Communicating Research to the General Public

The WISL Award for Communicating PhD Research to the Public launched in 2010, and since then over 100 Ph.D. degree recipients have successfully included a chapter in their Ph.D. thesis communicating their research to non-specialists. The goal is to explain the candidate's scholarly research and its significance—as well as their excitement for and journey through their area of study—to a wider audience that includes family members, friends, civic groups, newspaper reporters, program officers at appropriate funding agencies, state legislators, and members of the U.S. Congress.

WISL encourages the inclusion of such chapters in all Ph.D. theses everywhere, through the cooperation of PhD candidates, their mentors, and departments. WISL offers awards of $250 for UW-Madison Ph.D. candidates in science and engineering. Candidates from other institutions may participate, but are not eligible for the cash award. WISL strongly encourages other institutions to launch similar programs.

Wisconsin Initiative for Science Literacy

The dual mission of the Wisconsin Initiative for Science Literacy is to promote literacy in science, mathematics and technology among the general public and to attract future generations to careers in research, teaching and public service.

Contact: Prof. Bassam Z. Shakhashiri

UW-Madison Department of Chemistry

bassam@chem.wisc.edu

www.scifun.org

Progress in Understanding Kinetics of Homogeneous Catalytic Reactions

By

Katherine E. Ziebarth

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2022

Date of the final oral examination: 04/04/22

The dissertation is approved by the following members of the Final Oral Committee:

Clark R. Landis, Professor, Chemistry
John F. Berry, Professor, Chemistry
Ive Hermans, Professor, Chemistry
Tehshik P. Yoon, Professor, Chemistry

Chapter 1.

Introduction for a Non-Chemist

This chapter was written as part of the Wisconsin Institute for Scientific Literacy's program for Communicating PhD Research to the Public.

1.1 Introduction

My research revolves around studying the quantitative side of chemical reactions. I find this area to be fascinating, but I rarely get a chance to share the details of what I study to community members without scientific backgrounds. I would like to thank the Wisconsin Institute for Scientific Literacy at the University of Wisconsin-Madison for providing me with this opportunity to share my research in a way that is accessible to non-chemists. I hope that you take away from this chapter an understanding of why it is important to understand the quantitative side of how reactions proceed.

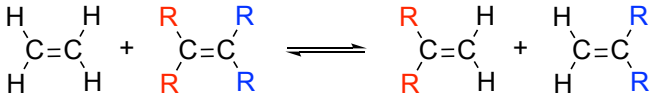
The main purpose of my research is to study how various chemical reactions work. In particular, I am interested in producing quantitative models to describe these reactions. These models can be used to predict what will happen when the reaction is run differently (for example, at a different temperature). This information guides chemists as they design reaction processes that are more efficient or produce different desired products.

1.2 What reactions am I interested in?

There are limited supplies of fossil fuels like petroleum available, so society cannot continue to depend on them indefinitely. When most people think about the problem of limited petroleum, they think about the need for alternative fuel sources for cars, trucks, and other machinery. However, petroleum is also an important source of chemical products used in everything from personal care products to clothing. Some people worry that these products are fundamentally unsafe because they use petroleum-derived components. While this is generally not a concern, it is still important to find alternative ways to make petroleum-derived chemicals before there are major shortages of petroleum.

As a result, chemists are developing ways to use plant-derived chemicals to make many products previously derived from petroleum. I am interested in ethenolysis, a reaction that allows

plant-based oils to be turned into useful products like lubricants and cleaning ingredients. Ethenolysis involves molecules that contain carbon-carbon double bonds. Chemical bonds involve sharing of electrons between two atoms and are classified based on how many electrons are shared. More shared electrons lead to a stronger bond, meaning it is harder to separate the atoms. Many bonds consist of two shared electrons and are called single bonds. Double bonds, like those needed here, involve the sharing of four electrons. Generally, reactions require certain types of bonds to be present; otherwise, there is no reaction. In ethenolysis, there is a reaction between ethylene (a particular molecule with a carbon-carbon double bond) and a different molecule with a carbon-carbon double bond (Scheme 1.1). You might have heard of ethylene as a gas that can help fruits ripen, or as the starting material to make the plastic called polyethylene. Derivatives of plant oils have carbon-carbon double bonds that allow them to react with ethylene using ethenolysis. Ethenolysis products are used as starting materials for reactions needed to make products like lubricants and cleaning ingredients.



Scheme 1.1 Ethenolysis reaction. Chemists use “R” to represent a generic group of atoms.

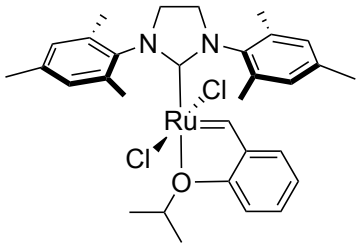
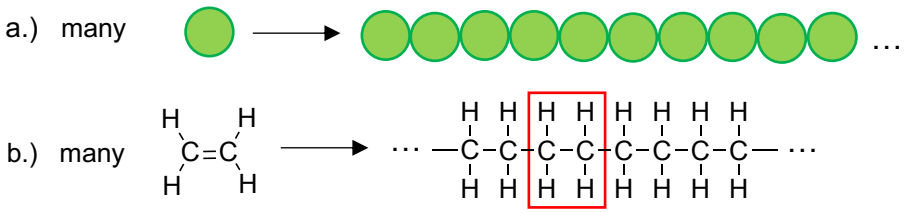


Figure 1.1 Second-generation Grubbs-Hoveyda catalyst. This is an abbreviated structure that does not show all the atoms. Unlabeled vertices and ends of line segments represent carbon atoms, and hydrogen atoms are not shown. Bolded lines represent parts of the structure coming forward, and dashed lines represent parts of the structure going back.

Ethenolysis uses a catalyst to speed up the reaction. Some reactions occur quickly at room temperature – think of the reaction of vinegar and baking soda as an example. Other reactions occur less readily. For example, to start a fire in your fireplace, you use a spark to start the reaction of burning wood. If you were to put wood in the fireplace and wait for it to burn without a spark, you would be disappointed. This reaction, like others, needs an input of energy to start re-arranging the bonds in the molecules. The amount of energy required depends on the reaction. Catalysts are chemicals that reduce the amount of energy required for a reaction to occur, so that it can occur faster. In some cases, like ethenolysis, catalysts allow a reaction to occur under temperature conditions where the reaction might otherwise not occur at all (like wood burning with no spark in your fireplace).

Many catalysts for ethenolysis have been developed. These catalysts typically contain ruthenium or molybdenum (two metallic elements) as a part of the molecule. I am interested in a particular ruthenium catalyst, called the second-generation Grubbs-Hoveyda catalyst (Figure 1.1). In my research, I carry out reactions involving this catalyst and ethylene to obtain quantitative information about how the catalyst is working.

The other main reaction I am interested in is polymerization. Polymers consist of long chains of molecular building blocks called monomers (Scheme 1.2). Plastics are one of the main examples of polymers. Depending on which monomers are used and what the molecular level structure is, different kinds of polymers can have very different properties. For example, plastic wrap has very different properties than disposable foam cups, even though both consist of polymers. To understand the origin of polymer properties and to make new types of polymers with desirable properties, we need to understand what is going on at the molecular level. Other members of my lab focus on studying these reactions in the lab, but my goal is to give my colleagues better ways to model their data so that they can extract more information from their experiments.



Scheme 1.2 Polymers consist of molecular building blocks called monomers. (a) Green spheres represent monomers. (b) Polyethylene consists of repeated ethylene-derived blocks, shown with the red square.

1.3 What are chemical mechanisms and kinetics?

At a molecular level, chemical reactions are often a lot more complicated than the starting material just turning directly into the product. To understand this, it is helpful to use a relay race as an analogy. Just as relay races have multiple legs, chemical reactions often involve multiple steps to turn the starting material into the product. Each of these steps can involve different chemical species, just as each leg of the race involves different runners. The set of steps involved in the reaction is called a mechanism, and the different chemical species that are produced on the way to the product are called intermediates.

A relay race coach wants his team to reach the finish line first. However, the overall rate at which the team reaches the finish line is determined by how long it takes each runner to finish his leg of the race. This means the coach needs to pay close attention to each runner and help each of them achieve his best time. By understanding how fast each runner travels and anything that affects the runner's speed, the coach can create a targeted plan for each runner. This will help the team's overall time to improve. Similarly, chemists are interested in how fast chemical reactions occur, an area of chemistry called kinetics. To really understand how fast a reaction will occur and what factors affect the rate, chemists need to understand the individual steps in the mechanism. This allows chemists to optimize the reaction conditions (temperature, pressure, amounts of starting materials, etc.). For example, increasing the pressure of certain reactions might cause them to finish faster.

Chemists create kinetic models to help them make predictions about what will happen in a reaction. For example, the model can show how much product will be produced in a certain amount of time, or the effect of changing the amount of a starting material. A kinetic model consists of a system of equations that describes how fast each chemical species in the reaction (starting materials, products, and any intermediates) appears or disappears. These equations include numbers called rate constants, which reflect the speed of the reaction. One of the challenges of creating a kinetic model is that you need to know the correct values of rate constants to use. If the values of the rate constants are wrong, the model's predictions will be incorrect. For complex reactions, there may be thousands of equations (or even more) in the model. The equations for a simple model can be solved using calculus. However, many models have sets of equations that are too complex to solve exactly, so a computer is used to estimate the solution.

1.4 How do chemists determine the mechanism?

When chemists are interested in understanding the mechanism and kinetics of a reaction, their first step is to collect experimental data. This data might show how fast the starting material in a reaction goes away or how fast the product forms. Sometimes, it is possible to measure when different intermediates form, and in what quantities. However, this can be difficult if the intermediates react quickly or if only small amounts of them form.

Once the experimental data are collected, the next step is to propose possible mechanisms. The best way to find ideas is to look at similar reactions that have been studied in the past. For each proposed mechanism, chemists can create a kinetic model consisting of the relevant rate equations. The next step is to find the best rate constants to use for each model. The goal is to have multiple proposed kinetic models, each including appropriate rate constants. At this point, each model is used to predict what will happen under certain experimental conditions, and the predictions are compared to the experimental data. The model with the most accurate prediction is chosen as the best model for the reaction. However, it is impossible to

"prove" a model. In the future, chemists might collect more data that suggest that a different model is more appropriate. Alternatively, someone might later think of a different possible mechanism that was not initially considered that is a better match for the experimental data.

A challenging step in developing kinetic models is finding the best set of rate constants. The best set of rate constants is the one that allows the model to predict most accurately what happens experimentally. The process starts out with an initial guess of the best rate constants, which are then gradually changed to improve how well the model's prediction matches the available experimental data (Figure 1.2). Special computer programs are used for this process, which is called gradient-based optimization.

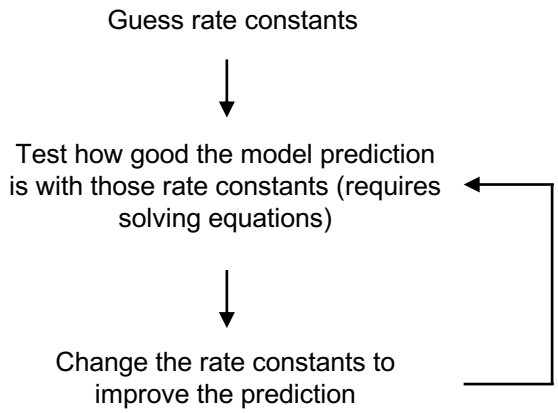


Figure 1.2 Summary of process to optimize rate constants

The chemist starts by guessing what a good set of rate constants might be. Then a computer program tests out how well the model's prediction matches the data. Next, the program changes the guess of the rate constants a little bit to improve the model's prediction. It then evaluates how well it is doing at fitting the data. The program goes through many iterations where it gradually varies the rate constants and then assesses how well it is doing. Each time the program evaluates its progress so far, it must solve the equations in the kinetic model. This process works well for simple models that have a small number of equations. However, for large

models, it can be time-consuming for the computer to approximate the solution to the set of equations. The time required is especially significant since this checking step must be repeated sequentially many times as the rate constants are gradually changed. When I was modeling my ethenolysis data, the model was simple, so I used this gradient-based optimization approach and was able to fit my data well.

The best way to visualize this is to think about a hiker trying to find the lowest spot in a mountainous region without a map (Figure 1.3). The hiker will check his altitude and the slope of the hillside, move down the slope a little, re-check his altitude and slope, and then repeat this process until he reaches the bottom of the valley. Depending on where the hiker starts, it might take him a long time to reach a valley. He also might end up in a valley that is not actually the lowest one in the region. In this analogy, different locations are like different sets of rate constants. The height of the land at each point corresponds to the error in the model prediction relative to the experimental data. Depending on how good the initial guess of rate constants is, it might take the computer a long time to find the best rate constants (with the lowest error). It is also possible that the program might miss the best set of rate constants.

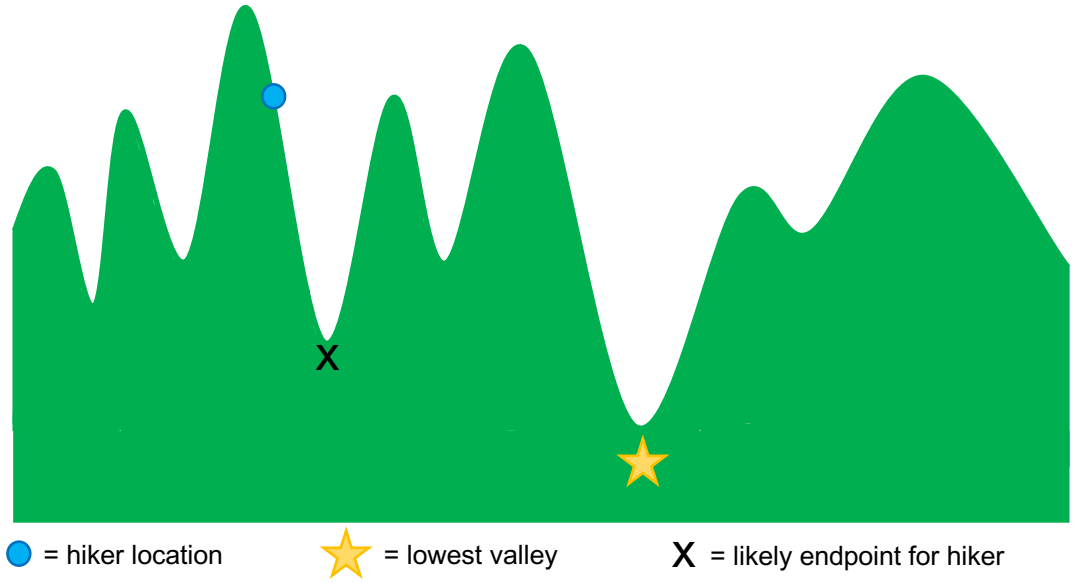


Figure 1.3 An example of a hiker trying to find the lowest valley in a mountainous region without a map.

1.5 How can the process of finding the rate constants be improved?

For some reactions, using gradient-based optimization as a modeling method does not work well. Polymerization reactions are especially complicated because a polymer sample typically contains many different chain lengths, and every chain length needs to be separately tracked with its own equation in the kinetic model. The more complicated a polymerization mechanism gets, the harder it is to effectively use gradient-based optimization. The process of determining the rate constants gets very slow and might not accurately determine the best rate constants. One of my major goals has been to improve the way that we model polymerization. As a result, I have written a computer program that uses a different approach to find the best set of rate constants.

Think back to the hiker in the mountainous region. Imagine that the hiker had a map with altitude measurements for many different spots in the area (Figure 1.4). This would make it significantly faster for him to find the lowest spot in the region. Furthermore, in the absence of this map, the hiker would likely end up in one of the nearest valleys regardless of whether it was the lowest, whereas the map could help him find the correct lowest valley, even if it was far from his current location. The computer program I wrote basically makes a map like this. It tests many different sets of rate constants to determine how well they allow the model to predict the experimental data. This provides a guide to finding the best rate constants for the model.

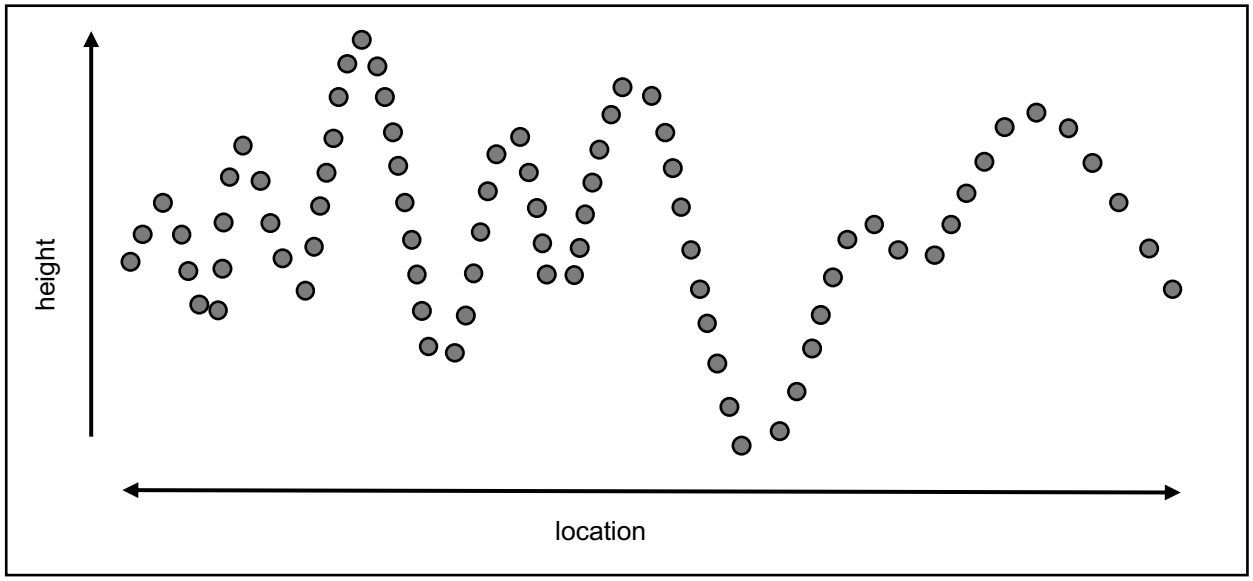


Figure 1.4 An example of a map that would help the hiker find the lowest valley.

This approach makes it easier to find the best set of rate constants. It can also be more efficient. In this day and age, there is a lot of computing power that is available. One resource I use is the Open Science Grid (OSG), which allows researchers to access computational resources across the country. The OSG consists of huge numbers of computer nodes located at various universities and National Labs. I can submit calculations to the OSG, and it will match up my calculations with available computer nodes in the network. This makes it possible to run many calculations at a time. As a result, creating the “map” of rate constants and their corresponding prediction errors can be done very efficiently. Different sets of rate constants are independent of one another, so they can be analyzed on different computer nodes. Dividing a big task (like creating the map) into smaller pieces that can be done concurrently is called parallelization. Before large numbers of computer processors were available, it would not have been possible to divide the map creation into many small parallel calculations. It might have been computationally infeasible to create this sort of map at all. Now that we live in an era of extensive computing power, this sort of approach to finding the best rate constants is practical and beneficial for complex systems.

The computer program I wrote makes it possible to study reactions with very complex mechanisms that might be too slow or too difficult to model otherwise. In my thesis, I demonstrate its application to polymerization. The hope is that this method and the code I wrote can be used to model more complicated reactions (especially different polymerization reactions) in the future. This will allow future researchers to better understand these reactions and guide their efforts to make current chemical products more efficiently or to create new products for different applications.