Computational Photoenzyme Discovery and Design

Julia Chen, YSP Student, *North Quincy High School*
Annie Stone-Peterson, YSP Student, *Cambridge Rindge & Latin School*
Gustavo Mondragón-Solórzano and Zheyu Zhang, Chemistry and Chemical Biology, *Northeastern University*
Professor Sijia Dong, Chemistry and Chemical Biology, *Northeastern University*

Northeastern University
College of Science

Northeastern University
Center for STEM Education

Abstract

Motivation:

Before enzymes are used to perform chemical reactions within molecules for biological and industrial uses, their molecular interactions and reactivity have to researched. The goal of the lab is to classify and categorize enzymes in a database. This information will then be used to create a model to predict substrates and reactions that the enzymes can perform. There will also be applications to see if these enzymes can carry out specific chemical reactions under the presence of light.

Personal Objectives:

- Analyze enzymes extracted from the Protein Data Bank (PDB) with the FMN (Flavin Mononucleotide) or NAD+ (Nicotinamide Adenine Dinucleotide) cofactor
- Determine the substrate(s) and cofactor(s) for the enzymes and update the FMN and NAD+ Excel Spreadsheets
- Create a new database for enzymes that contain more than one cofactor
- Develop a deeper understanding of enzyme catalysis and molecular interactions

Results:

Enzymes with the FMN or NAD+ cofactor were analyzed to identify their substrates, inhibitors, and cofactors. For the 662 enzymes containing the FMN cofactor, it was found that 74.2% of the enzymes had only one cofactor and 25.8% of the enzymes contained more than one cofactor.

Background

Enzymes are proteins that catalyze biochemical reactions. Within the protein structure, there are chemical substances (organic or inorganic) that substrates can bind to. Enzymes have specific functions in organisms, creating products not easily synthesized in the laboratory under mild temperature, pressure, or pH conditions. There are efforts to use enzymes to perform chemical reactions within molecules for biological and industrial uses as they assist with cellular metabolism, drug delivery, diagnosis of illnesses, and manufacture of medicine. However, in-depth structure description and understanding of the enzymes is necessary prior to such applications. Because of this, we focused on a group of photoenzymes extracted from the Protein Data Bank (PDB) that contains two organic cofactors, either Flavin Mononucleotide (FMN), or Nicotinamide Adenine Dinucleotide (NAD+).

This research will improve the accuracy of the prediction of substrates, inhibitors, and cofactors for enzymes while contributing data toward creating a model to predict the chemical reactions that different enzymes can perform. Some enzymes with the FMN cofactor were also characterized as having the potential to carry out reactions under the presence of light and further research will be conducted in this area.

Experimental Methods

Data Collection

- Using PDB, enzymes containing the FMN or NAD+ cofactor were researched to determine their substrate(s) and cofactor(s).

Data Analysis

- Information collected was transferred onto the FMN and NAD+ Excel Spreadsheets.
- A new database was created using Excel to document enzymes that were found to contain more than one cofactor.
- Statistical analysis was then performed to determine differences between the number of cofactors, substrates, and the types of cofactors for the enzymes with the FMN cofactor.

Graphical Representation

- After statistical analysis was performed, the data was converted into tables and pie charts through Google Sheets to provide a visual representation of the information collected.

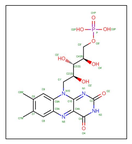

Figure 1: FMN Cofactor

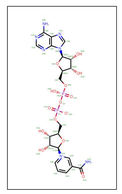

Figure 2: NAD+ Cofactor

Figure 3: 1UUM (example of an enzyme with the FMN cofactor)

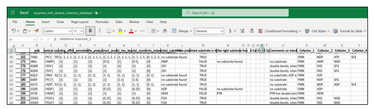

Figure 4: Enzymes with Multiple Cofactors Excel Spreadsheet

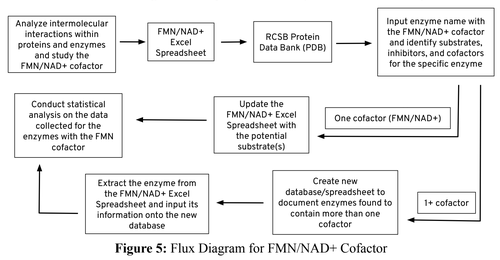

Figure 5: Flux Diagram for FMN/NAD+ Cofactor

Results

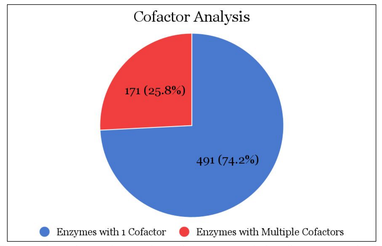

Figure 6: The pie chart shows the number of enzymes with 1 cofactor vs. multiple cofactors.

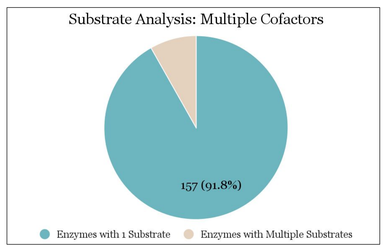

Figure 7: The pie chart shows the number of enzymes containing multiple cofactors with one substrate vs. multiple substrates.

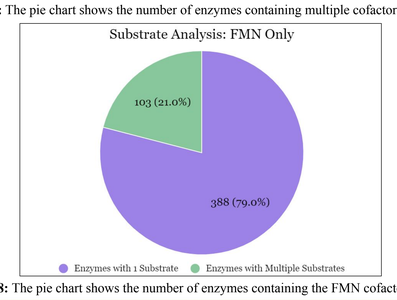

Figure 8: The pie chart shows the number of enzymes containing the FMN cofactor with one substrate vs. multiple substrates.

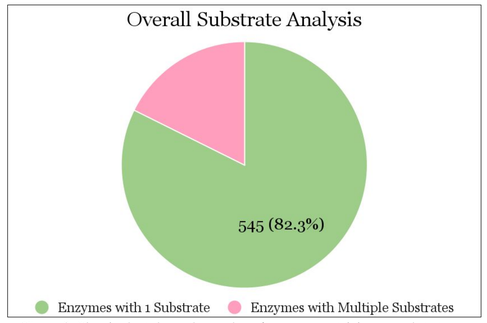

Figure 9: The pie chart shows the number of enzymes containing one substrate vs. multiple substrates.

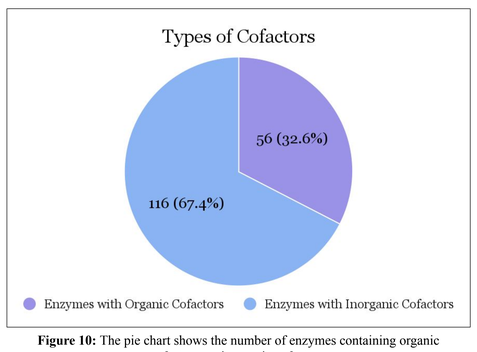

Figure 10: The pie chart shows the number of enzymes containing organic cofactors vs. inorganic cofactors.

Conclusion and Future Steps

Conclusion:

The analyzed data helps to improve the accuracy of the prediction of substrates, inhibitors, and cofactors for enzymes. Molecular interactions between substrates and the protein structure were also documented to foster a deeper understanding of the enzymes containing the FMN or NAD+ cofactor. For enzymes that were found to include more than one cofactor, a new database was created for further research regarding their intermolecular interactions and reactivity.

Broader Implications:

As future data is collected, these enzymes can be used in biological and industrial ways. The information published will be helpful for any future researchers as well as for students to understand the necessary conditions needed for enzymes, with specific cofactors, to carry out certain chemical reactions.

Future Steps:

The next steps would be to continue to look through the enzymes with the NAD+ cofactor and the other five cofactors to identify substrates and describe their intermolecular interactions. Steps will also be taken to research the chemical reactions that these enzymes can carry out.

References

- Cibulka, R., & Fraaije, M. W. (2021). Modes of flavin‐based catalysis. Flavin‐Based Catalysis, 97–124. https://doi.org/10.1002/9783527830138.ch4
- Nicholls, B. T., Oblinsky, D. G., Kurtoic, S. I., Grosheva, D., Ye, Y., Scholes, G. D., & Hyster, T. K. (2021). Engineering a non‐natural photoenzyme for Improved Photon Efficiency**. Angewandte Chemie, 134(2). https://doi.org/10.1002/ange.202113842
- Page, C. G., Cooper, S. J., DeHovitz, J. S., Oblinsky, D. G., Biegasiewicz, K. F., Antropow, A. H., Armbrust, K. W., Ellis, J. M., Hamann, L. G., Horn, E. J., Oberg, K. M., Scholes, G. D., & Hyster, T. K. (2020). Quaternary charge-transfer complex enables photoenzymatic intermolecular hydroalkylation of olefins. Journal of the American Chemical Society, 143(1), 97–102. https://doi.org/10.1021/jacs.0c11462
- Sandoval, B. A., Meichan, A. J., & Hyster, T. K. (2017). Enantioselective hydrogen atom transfer: Discovery of catalytic promiscuity in flavin-dependent 'ene'-reductases. Journal of the American Chemical Society, 139(33), 11313–11316. https://doi.org/10.1021/jacs.7b05468

Acknowledgements

The Dong Theoretical and Computational Chemistry Lab
Professor Sijia Dong, Assistant Professor
Dr. Gustavo Mondragón-Solórzano, Postdoctoral Scholar
Zheyu Zhang, PhD Candidate
Center for STEM Education
Claire Duggan, Executive Director
Jennifer Love, Associate Director
Claire Stipp & Yassine Souabny, YSP Coordinators
Nicolas Fuchs, Project Implementation Coordinator
Mary Howley. Administrative Officer

This work was supported by the National Science Foundation