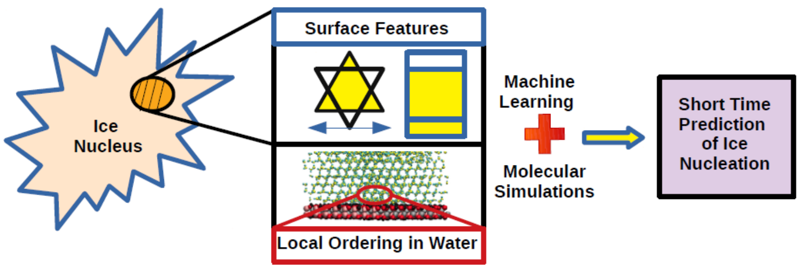
Predicting Heterogeneous Ice Nucleation via Short-Time Molecular Dynamics Simulation and Machine Learning Approaches

Abhishek Soni and G.N. Patey*

Department of Chemistry, The University of British Columbia, Vancouver, Canada, V6T 1Z1



Heterogeneous ice nucleation (HIN) has various applications in the fields of atmospheric science, food preservation, and nanotechnology. Pure water can be supercooled to ~−38 °C, and homogeneous ice nucleation at temperatures warmer than ~−20 °C has essentially zero probability, even on the time scale of the universe. Thus, much of the freezing that occurs on the Earth's surface or in the atmosphere occurs via a heterogeneous mechanism involving an ice nucleating particle (INP). These INPs can be mineral dust, soots, pollen, or bacteria. Generally, ice nucleation experiments identify substrates that act as efficient ice nuclei but lack sufficient spatial (nm) and/or temporal (ns) resolution to address basic mechanistic questions. Recently, molecular dynamics (MD) simulations of model systems have attempted to reveal the basic mechanism of ice nucleation and the fundamental molecular features of various good INPs. However, the large amount of computational cost required to cross the nucleation barrier and observe HIN in simulations is a current concern. Here, we employ information obtained from short MD simulations of water in contact with surfaces to predict the likelihood that surface would nucleate ice, or not, in sufficiently long simulations, or possibly in experiments. For prediction, we incorporated several supervised and unsupervised machine learning (ML) models. We considered various atomistic substrates with some surfaces differing from others, only in terms of lattice parameters, surface morphology, or surface charges. Various water features near the surface are extracted from MD simulations over a time interval where ice nucleation has not initiated. We find that the interplay of surface properties and local liquid water properties determines good/bad INPs, with the liquid water properties being dominant. The accuracy of our best ML classification model is 0.89 ± 0.05. Some of the important descriptors are interfacial icelike structures, hydrogen bonding to the surface, water density and water polarization near the surface, and the two-dimensional lattice match to ice. Taken altogether, we expect that this work will be a useful contribution in the field of HIN research and serve as a guide in the design of substrates that can promote or discourage ice nucleation.