Reducing Uncertainty Through Mutual Information in Structural and Systems Biology

Vincent D. Zaballa, Elliot E. Hui
Department of Biomedical Engineering, University of California, Irvine

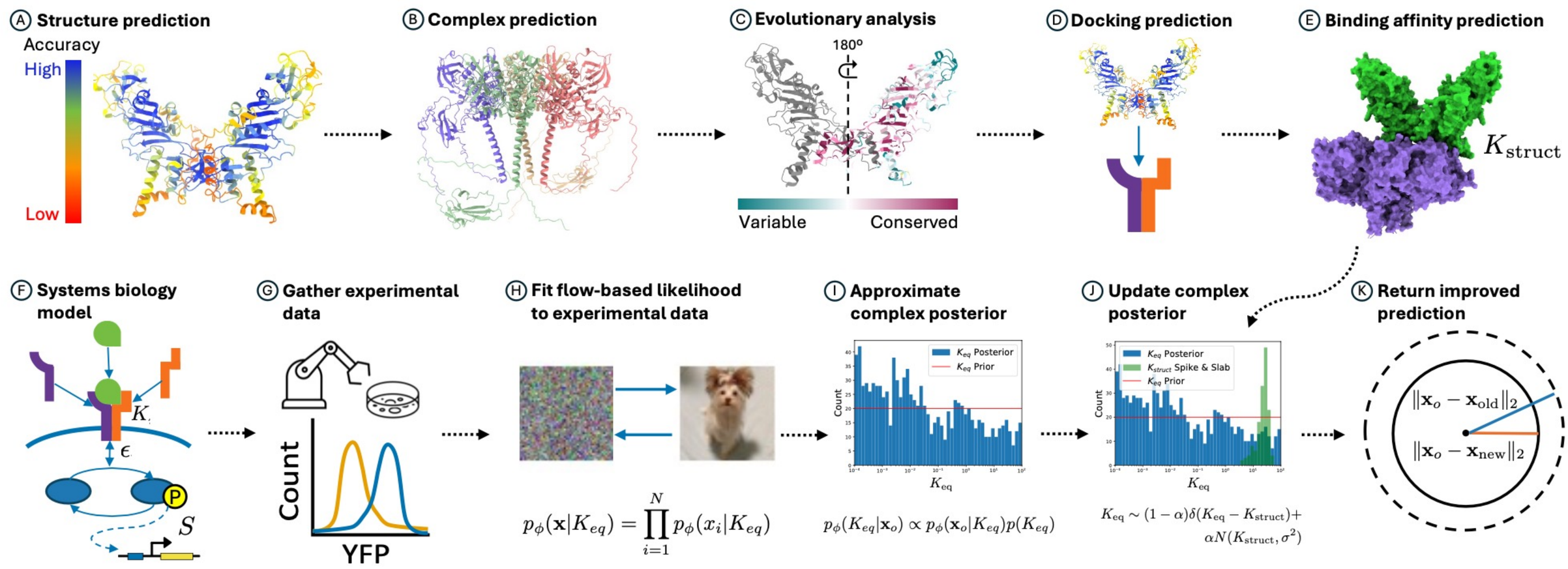
Systems Biology Models Robustly Extrapolate Predictions but Require Expensive Data to Fit

- Systems biology builds models of cell circuits all the way up to organ systems to model biology
- Using models based on physical and chemical principles can predict biological response outside of collected data distribution
- Can require *copious* amounts of data or compute to accurately infer latent parameters depending on the size of the model

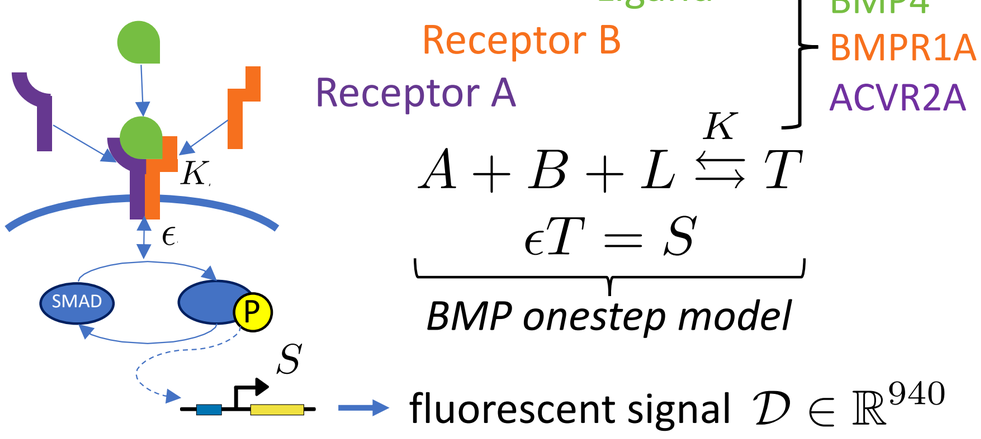
Protein Structure Prediction has Led to Unprecedented Biological Insights but Struggles to Extrapolate Out of Distribution

- Structure prediction software such as Alphafold, Openfold, and RosettaFold achieved unprecedented accuracy in single- and multi-chain prediction
- However, structure predictions are limited by their static descriptions of biology

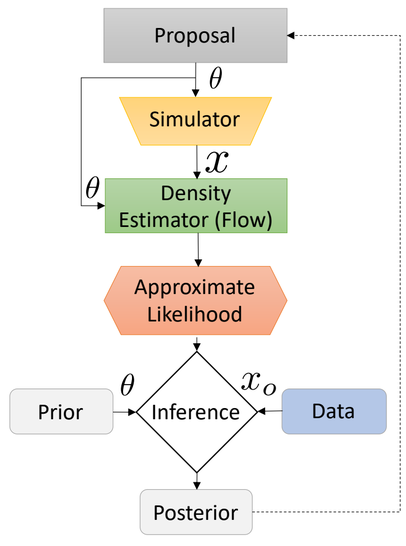
Systems Biology 🤝 Structural Biology



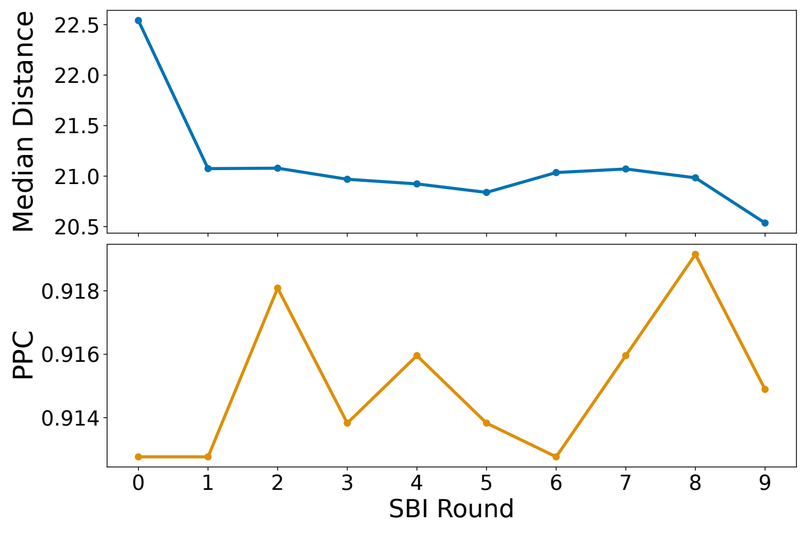
Simulators of Systems Biology: *The Bone Morphogenetic Protein (BMP) Pathway*



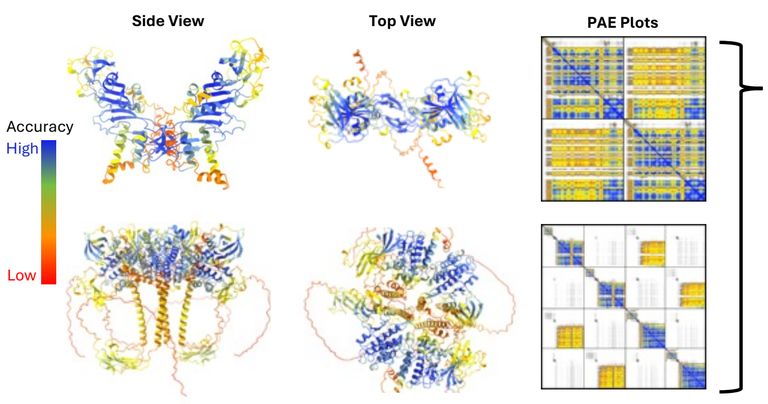
Normalizing Flows as Likelihoods in Simulation-Based Inference



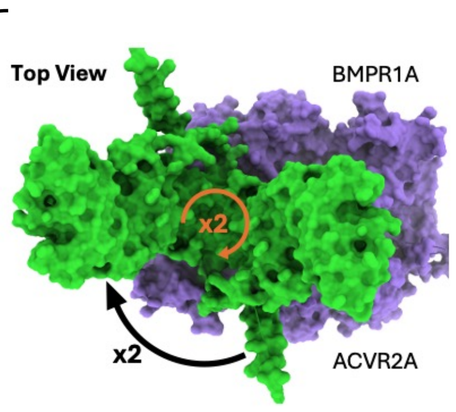
Structural Information Improves Systems Biology Predictions

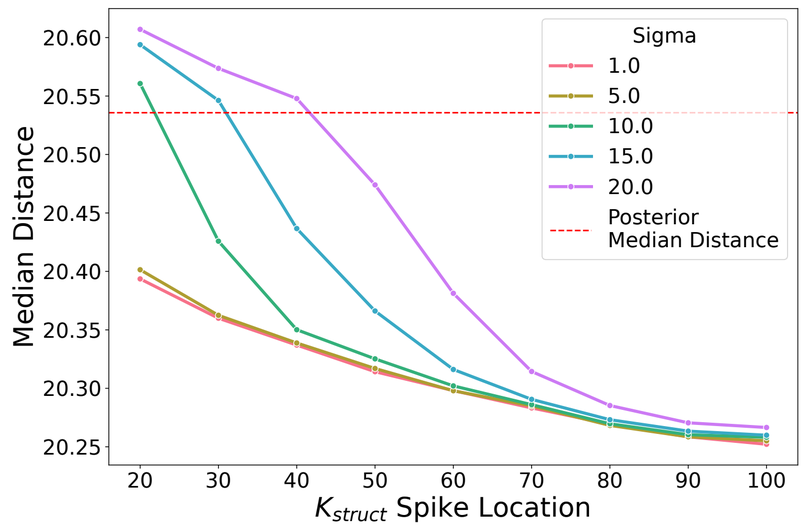


Systems Biology and Accurate Structure Predictions Help to Evaluate New Structural Hypotheses



- High accuracy in functionally-relevant regions (cell surface) enabled subsequent analysis
- Evaluating symmetries present in proteins allowed multiplying initial K_{struct} prediction by four times to 40.97





Discussion

- Demonstrated how to include structural information into systems biology predictions to improve systems biology predictions
- Introduced a novel method to cross validate structural hypotheses using systems biology models
- Choice of systems biology model has downstream implications in evaluation of predicted binding affinities
- Future work will include probabilistic implementations of the structural prediction pipeline to better capture uncertainty

Acknowledgements

This research was funded by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under award number 1F31GM145188-01. Coded in **JAX** and **LFIAX**. We thank Heidi Klumpe, Eric Bourgain, and Pieter Derdeyn for helpful feedback.