# Bayesian analysis of binding experiments

Full understanding of binding processes can require integrating large amounts of data collected using multiple analytical instruments and experimental protocols. Existing statistical methods and software do not fully integrate data from multiple sources to produce useful knowledge. In collaboration with [John Chodera of the Memorial Sloan Kettering Cancer Center](https://www.choderalab.org/) and [Lulu Kang from the applied mathematics department](http://math.iit.edu/~lkang2/), we are pioneering use of a new approach (a "Bayesian network") as a general framework for analyzing chemical measurement data from multiple instruments and protocols and for designing new experiments. The framework will be usable for both small laboratory experiments and the massive datasets generated by automated instrumentation. The software (including a straightforward user interface) will be used to teach the underlying principles in related courses, will be made freely available online, along with tutorials and clear documentation.

Our team is developing chemometric methods and software for analyzing data related to binding. We have developed Bayesian methods to analyze X-ray solution scattering [1] and isothermal titration calorimetry [2]. We are working to fuse data from diverse methods, including isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), absorbance, fluorescence, and X-ray solution scattering. Key features of the software include (1) automated parameter determination for physical binding models, (2) uncertainty propagation and quantification for model parameters, (3) automated and principled model selection and hypothesis testing, and (4) Bayesian experimental design to maximize acquisition of new information while minimizing cost. The software will automatically construct Bayesian networks that consider all sources of experimental error (e.g. dispensing, weighing, transfer, and measurement) for any experiment described in a machine-readable standard. It will then perform Bayesian inference to weigh evidence for competing physical models, obtain credible intervals for thermodynamic and kinetic parameters, and propose new experiments. Using this software, robotic experiments, statistical inference, and Bayesian experimental design may be efficiently iterated to reduce model ambiguity and improve parameter precision.

## Related References

[1] D. D. L. Minh and L. Makowski, “Wide-angle X-ray solution scattering for protein-ligand binding: multivariate curve resolution with bayesian confidence intervals.,” *Biophysical Journal*, vol. 104, no. 4, pp. 873–83, Feb. 2013, doi: 10.1016/j.bpj.2012.12.019.

[2] T. H. Nguyen *et al.*, “Bayesian analysis of isothermal titration calorimetry for binding thermodynamics,” *PloS One*, vol. 13, no. 9, p. e0203224, 2018, doi: 10.1371/journal.pone.0203224.