

**Abstract:**

GFP is a very popular protein used in the laboratory. Polymerase Chain Reaction was used to replicate the *gfp* genetic sequence and introduced into a plasmid to make pGLO. This plasmid was introduced into bacterial cells which were allowed to grow and in the process replicated the plasmid. The culture was grown on agarose plates exhibiting different conditions to determine whether or not the plasmid had been taken up. The plates containing arabinose caused the transformed cells to glow. It was predicted that the greater the amount of arabinose, the greater the number of glowing colonies. The few colonies that were successfully transformed were grown in mini cultures. These cells were lysated and then run through a Hydrophobic Interaction Column to separate the GFP protein from the other proteins present in the cell. A Western Blot was done to determine whether or not GFP was present and the blot did in fact exhibit a band the same size as the known size of the GFP protein (around 30kDaltons).

**Introduction:**

Since the late 1970's and early 1980's, scientists have been using a technique that is still widely used today called cell-based cloning. This consists of transforming a cell, usually a bacterial cell, by inserting a plasmid into it that contains a specific DNA sequence that codes for a gene that will give the cell a unique quality, such as antibiotic resistance. Scientists use vectors, which are circular pieces of DNA (also known as plasmids) to insert the foreign genetic sequence into a host organism (Piatelli & Holthaus, 2005).

A vector usually contains a selective marker that protects the host cell from antibiotics in order to differentiate transformed cells from untransformed cells. The vector

usually codes for another desired unique characteristic. A very common vector is pGLO, which contains the *gfp* gene. This gene produces the GFP protein which, in the presence of arabinose sugar, glows green under ultra violet light. This allows scientists to determine whether or not a bacterial cell has successfully taken up and replicated the desired vector.

In order to mass produce to GFP protein, a plasmid containing the *gfp* gene was inserted into *E. coli* bacterial cells. The plasmid replicates independently of the bacterial DNA. Since the pGLO plasmid replicates at a much faster rate than the DNA, this is called a “high copy number.” This permits scientists to produce many plasmids coding for the GFP protein from a single cell.

Even though the pGLO plasmid encodes for the *gfp* gene which makes the GFP protein, transformed colonies will only glow under ultra violet light in the presence of arabinose sugar. The ARAC inhibitor prevents the replication and translation of GFP. Arabinose binds to that inhibitor which causes a conformational change which causes the inhibitor to release the plasmid. This way RNA polymerase is free to bind at the origin and begin replication. It was predicted that the greater the quantity of arabinose added to the plates, the more colonies with GFP expression would grow on the plates and glow under ultra violet light.

### **Materials and Methods:**

A restriction digest was run in order to determine which of three different plasmids was pGLO. The plasmids present were pGLO, pBR322, and pUC19. The students used the NEB webcutter online by introducing the genetic sequence that codes for each of the unknown plasmids and then cutting it with various enzymes to determine

the size of the fragments the digestion would yield. The enzymes used were Hind III and EarI. In order to do the restriction, 5  $\mu\text{L}$  of the unknown DNA were added to a tube, along with 1  $\mu\text{L}$  of each enzyme, Hind III and EarI. Afterwards, 2  $\mu\text{L}$  of buffer were added and 11  $\mu\text{L}$  of distilled water were also added. After the restriction was done, the samples were run by Gel Electrophoresis in order to determine the different sized bands and which tube corresponded to which plasmid.

After determining which tube contained the pGLO plasmid, the contents of that same tube were used to run a polymerase chain reaction. First, the tube was incubated at  $94^{\circ}\text{C}$  for one minute. Then, it was incubated for 45 seconds with primer that coded for the second set of sequences on page 50 of the lab manual, and finally it was incubated at  $72^{\circ}\text{C}$  for 2 minutes with dNTPs. This was done with the other two tubes containing different plasmids as a control. A master mix was made using 1.5  $\mu\text{L}$  of pGLO, 7.5  $\mu\text{L}$  of PCR buffer, 1.5  $\mu\text{L}$  of dNTPs, 3  $\mu\text{L}$  of forward primer, 3  $\mu\text{L}$  of reverse primer, 1.5  $\mu\text{L}$  of Taq and 57  $\mu\text{L}$  of distilled water, This tube was run in the conditions mentioned earlier.

After the PCR was run that the transformed cells were grown, they were placed in agarose plates to grow colonies. Several conditions were varied in the plates. The concentration of arabanose utilized for the growth of these colonies was 5x. Two of the plates had arabanose, one of which had ampicillin and the other one did not. The other two plates had no arabanose and only one of those had ampicillin.

Afterwards, the colonies that glowed under ultra violet light were utilized to grow mini cultures. Two different concentrations exhibited colonies that glowed: 10x and 0.5x.

Different colonies were grown in a mini culture using LB broth and another tube containing sterile arabanose and were grown overnight in a 37°C incubator.

After this, the tube that glowed under ultraviolet light was used for the hydrophobic interaction column. The protocol can be found on page 89 and 90 of the laboratory manual. On step 17, however, 5 different tubes were utilized to collect the liquid approximately every 3 or 4 drops. A total of eight samples were collected.

A biuret assay was run using a spectrophotometer in order to determine the absorbance at different concentrations of BSA. A best fit curve was drawn from this information and the absorbencies of the fractions of the HIC were extrapolated against the curve to determine the total amount of protein present.

Two SDS-PAGE gels were run. One was transferred to a PVDF membrane, but only the GFP protein was transferred. The other one was stained with blue dye in order to determine the different sizes of the various proteins present in the cells. The protocol for the SDS-PAGE gel is found on page 69 of the lab manual. No practice lane was used and the molecular weight marker was loaded into lane 1. The protocol for the PVDF membrane can be found on pages 69-70. Immunoblotting was done to determine the presence of GFP. The protocol can be found on page 73 of the manual.

**Results:**

<b>Plasmid</b>	<b>Enzyme/s</b>	<b>Predicted length</b>	<b>Actual length</b>
pGLO	Hind III and Ear I	402, 510, 649, 1569, 2241	750, 1100, 1600, 1900, 2000, 2800, 3000, 9000
pBR322	Hind III and Ear I	229, 1804, 2328	2000, 2200, 2300
pUC19	Hind III and Ear I	151, 243, 488, 1804	600, 1900, 2200

**Table 1: DNA fragment size of digestion of varied plasmids with different enzymes**

It was predicted that the digestion of pGLO with Hind III and Ear I would yield five fragments of differing sizes. Ear I would cut at sites 1063, 2625, and 4865. Hind III would cut at 1465 and 2114. The fragment sizes that resulted from the digestion with both enzymes were 402, 510, 649, 1569, and 2241 base pairs.

From the digestion of pBR 322 with the same two enzymes it was predicted that HindIII would cut at on site and Ear I would cut at two. Hind III cut at site 30 while Ear I cut at sites 2358 and 4162. The three fragments formed were 2328, 2358 and 229 base pairs.

When pUC 19 was digested with the same two enzymes, it was predicted that Ear I would cut at three sites while Hind III would cut at only one. Ear I would cut at sites

297, 691, and 2494 while Hind II would only cut at 448. This would yield to four fragments 488, 151, 243, and 1804 base pairs long.

**Figure 1- Gel electrophoresis of the Restriction Enzyme Digest of pGLO, pUC19 and bPR322 with Hind III and Ear I**

**Figure 2- Gel electrophoresis of PCR of pGLO**

Actually, when the digested plasmids were run through gel electrophoresis it was observed that the sample containing pGLO actually contained 8 fragments of 9000, 3000, 2800, 2000, 1900, 1600, 1100, and 750 base pairs long. The pUC 19 plasmid yielded three fragments of 600, 1900, and 2200 base pairs long. The pBR 322 plasmid yielded three fragments 2000, 2200, 2300 base pairs long.

The PCR gel yielded one fragment of 650 base pairs for the control. The band for the cloned sample ran the same distance as the control, meaning the genes were the same size. However, the band of the cloned sample was much thicker signifying that several copies of the same gene were present. .

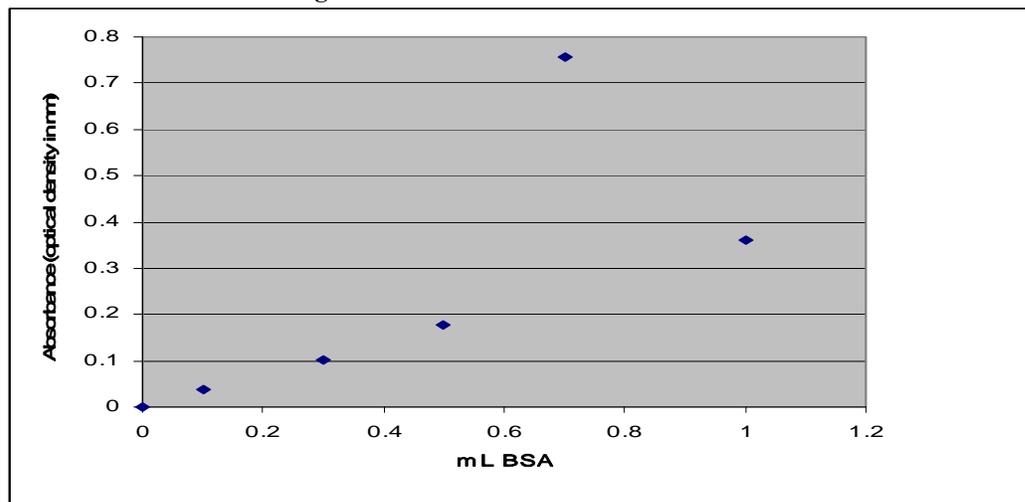
Media	(+)/ control	G 1 (5x)	G 2 (10x)	G 3 (10x)	G 4 (10x)	G 5 (0.5x)	G 6 (4x)	G 7 (5x)	G 8 (1x)	G 9 (5x)	G 10 (5x)
LB	(+)	-	lawn	lawn	lawn	lawn	lawn	lawn	lawn	lawn	lawn
	Control	-	lawn	lawn	lawn	lawn	lawn	lawn	lawn	lawn	lawn
LB/Arab	(+)	lawn	-	lawn	-	-	lawn	lawn	-	lawn	-
	Control	lawn	-	lawn	-	-	lawn	lawn	-	lawn	-
LB/amp	(+)	-	0	0	2	2	5	1	2	0	1
	Control	-	0	0	0	0	0	3	0	0	1
LB/Amp/Arab	(+)	2	0	-	1 glow	7 glow	1	1	0	0	0
	Control	0	0	-	0	0	0	0	0	0	1
LB/Amp/Arab II	(+)	-	0	0	(25x) 2	1	-	(10x) 0	(10x) 0	(10x) 0	-
	Control	-	0	0	0	0	-	0	0	0	-

**Table 2: Class Results for the growth of Bacterial Colonies in different mediums at different concentrations of arabinose**

In the bacterial growth of transformed bacteria it was seen in most groups that there was lawn growth on the plates containing LB or LB and arabinose alone. The majority of the controls showed no growth in the plates that had ampicillin except for group 7 that had three colonies in the control and group 10 that had one colony. Group 5 had the most growth of colonies in the plate containing LB, ampicillin and arabinose with 7 colonies. Only groups 4 and 5 had colonies that glowed in the LB/Amp/Arab plates. Since so little colonies grew, mini cultures were made from the colonies that glowed to make sure transformed cells were used and to see how fast these would reproduce. Since only a few colonies for groups 4 and 5 glowed under ultra violet light, those were used to make minicultures. Then, a hydrophobic interaction column was run to separate the GFP protein from other proteins that are present in *E. coli*.

Test Tube	mL BSA (5mg/mL)	Absorbance
1	0	0
2	0.1	0.039
3	0.3	0.101
4	0.5	0.177
5	0.7	0.756
6	1	0.361

**Table 3: Absorbance readings of different concentrations of BSA for standard curve**



**Graph 1: Standard Curve for Biuret Assay**

For the standard curve, there was no absorbance when no BSA was added. The absorbance was 0.039nm for 0.1mL of BSA, 0.101 nm for 0.3mL of BSA, 0.177 nm for 0.5mL, 0.756nm for 0.7mL, and 0.361 for 1mL. From the best fit curve of this graph it was possible to extract the amount of bacteria present from the HIC which allowed the students to separate the GFP protein from any other proteins present in the bacteria.

Unknown Sample	Abs. 1	Abs. 2	Avg.	mL of bacteria (5mg/mL)	g of bacteria
TCL control	0.006	-		0.0198	0.0975
TCL trans.	0.01	0.004	0.007	0.0226	0.113
wash control	-0.003	-		-0.0051	-
wash trans.	0.016	0.014	0.015	0.0447	0.224
GFP TE trans.	0	-0.004	-0.002	-0.0023	-
GFP TE trans. 2	0	-0.003	-0.015	-0.0382	-

**Table 4: Optical density(nm) of biuret Assay under different conditions (10  $\mu$ L of sample)**

For the standard curve, there was no absorbance when no BSA was added. The absorbance was 0.039nm for 0.1mL of BSA, 0.101 nm for 0.3mL of BSA, 0.177 nm for 0.5mL, 0.756nm for 0.7mL, and 0.361 for 1mL. From the best fit curve of this graph it was possible to extract the amount of protein present from the HIC which allowed the students to separate the GFP protein from any other proteins present in the bacteria.

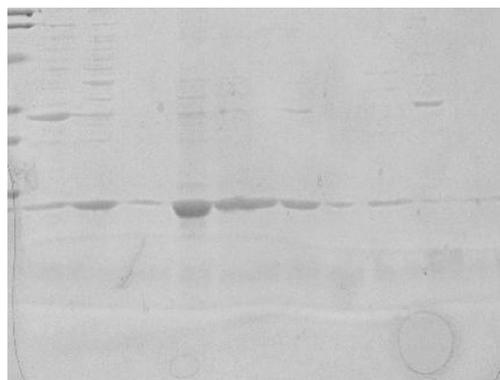
For the biuret the average absorbance of the TCL control was 0.006, 0.007 for the TCL transformed, -0.003 for the control wash, 0.015 for the transformed wash, -0.002 for the GFP TE transformed and -0.015 for the second run of GFP plus tris-edta buffer transformed. When compared to the standard curve it was seen that in the transformed run, the TCL had 0.0226 mL of bacteria (0.113g), the wash had 0.0447mL of bacteria (0.224 g). The GFP with the TE buffer had negative absorbencies on both runs and no bacteria present. The control had 0.0198 mL of bacteria in the TCL (0.0975g), and a negative absorbance for the wash (therefore there were no bacteria). From these samples,

two SDS-PAGE gels were run to determine how much protein was present, one exclusively for GFP and the other one for every protein.



**Fig. 3- Western blot gel**

An image of the SDS-PAGE gel can be seen above. On each side of the gel can be seen a ladder run, containing various proteins with sizes, from top to bottom of the image, of approximately 104, 81, 48, 36, 27, and 19 kDa. The first well to the right of the double ladder was the first wash with TE. The second well was the TCL with arabanose induced GFP. The fourth well was the wash with the arabanose. The fourth column was the arabanose elute. The fifth column was the TCL for the control and the sixth column was the elution fraction of the control. In the third well between the ladders is shown the band corresponding to the GFP protein in blue. It appeared to be around a size of 30 kDa.



**Fig 4. – Simply Blue Stain gel**

Two different groups used the same gel to run the simply blue stain portion of the experiment. To the far left is the ladder (half can be seen). The first well following the

ladder was first wash of the portion containing GFP. The second well after the ladder was the total crude lysate (TCL). On both runs there are bands of various different sizes, but in the first well the most notorious bands are around 2000 (the bottom one) and 4800 base pairs. The TCL shows various bands as well, but the most obvious one is around 2000 base pairs. The dark circle on the right, bottom corner is probably an air bubble.

### **Discussion:**

The goal of this experiment was to determine an effective way of creating many copies of the *gfp* plasmid by introducing it into foreign, bacterial cells. Then, the transformed cells were broken open and the various protein within were sorted in order to isolate GFP.

The results for the restriction enzyme experiment were inconclusive due to the fact that the enzymes used to digest the three different plasmids cut them into fragments of similar sizes. The students guided the experiment based on the number of fragments that the digestion would yield instead of the sizes. In order to obtain better results it is more appropriate to utilize enzymes that will cut the plasmids into fragments of very different sizes. This way, when the gel is run it will be obvious which fragments belong to which plasmid and it will be easier to recognize them. Also, the band from the blue tube (which contained the pGLO plasmid) had too many bands of DNA and some of the bands were extremely big, even bigger than the plasmid itself. This could have been due to different shapes of plasmids, such as coiled and super coiled plasmids.

The PCR gel showed that the correct segment of the DNA sequence was cloned, that which encodes for *gfp*. This can be seen because the band formed is almost 700 base pairs long, and this is the known length of the plasmid. Also, the thick band shows that

there are many more fragments of the same size for the cloned sample, rendering the PCR effective. The band for the control was also this length, but it was much thinner than that of the replicated band. There should have been no band present run in the control because no DNA was introduced to that culture at all. Perhaps the sample used had a few *gfp* plasmids present before the experiment was run..

Once the sample containing the pGLO is identified and amplified through PCR, it can be transformed into *E. coli* cells that will take in the plasmids and replicate with the plasmids. The bacteria can be utilized to grow colonies in agarose plates. Arabinose sugar can be added to the plates in order to identify if pGLO is present because any colonies that grow in the presence of arabinose that successfully took in the pGLO plasmid will produce GFP which is a protein that glows under UV light. It was predicted that colonies possessing cells containing the desired plasmid would glow more when arabinose was present at higher concentrations. The plates containing ampicillin and LB, but no arabinose exhibited no glowing colonies for the transformed sample and no growth at all for the control.

In order for the *E. coli* cell to take up the plasmids they must be submitted to cell heat shock in which various steps are followed in order to make the membrane more permeable to the plasmids. First,  $MgCl_2$  is added. The  $Mg^{2+}$  neutralizes the negative membrane making it easier for it to allow the negative DNA to trespass. Also, the cell takes up the  $Cl^-$  which swells up the cell and makes the membrane porous so the DNA can enter. Also, the cells are submitted to drastic temperature changes so the membrane becomes even more porous and then they are left to recuperate for about an hour.

Once the cells have absorbed the plasmid, this specimen is taken and cultured on various agarose plates with different variables. One plate contained only agarose and the essential nutrients required for cell growth in order to tell if the bacteria were not completely killed because of the shock treatment. Then, some of the sample is added to a plate with ampicillin on which only the colonies that contain the plasmid will grow because part of the plasmid codes for the resistance of this antibiotic. A third plate is used with arabinose and ampicillin on it, and if the colonies that grow on it glow under UV light, this indicates the cells are producing GFP protein and contain the pGLO plasmid.

For the class results it can be seen that there was uncontrolled, lawn growth in the plates that contained luria broth (LB) only or LB and arabanose, as it was expected, because these are the ideal conditions for the growth of bacteria. In the plates that contained ampicillin there was very limited growth because it was expected that not all the bacteria would be transformed, so only the ones that took up the plasmid with the *gfp* gene grew into colonies. However, those bacterium that did take up the plasmid became ampicillin resistant and that is why they grew on the plates containing the antibiotic. Two groups had growth for their control, untransformed bacteria on the plates that had ampicillin. This probably means that there were bacteria that were ampicillin resistant before the transformation or the plates were contaminated with transformed bacteria before the experiment was run. The LB/ampicillin plates with quantities of arabanose at 4x and 0.5x exhibited the most growth because they were the ones that contained the most colonies. The colonies that glowed under UV light, and therefore expressed the GFP protein, were at 10x and 0.5x. It was concluded that the best arabanose concentration for the growth and expression of GFP protein in bacterial cells was between 4x and 5x

because these concentrations exhibited the most growth in the presence of ampicillin. This is because the plasmid containing the *gfp* gene also contained a gene that exhibits ampicillin resistance. Since only two groups had colonies that glowed, these were used by all the groups to make the mini cultures for the Hydrophobic Interaction Chromatography (HIC) experiment because these colonies were the only ones expressing the GFP protein, even though there were many colonies that had the plasmid containing the gene because they were ampicillin resistant but did not glow under UV light.

The purpose of the HIC experiment was to isolate the GFP protein present in the *E. Coli* bacteria and separate it from all the other proteins in the cells. The cells were lysed open and were passed through a column with beads that interacted with hydrophobic proteins, such as GFP. After the mini cultures were passed through the column, the ones that glowed the most were the second and third wash with the TE buffer. This was because this caused the hydrophobic interaction between the GFP protein and the beads in the column to cease. These fractions were used for the Biuret assay in order to determine how much protein was actually present in the sample.

A standard curve was made by taking the absorbance reading of different amounts of bovine serum albumin (BSA). When protein is added to a certain amount of BSA, it changes its absorbance and then it can be compared to this curve in order to determine how much protein there is present. Some of the protein containing biurets displayed negative absorbances even though all the samples used in this part of the experiment glowed under UV light, proving that there was GFP protein present. This was probably because the Biuret assay is not sensitive enough to detect low concentrations of protein. Since the BSA was present in 5mg/ml of sample, the actual amount of protein was

calculated by multiplying by 5 the volume corresponding to the absorbance reading from the standard curve. The sample that contained the most protein was that of the total crude lysate (TCL) transformed bacteria, as it was expected. This is because the first run through the column caused the hydrophobic interaction between the beads and the GFP protein only, but all the other proteins present in the cell ran through. Then a first wash was used to wash off any hydrophobic proteins other than GFP that bound to the beads. This sample had 0.224 g of protein which makes sense because the cells have many other proteins other than GFP. Finally, a TE buffer was used to prevent the interaction of the GFP protein with the beads and the collected sample had negative absorbencies probably because there was such minute quantities of protein present.

The wash with the TE buffer was kept and used to make a western blot in order to determine if it was indeed GFP protein that was collected in this sample. This was run through gel electrophoresis along with a known sample of the protein. The western blot showed that which is between the values known to correspond to the actual size of the GFP protein (27-30 kDa). The band containing the GFP protein appeared to be around 30 kDa in size. Thus, SDS-PAGE was successfully determined the size and validated the identity of the GFP protein.

The Simply Blue Stain gell whow various bands of different sizes. The most notorious bands for the first lane were almost 50kDalton and 20kDaltons. This means that the most dominant proteins present in the first wash were lysozymes and ovalbumin. There were also trace amounts of BSA, which was around 80kDaltons, carbonic anhydrase, which are around 40kDaltons. The second column, which was the elution buffer after the GFP was collected, showed very trace amounts of almost very protein

except for a band at the 20 kDalton mark which accounts for a considerable amount of lysozymes. Perhaps this was because since they are so small, they easily bind to the beads and perhaps they have the opposite charge of the beads.

**Conclusion:**

Through various techniques the students were able to amplify a specific gene that coded for a desired protein, in this case, GFP. Even though the polymerase chain reaction seems to give good results for the transformed bacteria, very few colonies expressed the GFP protein, demonstrating that this process is extremely delicate and complicated and many factors can contribute to experimental error, like contamination of the samples by touching them and introducing foreign DNA. When the minicultures were grown, they did not glow under UV light which caused the students to believe that no transformed cells had grown in the culture. However, when the HIC column was run, the samples clearly glowed under UV light. The collected GFP protein could be used for further experimentation by using it to tag other proteins present in cells. The polymerase chain reaction can be utilized to make numerous copies of DNA sequences that code for different proteins. Perhaps these amplified proteins can be administered to people who have diseases due to the production of faulty proteins because of incorrect translation of DNA during development.

