

## Graphical models of interacting proteins

When studying a particular class of protein-protein interactions (e.g., PDZ domains and their ligands, or serine proteases and their inhibitors), a central goal is to model the underlying determinants of molecular recognition in a manner supporting prediction of and design for affinity and specificity. We have been addressing that goal by developing probabilistic graphical models for interacting proteins, one model centered on sequence information and one on structural information. Our sequence-based model integrates multiple sequence alignments and information about which pairs of proteins are known to interact, in order to extract residue "cross-coupling" constraints and thereby evaluate the plausibility of new interactions. We show that our approach uncovers biologically interesting constraints, successfully identifies known interactions, and makes explainable predictions about novel interactions. Our structure-based model employs molecular modeling and evaluates possible side-chain conformations, in order to compute partition functions and thereby predict changes in free energy of association for variants. We demonstrate that our approach is fast and accurate, and improves prediction error relative to gold-standard approaches due to its explicit computation of free energies.

Joint work with Alan Friedman (Purdue), Chris Langmead (CMU), and Naren Ramakrishnan (Va. Tech.)