Project Milestone 4 – Technical Brief

To: President Avery D. Lion

From: Team 08: Eric Rohweder, Blaine Ballard, Adrian Li, and Matthew Van Ausdal

RE: Enzyme Kinetics Analysis

Date: April 30th, 2021

Introduction

Our client, Natural Catalysts, is developing five new enzymes that would be implemented in detergents that break down organic substances that cause stains for the substances to be removed from the fabric. Natural Catalysts has tasked us with creating an algorithm that will analyze the kinetic enzyme test data used to determine the success and quality of the experimental enzymes. This algorithm will deliver a detailed analysis of the data, including charts and graphs of the data, as well as a description of any potential error in the data-analysis process. We will then be tasked with drafting a recommendation as to what the enzymes can truly accomplish, including any limitations with the enzyme itself.

The algorithm out team has developed is able to take any format of enzyme concentration data over time and calculate initial velocities for each concentration along with accurate Km and Vmax values for the enzyme. Along with this, our algorithm also produces appropriate concentration versus time graphs and a Michaelis Mensen model of the given enzyme.

Our first decision that we made was changing how we calculated the initial velocity. We originally used only the first two points to calculate the initial velocity. We figured that since the initial velocity was the closest reading to time zero that this would give us the most accurate reading on the initial velocity. However, the readings were wildly inaccurate, and the initial velocities of the enzyme did not increase as concentration increased, which was our first indication that we needed to change something in the way we calculated the initial velocity. We ultimately decided to use the first 20 points, as Henkel said to use the first 5% of the data points (2021). This decision ultimately led to very accurate readings of our initial velocity, with an SSE reading of 0.030.

When we first were developing the algorithm, we decided that we were going to use the Lineweaver-Burke method of linearization. We made this decision due to its popularity with past teams and because our research showed that it was popular among other companies in the enzyme industry. However, after running our algorithm with the data that we calculated, the Lineweaver-Burke model did not fit the data we calculated. After examining the code and finding no problem with the code itself, we decided to change the model to the Hanes-Woolf model, as it has found to be more reliable than the Lineweaver-Burke model according to an article done by Marasovic, Marasovic and Milos (2017). This one change improved our algorithm drastically, as this improved both how well the model fit the data, as well as the calculated maximum velocity and Michaelis-Menton constant (Km).

The final refinement that we made to the algorithm to improve the accuracy of the parameter identification was to improve the way we calculated SSE. Before the refinement, the Sum of Squared Errors for our data set was 3.502, which is extremely high. Even after changing the model to Hanes-Woolf, our SSE was calculated as 7.3. We figured out that our SSE needed to be improved, and all it took was adding a matrix calculation in order to improve this metric.

Parameter Identification Procedure

First, the algorithm imports the data set csv that was given to a matrix. After importing the data, the algorithm figures out which row the data starts on by checking which rows do not contain numbers since the data we were given usually had a few labels in the first few rows of the first column. Next, it figures out which row contains the enzyme concentration data, since we will need that for calculations later. Next, there is a loop that increments through each of the enzyme test groups and then each concentration within that group. The initial velocity of each enzyme at each concentration is found by taking the 20^{th} recorded velocity, which will create a more accurate initial velocity than just taking the 1^{st} recorded velocity. We then use MATLAB's built-in polynomial curve fitting function to find the coefficients of our Hanes-Woolf model. Utilizing the coefficients, we next calculate the maximum velocity attained by the enzyme as well as the Km of the enzyme at the given concentration. After calculating all of these values, it is just a matter of technically formatting the plots for each enzyme test group and calculating the SSE for the Michaelis-Menten model of each group. The duplicate data is handled at the end, where each initial velocity, max velocity, Km value, and SSE is averaged between the two duplicate tests.

Results

Table 1 shows the maximum velocity and Michaelis constants for each of the enzymes, as well as their sum of squared errors. Enzyme 5 had the fastest maximum velocity at 2.091 micro-molarity per second, while Enzyme 2 had the slowest maximum velocity at 1.344 micro-molarity per second. This means that Enzyme 5 will be able to break down its substrate at the quickest rate. Figure 1 shows our Michaelis-Menton curves for each of the five enzymes. Enzyme 5 had the steepest rate of change in our model as well as the fastest maximum velocity, while Enzyme 2 had the slowest maximum velocity and initial rate of change.

Interpretation

The error of our process is quantifiable by comparing the true values of initial velocities with that of our algorithm's calculated initial velocities. By taking the sum of the squared differences between each respective initial velocity and its calculated counterpart, we can obtain what is called an SSE (Sum of Squared Errors) values (Bourdin, 2021). This value quantifies the error of the data set and makes it easily comparable with other data sets. To interpret these SSE values, the goal is to obtain a value as close to 0 as possible as this would mean that there is zero difference between the actual data points and the calculated data points and thus no error in the data set.

If reaction velocity and concentration do have a relationship, then there should be a model that can accurately represent such a relationship. Although there are some differences within and between the predicted and actual data, it is safe to say that their products, for the most part, work consistently.

References

Henkel, J. (2021, March 25). The Initial Velocity of an Enzyme Reaction - Cell Biology. Hearts in Healthcare

https://www.heartsinhealthcare.com/cell-biology/the-initial-velocity-of-an-enzyme-reaction.html.

Maja Marasović, Tea Marasović, Mladen Miloš, "Robust Nonlinear Regression in Enzyme Kinetic Parameters Estimation", *Journal of Chemistry*, vol. 2017, Article ID 6560983, 12 pages, 2017. https://doi.org/10.1155/2017/6560983

Appendix: Figures and Tables

Table 1

	Maximum Velocity **(µM/s)**	Michaelis Constant (Km) **(µM)**
Enzyme 1	1.513	190.199
Enzyme 2	1.344	550.991
Enzyme 3	1.554	234.96
Enzyme 4	1.713	433.363
Enzyme 5	2.091	244.868

Figure 1