Understanding crystallization with atomistic machine learning and simulation

Pablo M. Piaggi[1,2]
[1]CIC nanoGUNE BRTA, Tolosa Hiribidea 76, Donostia-San Sebastián 20018, Spain
[2]Ikerbasque, Basque Foundation for Science, Bilbao 48013, Spain

pm.piaggi@nanogune.eu

Crystallization is a process of key importance for many modern technologies, such as the manufacturing of pharmaceuticals and 2D materials, and it also plays a central role in geological, planetary, and climate sciences. Over the years, molecular dynamics simulations have provided key insights into the microscopic mechanisms underlying crystal nucleation and growth. The recent advent of atomistic machine learning (ML), a class of methods based on learning atomistic properties from large datasets of *ab initio* electronic-structure calculations, is dramatically improving the accuracy and predictive power of these simulations. In particular, learning the potential energy surface has provided access to much larger system sizes and longer simulation times, greatly expanding the scope of molecular simulations and providing access to reactive crystallization processes.

In this talk, I will discuss the use of this technique to drive large-scale, highly-accurate and reactive molecular dynamics simulations of crystallization. I will illustrate this approach with several examples. First, I will show that we can leverage this tool to compute homogeneous ice nucleation rates from first principles which are in remarkable agreement with experiment [1]. Furthermore, I will present some results about the formation of ice on feldspar, the most important ice nucleating particle in the atmosphere [2]. Finally, I will discuss the application of this technique to study the crystallization of calcium carbonate from aqueous solution [3], a process where reactivity plays an essential yet poorly understood role (see Figure 1 for an illustration of proton transfer during ion pairing). I will also highlight the role of long-range interactions captured via machine-learned Wannier centroids, which act as surrogates for the center of the electronic charge distributions.

Taken together, these results show the great promise of atomistic machine learning as a tool to bridge time and length scales, and to provide insight into complex crystallization phenomena, which were thought to be out of reach for molecular simulation.

References

[1] Piaggi, Weis, Panagiotopoulos, Debenedetti, and Car, *Proc. Natl. Acad. Sci.* 119 (2022) 33.
[2] Piaggi, Selloni, Panagiotopoulos, Car, and Debenedetti, Faraday Discuss. 249 (2024) 98.
[3] Piaggi, Gale, and Raiteri, *Proc. Natl. Acad. Sci.* (2025)

Figures

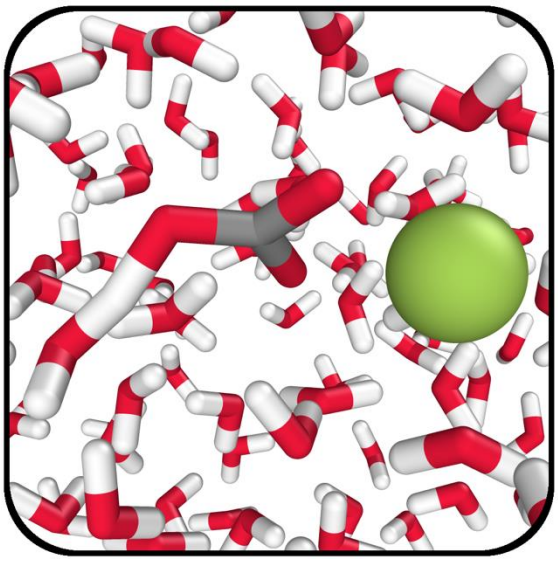

Figure 1. Proton transfer during Ca and carbonate ion pairing in water, captured with ab initio machine learning simulations [3]. Oxygen, hydrogen, carbon, and calcium atoms are shown in red, white, gray, and green, respectively.