

Wen Ma EXPAND Mentor



Applying machine learning to determine biomolecular dynamics from cryo-EM images

J. Andrew McCammon/McCammon Lab (2022-2023)

What you will do

One of the most remarkable inventions of biological evolution is molecular machines that participate in various vital processes, such as cellular force generation and genome maintenance. A molecular picture of how these nanomachines work will represent a milestone in advancing drug discovery for treating related diseases. A particularly exciting imaging technique is Cryo-electron microscopy (cryo-EM), which is now capable of achieving atomic resolution for large biomolecular structures and was awarded the Nobel Prize in Chemistry in 2017. In practice, a raw cryo-EM dataset contains hundreds of thousands of molecular images, which are highly heterogeneous due to the different orientations and shapes of a biomolecule of interest. Conventional cryo-EM data processing involves clustering these images into a small number of homogeneous classes. However, the signals of functionally relevant intermediate states are usually averaged out or not used for the final structure reconstruction, limiting our understanding of highly dynamic systems. In this project, we would like to apply machine learning techniques based on neural networks to extract the dynamics of molecular machines such as viral polymerase and human proteasome, which are important drug targets. I will guide you to study related literature and come up with related biological questions that will be addressed using the above approach. You will have the opportunity to participate in cutting-edge research and contribute to potential research papers.

Skills you will acquire

- Understanding of principles of molecular machines and concepts of drug discovery
- Fluency in a basic programming language, such as Python
- Experience in applying machine learning to solve a scientific problem
- Familiarity with cryo-EM imaging analysis, as well as common structural biology software
- Experience in academic writing

