Expression and Characterization of the Luciferase of *Noctiluca scintillans*

Sean P. O'Hare[1,], Steven O. Mansoorabadi[2]*

[1]Undergraduate Student, Department of Chemistry and Biochemistry, Auburn University
[2]Associate Professor, Department of Chemistry and Biochemistry, Auburn University

Dinoflagellates are an important group of microorganisms that are ubiquitous in the world's oceans. Many species can produce bioluminescent light, which occurs through the oxidation of dinoflagellate luciferin by the dinoflagellate luciferase (LCF), as depicted in Figure 1. LCF is an enzyme of great interest because it has the potential to be utilized as a reporter enzyme, a useful tool for detecting gene expression, studying infectious diseases, and monitoring cancer growth.

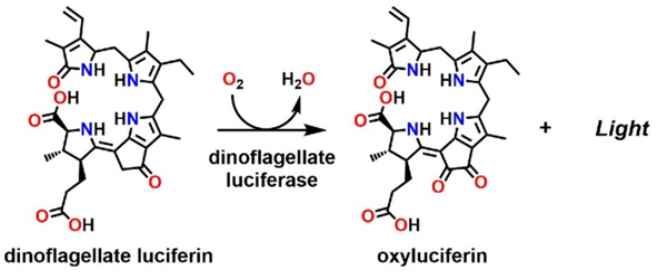


Fig. 1. Dinoflagellate bioluminescence reaction

We are specifically studying the LCF of *Noctiluca scintillans* (NS-LCF) due to its unique structure. While LCF typically contains three homologous domains, NS-LCF has two distinct domains (refer to structural comparison in Figure 2) (Liu & Hastings, 2007). As depicted in Figure 3, the first domain shares 65% sequence identity with domain 1 of the LCF found in the luminous dinoflagellate *Lingoludinium polyedra* (LP-LCF), and thus is labeled the LCF-like region (Liu & Hastings, 2007). The second domain, however, shares 40% sequence identity with the separately encoded luciferin-binding protein (LBP) of *L. polyedra* (LP-LBP), another essential component of the bioluminescence system, so it is denoted as the LBP-like region (Liu & Hastings, 2007). This distinctive structure of NS-LCF has led to the hypothesis that all other forms of LCF and LBP evolved from a NS-LCF-like progenitor (Liu & Hastings, 2007). Therefore, elucidating the structure and catalytic regions of NS-LCF and comparing them to other forms of LCF and LBP could explain how the functions of these proteins evolved over time.

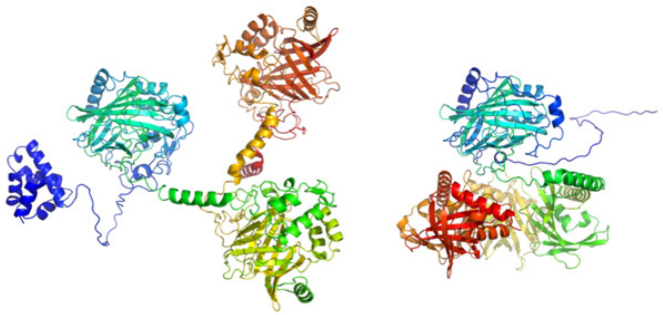


Fig. 2. Predicted structures of the canonical LCF from *L. polyedra* (left) and of NS-LCF (right)

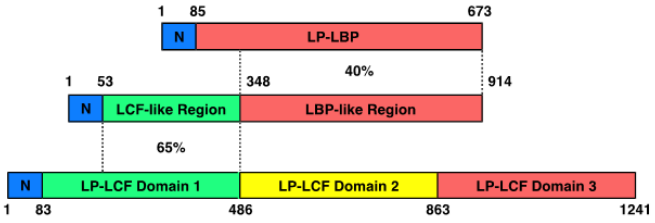


Fig. 3. NS-LCF domain structure in comparison to LP-LCF and LP-LBP

In this research, the focus was on expressing and purifying NS-LCF. To achieve this, a technique called heterologous expression was used. Heterologous expression involves inserting the gene of interest into a different host organism, in this case, *Echerichia coli*, which is a common bacterium used for protein expression (Kaur

*Corresponding author: spo0007@auburn.edu

et al., 2018). To accomplish this, a technique called polymerase chain reaction (PCR) was first used to amplify the NS-LCF gene. Then, a technique called Gibson Assembly was performed to insert the NS-LCF gene into the Champion™ pET-SUMO vector. A vector is a DNA molecule that acts as a carrier to introduce the target gene into the host organism (Nora et al., 2019). The pET-SUMO vector is designed to facilitate the expression of the target protein by adding a SUMO (small ubiquitin-like modifier) tag to the protein, which enhances protein solubility levels (Butt et al., 2005). After the construction of the plasmids containing the NS-LCF gene in the pET-SUMO vector, they were introduced into *E. coli* BL21 (DE3) cells for expression. BL21 (DE3) is a commonly used strain of *E. coli* that is known for its ability to produce large amounts of recombinant proteins (Robichon et al., 2011). Following expression, the SUMO-tagged NS-LCF protein was purified using a technique called immobilized metal affinity chromatography, which takes advantage of part of the SUMO tag to specifically capture and purify the NS-LCF protein from the mixture of proteins produced by the *E. coli* BL21 (DE3) cells (Falke & Corbin, 2013). Finally, the purified sample was analyzed using sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE), which is a technique that separates proteins based on their size using a gel and an electric field.

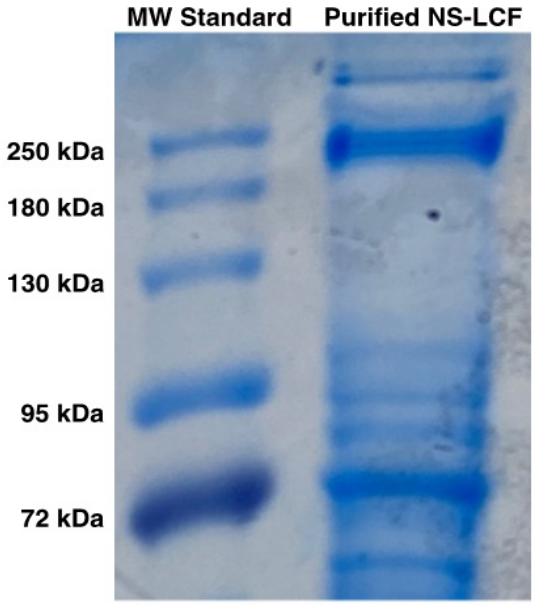


Fig. 4. SDS-PAGE analysis of IMAC-purified NS-LCF

SUMO-tagged NS-LCF has a molecular weight (MW) of 111 kDa, but looking at the SDS-PAGE gel (see Figure 4), no band corresponding to the expected molecular weight was observed. Instead, a prominent band was detected at the 72 kDa marker, which is likely indicative of the presence of DnaK, a common *E. coli* chaperone protein contaminant (Rial & Ceccarelli, 2002). This inference aligns with previous findings from the lab, although further confirmation will be sought using mass spectrometry. Additionally, a band was observed between the 180 and 250 kDa molecular weight markers, suggesting the possible presence of NS-LCF dimers, which would have a molecular weight of 222 kDa. This hypothesis is backed by data, as the structural prediction program AlphaFold indicates that the N-terminal region (front-most portion) of NS-LCF, which appears intrinsically disordered, may group together to form higher order oligomer complexes like dimers, trimers or dodecamers (see Figure 5). However, this assignment has not yet been confirmed.

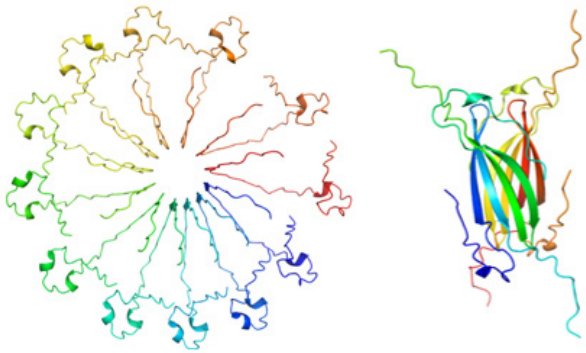


Fig. 5. Potential NS-LCF N-terminal trimer (right) and dodecamer (left) complexes

It is still unclear if NS-LCF has been successfully expressed and purified. It is possible that dimers have been isolated, but this needs to be further investigated. For future steps, we first plan to perform a western blot, an analytical technique used to identify specific proteins. This will allow us to validate that the protein we purified is NS-LCF. If successful, then we plan to run multiple different experiments to study NS-LCF. Structural analysis will first be conducted using X-ray crystallography to elucidate its three-dimensional structure. Mutational analysis will then be performed to determine the active site residues by selectively mutating specific amino acids and studying their functional impact. Finally, the pH dependence of the enzyme will be studied using fluorescence spectroscopy. Through

these investigations we hope to gain an understanding of the enzyme's mechanism of action, laying the foundation for further research into the enzyme's potential use as a reporter enzyme.

Statement of Research Advisor

Sean's undergraduate research project is focused on understanding the mechanism of dinoflagellate bioluminescence. Dinoflagellates are a diverse group of freshwater and marine microorganisms that are simultaneously responsible for harmful algal blooms (red tides) and the 'phosphorescence of the sea' via a unique bioluminescence system consisting of a luciferase enzyme (LCF), a luciferin substrate, and a luciferin-binding protein (LBP). Sean began volunteering in the lab and started a new project focusing on a primitive luciferase that combines the functionalities of LCF and LBP in one protein in collaboration with a senior graduate student. After receiving an Undergraduate Research Fellowship, Sean has continued to work on this project independently, as his graduate student mentor left for a position at a biotechnology startup company, and he has made significant progress that sets the stage for future research in the laboratory.

- Dr. Steven O. Mansoorabadi, Department of Chemistry and Biochemistry, College of Math and Science

References

[1]Butt, T. R., Edavettal, S. C., Hall, J. P., Mattern, M. R. SUMO Fusion Technology for Difficult-to-Express Proteins. Protein Expr. Purif. 43 (1), 1–9, (2005).

[2]Falke, J. J. and Corbin, J. A. Methods | Affinity Tags for Protein Purification. Encyclopedia of Biological Chemistry III, 701–705 (2013).

[3]Kaur, J., Kumar, A., Kaur, J. Strategies for Optimization of Heterologous Protein Expression in E. Coli: Roadblocks and Reinforcements. Int. J. Biol. Macromol., 106, 803–822 (2018).

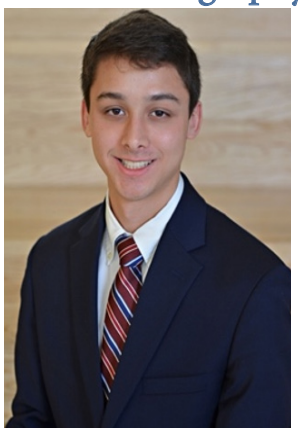
[4]Liu, L. and Hastings, J. W. Two Different Domains of the Luciferase Gene in the Heterotrophic Dinoflagellate Noctiluca scintillans Occur as Two Separate Genes in Photosynthetic Species. Proceedings of the National Academy of Sciences, 104 (3), 696–701, (2007).

[5]Nora, L. C., Westmann, C. A., Martins-Santana, L., Alves, L. de F., Monteiro, L. M. O., Guazzaroni, M.-E., Silva-Rocha, R. The Art of Vector Engineering: Towards the Construction of next-Generation Genetic Tools. Microb. Biotechnol, 12 (1), 125–147 (2019).

[6]Rial, D. V. and Ceccarelli, E. A. Removal of DnaK ontamination during Fusion Protein Purifications. Protein Expr. Purif., 25 (3), 503-507 (2002)

[7]Robichon, C., Luo, J., Causey, T. B., Benner, J. S.; Samuelson, J. C. Engineering Escherichia Coli BL21(DE3) Derivative Strains to Minimize E. Coli Protein Contamination after Purification by Immobilized Metal Affinity Chromatography. Appl. Environ. Microbiol., 77 (13), 4634–4646 (2011).

Authors Biography



Sean P. O'Hare is a junior-year student pursuing a B.S. degree in Biochemistry at Auburn University. As a 2022-2023 Undergraduate Research Fellow, he played a key role in the research design, data collection, and analysis.

Dr. Steven O. Mansoorabadi is the J. Milton Harris Associate Professor in the Department of Chemistry and Biochemistry at Auburn University. He oversaw all aspects of the fellowship project.