



Cite this: *Phys. Chem. Chem. Phys.*,
2025, 27, 21553

Finding heterogeneous nucleating agents for ice using a data-driven approach

Zixuan Wang,^a David E. Oliver,^b Andrew J. Bissell,^b Gylan Odling,^b
Colin R. Pulham^{id}^a and Carole A. Morrison^{id}^{*a}

We present a high-throughput data-driven workflow to identify potential heterogeneous nucleating agents from structural databases for phase change materials, such as ice. Our model evaluates the fit between ice I_h and nucleator docked slabs, considering Miller index planes up to (333), thus addressing some of the structural complexities in nucleation by examining crystal morphology features. Bulk water immersion experiments on a set of ten known nucleators set a delineating temperature to distinguish between good and poor nucleation behaviour, which helped derive numerical tolerance limits to allow reliable differentiation on the basis of the number of predicted matching interface models. We then used our algorithm to screen 3500 simple metal oxides and halides taken from the inorganic chemistry structural database (ICSD), and show that just 7% of the former and 3% of the latter were predicted to nucleate ice on the basis of geometric slab matching alone. Subsequent experimental testing of 22 compounds suggested a 64% correct prediction rate, and identified four new ice nucleators (CeO_2 , WO_3 , Bi_2O_3 , Ti_2O_3). Inspired by the ice-nucleating efficiency of copper oxides, we also tested copper tubing with local tap water, and observed sub-cooling suppression, most likely due to copper oxide buildup. Although based on a simple geometric interface matching model, this approach offers an efficient route as a first stage high throughput screen for potential heterogeneous nucleating agents.

Received 15th July 2025,
Accepted 22nd September 2025

DOI: 10.1039/d5cp02701b

rsc.li/pccp

1. Introduction

Super-cooling is a well-known phenomenon whereby a liquid exists in a metastable state below its freezing point. It can affect phase change materials, such as salt hydrates and ice banks, that are used for thermal energy storage.^{1,2} For ice formation, which is the subject of this study, the control of crystal nucleation also has applications in cloud seeding and the production of artificial snow.^{3–5} One pathway to mitigate against super-cooling is through the addition of heterogeneous nucleating agents; these are insoluble materials that present solid templating surfaces to facilitate the nucleating process. This differentiates from freezing point depression, which is a thermodynamic effect driven through the addition of solutes.⁶ Herein our focus is on the former: the onset of freezing induced by surface-facilitated nucleation. Particles that are known to be effective heterogeneous nucleators for ice span a broad range of materials,^{5,7} from ionic salts,^{8–10} to minerals,¹¹ carbonaceous materials¹² and organic matter.¹³

The available literature testifies to the complex nature of heterogeneous ice nucleation, with papers citing the importance of crystallographic similarities,^{14,15} surface chemistries,^{7,14,16–18} topologies,¹⁹ water structuring effects and adsorption strengths,^{20–22} and suspended solid or liquid particle sizes,^{23–25} with experimental conditions that range from macroscopic observations on ice formation in the atmosphere,^{18,25} to those performed in ultra-clean materials chemistry labs.²⁶ Moreover, it is known that different nucleation pathways exist, depending on whether the nucleating particles are immersed in liquid water or suspended in a supersaturated vapour.^{27,28} The broadness of the field, combined with the variations in experiments conducted, results in variable reporting of ‘good’ or ‘poor’ ice-nucleating ability.

From a theoretical perspective, early reports^{8,29} often attributed ice-nucleating ability with a zero-lattice mismatch registry, *i.e.* a close similarity between the unit-cell dimensions of the nucleator and a particular face of the hexagonal phase of ice (ice I_h), typically defined as the basal plane (0001).³⁰ While this has proven effective to account for the well-known ice-nucleating properties of simple compounds such as AgI ,⁸ it is now widely accepted as an over-simplification,³¹ not least because it does not take into account the chemistry of the nucleator/ice-forming interface. Computer simulation has made significant inroads into providing insights at the atomic

^aSchool of Chemistry and EaStCHEM Research School, University of Edinburgh, The King's Buildings, David Brewster Road, Edinburgh, EH9 3FJ, UK.
E-mail: c.morrison@ed.ac.uk

^bSunamp Ltd., Macmerry Business Park, 1 Satellite Park, Tranent, East Lothian, EH33 1RY, UK



level for heterogeneous³² and homogeneous³³ ice nucleation. In particular, work by Michaelides *et al.* has highlighted the importance of understanding surface hydrophobicity, morphology and the variation in the adsorption energy landscape,^{14,17} as well as considering how ordered water molecule layers build up on a nucleating substrate,^{5,17} and how the density of the liquid water reduces near the surface.²² More recently, machine-learning techniques trained on images of water-contact layers and the resulting prediction model (IcePic) have demonstrated success at accurately and rapidly predicting heterogeneous ice-nucleating behaviour.²⁴

Herein we have taken a different approach that looks to take advantage of the wealth of potential heterogeneous nucleators available through databases such as the international centre for diffraction data (ICDD)³⁴ and the inorganic chemistry structural database (ICSD).³⁵ We seek to generate a geometric docking model that assesses the quality of fit between ice I_h /nucleator docked slabs cleaved along Miller index planes from the respective bulk crystal lattices. While this has similarities to the zero-lattice mismatch approach, it goes beyond the low-index planes to consider the docking of all interfaces (both nucleator and ice I_h) described by the Miller indices up to (333). In this way, we are addressing some of the structural complexity of the nucleation process by considering crystal morphology, where ice crystallites could seed on the faces, edges, corners, defects or other surface features of the nucleating crystal that could be described by these higher Miller-index planes. While this study focuses on ice nucleation in bulk water, our overarching goal is to build a generalisable high-throughput framework for predicting heterogeneous nucleation agents for any given phase change material.

Given the variation in the literature regarding experimental set up, our study began with establishing our own experimental benchmarking, *via* bulk water immersion experiments, on a set of ten widely known effective or poor nucleators for ice that we could readily source. We then derived a data-driven approach capable of identifying new heterogeneous crystal nucleators using geometric interface matching, where the quality of fit between ice I_h /nucleator docked slabs cleaved along Miller index planes from the respective bulk crystal lattices are assessed and ranked. By tightening a set of geometric criteria that describe the fit of the docked nucleator and ice I_h cut planes, the number of matching slab interfaces that remain can act as a guide to the likely classification of a good or poor nucleator for ice I_h . On this basis, we then screened the ICSD for several thousand simple metal oxide and halide structures. Testing the predicted outcomes for 22 compounds showed a 64% success rate. Our procedure has also led to the discovery of four compounds, along with standard copper tubing, that can act as ice nucleators under immersion conditions.

2. Bulk water nucleation study

Ten compounds were selected for the heterogeneous immersion nucleation study based on existing literature reports of their ice-nucleating properties and their ready accessibility in

our laboratory. The compounds chosen for their known effective behavior were MnO ,³⁶ FeO (Wüstite),³⁶ AgI ,^{30,37} Cu_2O ,³⁸ AgCl ,³⁷ CuO ³⁹ and SiO_2 (quartz).²⁷ For the poor nucleating agents, we chose BaF_2 ,⁴⁰ CaCO_3 (calcite),⁴¹ and $\text{Al}(\text{OH})_3$ (gibbsite).⁴²

The identity and phase purity of all compounds were verified by powder X-ray diffraction measurements, through cross-referencing against known structures in the ICDD³⁴ database, PDF-5+ (see SI, Section S1). These crystal structures from the database were extracted as crystallographic information files (CIFs) to be used as input structures for our interface matching workflow, which is described in the following section.

Results from the immersion experiments, conducted according to the Experimental Methods, are shown in Fig. 1(a), with further data presented in the SI (Section S3). All measurements are in line with expectations of effective or poor ice nucleation capability according to literature precedents.^{10,19} Ice nucleation onset temperatures are known to vary significantly, even for well-documented compounds like AgI ,^{10,43–45} owing to the inherent stochastic nature of the nucleation process and the variability of different experimental protocols. Our set-up affords sufficient reliability to differentiate between effective nucleators (*e.g.* AgI , Cu_2O) and weak or inactive ones (*e.g.* $\text{Al}(\text{OH})_3$, BaF_2). This is sufficient for our means, as we look only to define a boundary temperature to delineate between the two behaviours. According to the accuracy limitations afforded by the Polar Bear apparatus and our sample preparations (1 wt% solid loading in 10 mL ultra-pure water), we set the boundary temperature to -4°C ; this temperature distinction will be used to experimentally classify all further compounds as either a good or poor ice nucleator under these immersion conditions. We note that our binary classification is an operational choice specific to the accuracy limits of our experimental set up, and does not replace the droplet-freezing $T_{1/2}$ values commonly reported in the literature.⁴⁶ In the absence of a nucleator, our experimental set-up achieves reliable sub-cooling to $-12 \pm 3^\circ\text{C}$ (see SI, Section S3). We fully acknowledge that *e.g.* nucleation chamber experiments⁴⁷ and drop-freezing assays⁴⁸ would yield far more reliable nucleation temperatures than we have achieved here.

3. Ice/nucleator interface matching prediction workflow

With an experimental data set thus defined, the crystallographic information files (CIFs) for the ten compounds discussed in the previous section were used as input models for the interface-matching process. An overview of the general workflow, which was constructed in Python 3, underpinned by ASE⁴⁹ and Pymatgen⁵⁰ is presented in Fig. 2. Throughout this work, we refer to crystal surfaces using Miller indices ($hk(i)l$), which define specific cut planes in the unit cell. Interface-matching refers to the geometric alignment of these cut planes, evaluated according to a set of criteria based on area overlap, angle and unit cell mismatch. Further details are provided in the SI, Section S4.



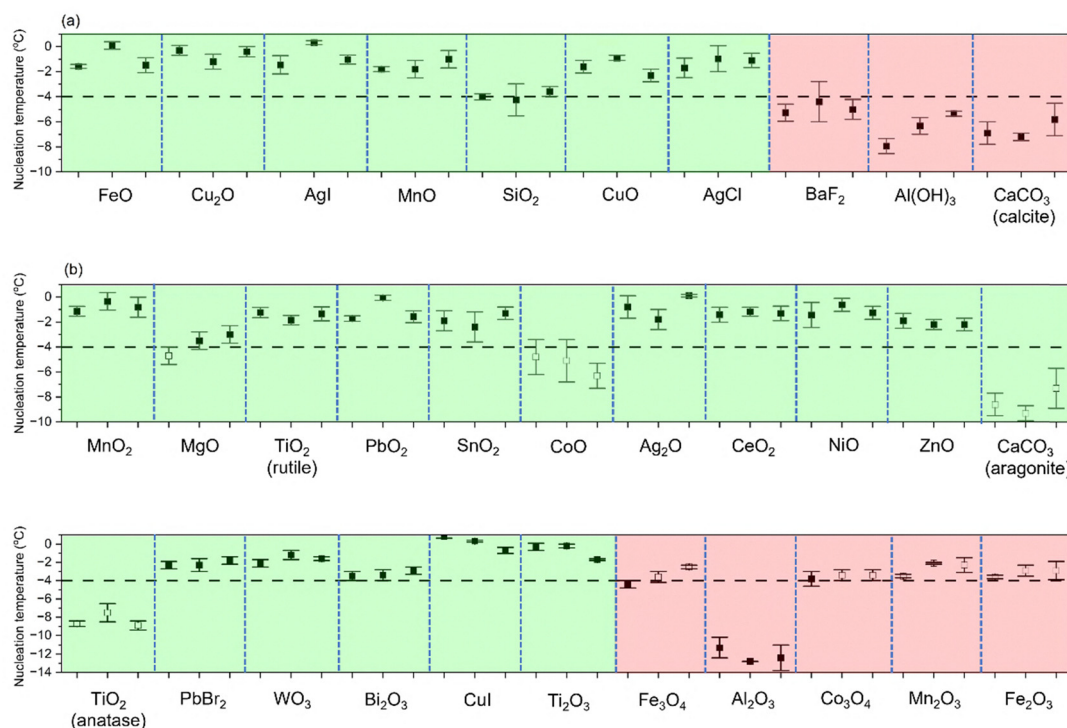


Fig. 1 Freezing onset temperatures from bulk water immersion experiments conducted using the Polar Bear apparatus. (a) Known nucleator dataset that defines the -4 °C decision boundary temperature for classifying good vs. poor nucleators. (b) Experimental results for 22 test compounds predicted via the geometric interface-matching workflow. Compounds are ordered by descending number of predicted matching interfaces with ice I_h (green: ≥ 10 matches; red: < 10 matches). Filled symbols indicate agreement between prediction and experiment; empty symbols indicate disagreement. Error bars represent the standard deviations recorded from the four temperature cycles obtained for each compound (full data set in SI Section S3).

In the first instance, sets of surfaces were created by cleaving the corresponding bulk crystal lattices along the Miller index planes hkl for ice I_h and for hkl for the nucleators to create a pool of 64 non-duplicated surfaces for each crystal lattice. Allowing all ice I_h surfaces to dock on all nucleator surfaces generated a total of 2401 non-duplicated interface models per nucleator.

The geometric matching of two docked surfaces was assessed by searching through integer multiples of the vectors of each surface to find the supercell models that present the smallest unit-cell mismatch. The geometry of each surface is described by two vectors parallel to each slab edge, \vec{a} and \vec{b} , expressed as (m, n) supercells, such that $m\vec{a} \approx n\vec{b}$. These define the new vectors \vec{u} and \vec{v} , respectively (see Fig. 3). A reduction scheme was then used to express the vectors in the slab frame of reference, to negate the effects of translation, rotation or reflection of the individual surfaces. The two slabs were aligned

by minimising $\left| \frac{\|\vec{u}_1\|}{\|\vec{u}_2\|} - 1 \right|$ and $\left| \frac{\|\vec{v}_1\|}{\|\vec{v}_2\|} - 1 \right|$ where the subscripts 1 and 2 denote ice I_h and nucleator supercell surfaces, respectively. Following translational alignment, the slabs were then rotated by transformational matrices $R_i(\theta)$ to lie parallel to each other. The surface generation, alignment and subsequent docking procedures are summarized in Fig. 3.

Five features were then defined to quantify the quality of fit for the resulting bank of ice I_h /nucleator interfaces (Fig. 4).

These were the (i) maximum area overlap, (ii) angle mismatch, (iii) supercell vector mismatch $[0]$ for vector \vec{u} , and the (iv) supercell vector mismatch $[1]$ for vector \vec{v} , according to eqn (1)–(4). Finally, to temper the (m, n) supercell generation to sensible outcomes compared to the maximum area overlap, we set a maximum value of tolerance for variables m and n , according to eqn (5). Note that the maximum area overlap and the m_n tolerance are defined to strike a balance between achieving close lattice registry and avoiding unphysically large super-cells. Increasing the size of the super-cell makes geometric matching easier, but also less physically meaningful, since epitaxial interactions become ineffective at very large interface areas. To prevent this, the maximum area overlap feature sets an upper bound on the permissible supercell size, while the m_n tolerance parameter restricts the integer scaling factors (m, n) used to construct commensurate supercells. Together, these criteria ensure that the generated interface models are both computationally tractable and physically representative of realistic epitaxial relationships, rather than artefacts of excessive cell scaling.

$$\text{Maximum area overlap} = |\vec{u}_1 \cdot \vec{v}_1| \approx |\vec{u}_2 \cdot \vec{v}_2| \quad (1)$$

$$\text{Angle mismatch} = \angle(\vec{u}_1, \vec{u}_2) - \angle(\vec{v}_1, \vec{v}_2) \quad (2)$$

$$\text{Supercell vector mismatch } [0] = \left| \frac{\|\vec{u}_1\|}{\|\vec{u}_2\|} - 1 \right| \quad (3)$$



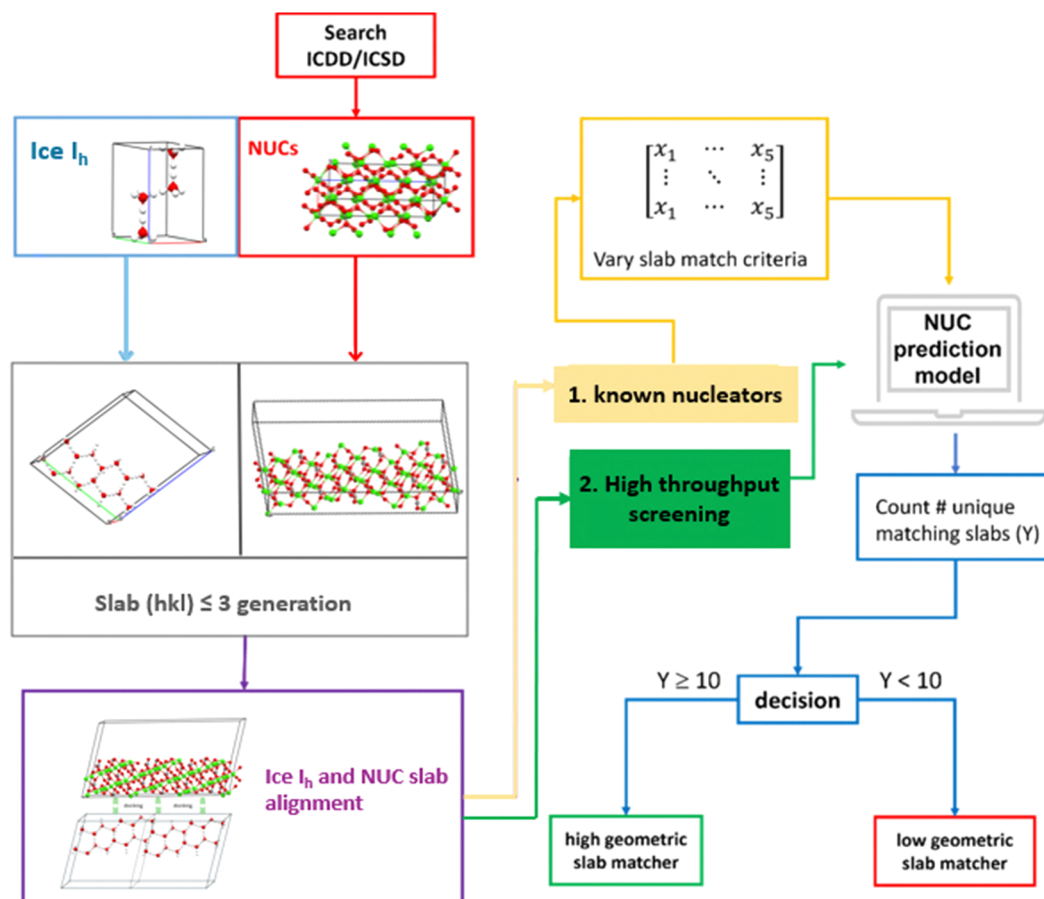


Fig. 2 Workflow for the data-driven geometric-matching model used to predict heterogeneous ice nucleation agents. Input CIFs - ice I_h (blue box) and nucleators ('NUCs', red box). Each crystal structure is then cleaved to generate slabs for all Miller indices with $hk(l) \leq 3$ (grey box). All ice I_h /NUC slab pairings are then docked and supercells explored (purple box). The threshold values for the five features that assess the quality of geometric matching are determined via pipeline 1 (yellow) for the ten literature-benchmarked nucleators, while pipeline 2 (green) then applies these feature thresholds for the high-throughput screening of CIFs downloaded from the ICSD. The output pipeline (blue) counts the number of unique matching slabs that fulfill the defined numerical criteria and classifies nucleator candidates according to their high or low geometric slab matching ability with ice I_h .

$$\text{Supercell vector mismatch [1]} = \left| \frac{\|\vec{v}_1\|}{\|\vec{v}_2\|} - 1 \right| \quad (4)$$

$$m_n\text{-tolerance} = \frac{|\vec{u}_1 \cdot \vec{v}_1| - |\vec{u}_2 \cdot \vec{v}_2|}{\text{maximum area overlap}} \quad (5)$$

A sensible value for the maximum-area overlap feature (eqn (1)) can be deduced based on the maximum surface area of possible surfaces generated from the bulk ice I_h lattice. This is the largest possible area of vector overlap achieved during slab docking. For this, the upper limit value of \vec{a} (4.5193 \AA) \times \vec{b} (7.3595 \AA) \times $10 \approx 330 \text{ \AA}^2$ (where the multiplier by 10 ensures ample tolerance of surface size differences) was chosen for eqn (1), and was held fixed while the criteria for eqn (2)–(5), which by definition are assumed to adopt values close to zero, were allowed to vary. Tightening the parameters refers to applying stricter numerical thresholds for lattice vector length mismatch and angular deviation. The 'loose' and 'medium' thresholds were chosen empirically based on literature tolerance ranges observed in epitaxial lattice matching.⁵¹

Starting with pre-defined loose values of 0.2° , 0.2 , 0.2 and 0.1 for eqn (2)–(5) respectively, the numbers of ice I_h /nucleator surface pairings from the known nucleator dataset that conform to these numerical limits are given in Fig. 5(a). Applying medium-level values of 0.1° , 0.01 , 0.01 and 0.1 generates the plot given in Fig. 5(b), while tightening the $m_n\text{-tolerance}$ parameter further to 0.01 gives the plot shown in Fig. 5(c). At this point the number of matching interfaces for the effective nucleators remains over ten, while those for the poor nucleators fall below ten (see SI Section S5 for the Miller indices of the matching interface models). Tightening the parameters any further results in a substantial and comprehensive loss in matching interface models; thus these fit criteria represent a probability boundary to differentiate between predicted good and poor ice nucleation behavior.

4. High throughput nucleator screening

Our algorithm was then applied in a high throughput study to identify potential new ice nucleators from crystal structures



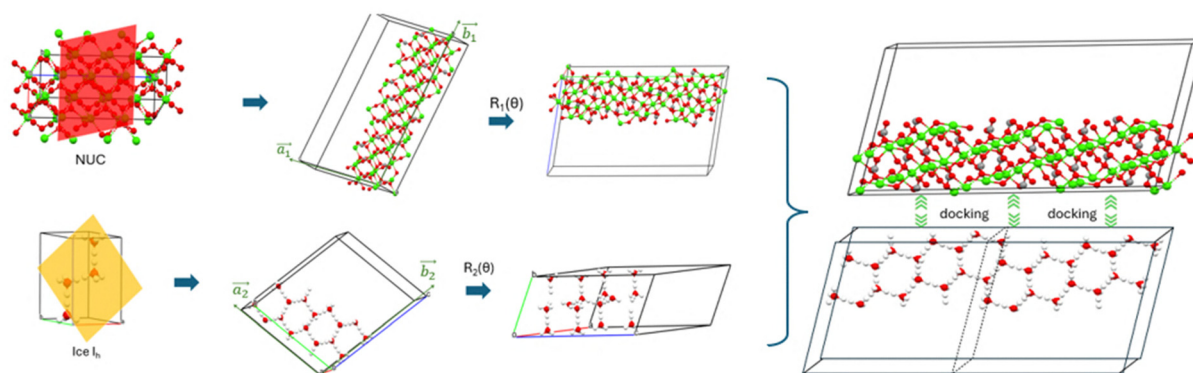


Fig. 3 Schematic representation of slab generation and alignment between ice I_h and nucleator (NUC) surfaces. For each bulk crystal, a chosen (hkl) plane is cleaved to create a surface with in-plane lattice vectors \vec{a} and \vec{b} , and grown to create integer (m , n) supercells. Slabs are docked by (i) translation to maximise in-plane overlap, and (ii) rotation by $R_1(\theta)$ and $R_2(\theta)$ to complete slab alignment.

deposited in the ICSD. Our search focused on simple binary metal halides (1257 compounds) and oxides (2267 compounds) to maximize the probability that compounds with comparable unit-cell lattice sizes to ice I_h were explored, and to generate candidates for experimental testing that were more likely to be easily sourced. Of the *ca.* 3500 compounds studied, just *ca.* 7% of the metal halides and *ca.* 3% of the metal oxides were classified as potential ice nucleators, defined as presenting ten or more matching interface models (see SI, Section S8). We then selected 22 compounds that displayed a range of slab-matching behavior (seventeen predicted as effective nucleators, five as poor nucleators, see Fig. 5(d)) that were also known to have very low aqueous solubilities and which could be readily obtained. Experimental testing results are given in Fig. 1(b). Fourteen out of the 22 compounds were correctly categorised as being a good or poor ice nucleator according to our criteria, while five were wrongly assigned (CoO, CaCO_3 (aragonite), TiO_2 (anatase), Mn_2O_3 , and Fe_2O_3), and three were ambiguous (MgO , Fe_3O_4 and Co_3O_4). This is a 64% successful prediction rate.

Of the correct predictions, some have previously been reported as ice nucleators in cloud-chamber experiments. We highlight the very early work by Fukuta,⁵² who investigated the behavior of many metal salts under vapor-deposition conditions, including MnO_2 , MgO , CoO , Ag_2O , NiO , ZnO , and PbBr_2 . NiO was trialed for artificial snow production as far back as 1956,⁵³ while more recently lead oxide was highlighted as an anthropogenic climate modifier.⁵⁴ Early reports by Vonnegut⁵⁵ suggested that solid solutions of CuI with AgI improve the nucleating ability of the latter under immersion conditions, which was attributed to improvements in the lattice mismatch with the crystal structure of ice I_h . MgO , TiO_2 , AgI , Al_2O_3 , and SiO_2 have attracted attention from materials scientists and computational modelers, who have studied the ice-nucleating abilities of individual faces. For instance, an experimental study on pristine MgO (100) and TiO_2 (100) (rutile) suggested the absence of a templating effect,⁵⁶ which matches our observation that these surfaces do not interface match with ice I_h (see SI, Section S5). The same paper reports that the (110) face of TiO_2 (rutile) supports the growth of cubic ice.⁴¹ While we

specifically matched against ice I_h , our modelling suggests that the (110) TiO_2 (rutile) surface only matches against one cleaved surface from ice I_h ; instead, the (001), (010), and (011) feature more heavily, geometrically matching with the basal and primary ice I_h faces (see SI, Section S5). Molecular dynamics simulations performed on the (0001) and (10 $\bar{1}$ 0) faces of the β -polymorph of AgI (as studied here) concluded that ice I_h nucleates on the former, but not the latter.⁵⁷ This also matches the outcome of our study, with the (0001) face pairing with both the basal and primary ice I_h faces, whereas none of the ice surface models interface with the (10 $\bar{1}$ 0) face of AgI . For Al_2O_3 , the γ -polymorph (as studied here, although with fairly low crystallinity, see SI, Section S1) is known to be the oxide that forms on aluminum surfaces when exposed to the atmosphere.⁵⁸ Immersion studies have previously shown that α - Al_2O_3 is a more effective ice nucleator, and that any effect by γ - Al_2O_3 is weak.⁵⁹ This is borne out in our work, where we classed γ - Al_2O_3 as a weak nucleator (Fig. 1b), which was substantiated by the low number of matching interfaces it presents with ice I_h (Fig. 5(d)); re-running the slab-matching process with the α -polymorph resulted in considerably more matching interfaces, which may be indicative of a higher ice nucleating ability (see SI, Section S5). SiO_2 (10 $\bar{1}$ 0) has been observed to template ice I_h under immersion freezing conditions.²⁷ While we initially paired this face with multiple ice I_h faces under our loose geometric criteria (Fig. 5(a)), upon tightening the criteria other Miller planes of SiO_2 (notably (0001)) were found to match more closely with those of ice I_h (see SI, Section S5). We note that thin films of the nanocomposite SnO_2 (cassiterite)/ TiO_2 (anatase), spin coated with a Krytox grease lubricant have been shown to display anti-icing properties,⁶⁰ suggesting that ice templates poorly on this substrate. While geometric slab matching for SnO_2 and TiO_2 (anatase) returns high numbers of matching interfaces (Fig. 5(d) and SI, Section S5), suggesting that both should template for ice I_h , the latter was one of our five wrong assignments, as our bulk water immersion experiments showed that TiO_2 (anatase) does not nucleate ice (Fig. 1b). While there are likely to be many reasons why the $\text{SnO}_2/\text{TiO}_2$ nanocomposite inhibits ice growth, it would be of interest to explore if this could be attributed to



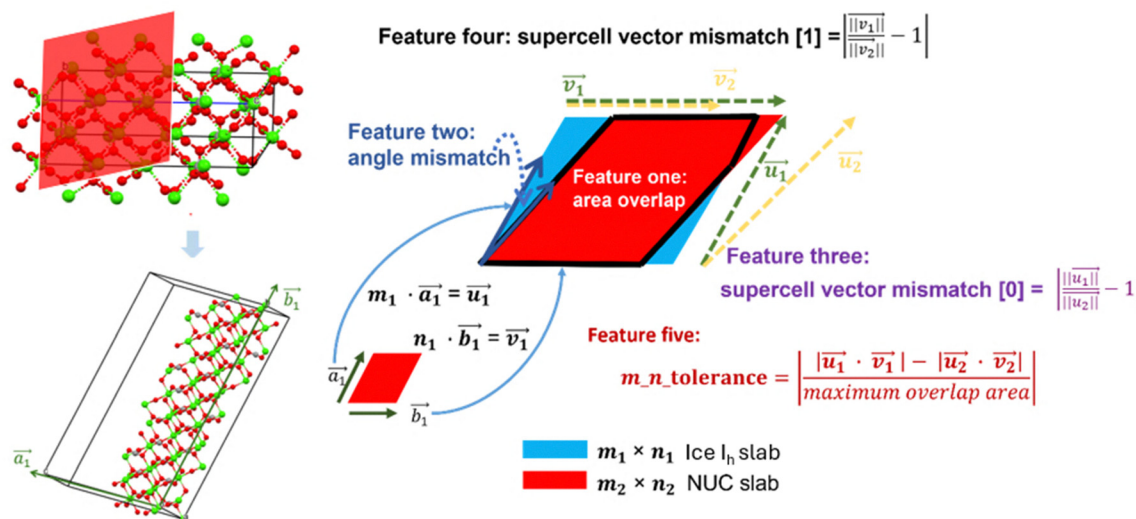


Fig. 4 Summary of the five geometric descriptors used to quantify interface compatibility between ice I_h and nucleator surfaces: (1) maximum area overlap; (2) angular mismatch between slab vectors; (3–4) length mismatch of supercell vectors \vec{a} and \vec{b} ; and (5) supercell (m , n) generation tolerance threshold to avoid exaggerated surface scaling.

TiO₂ (anatase) dominating the suppression of ice formation. To the best of our knowledge no prior reports have attributed ice nucleation properties to CeO₂, WO₃, Bi₂O₃ or Ti₂O₃, suggesting these could be new ice I_h nucleators under immersion conditions.

While these findings are generally encouraging for a high throughput screening approach for identification of heterogeneous nucleating agents based purely on interface matching, it is important to note that predictions will miss any potential nucleators that do not fulfil the matching criterion, as indeed illustrated by Co₃O₄, Mn₂O₃ and Fe₂O₃. It also does not consider any surface chemistry effects (such as surface polarity), allow for any surface reactions or reconstructions, variation in surface termination, slab stretching/compression, or rank the relative stabilities of the surface models. Our data set highlights a number of false negatives (CoO, CaCO₃ (aragonite) and TiO₂ (anatase)). MgO displays ambiguous behaviour, despite presenting with one of the highest number of matching interface models with ice I_h . One possible explanation is that the solid is undergoing a surface reaction in water to form Mg(OH)₂ (brucite) or a hydrate. Repeating our slab matching approach with Mg(OH)₂ (ICSD code 34401), and Mg(OH)₂·2H₂O (ICSD code 118781) return a total of 45 and 4 slab matching interfaces, respectively, with ice I_h – down considerably from the 81 predicted interfaces with MgO. Without an in-depth experimental validation that explores ice nucleation onto defined nucleator surfaces, it remains unknown whether nucleation actually proceeds *via* these geometrically-matching interfaces. Nevertheless we note that, reassuringly, the Miller index of the basal face of ice I_h features heavily on our paired slabs list, as do the primary and secondary prismatic faces (see SI, Section S5). Finally, the data presented in Fig. 1 and 5 illustrates an important point that a higher number of matching interface models does not correlate with a greater extent of

suppression of supercooling; rather this demonstrates that the given crystal nucleator morphology (the edges, corners and potential defect sites captured by the Miller index planes up to $hkl = (333)$) are more likely to geometrically match with a corresponding Miller index plane for ice I_h .

In an attempt to show whether interface matching offers new information beyond the zero-lattice mismatch approach,³⁰ we have also calculated the mismatch registry parameter for the unit-cell parameters and for each of the matching interface models (see SI, Section S5). The analysis shows that, based on similarities of unit-cell dimensions between the nucleator and ice I_h basal face, just seven of the 32 compounds we explored experimentally would be correctly predicted as an effective or poor nucleator. For the matched interface models, while some of the effective nucleators do return low lattice mismatch values, the majority of the interface pairings do not. Overall, this suggests that interface matching offers a broader search criterion for potential nucleation behavior than the lattice mismatch approach.

It is important to further acknowledge that our geometric slab-matching approach, while efficient as a high-throughput screening tool, does not capture several features that are likely to play a significant role in heterogeneous ice nucleation. For example, the consideration of higher Miller index surfaces in our workflow inevitably includes surface terminations that are energetically costly and thus may be unstable under realistic conditions. In aqueous environments, such surfaces may undergo reconstructions or hydroxylation to alter the interfacial chemistry, thereby changing their propensity to nucleate ice. More sophisticated computational methods, such as density functional theory (DFT) calculations of surface energetics and adsorption geometries, or molecular dynamics (MD) simulations of hydration layer structure and dynamics, would be required to capture these effects in detail. Finally, our



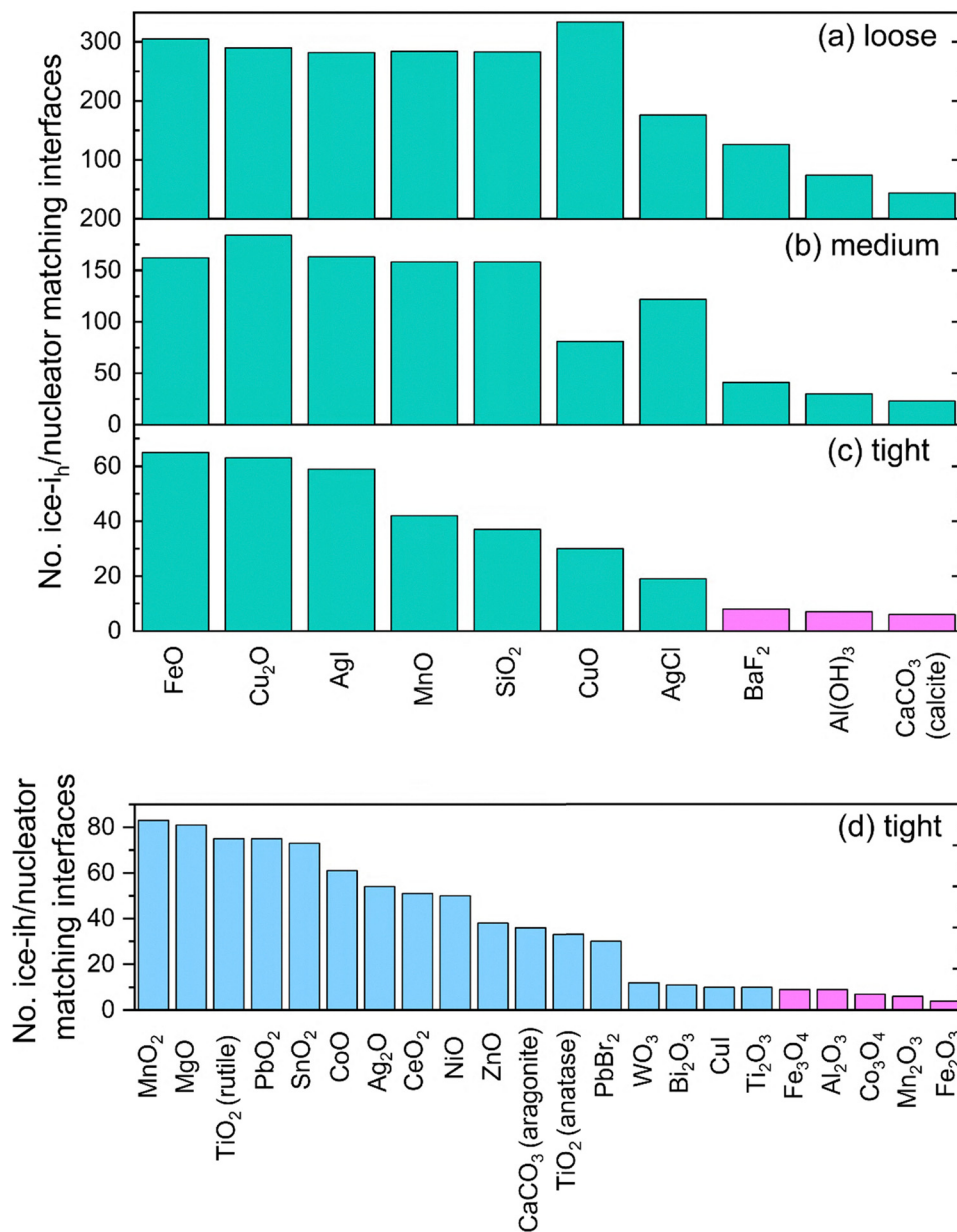


Fig. 5 Interface-matching results for ice I_h /nucleator pairs based on progressively stricter geometric thresholds: (a) loose, (b) medium, and (c) tight tolerance levels, along with results for (d) the 22 test compounds according to the tight tolerance criteria. Compounds shown in pink return less than 10 slab matching interfaces.

current model cannot assess the role of surface polarity, hydration dynamics, or specific adsorption energetics, all of which have been identified as key descriptors in recent simulation studies. Despite these limitations, our results show that geometric matching alone provides a surprisingly effective first filter for identifying candidate nucleators, which can then be prioritised for more detailed computational and experimental investigation.

Finally, given the high performance of copper oxides in our validation list, combined with the recent report that copper oxide nanoparticles act as ice nucleation sites,^{60,61} we decided to test 'off-the-shelf' copper tubing for ice nucleation. This

follows given that a copper surface will readily oxidize in contact with air and water, and thus could promote the nucleation of ice in bulk water. This is particularly relevant given that copper is widely used in plumbing, heating ventilation and cooling applications, and electrical transmission power lines, for which the formation of ice contributes to burst water pipes and power outages.^{62,63} The resulting data show that the copper tubing did indeed induce ice nucleation at -2.3 ± 0.2 °C (compared to -10.3 ± 0.9 °C in its absence). Moreover, the nucleation temperature increased over multiple cycles, potentially correlating with increased oxidation of the metal surface (see SI, Section S7).



5. Experimental methods

For the immersion experiments, samples were prepared in screw-top, borosilicate glass vials, which were previously washed in hot soapy water, and repeatedly rinsed with ultra-pure water passed through a Millex (33 mm) sterile filter unit with Millipore Express PES membrane. 1 wt% of lightly ground nucleator was placed at the bottom of the vial, onto which 10 ml of ultra-pure, filtered water was added. On top of this, a 1 ml layer of silicone oil was carefully applied, into which an RS PRO Type K thermocouple (exposed junction 0.2 mm diameter) encased in a stainless-steel sheath, that had also previously washed in ultra-pure filtered water was carefully positioned (see SI, Section S2). This was to limit the potential for contamination, either from the air or from the thermocouple itself. Samples were cycled from +20 to −20 °C in a Polar Bear Plus instrument, using a cooling/heating rate of 0.5 °C min^{−1} for four cycles. The efficacy of nucleation was assessed through the degree of sub-cooling suppression, with values closer to 0 °C indicating a better performing nucleator. In the absence of any nucleating agent, the ultra-pure filtered water would reliably sub-cool to -12 ± 3 °C (see SI, Section S3). All samples were measured in triplicate. It is important to note that this captures the freezing onset temperature of the bulk solution, and as such does not represent an average behaviour across all nucleation sites, but rather the performance of the most active under the given conditions. This introduces a statistical limitation: subtle variations in particle dispersion or surface exposure may influence which site initiates freezing first. Despite this, the method offers a reproducible comparative basis across materials when sample loading and conditions are standardised.

For the copper tubing experiments, 1 cm samples of 10-mm diameter copper pipe (BS EN 1057 standard) were added to 20 ml samples of tap water (East Lothian, CaCO₃ concentration = 74.23 ppm⁶⁴) in sealed sample vials and thermally cycled from −20 to +20 °C in the Polar Bear Plus Crystal apparatus using a heating/cooling rate of 1 °C min^{−1}, for 25 cycles. Blank samples in the absence of copper tubing were run alongside, and measurements were repeated in triplicate.

6. Conclusions

Herein, we present a high throughput data-driven workflow for identifying potential heterogeneous nucleating agents, like ice, from structural chemistry databases, such as the ICSD. Our model is based on assessing the quality of fit between ice *I*_h and nucleator docked slabs, formed from cleaving along Miller index planes from the respective bulk crystal lattices. While this has similarities to the zero-lattice mismatch approach, it goes beyond the low-index planes to consider the docking of all interfaces described by the Miller indices up to (333). In this way, we are addressing some of the structural complexity of the nucleation process by considering crystal morphology, where ice crystallites could seed on the faces, edges, corners, defects

or other surface features of the nucleating crystal that could be described by these higher Miller-index planes.

Numerical tolerance limits for the docking model were derived from a training set of ten compounds. Our bulk water freezing experiments were sufficiently reliable to classify each compound correctly as an effective or poor ice nucleator, based on expectations from the literature. Tightening the geometric matching criteria resulted in a fall in the number of matching interface models until differentiation between the two classes was obtained.

The optimized model screened approximately 3500 simple metal oxides and halides from the ICSD for predicted nucleation behaviour. Subsequent experimental measurements of 22 compounds showed a 64% prediction success rate, as defined by the freezing temperature boundary obtained from the experimental training set data. Our workflow also identified four previously unreported ice nucleating agents (CeO₂, WO₃, Bi₂O₃, Ti₂O₃).

Given the high ice nucleating ability demonstrated for copper oxides, we were also inspired to test the nucleating ability of standard copper tubing immersed in samples of local tap water. This was also found to suppress sub-cooling, likely due to the build-up of copper oxides over the timescale of the experiment.

While the approach we have taken here is undoubtedly simplistic, and does not account for many important aspects, such as surface chemistry effects, reactions and reconstructions, it nevertheless demonstrates an acceptable level of success to form the basis for a high throughput computational screening approach to locate potential heterogeneous nucleating agents for further investigation.

Author contributions

ZW, CAM, AJB, DEO and CRP provided the intellectual framework for the slab-matching algorithm. ZM created the algorithm. Experiments were performed by ZM, CAM and GO. Data analysis was performed by ZW and CAM. ZW and CAM wrote the manuscript, with inputs from all authors.

Conflicts of interest

There are no conflicts to declare.

Data availability

Data sets generated during the current study are given in the supplementary information (SI). See DOI: <https://doi.org/10.1039/d5cp02701b>.

Further information is available from the corresponding author on reasonable request.

Acknowledgements

ZW is grateful to Sunamp Ltd for the award of a PhD scholarship. We thank Prof J. P. Attfield and Prof J. B. Love for access to chemicals in their laboratories.



References

- 1 S. Kiyabu, P. Girard and D. J. Siegel, *J. Am. Chem. Soc.*, 2022, **144**, 21617–21627.
- 2 S. Hasnain, *Energy Convers. Manage.*, 1998, **39**, 1139–1153.
- 3 O. Boucher, D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Förster, V.-M. Kerminen, Y. Kondo, H. Liao and U. Lohmann, *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*, Cambridge University Press, 2013, pp. 571–657.
- 4 W. K. Tao, J. P. Chen, Z. Li, C. Wang and C. Zhang, *Rev. Geophys.*, 2012, **50**.
- 5 B. Murray, D. O'Sullivan, J. Atkinson and M. Webb, *Chem. Soc. Rev.*, 2012, **41**, 6519–6554.
- 6 G. Vali, P. J. DeMott, O. Möhler and T. F. Whale, *Atmos. Chem. Phys.*, 2015, **15**, 10263–10270.
- 7 Z. A. Kanji, L. A. Ladino, H. Wex, Y. Boose, M. Burkert-Kohn, D. J. Cziczo and M. Krämer, *Meteorol. Monogr.*, 2017, **58**, 1.1–1.33.
- 8 B. Vonnegut, *J. Appl. Phys.*, 1947, **18**, 593–595.
- 9 B. Vonnegut and H. Chessin, *Science*, 1971, **174**, 945–946.
- 10 C. Marcolli, B. Nagare, A. Welti and U. Lohmann, *Atmos. Chem. Phys.*, 2016, **16**, 8915–8937.
- 11 A. Kiselev, F. Bachmann, P. Pedevilla, S. J. Cox, A. Michaelides, D. Gerthsen and T. Leisner, *Science*, 2017, **355**, 367–371.
- 12 B. Friedman, G. Kulkarni, J. Beránek, A. Zelenyuk, J. A. Thornton and D. J. Cziczo, *J. Geophys. Res.: Atmos.*, 2011, **116**.
- 13 T. W. Wilson, L. A. Ladino, P. A. Alpert, M. N. Breckels, I. M. Brooks, J. Browse, S. M. Burrows, K. S. Carslaw, J. A. Huffman and C. Judd, *Nature*, 2015, **525**, 234–238.
- 14 M. Fitzner, G. C. Sosso, S. J. Cox and A. Michaelides, *J. Am. Chem. Soc.*, 2015, **137**, 13658–13669.
- 15 C. Hoose and O. Möhler, *Atmos. Chem. Phys.*, 2012, **12**, 9817–9854.
- 16 P. Pedevilla, M. Fitzner and A. Michaelides, *Phys. Rev. B*, 2017, **96**, 115441.
- 17 J. Carrasco, A. Hodgson and A. Michaelides, *Nat. Mater.*, 2012, **11**, 667–674.
- 18 T. Koop, B. Luo, A. Tsias and T. Peter, *Nature*, 2000, **406**, 611–614.
- 19 T. F. Whale, M. A. Holden, A. N. Kulak, Y.-Y. Kim, F. C. Meldrum, H. K. Christenson and B. J. Murray, *Phys. Chem. Chem. Phys.*, 2017, **19**, 31186–31193.
- 20 G. C. Sosso, T. F. Whale, M. A. Holden, P. Pedevilla, B. J. Murray and A. Michaelides, *Chem. Sci.*, 2018, **9**, 8077–8088.
- 21 H. R. Pruppacher, J. D. Klett and P. K. Wang, *Microphysics of clouds and precipitation*, Taylor & Francis, 1998.
- 22 M. Fitzner, P. Pedevilla and A. Michaelides, *Nat. Commun.*, 2020, **11**, 4777.
- 23 J. Chen, Z. Wu, J. Chen, N. Reicher, X. Fang, Y. Rudich and M. Hu, *Atmos. Chem. Phys.*, 2021, **21**, 3491–3506.
- 24 M. B. Davies, M. Fitzner and A. Michaelides, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2205347119.
- 25 G. Vali, *J. Atmos. Sci.*, 1971, **28**, 402–409.
- 26 G. Vali, *Atmos. Chem. Phys.*, 2008, **8**, 5017–5031.
- 27 M. A. Holden, J. M. Campbell, F. C. Meldrum, B. J. Murray and H. K. Christenson, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2022859118.
- 28 C. Marcolli, *Atmos. Chem. Phys.*, 2014, **14**, 2071–2104.
- 29 B. Mason, *Adv. Phys.*, 1958, **7**, 221–234.
- 30 D. Turnbull and B. Vonnegut, *Ind. Eng. Chem.*, 1952, **44**, 1292–1298.
- 31 G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen and A. Michaelides, *Chem. Rev.*, 2016, **116**, 7078–7116.
- 32 M. Fitzner, G. C. Sosso, S. J. Cox and A. Michaelides, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 2009–2014.
- 33 T. Li, D. Donadio and G. Galli, *Nat. Commun.*, 2013, **4**, 1887.
- 34 S. N. Kabekkodu, A. Dosen and T. N. Blanton, *Powder Diffr.*, 2024, **1**–13.
- 35 D. Zagorac, H. Müller, S. Ruehl, J. Zagorac and S. Rehme, *J. Appl. Crystallogr.*, 2019, **52**, 918–925.
- 36 E. Chong, K. E. Marak, Y. Li and M. A. Freedman, *Phys. Chem. Chem. Phys.*, 2021, **23**, 3565–3573.
- 37 W. G. Finnegan and S. K. Chai, *J. Atmos. Sci.*, 2003, **60**, 1723–1731.
- 38 C. X. Kronawitter, C. Riplinger, X. He, P. Zahl, E. A. Carter, P. Sutter and B. E. Koel, *J. Am. Chem. Soc.*, 2014, **136**, 13283–13288.
- 39 X. Zhang, X. Meng, Q. Zhao and C. Li, *CIESC J.*, 2014, **65**, 4321.
- 40 P. Conrad, G. E. Ewing, R. L. Karlinsey and V. Sadtschenko, *J. Chem. Phys.*, 2005, **122**.
- 41 L. Kaufmann, C. Marcolli, J. Hofer, V. Pinti, C. R. Hoyle and T. Peter, *Atmos. Chem. Phys.*, 2016, **16**, 11177–11206.
- 42 A. Kumar, C. Marcolli and T. Peter, *Atmos. Chem. Phys.*, 2019, **19**, 6059–6084.
- 43 A. J. Miller, C. Fuchs, F. Ramelli, H. Zhang, N. Omanovic, R. Spirig, C. Marcolli, Z. A. Kanji, U. Lohmann and J. Henneberger, *Atmos. Chem. Phys.*, 2025, **25**, 5387–5407.
- 44 A. Soni and G. Patey, *J. Phys. Chem. C*, 2022, **126**, 6716–6723.
- 45 Y. Yu, M. Chen, Y. Lei and H. Niu, *Nat. Commun. Phys.*, 2025, **8**, 7.
- 46 N. Shardt, F. N. Isenrich, B. Waser, C. Marcolli, Z. A. Kanji, A. J. deMello and U. Lohmann, *Phys. Chem. Chem. Phys.*, 2022, **24**, 28213–28221.
- 47 B. Nagare, C. Marcolli, A. Welti, O. Stetzer and U. Lohmann, *Atmos. Chem. Phys.*, 2016, **16**, 8899–8914.
- 48 J. C. Lee, T. Hansen and P. L. Davies, *Cryobiology*, 2023, **113**, 104584.
- 49 A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer and C. Hargus, *J. Phys.: Condens. Matter*, 2017, **29**, 273002.
- 50 S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson and G. Ceder, *Comput. Mater. Sci.*, 2013, **68**, 314–319.
- 51 A. Zur and T. McGill, *J. Appl. Phys.*, 1984, **55**, 378–386.
- 52 N. Fukuta, *J. Atmos. Sci.*, 1958, **15**, 17–26.



- 53 I. Sano, Y. Fujitani, K. Ito and S. Kitani, *J. Meteorol. Soc. Jpn.*, 1956, **34**, 185–189.
- 54 D. J. Cziczo, O. Stetzer, A. Worringer, M. Ebert, S. Weinbruch, M. Kamphus, S. J. Gallavardin, J. Curtius, S. Borrmann and K. D. Froyd, *Nat. Geosci.*, 2009, **2**, 333–336.
- 55 R. E. Passarelli Jr, H. Chessin and B. Vonnegut, *Science*, 1973, **181**, 549–551.
- 56 R. Souda, T. Aizawa, N. Sugiyama and M. Takeguchi, *Phys. Chem. Chem. Phys.*, 2020, **22**, 20515–20523.
- 57 S. A. Zielke, A. K. Bertram and G. N. Patey, *J. Phys. Chem. B*, 2015, **119**, 9049–9055.
- 58 A. Raveh, Z. Tsameret and E. Grossman, *Surf. Coat. Technol.*, 1997, **88**, 103–111.
- 59 E. Chong, M. King, K. E. Marak and M. A. Freedman, *J. Phys. Chem. A*, 2019, **123**, 2447–2456.
- 60 F. L. Heale, I. P. Parkin and C. J. Carmalt, *ACS Appl. Mater. Interfaces*, 2019, **11**, 41804–41812.
- 61 Z. R. Schiffman, M. S. Fernanders, R. D. Davis and M. A. Tolbert, *ACS Earth Space Chem.*, 2023, **7**, 812–822.
- 62 J. C. Pohlman and P. Landers, *IEEE Trans. Power Appar. Syst.*, 1982, 2443–2450.
- 63 M. Farzaneh, C. Volat and A. Leblond, *Atmospheric icing of power networks*, Springer, 2008, pp. 229–268.
- 64 *Scottish Water, Water Hardness Data 2024*, <https://www.scottishwater.co.uk/help-and-resources/document-hub/key-publications/water-quality>, (accessed 15 July, 2025).

