# SEEING THE INVISIBLE: LEARNING PATHWAYS TO POLYMORPHS THROUGH MACHINE LEARNING ANALYSIS OF ATOMIC TRAJECTORIES

SAPNA SARUPRIA

Crystallization is consequential to many applications, including pharmaceutical production, flow assurance, and climate modeling. Effective control over crystallization relies on understanding the possible structures that can form during nucleation and growth. Molecular simulations allow a more fine-grained approach to discovering important, though possibly short-lived, intermediate structures, but their characterization from atomic coordinates is often difficult. We combine general features of the local atomic arrangements with a deep learning model to discover the unique structures that form during crystal nucleation. While many previous mechanistic studies have relied on features that describe the entire crystal nucleus, such as its size, shape, and composition, we focus on the evolution of the atoms involved in the formation of the nucleus in the feature space to describe nucleation processes. We find that this approach reveals structures not found by traditional approaches, and also enables us to classify nucleation pathways. Hence, we are able to shed light on previously “invisible” structures. This has implications in understanding and controlling the long-standing problem of polymorph selection during crystallization.