the IN ability of all substrates. The same behavior is found for the anisotropy as illustrated in Fig. 4(d). Each of our OH symmetries is associated with an anisotropy that is independent of the OH density. However, as we already saw from Fig. 2(a), the same symmetry (and hence

anisotropy) can make for both good and bad ice nucleators. Overall, our findings agree with those of Qiu *et al.*: If the OH arrangement resembles the structure of ice very closely, area match and anisotropy describe the nucleation enhancement well. We show, however, that for OH symmetries that do not belong to this specific class, these descriptors do not correlate in general with the IN ability.

To get a more general measure of the matching quality, we compute the match between a substrate and ice using the root-mean-square deviation (RMSD) of various ice crystallites on top of a surface and report the RMSD that is associated with the ice face that matches the substrate most closely. Details of the approach are outlined in the SM [48]. Here it is sufficient to note that a small value of the mismatch signifies a good match between the substrate and ice, and a large value of
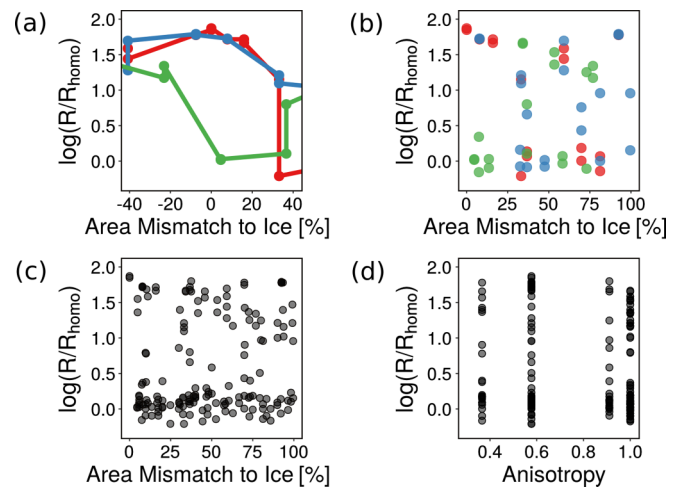

FIG. 4. Capability of area mismatch and lattice anisotropy to describe IN ability as suggested by Qiu *et al.* [24]. (a) Our results agree with Qiu *et al.* [24], in that a subset of structures (hexagonal, red and triangular, blue) can be described well using area match, whereas others (such as a square arrangement of OH groups, green) fall outside the scope of this descriptor. (b) The IN ability of the same structures described in (a) does not follow the expected behavior once larger deviations of lattice match are considered. (c, d) The full set of substrates considered in this study as a function of area mismatch and anisotropy, respectively.
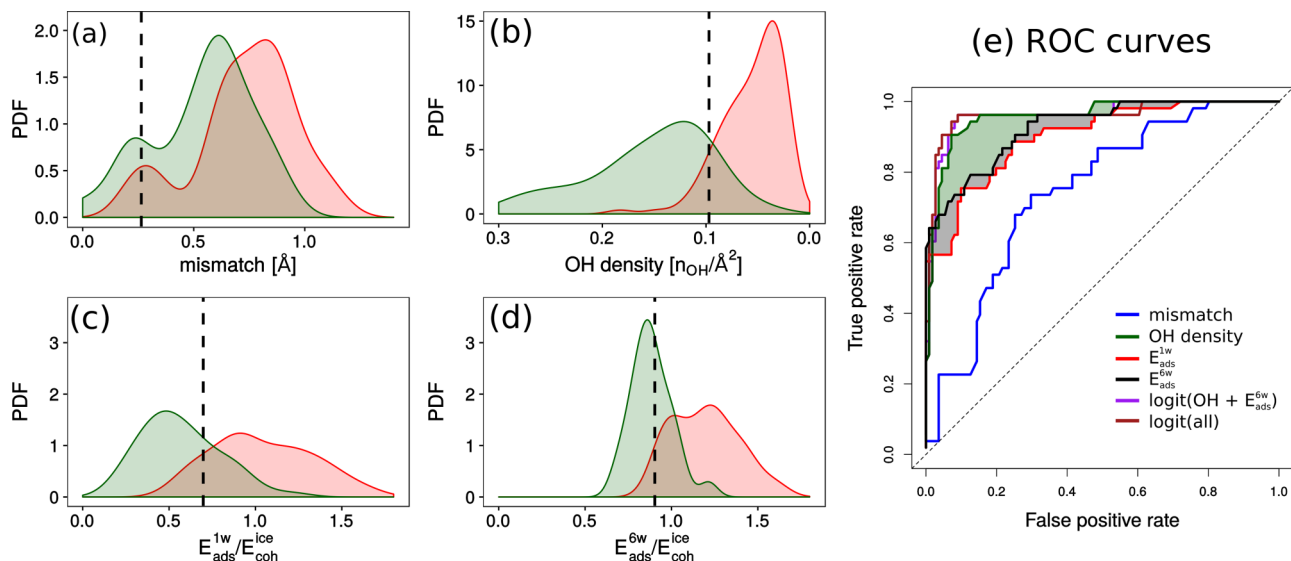
FIG. 5. Performance of various descriptors for heterogeneous ice nucleation. Probability density function (PDF) of good (green) and bad (red) ice nucleators as a function of the mismatch between the substrates and ice (a), OH density (b), the adsorption energy of a single water (E_{ads}^{1w}) (c), and of six waters (E_{ads}^{6w}) (d). Adsorption energies are reported relative to the cohesive energy of ice (E_{coh}^{ice}). The black dashed line shows the decision boundary that corresponds to the highest classification accuracy. (e) The ROC curves for the features in (a)–(d) as well as combinations thereof in the form of logit models.

the mismatch means that the substrate and ice do not match well. Figure 5(a) shows the performance of this mismatch in distinguishing good and bad ice nucleators. On the *x* axis, the classifier is shown, and on the *y* axis the probability density function (PDF) is shown for good (green) and bad (red) ice nucleators. The smaller the overlap between the two densities, the better the classifier. An ideal classifier would completely separate the two densities. It can be seen that the PDFs for the good and bad nucleating substrates overlap considerably, hence making lattice mismatch a poor descriptor in general. Clearly, therefore, the similarity between a substrate and ice is not a good way to think about IN ability for the hydroxylated surfaces considered here. This is in line with previous work, e.g., simulations on Lennard-Jones crystals exposing different surfaces [29].

Figure 5(b) shows that OH density separates IN ability significantly better than mismatch does. This is particularly clear when comparing the overlap between good and bad nucleators in both cases. The PDF for good nucleators is close to zero when the PDF of bad nucleators reaches its maximum and vice versa. This in turn means that a classification based on OH density will have low false positive and false negative classification rates. As a consequence, for the range of hydroxylated surfaces considered here, the OH density for regular tiling patterns is a good descriptor for IN ability: high OH densities (small d_{OH-OH}) and low OH densities (large d_{OH-OH}) correspond to good and bad IN ability, respectively. This is consistent with the work of Qiu *et al.* [24], which shows that having a close area match to ice (high OH density) nucleated better than substrates with a larger mismatch.

In previous studies, water monomer adsorption energies were shown to correlate with IN ability, however this was system dependent [28,29]. Here we find as well that the monomer can be a good descriptor. Specifically, small adsorption energies yield good nucleators whereas large adsorption

energies generally lead to poor nucleating substrates. However, we also find here that adsorption energies that take into account more water molecules are consistently better. With this in mind we show in Figs. 5(c) and 5(d) how the adsorption energy of a single water molecule and six water molecules perform in discriminating between good and bad IN ability. The PDF range of good and bad IN ability is considerably broader when using only a single water molecule compared to the case when using six water molecules, and densities between good and bad ice nucleators overlap more substantially in the former case. Moreover, the maximum of the density for bad ice nucleators overlaps more with the good ice nucleators when using only one water molecule, which also makes the classification in the former case worse. The adsorption energy that most accurately discriminates between good and bad IN ability is less than the bulk cohesive energy of ice [around 90%, see Fig. 5(d)]. If water molecules adsorb more strongly on the surface than they would in ice, it is not favorable energetically to rearrange the contact layer into ice. If, however, adsorption is weaker, then this rearrangement of the contact layer can and does happen, and hence ice can form.

Figure 5(e) shows a quantitative comparison between different classifiers (and also combinations of them) in the form of receiver operator curves (ROCs). ROCs are a common way to quantify the quality of a descriptor or set of descriptors, and show the true positive rate of classification against the false positive rate. An ideal classifier would have a 100% true positive and 0% false positive classification, and a completely random classifier would follow the *y* = *x* line (area of 0.5), shown as a dashed line in Fig. 5(e). The performance of mismatch (blue), OH density (green), monomer adsorption energy (red), and hexamer adsorption energy (black) is shown. Additionally, we highlight the area between E_{ads}^{1mW} and E_{ads}^{6mW} as well as the area between E_{ads}^{6mW} and the OH density to show the respective improvement. We also looked at

how combinations of descriptors perform compared to single classifiers using a logistic regression model (logit model). Figure 5(e) shows that the improvement from combining classifiers is only a minor one, which is in line with the fact that OH density and adsorption energy are related with each other (see SM [48]). Combining them will therefore not improve the classification substantially.

IV. DISCUSSION AND CONCLUSIONS

Having established a simple means of gauging the IN ability of model systems, we now show that the insights can be used to rationalize various experimental observations on hydroxylated surfaces. One of the most extensively studied systems for heterogeneous ice nucleation is the mineral kaolinite. The basal surface of kaolinite has received the most attention and is, to the best of our knowledge, the only system for which the heterogeneous ice nucleation rate has been calculated quantitatively with an all-atom force field [32]. Interestingly, water hexamers bind to kaolinite with an adsorption energy that is ∼90% the cohesive energy of ice, implying that kaolinite is an effective ice nucleating agent. In addition, the prism face nucleated on kaolinite [20,32,54], which is worth pointing out because water structures found by Hu and Michaelides did not resemble the prism face of ice [38]. Thus, in order for ice to form on the basal face of kaolinite, water molecules had to rearrange their structure in the contact layer. This example illustrates a key finding of this work: if the interaction between the substrate and water is not too strong, the contact layer can rearrange, in this particular case to form the prismatic face of ice.

A wide-ranging implication of this work is that edges and cracks for minerals in general might be more important for IN ability than previously thought. At edges and cracks there will be more broken covalent bonds, which will be saturated with hydroxyl groups under ambient conditions, which implies that there will often be a higher OH density at defect sites. Therefore, they might play a more pronounced role than the surfaces that are exposed predominantly, and recent experimental evidence is supporting this hypothesis [12,13]. A recent experimental study demonstrated, for example, that ice does not form on the basal face of kaolinite but instead nucleates on the edges between stacked basal platelets [13]. Similarly, feldspar's outstanding IN ability was attributed to the formation of ice in a defect site [12].

To summarize, this work is a systematic study of how ice nucleation ability relates to OH patternation on hydroxylated surfaces. The striking result was that the ice nucleation ability depends less on the symmetry of the pattern itself but rather on the OH group density. An interpretation that explains this was developed based on the interaction strength between water and the substrate relative to the bulk cohesive energy of ice. Substrates that template icelike contact layers were found to be equally efficient ice nucleators at the supercooling in this study than non-icelike substrates, if they do not adsorb water too strongly. The rationale behind this interpretation is physically well motivated: if the surface is able to accommodate water not too strongly, the contact layer can rearrange easily to reach the basin of ice. If, however, water adsorbs more strongly than the energetic gain from this, the energy cost associated with rearranging the structure is too high and ice cannot form efficiently on such a substrate. The information required to estimate IN ability is accessible experimentally and computationally, and hence we believe our interpretation shows ways forward to understanding heterogeneous ice nucleation in the future.

[1] G. He, T. Dahl, A. Veis, and A. George, Nat. Mater. **2**, 552 (2003).

[2] P. Jonkheijm, P. van der Schoot, A. P. H. J. Schenning, and E. W. Meijer, Science **313**, 80 (2006).

[3] Y. Chen and J. Washburn, Phys. Rev. Lett. **77**, 4046 (1996).

[4] B. Yang, F. Liu, and M. G. Lagally, Phys. Rev. Lett. **92**, 025502 (2004).

[5] J. A. Sellberg, C. Huang, T. A. McQueen, N. D. Loh, H. Laksmono, D. Schlesinger, R. G. Sierra, D. Nordlund, C. Y. Hampton, D. Starodub *et al.*, Nature (London) **510**, 381 (2014).

[6] B. J. Murray, D. O'Sullivan, J. D. Atkinson, and M. E. Webb, Chem. Soc. Rev. **41**, 6519 (2012).

[7] K. A. Pratt, P. J. DeMott, J. R. French, Z. Wang, D. L. Westphal, A. J. Heymsfield, C. H. Twohy, A. J. Prenni, and K. A. Prather, Nat. Geosci. **2**, 398 (2009).

[8] T. Bartels-Rausch, Nature (London) **494**, 27 (2013).

[9] B. J. Murray, Science **355**, 346 (2017).

[10] J. Carrasco, A. Michaelides, M. Forster, S. Haq, R. Raval, and A. Hodgson, Nat. Mater. **8**, 427 (2009).

[11] M. Forster, R. Raval, A. Hodgson, J. Carrasco, and A. Michaelides, Phys. Rev. Lett. **106**, 046103 (2011).

[12] A. Kiselev, F. Bachmann, P. Pedevilla, S. J. Cox, A. Michaelides, D. Gerthsen, and T. Leisner, Science **355**, 367 (2017).

[13] B. Wang, D. A. Knopf, S. China, B. W. Arey, T. H. Harder, M. K. Gilles, and A. Laskin, Phys. Chem. Chem. Phys. **18**, 29721 (2016).

[14] J. D. Atkinson, B. J. Murray, M. T. Woodhouse, T. F. Whale, K. J. Baustian, K. S. Carslaw, S. Dobbie, D. O'Sullivan, and T. L. Malkin, Nature (London) **498**, 355 (2013).

[15] T. F. Whale, M. Rosillo-Lopez, B. J. Murray, and C. G. Salzmann, J. Phys. Chem. Lett. **6**, 3012 (2015).

[16] D. Zahn, Phys. Rev. Lett. **92**, 040801 (2004).

[17] D. Moroni, P. R. T. Wolde, and P. G. Bolhuis, Phys. Rev. Lett. **94**, 235703 (2005).

[18] A. Barducci, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **100**, 020603 (2008).

[19] G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen, and A. Michaelides, Chem. Rev. (Washington, DC, US) **116**, 7078 (2016).

[20] S. A. Zielke, A. K. Bertram, and G. N. Patey, J. Phys. Chem. B **120**, 1726 (2016).

[21] A. Haji-Akbari and P. G. Debenedetti, Proc. Natl. Acad. Sci. USA **112**, 10582 (2015).

[22] L. Lupi and V. Molinero, J. Phys. Chem. A **118**, 7330 (2014).

[23] L. Lupi, A. Hudait, and V. Molinero, J. Am. Chem. Soc. **136**, 3156 (2014).

[24] Y. Qiu, N. Odendahl, A. Hudait, R. Mason, A. K. Bertram, F. Paesani, P. J. DeMott, and V. Molinero, J. Am. Chem. Soc. **139**, 3052 (2017).

[25] X. Zhang, M. Chen, and M. Fu, J. Chem. Phys. **141**, 124709 (2014).

[26] A. Reinhardt and J. P. K. Doye, J. Chem. Phys. **141**, 084501 (2014).

[27] S. J. Cox, S. M. Kathmann, B. Slater, and A. Michaelides, J. Chem. Phys. **142**, 184705 (2015).

[28] S. J. Cox, S. M. Kathmann, B. Slater, and A. Michaelides, J. Chem. Phys. **142**, 184704 (2015).

[29] M. Fitzner, G. C. Sosso, S. J. Cox, and A. Michaelides, J. Am. Chem. Soc. **137**, 13658 (2015).

[30] Y. Bi, B. Cao, and T. Li, Nat. Commun. **8**, 15372 (2017).

[31] J. R. Espinosa, C. Vega, C. Valeriani, and E. Sanz, J. Chem. Phys. **144**, 034501 (2016).

[32] G. C. Sosso, T. Li, D. Donadio, G. A. Tribello, and A. Michaelides, J. Phys. Chem. Lett. **7**, 2350 (2016).

[33] A. Haji-Akbari and P. G. Debenedetti, Proc. Natl. Acad. Sci. USA **114**, 3316 (2017).

[34] V. Molinero and E. B. Moore, J. Phys. Chem. B **113**, 4008 (2009).

[35] E. B. Moore, E. de La Llave, K. Welke, D. A. Scherlis, and V. Molinero, Phys. Chem. Chem. Phys. **12**, 4124 (2010).

[36] Y. Bi, R. Cabriolu, and T. Li, J. Phys. Chem. C **120**, 1507 (2016).

[37] L. Lupi, B. Peters, and V. Molinero, J. Chem. Phys. **145**, 211910 (2016).

[38] X. L. Hu and A. Michaelides, Surf. Sci. **602**, 960 (2008).

[39] G. E. Ewing, Chem. Rev. (Washington, DC, US) **106**, 1511 (2006).

[40] G. Algara-Siller, O. Lehtinen, F. C. Wang, R. R. Nair, U. Kaiser, H. A. Wu, A. K. Geim, and I. V. Grigorieva, Nature (London) **519**, 443 (2015).

[41] J. Chen, G. Schusteritsch, C. J. Pickard, C. G. Salzmann, and A. Michaelides, Phys. Rev. Lett. **116**, 025501 (2016).

[42] B. M. Craven, Nature (London) **260**, 727 (1976).

[43] P. Pedevilla, S. J. Cox, B. Slater, and A. Michaelides, J. Phys. Chem. C **120**, 6704 (2016).

[44] J. Yang, S. Meng, L. F. Xu, and E. G. Wang, Phys. Rev. Lett. **92**, 146102 (2004).

[45] B. Yang, E. Emmez, W. E. Kaden, X. Yu, J. A. Boscoboinik, M. Sterrer, S. Shaikhutdinov, and H. J. Freund, J. Phys. Chem. C **117**, 8336 (2013).

[46] B. Yang, S. Shaikhutdinov, and H. J. Freund, J. Phys. Chem. Lett. **5**, 1701 (2014).

[47] X. Yu, E. Emmez, Q. Pan, B. Yang, S. Pomp, W. E. Kaden, M. Sterrer, S. Shaikhutdinov, H.-J. Freund, I. Goikoetxea, R. Wlodarczyk, and J. Sauer, Phys. Chem. Chem. Phys. **18**, 3755 (2016).

[48] See Supplemental Material http://link.aps.org/supplemental/10.1103/PhysRevB.96.115441 for details about the (1) contact layer rearrangement, (2) mismatch calculation, (3) computational setup, (4) structure search method, (5) adsorption energy dependence with OH density, and (6) convergence of the decision boundary, which includes Refs. [55–59].

[49] T. L. Malkin, B. J. Murray, A. V. Brukhno, J. Anwar, and C. G. Salzmann, Proc. Natl. Acad. Sci. USA **109**, 1041 (2012).

[50] W. F. Kuhs, C. Sippel, A. Falenty, and T. C. Hansen, Proc. Natl. Acad. Sci. USA **109**, 21259 (2012).

[51] T. L. Malkin, B. J. Murray, C. G. Salzmann, V. Molinero, S. J. Pickering, and T. F. Whale, Phys. Chem. Chem. Phys. **17**, 60 (2015).

[52] A. Hudait and V. Molinero, J. Am. Chem. Soc. **138**, 8958 (2016).

[53] H. R. Pruppacher and J. D. Klett, *Microphysics of Clouds and Precipitation* (Springer, The Netherlands, 1997).

[54] S. J. Cox, Z. Raza, S. M. Kathmann, B. Slater, and A. Michaelides, Faraday Discuss. **167**, 389 (2013).

[55] S. J. Cox, S. M. Kathmann, J. A. Purton, M. J. Gillan, and A. Michaelides, Phys. Chem. Chem. Phys. **14**, 7944 (2012).

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

[57] S. Nosé, J. Chem. Phys. **81**, 511 (1984).

[58] G. J. Martyna, M. L. Klein, and M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992).

[59] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (CRC Press, 1994).