Freedom of design: towards in silico design of molecules with desired quantum-mechanical properties

Abstract:

The rational in silico design of chemical compounds requires a deep understanding of both the structure–property and property–property relationships that exist across chemical compound space (CCS), as well as efficient methodologies for defining an inverse property-to-structure mapping. This presentation will discuss these relationships in the CCS sector spanned by small [Sci. Data 8, 43 (2021)] and large [Sci. Data 11, 742 (2024)] drug-like molecules, highlighting the existence of the "freedom of design" principle [Chem. Sci. 14, 10702 (2023)]. The insights gained are subsequently leveraged to design molecules with desired properties. To this end, we first developed a variational autoencoder (VAE) approach and demonstrated that CCS can be parameterized using a finite set of quantum-mechanical (QM) properties [Nat. Commun. 15, 6061 (2024)]. We showcased the capabilities of this method by conditionally generating de novo molecular structures, interpolating transition paths for chemical reactions, and providing insightful insights into property–structure relationships. Then, a diffusion generative model on both small and large drug-like molecules was trained to examine the limits of scalability and chemical diversity when targeting diverse sets of QM properties. We expect our work will contribute to the development of advanced generative frameworks that enhance the in silico design and identification of molecules for specific chemical processes.