

Nathanael P. Kazmierczak  
Dr. Douglas A. Vander Griend

## Analyzing Chemical Self-Assembly Systems

When you think about a chemical reaction, you may picture molecule A running into molecule B to form a completely new compound, molecule C. But in a large and important class of chemical reactions, molecule A and B merely 'stick' to each other rather than forming something completely new. This 'sticky' interaction is known as a coordinative bond, and under certain conditions it allows individual molecules to spontaneously assemble into larger objects such as grids, tetrahedrons, and jars. This phenomenon is known as chemical self-assembly, and it forms one of the principle methods for chemists to build structures and ultimately machines on the molecular scale.

Our research seeks to find the best data analysis methods for studying how chemical systems self-assemble in solution. This is challenging because self-assembly is almost always an equilibrium process. Thus, if we put two starting molecules in a solution in the hopes of forming a tetrahedron through self-assembly, the solution will actually be a mixture of the tetrahedron, the starting molecules, and a variety of sub-assemblages that form on the way to making the full tetrahedron. It can be quite challenging to figure out how many different assemblages are present in solution, as well as which assemblages are most prevalent. The key to answering this last question is to determine the equilibrium constants governing the assembly process, and so our lab focuses on obtaining the most accurate equilibrium constants possible. To do this, we conduct careful titrations in a UV-Vis spectrophotometer, which is an instrument that functions like a mechanical eyeball, telling us the precise color of the solution. By analyzing the color with mathematical techniques, we can determine the number of assemblages and the relevant equilibrium constants.

I worked on three main projects this summer to help us get more accurate information from our titrations. First, I studied how different ways of processing the titration data affect the error on the calculated equilibrium constants. Our work in this area showed through algebraic proofs that baseline shifting to correct for negative absorbance values is a reliable technique as long as we conduct the titration through two mutually diluting stock solutions. Second, I studied different methods of estimating the uncertainty on the calculated equilibrium constants from a single dataset. This is important because it tells us how much confidence we can have in our experimental conclusions. However, using computer simulations in the scientific computing language Matlab, I demonstrated that the usual (linearized) method of estimating equilibrium constant uncertainty in fact greatly underestimates the true uncertainty, while another technique known as bootstrapping provides more accurate estimates (though it takes longer to calculate). Finally, I programmed a method by which the computer can determine the number of each starting molecule present in the self-assembled structures. This is a substantial time-saver compared to the previous manual guess-and-check technique that researchers have employed.

This summer taught me a lot about interdisciplinary thinking, as my research includes ideas from chemistry, computer science, statistics, and applied mathematics. I became familiar with the technical language used in each of these disciplines so that I could understand all the scientific articles relevant to my research. I have also deepened my appreciation for the challenges and rewards of scientific research. Sometimes I thought a project had hit a wall, but drawing conclusions at the end was very satisfying.