Engineering the Universal Donor: Lessons Learned from Attempting to Express GH109 Enzymes in *E.Coli*

Benjamin Zaidel, Ahmed Yousif, Lorenzo Del Rosario
Team 2, BIOE 44, Fall 2022
December 15, 2022

Results

The overarching goal of this BIOE 44 research project was to characterize and measure the expression of the enzyme GH109 in *E.Coli.* The GH109 class of enzymes possess catalytic activity that can cleave the A antigen off of red blood cells to transform them into the universal donor, blood type O. To express this enzyme *in vitro*, our research group ordered the pET-21(+) plasmid vector on TWIST biosciences, with an inserted region that encoded for GH109, tagged with Juniper GFP via a Gly4Ser peptide linker sequence. In terms of our experimental design process, we completed a bacterial transformation of pET-21_GH109-GFP into three cell lines of *E.coli* (BL21, NEBa high efficiency, and NEBa subcloning). Afterwards, we characterized its expression using fluorescence microscopy, as well as Western blot.

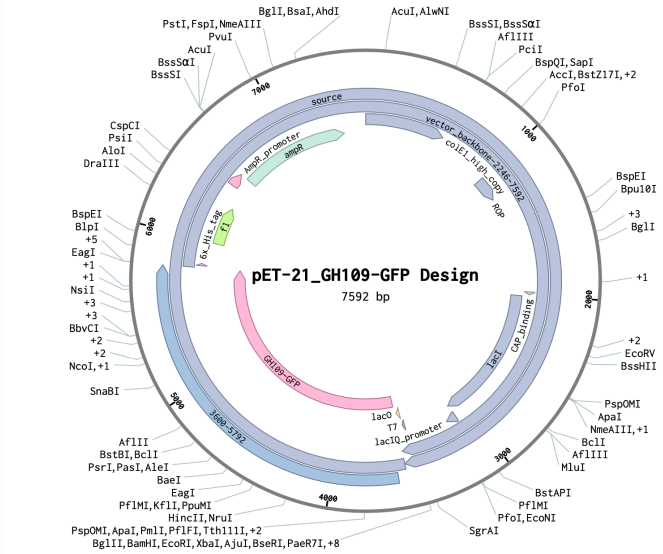


Figure 1: Image of pET-21_GH109-GFP, ordered from TWIST Biosciences.
This plasmid contains GH109-GFP (pink), ampR (turquoise), an origin of replication (ori), T7 promoter region, ribosome binding sites, and the lac operon. Additionally, it holds various restriction enzyme sites and its length is 7592 bp.

Part 1: Transformation

Transformations enable the uptake of DNA sequences into cells, including bacteria. Within the context of this project, our plasmid of interest, pET-21_GH109-GFP, was transferred into E.coli cells. The pET-21_GH109-GFP sample initially came in dry, but was resuspended in TE buffer to a final concentration of 50 ng/uL. To optimize the transformation, 75 ng of plasmid DNA was transferred into three different cell lines that we used (BL21, NEB-alpha high efficiency, and NEB-alpha subcloning) via heat shock. These transformed *E.coli* samples were added into SOC media and streaked on plates to be incubated overnight.

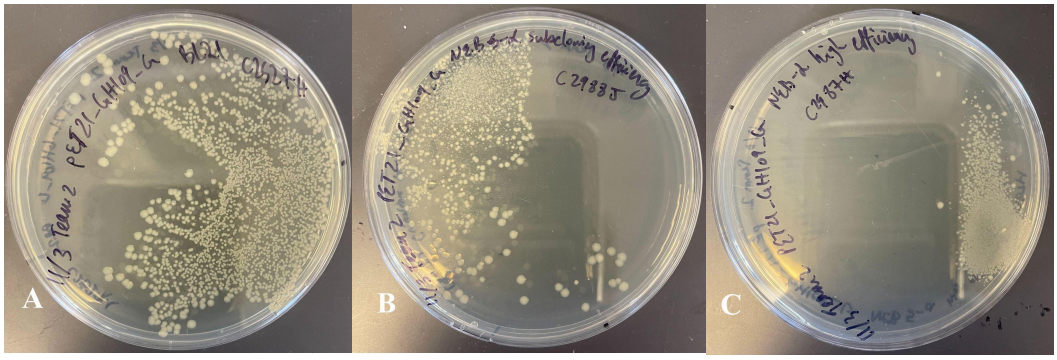


Figure 2: Transformation plates with the three different E.coli strains: Figure 2A shows BL21, 2B shows NEBa subcloning, and 2C shows NEBa high efficiency.

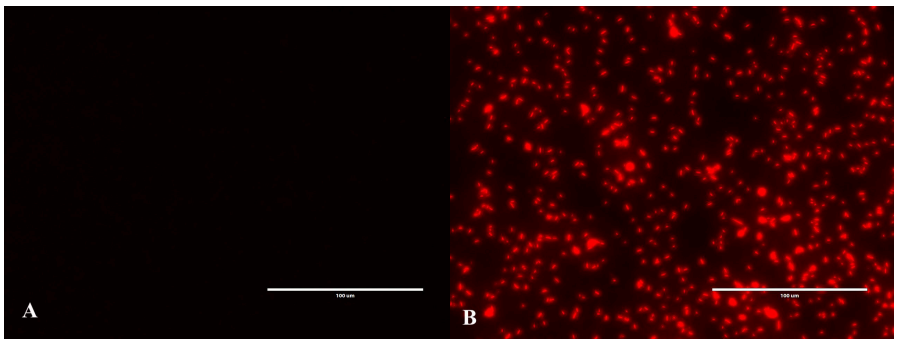
Based on Figure 2, all three samples demonstrated a fairly successful transformation, with hundreds of individual colonies being visible on each plate. To quantify the success of this experiment, we also calculated the transformation efficiencies of each plate, which are used to measure how well cells can uptake extracellular DNA during transformation. The BL21 sample ultimately resulted in the highest estimated transformation efficiency at 115.1 CFU/ng, with the NEB-alpha subcloning efficiency being slightly lower at 95.92 CFU/ng and the NEB-alpha high efficiency efficiency being an entire magnitude smaller at 15.4 CFU/ng.

Transformed E.Coli Sample	***Transformation Efficiency (CFU/ng)***
BL21	115.1
NEB-alpha subcloning	95.92
NEB-alpha high efficiency	15.4

Figure 3: Transformation efficiencies for each bacteria sample. CFU/ng was found by multiplying the number of transformants per ug of DNA by the final volume, divided by volume plated.

Part 2: Fluorescence Microscopy

Following transformation of the bacteria, our research team did fluorescence microscopy to assess whether or not GH109-GFP would be successfully expressed among the *E.coli* colonies. The team completed two rounds of fluorescence microscopy, and decided to place less weight on the interpretation of the first microscopy because the team forgot to include a positive control group. To test the expression of GH109-GFP, we inoculated two liquid cultures of the transformed BL21 bacteria overnight (for 22 hours), one with a 1mM concentration of IPTG and one with no IPTG at all. The BL21 cell line was chosen because it contains the T7 RNA polymerase, and our GH109-GFP insert is downstream from an IPTG-inducible T7 promoter region. The IPTG induction concentration was chosen to be 1mM because this concentration was known to cause full induction of the T7 promoter region. For our positive control, we used pColi_red because it was known to display large amounts of fluorescence after induction with IPTG.



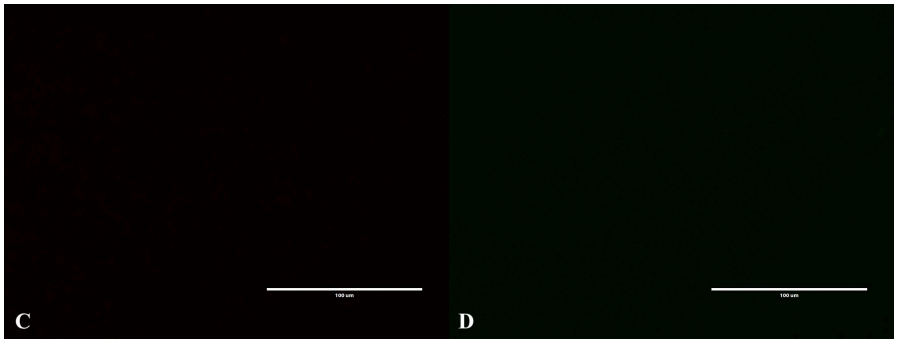


Figure 4: Results of fluorescence microscopy. The top two images represent the pColi_red samples, with Figure 4A showing the sample inoculated in 1mM of IPTG and Figure 4B without IPTG. The bottom two represent pET-21_GH109-GFP in green light, with Figure 4C being inoculated with 1mM IPTG Figure 4D and with no IPTG. All images were taken with the 100x objective lens.

Based on the results in Figure 4, it was observed that our protein of interest, GH109, was not being expressed by the BL21 bacterial cell line. While the pColi_red was successfully able to demonstrate fluorescence as expected, the transformed BL21 inoculated with 1mM IPTG overnight showed essentially no fluorescence, as did the sample that was inoculated without IPTG. With this in mind, we wanted to test whether or not the plasmid was actually present in the bacteria, as well as whether or not our protein was being expressed but there was an issue with the fluorescence component.

Part 3: Miniprep, Nanodrop, and Sequencing

Following our fluorescence assays, questions were raised regarding the purity of our plasmid strains. Did they contain our specified DNA at all? To test this, all three strains of our plasmid – high efficiency, subcloning, BL21 – were miniprepped to lyse the cells and extract only plasmid DNA from our sample. A subsequent run in a NanoDrop spectrophotometer yielded both a concentration and a 260/280 ratio of our sample (Figure 5), both of which fell within acceptable ranges[1]. Knowing that our plasmid DNA was both present and pure within our samples validated our previous transformations and fluorescence assays and helped the team rule out an explanation for a lack of fluorescence.

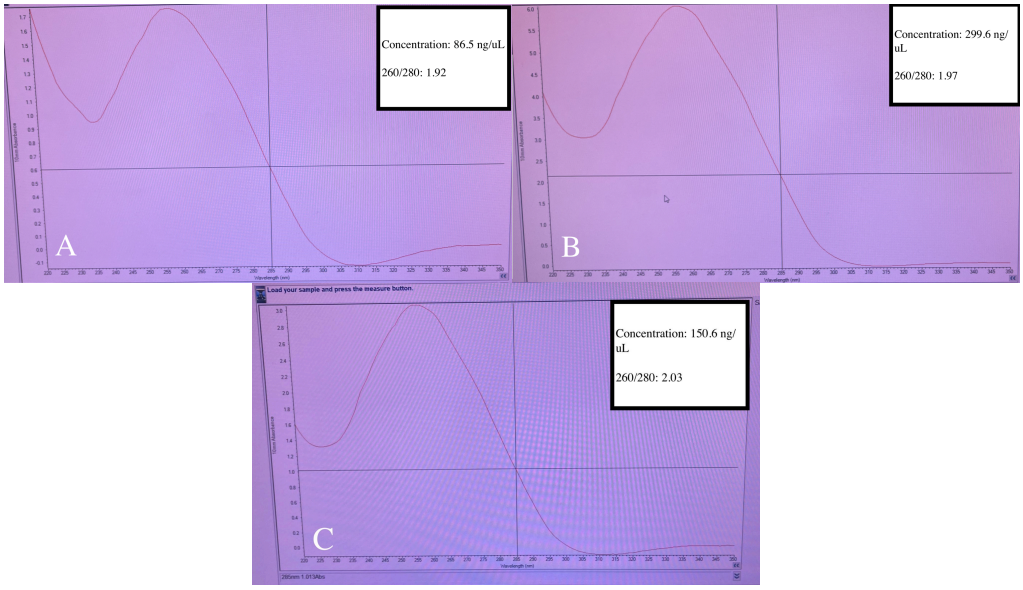


Figure 5: Results of testing three strains of pET-21_GH109-GFP via Nanodrop spectrophotometry. Figure 5A corresponds to the subcloning strain, 5B corresponds to the high efficiency strain, and 5C corresponds to the BL21 strain. Graph is plotting absorbance vs wavelength (nm). Tests also indicated both concentration and the 260/280 ratio, listed on the right of each figure.

Having established that our samples contain some pure concentration of plasmid DNA, the logical process that followed was validating whether or not said DNA matched the team's original design of the synthetically engineered pET-GH109-GFP plasmid. Our miniprepped samples were sent into Primordium, a sequencing company. Results indicated a 100% sequence alignment with the original plasmid sent to Twist Biosciences. This result demonstrated that the plasmid present inside of our BL21 cell line was the exact same plasmid as the one from TWIST biosciences that we completed a transformation with.

Part 4: Western Blot

The purpose of a Western blot is to separate a mixture of proteins based on their molecular weight, allowing the visualization of specific proteins of interest. Given the lack of conclusive results from previous fluorescence microscopy assays, Western blot served as a final method to see whether our protein, GH109-GFP, was being expressed by the BL21 strain of the plasmid. Gel electrophoresis separated the proteins, which were then transferred to a PVDF membrane using electropheoretic transfer; an anti-comet GFP antibody was used along with a secondary antibody to visualize the target proteins. It is difficult to draw definitive conclusions, but the faint band that exists at the molecular weight of GH109-GFP (~75 kD) indicates that our protein may be expressed by our plasmid. Moreover, a few darker bands can be observed at 20-25 kD in the BL21 IPTG- lane, indicating possible fragmentation of our target protein as this molecular weight range is far less than what we expected to observe. The absence of bands down the pColi Red ladder (negative control) and dark staining on the Rhamnose-inducible lysate (positive control) indicate that the Western blot was executed properly.

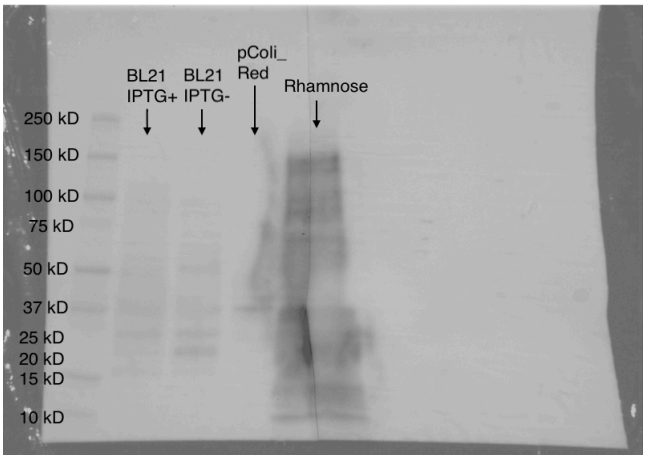


Figure 6: Western blot of BL21 bacterial samples. From left to right, the lanes contain a ladder, BL21 IPTG+, BL21 IPTG-, pColi_Red (negative control), and Rhamnose (positive control)

Discussion

It remains difficult to draw specific conclusions regarding the characterization of GH109 enzyme expression due to ambiguous results from our characterization assays. To start, we were unable to observe fluorescence upon conducting the GFP assay, despite verifying the transformation of our designed plasmid into the BL21 cell line using Primordium and sequence alignment on Benchling. This result may be caused by toxicity of our expressed GH109 enzyme and Juniper GFP, although this seems somewhat unlikely due to existing literature documenting the expression of both these proteins in the E. coli chassis we used. Another possibility is that our GFP was unstable, misfolded, or nonfunctional when expressed in BL21. Given that no literature exists on tracking GH109 expression using GFP, the robustness of the plasmid design we came up with in vivo is unclear, making it difficult to explain the lack of fluorescence. Similarly, the Western blot did not provide definitive evidence that GH109-GFP, with molecular weight ~80 kD, was expressed. Although a faint band can be seen at approximately 75 kD for BL21 IPTG+, the experimental group we expected to express our protein of interest, the faintness of the band makes the result inconclusive. Additionally, two significantly darker bands can be seen in the 20-25 kD range in the BL21 IPTG- lane, presumably due to the leakiness of our plasmid's T7 promoter, as no protein expression should be observed in the absence of IPTG activation. The lower weight range at which these bands can be observed may indicate misfolding or degradation of a portion of the expressed protein, so that our primary antibody could only bind a portion of the total protein. Our unclear Western blot results can also be attributed to the specificity of our antibody, where we were only able to obtain an anti-comet GFP antibody that did not perfectly match our Juniper GFP, which shares only 76% of comet GFP's sequence, as found by sequence alignment.

Making sense of mixed results elucidated above was one of the team's most pressing challenges over the course of our project. Despite being familiar with the innate complexity of biological systems and the difficulties associated with characterizing a synthetic system – two of the four challenges that Drew Endy labeled as the biggest obstacles to engineering biology – the team still sought out unifying theories that explained *why* what was going wrong was going wrong[2]. As such, a question arose: when our plasmid has this many moving biological parts, all interacting with each other in an uncountable number of ways, how is one supposed to isolate singular causes of error? Could our lack of fluorescence be a result of our general design? Could our GFP-linked protein have fallen off, or have simply been faulty?

The good news is that we had more tangible questions to seek answers to in the meantime. One such set of questions involved our protein's actual implementation into the human body, and the various endogenous factors that could complicate its use. An obvious first consideration was the other surface proteins on a red blood cell (RBC) beyond the A-antigen that we were attempting to cleave. The rhesus factor (Rh), for example, is one well-known RBC surface protein. This would limit our ability of conversion, breaking up A→O transfers to either A+→O+ or A-→O-. Our more recent research into RBC's and blood group systems revealed that ABO categorization and the Rh factor are not the only way to categorize blood. One review on the structure and function of RBC's listed 27 other blood grouping systems, all with their own antigens and component makeups[3], meaning that even a perfectly working GH109 protein might only make minor progress in making blood that is safe for cross-type transfusions. There could be other unmeasured factors that differentiate type A blood from the rest, which our team would consider if we moved the project forward.

Another overarching question we had is the extent to which research was previously done on synthetically engineering the universal donor. If this project was semi-feasible in a 10-week, biosafety level 1 (BSL-1) setting, what has been done at more advanced levels? One 2003 paper referenced the conversion of A and B antigens to H (equivalently O-type blood) via glycosides in a process somewhat adjacent to our own. But despite seeing successful clinical trials and transfusions of converted B antigen RBC's, the paper cites complications with type A blood due to the more complex nature of the antigens[4]. With this paper in hand, we questioned whether type A blood was the best choice as a starting point for our project. Perhaps another glycoside targeted for type B cleavage would have been a more pragmatic choice.

Given the results of our experiment and the subsequent questions that followed, our potential followup experiments would revolve mainly around repeated characterization assays, with the primary goal of figuring out if our protein of interest is being expressed at all. Our fluorescence assays, along with the Western blot, did not provide sufficient conclusive evidence in this regard. If feasible, we would consider taking time to redesign and reorder the plasmid from Twist Biosciences. Hopefully, our redesigned plasmid would come with multiple detection methods in addition to GFP, e.g a FLAG tag. As addressed previously, we would likely also consider using a more common, non IP-free GFP, e.g. standard comet GFP. Following characterization assays and hypothetical success in expressing our protein, the goal in the long term would be to move to blood testing, which presents a litany of barriers. For one, we would need to move to a BSL-2 space. However, testing for the in vitro success of our expressed protein would be easier; a simple antibody test would work. But once again, it's nearly impossible to control for the variety of different endogenous factors that complicate blood types without some kind of in vivo testing, which is likely beyond the scope of our project.

Despite the many hurdles our team had to overcome, this project still showed some promise with regard to the original goal that it tried to accomplish, which was to purify the GH109 enzyme and characterize its expression through various assays. The results may be inconclusive, but are not entirely dismissable, and thus the next best step would be to troubleshoot by trying new methods as mentioned before. As research on the topic of creating a universal donor expands, it will also be important to consider what the utility of an isolated GH109 sample would be. Amidst the complex biological interactions that occur in human blood, how would the implementation of this enzyme to convert A-type into O-type actually work? Could GH109 transformed bacteria simply be dropped into a sample of blood? After successful purification of GH109 in E.Coli is achieved, research would need to consider other aspects involved in successful blood transfusion before testing this method in animals and humans.

With the broader scope of this project in mind, there are also many ethical and practical considerations that would need to be made should the expression of GH109 and its subsequent implementation to convert A-type blood into O-type blood be successful. Our initial project proposal stated that low- and middle-income countries experience the largest public health impacts due to a lack of blood transfusion regulations[5]. How realistic would it be to actually carry out this solution to the countries that need it the most? Given that this relatively small experiment cost hundreds of dollars and ultimately produced ambiguous results, what would the cost of converting blood on a large scale look like? Although this research project has exciting implications for further research at the interface of biochemistry, cardiology and translational research, its potential impact and practicality are also hard to gauge given what we currently know. Regardless, understanding the process by which we can cleave antigens off of blood would contribute plenty to the basic science literature surrounding blood transfusions.

Author Contributions

The original inspiration for this project came from Lorenzo Del Rosario, who wrote his project proposal on the glycoside hydrolase family and its A-antigen cleaving properties. Most plasmid parts were compiled collectively; Lorenzo focused primarily on the GFP-linker sequence while Benjamin Zaidel and Ahmed Yousif focused on optimizing the codon sequence for GH109. The original transformation of our plasmid was conducted by all three team members. Fluorescence assays were executed mostly by Lorenzo, with calculational support and write-ups completed by Benjamin and Ahmed. Benjamin and Ahmed both handled the miniprep and nanodrop procedures while Lorenzo took these results and implemented the proper sequencing methods with Primordium. Finally, Ahmed made glycerol stocks of our solution for later use and Benjamin focused on the creation of a data sheet detailing the process behind our plasmid's creation. All members contributed equally to both the analysis of data and the writing of this paper.

Benjamin Zaidel: Supervision, Resources. **Ahmed Yousif**: Investigation, Project administration. **Lorenzo Del Rosario**: Conceptualization, Validation.

References

1. Interpreting nanodrop (spectrophotometric) results. *Interpreting Nanodrop (Spectrophotometric) Results*. (Accessed: 16th December 2022) http://www.u.arizona.edu/~gwatts/azcc/InterpretingSpec.pdf.
2. Endy, D. Foundations for engineering biology. 2005. Nature 438, 449–453. https://doi.org/10.1038/nature04342
3. Marion E Reid, Narla Mohandas. 2004. Red blood cell blood group antigens: structure and function. Seminars in Hematology, Volume 41, Issue 2, Pages 93-117. ISSN 0037-1963. https://doi.org/10.1053/j.seminhematol.2004.01.001.
4. Martin L Olsson, Cheryl A Hill, Humberto de la Vega, Qiyong P Liu et. al. 2004. Universal red blood cells—enzymatic conversion of blood group A and B antigens. Transfusion Clinique et Biologique, Volume 11, Issue 1, Pages 33-39. ISSN 1246-7820. https://doi.org/10.1016/j.tracli.2003.12.002.
5. Hillary E Jenny, et al. 2017. Access to safe blood in low-income and middle-income countries: lessons from India. BMJ Global Health. 2(2) http://dx.doi.org/10.1136/bmjgh-2016-000167