A122 · Use of Computed Electrostatic and Geometric Information to Investigate Protein Functions and Functional Sites

Presenter: Tina Harati & Mary Jo Ondrechen — Northeastern University and The University of Illinois Chicago

Authors: Trang Tu, Tina Harati, Gayathri Ayyar Manjula, Geordie Emberling, Amman Hossain, Ruby Renfrow, Alquama Lokhandwala, Kyle Zhang, Ganesan Murugan, Nicole J. Curtis, Mary Jo Ondrechen, Constance

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

Partial Order Optimum Likelihood (POOL) is a machine learning method that combines computed electrostatic and geometric information for high-performance prediction of catalytic residues in enzyme structures. But many proteins are not enzymes. We are developing new methods using POOL to study noncatalytic protein functions. Two of these methods target RNA binding proteins and pseudoenzymes. RNA binding proteins play vital roles in RNA metabolism and function, including splicing, translation, localization, stability and degradation. In addition to canonical RNA binding proteins where RNA binding is an aspect of their main function, dozens of moonlighting proteins have been found that combine an enzymatic function in sugar, lipid, or amino acid metabolism with an RNA binding function. We developed RNABinderFinder, a novel machine learning method that combines POOL results with additional sequence and structural information to identify RNA binding sites in canonical and moonlighting RNA binding proteins. Pseudoenzymes are proteins or domains that have three-dimensional folds and amino acid sequences that are similar to conventional catalytically active enzymes, but have no catalytic activity. They serve in allosteric regulation of active enzymes, signal integration, competitive inhibition or scaffolding protein complexes. Some proteins that were presumed to be catalytically inactive based on amino acid sequence analysis have been found to have an alternative catalytic function due to the use of other amino acids in the active site. Our new method using POOL distinguishes between active enzymes (canonical or noncanonical) and inactive pseudoenzymes within enzyme superfamilies.

Use of Computed Electrostatic and Geometric Information to Investigate Protein Functions and Functional Sites

Trang Tu[1], Tina Harati[1], Gayathri Ayyar Manjula[1], Geordie Emberling[1], Amman Hossain[2], Ruby Renfrow[2], Alquama Lokhandwala[1], Kyle Zhang[1], Ganesan Murugan[1], Nicole J. Curtis[2], Mary Jo Ondrechen[1], and Constance (Connie) Jeffery[1,2]

[1] Dept. Chemistry and Chemical Biology, Northeastern University, Boston, MA 02115 USA

[2] Dept. Biological Sciences, University of Illinois at Chicago, MC066, Chicago, IL, 60607, USA

Partial Order Optimum Likelihood (POOL) is a machine learning method that combines computed electrostatic and geometric information for high-performance prediction of catalytic residues in enzyme structures. But many proteins are not enzymes. We are developing new methods using POOL to study noncatalytic protein functions. Two of these methods target RNA binding proteins and pseudoenzymes.

RNA binding proteins play vital roles in RNA metabolism and function, including splicing, translation, localization, stability and degradation. In addition to canonical RNA binding proteins where RNA binding is an aspect of their main function, dozens of moonlighting proteins have been found that combine an enzymatic function in sugar, lipid, or amino acid metabolism with an RNA binding function. We developed RNABinderFinder, a novel machine learning method that combines POOL results with additional sequence and structural information to identify RNA binding sites in canonical and moonlighting RNA binding proteins.

Pseudoenzymes are proteins or domains that have three-dimensional folds and amino acid sequences that are similar to conventional catalytically active enzymes, but have no catalytic activity. They serve in allosteric regulation of active enzymes, signal integration, competitive inhibition or scaffolding protein complexes. Some proteins that were presumed to be catalytically inactive based on amino acid sequence analysis have been found to have an alternative catalytic function due to the use of other amino acids in the active site. Our new method using POOL distinguishes between active enzymes (canonical or noncanonical) and inactive pseudoenzymes within enzyme superfamilies.