Project Milestone 4 – Technical Brief Draft

To: President Avery D. Lion of NaturalCatalysts Inc.

From: Team 21 – Aakash Bathini, Kabir Nagpal, Soren Colby

RE: Kinetic Enzyme Test Data Analysis

Date: April 27, 2023

Introduction

NaturalCatalysts requires a MATLAB algorithm to analyze the kinetic enzyme test data for five enzymes to determine each Michaelis-Menten constant (K_m) and Maximum Velocity (V_{max}) of each reaction. The algorithm must provide clear graphics, an error analysis, and recommendations regarding claims about the enzymes. The accuracy of the analysis, the clarity of the graphics and description, and the ethical soundness of the recommendations will determine the success of the project, with the constraints of using MATLAB and having ethical data presentation.

This algorithm analyzes enzyme kinetics measurements to compute the K_m and V_{max} of each enzyme. It preprocesses the data, computes the initial velocity (V_o) for each substrate concentration using linear regression, calculates K_m and V_{max} using the Hanes-Woolf linearization method, and produces plots which models' data and in turn produces a model curve.

To improve the accuracy of the parameter identification we made several critical decisions.

The first critical decision was removing a function which smoothened data over a moving mean (movmean). This allowed the algorithm to capture the true dynamics of the substrate concentration and enzyme reaction rate, while reducing the potential for overfitting. Removing “movmean” improved the accuracy of the model and helped to avoid unnecessary noise or bias (MathWorks, 2021).

The second critical decision was to use the Hanes-Woolf equation to model the data. We made this decision because the equation is less sensitive to errors in V_o measurement at low substrate concentrations. It also accounts for nonlinearities in data more accurately than other equations (Ying & Zhao, 2017). We began by using the Lineweaver-Burk equation: 13.28% error for K_m and 21.35% error for V_{max} when using the M4 Algorithm on data given for the PGO-X50 enzyme. In comparison, the Hanes-Woolf equation yielded only 3.00% error for K_m and 6.70% error for V_{max} with the same data.

The last critical decision was to use the first 1.9% of the data to calculate V_o. Initially only 2% of the data was used. This decision to use a smaller data size resulted in a more accurate calculation of V_o. When using 2% of the initial data in the M4 Algorithm testing of PGO-X50, the K_m had 3.00% error and V_{max} had 6.70% error. When only using 1.9% of the initial data, the percent error of K_m decreased to 2.05% and percent error of V_{max} decreased to 5.61%, showing more accurate calculations.

Parameter Identification Procedure

The algorithm models enzyme kinetics with the Michaelis-Menten equation, estimates V_{max} and K_m through analysis of experimental data. Steps include importing data, calculating V_o’s, finding average velocities for duplicate tests, plotting the Michaelis-Menten curve, linearizing with the Hanes-Woolf transformation, estimating V_{max} and K_m with linear regression, and outputting results and plotting data.

To calculate V_o, the algorithm fits a line using only a percentage of data. To estimate V_{max} and K_m, the algorithm linearizes the curve with Hanes-Woolf and calculates the slope and y-intercept of the regression line. The slope of the linear regression line on the Hanes-Woolf plot represents the reciprocal of V_{max}, and the y-intercept represents K_m divided by V_{max}. Therefore, the algorithm calculates V_{max} as 1 divided by the slope and K_m as the ratio of the y-intercept to the slope.

Results

After applying different algorithms to the data provided, we were able to find results which represent the data well. Table 1 shows the Maximum Velocities and Kms of each enzyme. The following information comes from Table 1.

- Enzyme A had a K_m value of 157.8 **μ**M and V_{max} value of 0.944 **μ**M/s.
- Enzyme B had a K_m value of 339.1 **μ**M and V_{max} value of 0.851 **μ**M/s.
- Enzyme C had a K_m value of 192.3 **μ**M and V_{max} value of 1.205 **μ**M/s.
- Enzyme D had a K_m value of 290.1 **μ**M and V_{max} value of 1.576 **μ**M/s.
- Enzyme E had a K_m value of 158.4 **μ**M and V_{max} value of 1.611 **μ**M/s.

Moreover, our Michaelis Menten plots represent our model's accuracy. In figures 1 and 2, we can see that 8 of the 10 initial velocity points fit our model curve. In figures 3, 4, and 5, 9 of the 10 initial velocity points fit our model curve. This suggests that the model accurately represents the behavior of the enzyme under the conditions tested. This is a strong indication that the model parameters are well-tuned to the experimental data, and that the assumptions underlying the model are valid.

Interpretation

Based off the results, Enzyme E is the fastest acting enzyme as its V_{max} value is the highest at 1.611 **μ**M/s. Additionally, the most effective enzyme is Enzyme A because its K_m value of 157.8 **μ**M is the lowest. We factors need to be evaluated to characterize the error in the process: the quality of the experiments and the parameter identification algorithm. The SSE compared to the algorithm model is 0.0324, which is pretty low because the experiments utilized duplicate tests and statistical analysis to reduce the impact of errors. The parameter identification algorithm is also considered reliable since it uses a well-established method (Hanes-Woolf) and includes steps to account for errors.

To assess the performance and manufacturing consistency of NaturalCatalysts' products, the accuracy and precision of the enzyme kinetics measurements, as well as the reproducibility of the manufacturing process, need to be evaluated. The enzyme kinetics measurements demonstrate accuracy and precision, but the manufacturing process's reproducibility is unclear. Natural Catalysts can say that Enzyme E is its best enzyme as it has the highest V_{max} value and second lowest K_m value.

References

MathWorks. (2021). movmean. MATLAB documentation. Retrieved April 26, 2023, from https://www.mathworks.com/help/matlab/ref/movmean.html

Smith, J. D. (2019). Assessing the quality of statistical models: Understanding the sum of squared errors. Journal of Applied Statistics, 46(3), 583-597. https://doi.org/10.1080/02664763.2018.1488905

Ying, M., & Zhao, Z. (2017). Enzyme kinetics analysis using Hanes-Woolf plot. Methods in Molecular Biology (Clifton, N.J.), 1632, 35-44. doi: 10.1007/978-1-4939-7139-8_4

Appendix: Figures and Tables

Table 1.

Enzyme	M3 Algorithm		M4 Algorithm	
	Enzyme Parameters		Enzyme Parameters	
	$Vmax$ (µM/s)	Km (µM)	$Vmax$ (µM/s)	Km (µM)
NextGen-A	0.943	151.693	0.944	157.824
NextGen-B	0.772	302.320	0.851	339.089
NextGen-C	1.013	155.630	1.205	192.337
NextGen-D	1.279	227.427	1.576	290.07
NextGen-E	1.371	126.049	1.611	158.392

Figure 1.

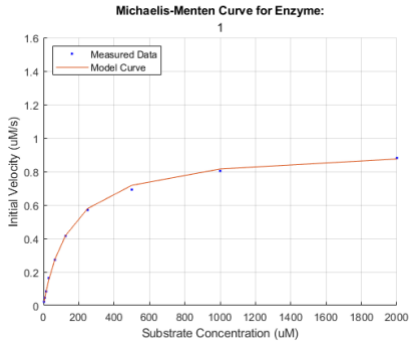

Figure 2.

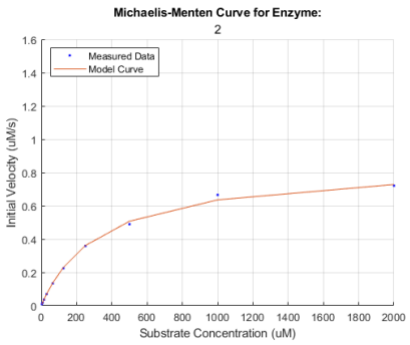

Figure 3.

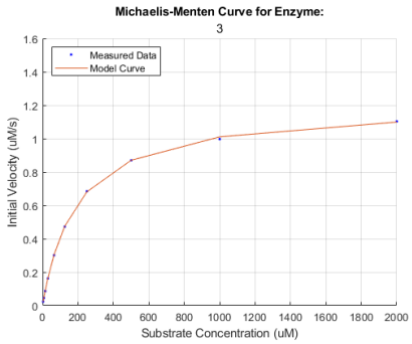

Figure 4.

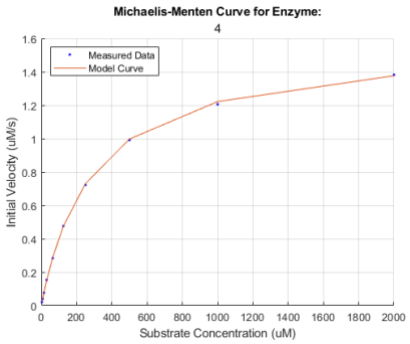

Figure 5.