

Predicting Chemical Shifts with Graph Neural Networks

Ziyue Yang¹, Maghesree Chakraborty¹, Andrew White¹

¹*University of Rochester*

zyang43@ur.rochester.edu

Nuclear Magnetic Resonance (NMR) chemical shifts of a molecule can provide detailed structural information at room temperature and reasonable concentrations, in a physiologically relevant ensemble of conformations and even *in situ*. There has been continued interest in methods to resolve protein structure from NMR due to the accessibility. Inferring molecular structure from NMR measurements requires an accurate forward model that can predict chemical shifts from 3D structure. However, current state-of-the-art forward models are not differentiable. Thus they cannot be used with gradient methods like biased molecular dynamics. In this work we use graph neural networks (GNNs) for NMR chemical shift prediction. Our GNN can model chemical shifts accurately and capture important phenomena like hydrogen bonding induced downfield shift between multiple proteins, secondary structure effects, and predict shifts of organic molecules. Previous empirical NMR models of protein NMR have relied on careful feature engineering with domain expertise.

These GNNs are trained from data alone with no feature engineering yet are as accurate and can work on arbitrary molecular structures. The models are also efficient, able to compute one million chemical shifts in about 5 seconds. Our model can be further combined with other gradient based methods to infer protein structures due to the differentiability and also can serve as guidance for diffusion models to generate protein structures due to its high efficiency and accuracy.