A129 · Applications of AI to biomolecules for both answers and insight into enzyme function

Presenter: Mary Jo Ondrechen — Northeastern University

Authors: Mary Jo Ondrechen

Session: Day 2 · Fri Oct 2, 2026 · 2:15–4:15 PM

Artificial Intelligence (AI) has moved computational biology forward in recent years with dramatic breakthroughs, including good-performing solutions to the ab initio protein folding problem. We developed a machine learning (ML) method Partial Order Optimum Likelihood (POOL), first reported in 2009, to predict the amino acids in a protein structure that are biochemically active in catalysis or ligand binding. POOL has been used to identify catalytically active residues and to establish how distal residues contribute to activity in enzymes. POOL has been used to predict the function of protein structures of unknown function, using a local structure match, with subsequent experimental testing and verification. Recently POOL has enabled the uncovering of insights into how the amino acids in enzyme active sites achieve their catalytic power. Noting that the side chains of lysine, aspartic acid and glutamic acid are weak Brønsted acids and bases for the free amino acids in solution, we have shown how specific types of interactions with nearby amino acids in the local region of an enzyme active site can increase the acidity, basicity, or nucleophilicity of the catalytic residues and thus enable catalysis. Our ML approach has also successfully predicted whether specific missense mutations impair catalysis in an enzyme. A common criticism of AI in fields like chemistry and biology is that it gives you “answers but not insight.” Here it is shown how AI methods can be constructed to give both answers and insight.

Applications of AI to biomolecules for both answers and insight into enzyme function

Mary Jo Ondrechen*

Department of Chemistry and Chemical Biology and Department of Bioengineering,

Northeastern University, Boston, MA 02115 USA

Artificial Intelligence (AI) has moved computational biology forward in recent years with dramatic breakthroughs, including good-performing solutions to the *ab initio* protein folding problem. We developed a machine learning (ML) method Partial Order Optimum Likelihood (POOL), first reported in 2009, to predict the amino acids in a protein structure that are biochemically active in catalysis or ligand binding. POOL has been used to identify catalytically active residues and to establish how distal residues contribute to activity in enzymes. POOL has been used to predict the function of protein structures of unknown function, using a local structure match, with subsequent experimental testing and verification. Recently POOL has enabled the uncovering of insights into how the amino acids in enzyme active sites achieve their catalytic power. Noting that the side chains of lysine, aspartic acid and glutamic acid are weak Brønsted acids and bases for the free amino acids in solution, we have shown how specific types of interactions with nearby amino acids in the local region of an enzyme active site can increase the acidity, basicity, or nucleophilicity of the catalytic residues and thus enable catalysis. Our ML approach has also successfully predicted whether specific missense mutations impair catalysis in an enzyme. A common criticism of AI in fields like chemistry and biology is that it gives you "answers but not insight." Here it is shown how AI methods can be constructed to give both answers and insight.