Green Fluorescent Protein (GFP) Purification

Student Manual

"Bioengineered DNA was, weight for weight, the most valuable material in the world. A single microscopic bacterium, too small to see with the human eye, but containing the gene for a heart attack enzyme, streptokinase, or for "ice-minus" which prevented frost damage to crops, might be worth 5 billion dollars to the right buyer."

Michael Crichton - Jurassic Park

Contents

Lesson 1 Finding the Green Fluorescent Molecule

Genetic Transformation Review

In Bio-Rad Kit 1, you performed a genetic transformation of *E. coli* bacterial cells. The results of this procedure were colonies of cells that fluoresced when exposed to ultraviolet light. This is not a normal phenotype (characteristic) for *E.coli*. You were then asked to figure out a way to determine which molecule was becoming fluorescent under UV light. After determining that the pGLO plasmid DNA was not responsible for the fluorescence under the UV light, you concluded that it was not the plasmid DNA that was fluorescing in response to the ultraviolet light within the cells. This then led to the next hypothesis that if it is not the DNA fluorescing when exposed to the UV light, then it must be a protein that the new DNA produces within the cells.

1. Proteins.
   a. What is a protein?

   b. List three examples of proteins found in your body.

   c. Explain the relationship between genes and proteins.

2. Using your own words, describe cloning.

3. Describe how the bacterial cloned cells on your LB/amp plate differ from the cells on your LB/amp/ara plate. Can you design an experiment to show that both plates of cloned cells behave similarly and do contain the same DNA?

4. Describe how you might recover the cancer-curing protein from the bacterial cells.

Laboratory Procedure for Lesson 2

Picking Colonies and Growing a Cell Culture

Examine your two transformation plates under the ultraviolet (UV) lamp. On the LB/amp plate pick out a single colony of bacteria that is well separated from all the other colonies on the plate. Use a magic marker to circle it on the bottom of the plate. Do the same for a single green colony on the LB/amp/ara plate. Theoretically both white and green colonies were transformed with the pGLO plasmid? How can you tell?

Both colonies should contain the gene for the Green Fluorescent Protein. To find out, you will place each of the two different bacterial colonies (clones) into two different culure tubes and let them grow and multiply overnight.

Your Task

In this lab, you will pick one white colony from your LB/amp plate and one green colony from your LB/amp/ara plate for propagation in separate liquid cultures. Since it is hypothesized that the cells contain the Green Fluorescent Protein, and it is this protein we want to produce and purify, your first consideration might involve thinking of how to produce a large number of cells that produce GFP.

You will be provided with two tubes of liquid nutrient broth into which you will place cloned cells that have been transformed with the pGLO plasmid.

Workstation Daily Inventory Check (✔) List

Your Workstation. Materials and supplies that should be present at your student workstation site prior to beginning this lab activity are listed below.

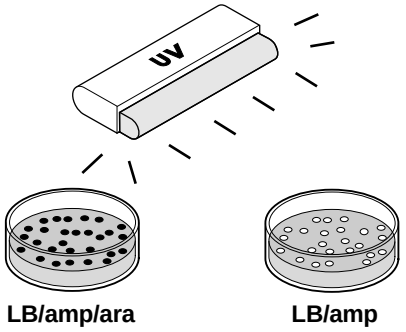
Instructors (Common) Workstation. Materials, supplies, and equipment that should be present at a common location that can be accessed by your group during each lab activity are also listed below.

Your workstation	**Number**	**(✔)**
Transformation plates from Kit 1 (LB/amp/ara and LB/amp)	2	❑
Inoculation loops	2	❑
Culture tubes, containing 2 ml growth media	2	❑
Marking pen	1	❑
Test tube holder	1	❑
Instructors workstation		
Shaking incubator or platform (optional)	1	❑
UV light	1	❑

Laboratory Procedure for Lesson 2

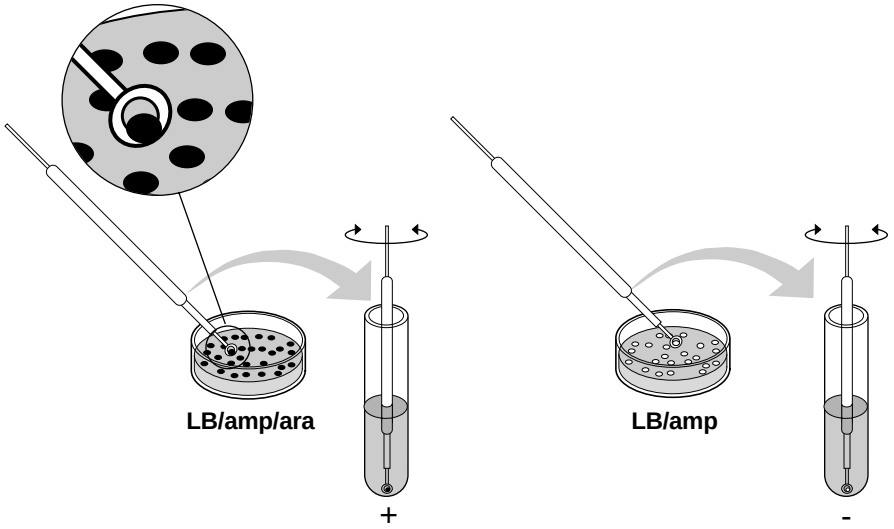
1. Examine your LB/amp and LB/amp/ara plates from the transformation lab. First use normal room lighting, then use an ultraviolet light in a darkened area of your laboratory. Note your observations.

   To prevent damage to your skin or eyes, avoid exposure to the UV light. Never look directly into the UV lamp. Wear safety glasses whenever possible.

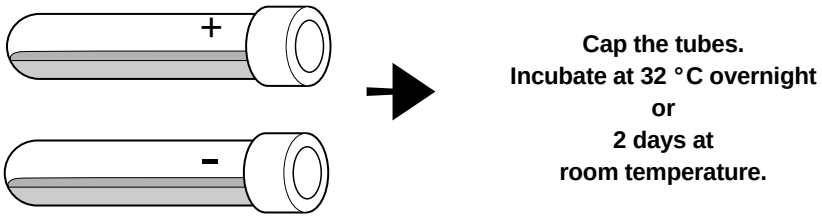


2. Identify several green colonies that are not touching other colonies on the LB/amp/ara plate. Turn the plate over and circle several of these green colonies. On the other LB/amp plate identify and circle several white colonies that are also well isolated from other colonies on the plate.

3. Obtain two 15 milliliter culture tubes containing 2 milliliters of nutrient growth media and label one tube "+" and one tube "-". Using a sterile inoculation loop, lightly touch the "loop" end to a circled single green colony and scoop up the cells without grabbing big chunks of agar. Immerse the loop in the "+" tube. Spin the loop between your index finger and thumb to disperse the entire colony. Using a new sterile loop, repeat for a single white colony and immerse it in the "-" tube. It is very important to pick cells from a single bacterial colony.



4. Cap your tubes and place them in the shaker or incubator. Let the tubes incubate for 24 hours at 32 °C or for 2 days at room temperature. If a shaker is not available, shake your two tubes vigorously, like you would shake a can of spray paint, for about 30 seconds. Then place them in an incubator oven for 24 hours. Lay the tubes down horizontally in the incubator. (If a rocking table is available, but no incubator, tape the tubes to the platform and let them rock at maximum speed for 24 hours at 32° or at room temperature for 48 hours. We do not recommend room temperature incubation without rocking or shaking.)



Culture Condition	Days Required
32 °C—shaking	1 day
32 °C—no shaking	1–2 days*
Room Temperature—shaking	2 days
Room Temperature—no shaking	Not recommended

* Periodically shake by hand and lay tubes horizontally in incubator.

Lesson 2 Name_____

Review Questions

1. What is a bacterial colony?

2. Why did you pick one green colony and one white colony from your agar plate(s)? Why do you think you picked one of each color? What could this prove?

3. How are these items helpful in this cloning experiment?
   a. ultraviolet (UV) light -

   b. incubator -

   c. shaking incubator -

4. Explain how placing cloned cells in nutrient broth to multiply relates to your overall goal of purifying the fluorescent protein.

Lesson 3

Purification Phase 1
Bacterial Concentration and Lysis

So far you have mass produced living cultures of two cloned bacterium. Both contain the gene which produces the green fluorescent protein. Now it is time to extract the green protein from its bacterial host. Since it is the bacterial cells that contain the green protein, we first need to think about how to collect a large number of these bacterial cells.

A good way to concentrate a large number of cells is to place a tube containing the liquid cell culture into a centrifuge and spin it. As you spin the cell culture, where would you expect the cells to concentrate, in the liquid portion or at the bottom of the tube in a pellet?

Workstations Check (✔) List

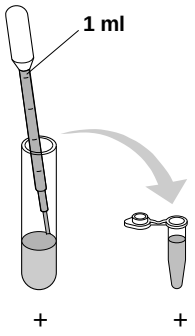
Your Workstation. Make sure the correct materials listed below are present at your workstation prior to beginning this lab experiment.

Instructors (Common) Workstation. Materials that should be present at a common location to be accessed by your group are also listed below.

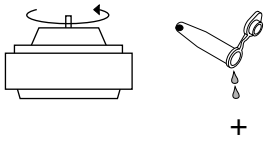
Your workstation	**Number**	**(✔)**
Microtubes	1	❑
Pipette	1	❑
Microtube rack	1	❑
Marker	1	❑
Beaker of water for rinsing pipettes	1	❑
Instructors workstation		
TE solution	1 vial	❑
Lysozyme (rehydrated)	1 vial	❑
Centrifuge	1	❑
UV light	1–4	❑

Laboratory Procedure for Lesson 3

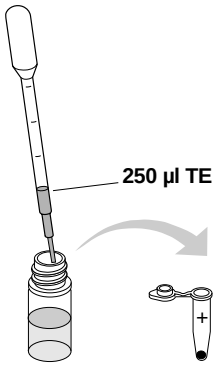
1. Using a marker, label one new microtube with your name and period.
2. Remove your two liquid cultures from the shaker or incubator and observe them in normal room lighting and then with the UV light. Note any color differences that you observe. Using a clean pipette, transfer the entire contents of the (+) liquid culture into the 2 milliliter microtube also labeled (+), then cap it. You may now set aside your (-) culture for disposal.



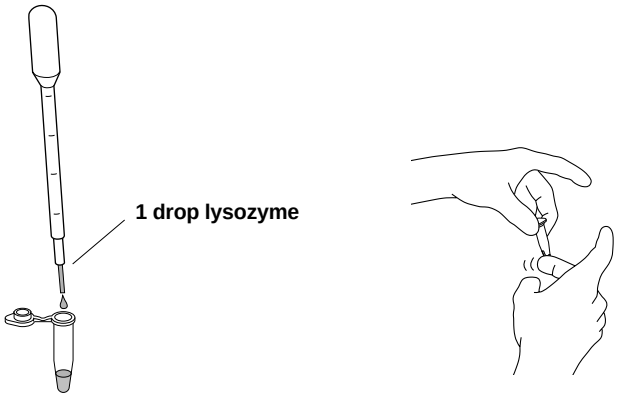
3. Spin the (+) microtube for 5 minutes in the centrifuge at maximum speed. Be sure to balance the tubes in the machine. If you do not know how to balance the tubes, do not operate the centrifuge.
4. After the bacterial liquid culture has been centrifuged, open the tube and slowly pour off the liquid supernatant above the pellet. After the supernatant has been discarded, there should be a large bacterial pellet remaining in the tube.



5. Observe the pellet under UV light. Note your observations.
6. Using a new pipette, add 250 µl of TE Solution to each tube. Resuspend the bacterial pellet thoroughly by rapidly pipetting up and down several times with the pipette.



7. Using a rinsed pipette, add 1 drop of lysozyme to the resuspended bacterial pellet. Cap and mix the contents by flicking the tube with your index finger. The lysozyme will start digesting the bacterial cell wall. Observe the tube under the UV light. Place the microtube in the freezer until the next laboratory period. The freezing will cause the bacteria to explode and rupture completely.



Lesson 3 Name _____

Review Questions

1. You have used a bacterium to propagate a gene that produces a green fluorescent protein. Identify the function of these items you need in Lesson 3.
   a. Centrifuge -
   b. Lysozyme -
   c. Freezer -

2. Can you explain why both liquid cultures fluoresce green?

3. Why did you discard the supernatant in this part of the protein purification procedure?

4. Can you explain why the bacterial cells' outer membrane ruptures when the cells are frozen. What happens to an unopened soft drink when it freezes?

5. What was the purpose of rupturing or lysing the bacteria?

Lesson 4

Purification Phase 2
Bacterial Lysis

The bacterial lysate that you generated in the last lab contains a mixture of GFP and endogenous bacterial proteins. Your goal is to separate and purify GFP from these other contaminating bacterial proteins. Proteins are long chains of amino acids, some of which are very hydrophobic or "water-hating". GFP has many patches of hydrophobic amino acids, which collectively make the entire protein hydrophobic. Moreover, GFP is much more hydrophobic than most of the other bacterial proteins. We can take advantage of the hydrophobic properties of GFP to purify it from the other, less hydrophobic (more hydrophilic or "water-loving") bacterial proteins.

Chromatography is a powerful method for separating proteins and other molecules in complex mixtures and is commonly used in biotechnology to purify genetically engineered proteins. In chromatography, a column is filled with microscopic sperical beads. A mixture of proteins in a solution passes through the column by moving downward through the spaces between the beads.

You will be using a column filled with beads that have been made very hydrophobic—the exact technique is called hydrophobic interaction chromatography (HIC). When the lysate is applied to the column, the hydrophobic proteins that are applied to the column in a high salt buffer will stick to the beads while all other proteins in the mixture will pass through. When the salt is decreased, the hydrophobic proteins will no longer stick to the beads and will drip out the bottom of the column in a purified form.

Workstations Check (✔) List

Student Workstations. Make sure the materials listed below are present at your workstation prior to beginning this lab experiment.

Instructors (Common) Workstation. Materials that should be present at a common location to be accessed by your group are also listed below.

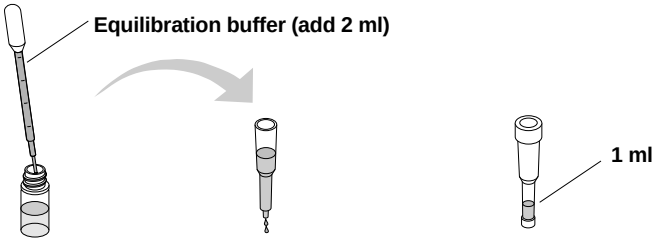
Student workstation items	**Number**	**(✔)**
Microtubes	1	❑
Pipette	1	❑
Microtube rack	1	❑
Marker	1	❑
Beaker of water for rinsing pipettes	1	❑
HIC chromatography column	1	❑
Column end cap	1	❑
Waste beaker or tube	1	❑
Instructors workstation items		
Binding Buffer	1 vial	❑
Equilibration Buffer	1 vial	❑
Centrifuge	1	❑
UV light	1–4	❑

Laboratory Procedure for Lesson 4

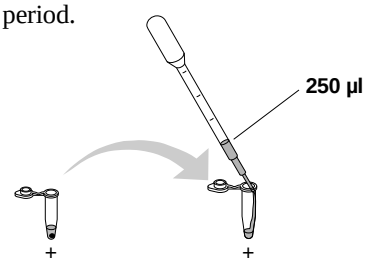
1. Remove your microtube from the freezer and thaw it out using hand warmth. Place the tube in the centrifuge and pellet the insoluble bacterial debris by spinning for 10 minutes at maximum speed. Label a new microtube with your team's initials.

2. While you are waiting for the centrifuge, prepare the chromatography column. Before performing the chromatography, shake the column vigorously to resuspend the beads. Then shake the column down one final time, like a thermometer, to bring the beads to the bottom. Tapping the column on the table-top will also help settle the beads at the bottom. Remove the top cap and snap off the tab bottom of the chromatography column. Allow all of the liquid buffer to drain from the column (this will take ~3–5 minutes).

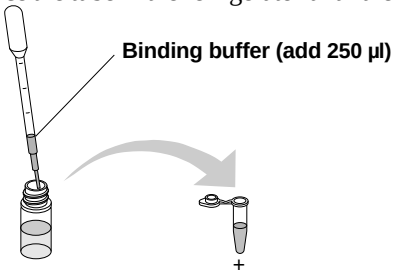
3. Prepare the column by adding 2 milliliters of Equilibration Buffer to the top of the column, 1 milliliter at a time using a well rinsed pipette. Drain the buffer from the column until it reaches the 1 milliliter mark which is just above the top of the white column bed. Cap the top and bottom of the column and store the column at room temperature until the next laboratory period.



4. After the 10 minute centrifugation, immediately remove the microtube from the centrifuge. Examine the tube with the UV light. The bacterial debris should be visible as a pellet at the bottom of the tube. The liquid that is present above the pellet is called the supernatant. Note the color of the pellet and the supernatant. Using a new pipette, transfer 250 µl of the supernatant into the new microtube. Again, rinse the pipette well for the rest of the steps of this lab period.



5. Using the well-rinsed pipette, transfer 250 µl of Binding Buffer to the microtube containing the supernatant. Place the tube in the refrigerator until the next laboratory period.



Lesson 4 Name _____

Review Questions

1. What color was the pellet in this step of the experiment? What color was the supernatant? What does this tell you?

2. Why did you discard the pellet in this part of the protein purification procedure?

3. Briefly describe hydrophobic interaction chromatography and identify its purpose in this lab.

Lesson 5

Purification Phase 3
Protein Chromatography

In this final step of purifying the Green Fluorescent Protein, the bacterial lysate you prepared will be loaded onto a hydrophobic interaction column (HIC). Remember that GFP contains an abundance of hydrophobic amino acids making this protein much more hydrophobic than most other bacterial proteins. In the first step, you will pass the supernatant containing the bacterial proteins and GFP over an HIC column in a highly salty buffer. The salt causes the three-dimensional structure of proteins to actually change so that the hydrophobic regions of the protein move to the exterior of the protein and the hydrophilic ("water-loving") regions move to the interior of the protein.

The chromatography column at your workstation contains a matrix of microscopic hydrophobic beads. When your sample is loaded onto this matrix in very salty buffer, the hydrophobic proteins should stick to the beads. The more hydrophobic the proteins, the tighter they will stick. The more hydrophilic the proteins, the less they will stick. As the salt concentration is decreased, the three-dimensional structure of proteins change again so that the hydrophobic regions of the proteins move back into the interior and the hydrophilic ("water-loving") regions move to the exterior.

You will use these four solutions to complete the chromatography:

Equilibration Buffer—A high salt buffer (2 M $(NH_4)_2SO_4$)

Binding Buffer—A very high salt buffer (4 M $(NH_4)_2SO_4$)

Wash Buffer—A medium salt buffer (1.3 M $(NH_4)_2SO_4$)

Elution Buffer—A very low salt buffer (10 mM Tris/EDTA)

Workstation Check (✔) List

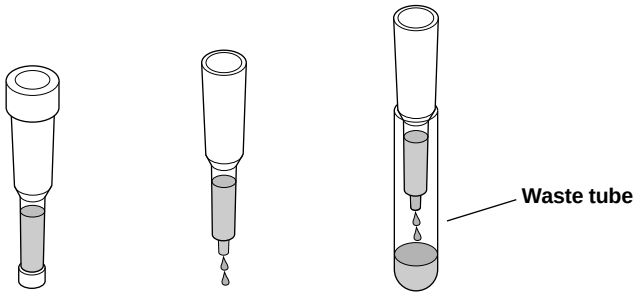
Your Workstation. Make sure the materials listed below are present at your workstation prior to beginning this lab experiment.

Instructors (Common) Workstation. Materials that should be present at a common location to be accessed by your group are also listed below.

Your workstation	**Number**	**(✔)**
Collection tubes	3	❑
Pipette	1	❑
Microtube rack	1	❑
Marker	1	❑
Beaker of water for rinsing pipettes	1	❑
HIC chromatography column	1	❑
Column end cap	1	❑
Test tube or beaker to collect waste	1	❑
Instructors workstation		
Wash Buffer	1 vial	❑
Equilibration Buffer	1 vial	❑
TE Buffer	1 vial	❑
UV light	1–4	❑

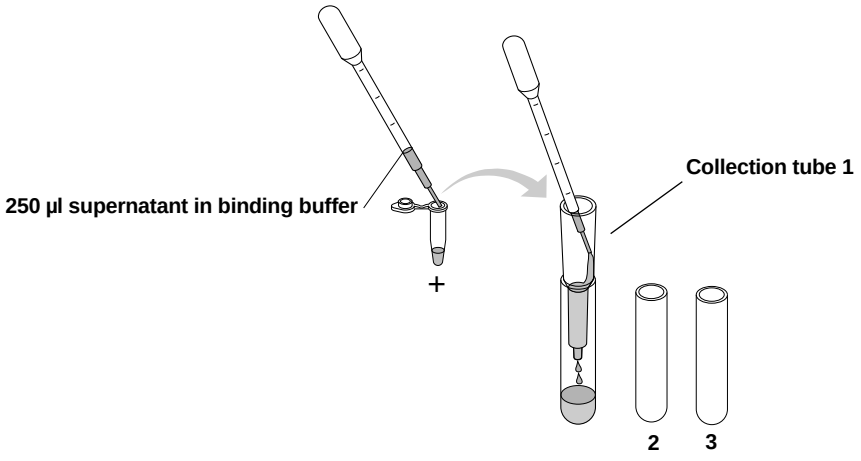
Lesson 5 Laboratory Procedure

1. Obtain 3 collection tubes and label them 1, 2, and 3. Place the tubes in a rack. Remove the cap from the top and bottom of the column and let it drain completely into a liquid waste container (an extra test tube will work well). When the last of the buffer has reached the surface of the HIC column bed, gently place the column on collection tube 1. Do not force the column tightly into the collection tubes—the column will not drip.

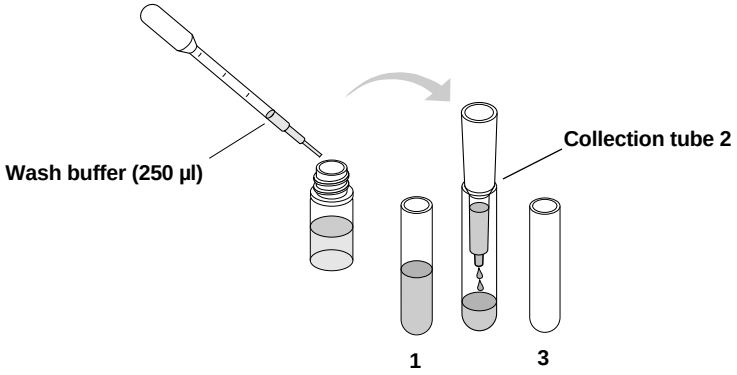


2. Predict what you think will happen for the following steps and write it along with your actual observations in the data table on page 42.

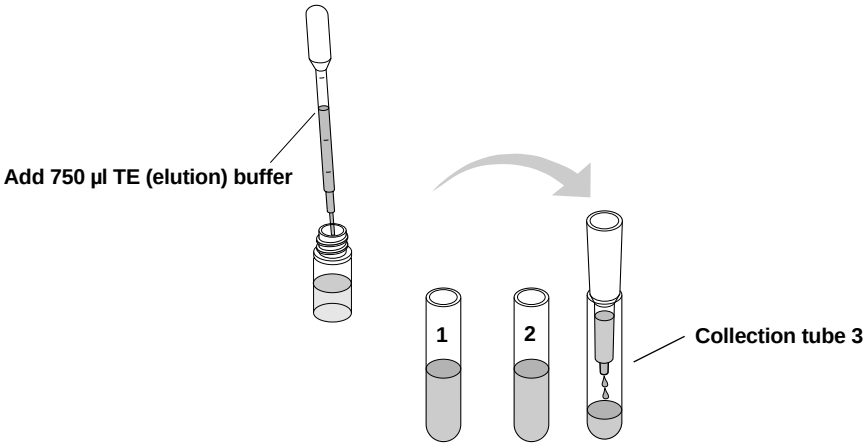
3. Using a new pipette, carefully load 250 µl of the supernatant (in Binding Buffer) into the top of the column by resting the pipette tip against the side of the column and letting the supernatant drip down the side of the column wall. Examine the column using the UV light. Note your observations in the data table. Let the entire volume of supernatant flow into tube 1.



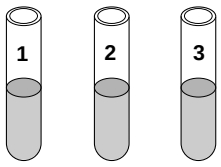
4. Transfer the column to collection tube 2. Using the rinsed pipette and the same loading technique described above, add 250 µl of Wash Buffer and let the entire volume flow into the column. As you wait, predict the results you might see with this buffer. Examine the column using the UV light and list your results on page 42.



5. Transfer the column to tube 3. Using the rinsed pipette, add 750 µl of TE buffer (Elution Buffer) and let the entire volume flow into the column. Again, make a prediction and then examine the column using the UV light. List the results in the data table on page 42.



6. Examine all of the collection tubes using the UV lamp and note any differences in color between the tubes. Parafilm or Saran Wrap the tubes and place in the refrigerator until the next laboratory period.



Lesson 5 Name _____

Review Questions

1. List your predictions and observations for the sample and what happens to the sample when the following buffers are added to the HIC column.

Collection Tube Number	Prediction	**Observations Under UV Light** (column and collection tube)
Tube 1 Sample in Binding Buffer		
Tube 2 Sample with Wash Buffer		
Tube 3 Sample with Elution Buffer		

2. Using the data table above, compare how your predictions matched up with your observations for each buffer.

   a. Binding Buffer-

   b. Wash Buffer-

   c. Elution Buffer-

3. Based on your results, explain the roles or functions of these buffers. Hint: how does the name of the buffer relate to its function.

   a. Equilibration Buffer-

   b. Binding Buffer-

   c. Wash Buffer-

   d. TE (Elution) Buffer-

4. Which buffers have the highest salt content and which have the least? How can you tell?

5. Were you successful in isolating and purifying GFP from the cloned bacterial cells? Identify the evidence you have to support your answer.

Appendix A Glossary of Terms

Agar — Provides a solid matrix to support bacterial growth. Contains nutrient mixture of carbohydrates, amino acids, nucleotides, salts, and vitamins.

Antibiotic Selection — The plasmid used to move the genes into the bacteria also contain the gene for beta-lactamase which provides resistance to the antibiotic ampicillin. The beta-lactamase protein is produced and secreted by bacteria which contain the plasmid. The secreted beta-lactamase inactivates the ampicillin present in the LB/agar, which allows for bacterial growth. Only bacteria which contain the plasmids, and express beta-lactamase can survive on the plates which contain ampicillin. Only a very small percentage of the cells take up the plasmid DNA and are transformed. Non-transformed cells, cells that do not contain the plasmid, can not grow on the ampicillin selection plates.

Arabinose — A carbohydrate, normally used as source of food by bacteria.

Bacterial Library — A collection of *E. coli* that has been transformed with recombinant plasmid vectors carrying DNA inserts from a single species.

Bacterial Lysate — Material released from inside a lysed bacterial cell. Includes proteins, nucleic acids, and all other internal cytoplasmic constituents.

Beta-Lactamase — Beta-lactamase is a protein which provides resistance to the antibiotic ampicillin. The beta-lactamase protein is produced and secreted by bacteria which have been transformed with a plasmid containing the gene for beta-lactamase (*bla*). The secreted beta-lactamase inactivates the ampicillin present in the growth medium, which allows for bacterial growth and expression of newly acquired genes also contained on the plasmid *i.e.* GFP.

Biotechnology — Applying biology in the real world by the specific manipulation of living organisms, especially at the genetic level, to produce potentially beneficial products.

Chromatography — A process for separating complex liquid mixtures of proteins or other molecules by passing a liquid mixture over a column containing a solid matrix. The properties of the matrix can be tailored to allow for the selective separation of one kind of molecule from another. Properties include solubility, molecular size, and charge.

Cloning

When a population of cells is prepared by growth from a single cell, all the cells in the population will be genetically identical. Such a population is called "clonal". The process of creating a clonal population is called "cloning". Identical copies of a specific DNA sequence, or gene, can be accomplished following mitotic division of a transformed host cell.

Colony

A clump of genetically identical bacterial cells growing on an agar plate. Because all the cells in a single colony are genetically identical, they are called clones.

Centrifugation

Spinning a mixture at very high speed to separate heavy and light particles. In this case, centrifugation results in a "pellet" found at the bottom of the tube, and a liquid "supernatant" that resides above the pellet.

Culture Media

The liquid and solid media are referred to as LB (named after Luria-Bertani) broth and agar are made from an extract of yeast and an enzymatic digest of meat byproducts which provides a mixture of carbohydrates, amino acids, nucleotides, salts, and vitamins, all of which are nutrients for bacterial growth. Agar, which is from seaweed, polymerizes when heated to form a solid gel (very analogous to Jell-O), and functions to provide a solid support on which to culture the bacteria.

DNA Library

When DNA is extracted from a given cell type, it can be cut into pieces and the pieces can be cloned en masse into a population of plasmids. This process produces a population of hybrid *i.e.* recombinant DNAs. After introducing these hybrids back into cells, each transformed cell will have received and propagated one unique hybrid. Every hybrid will contain the same vector DNA but a different "insert" DNA.

If there are 1,000 different DNA molecules in the original mixture, 1,000 different hybrids will be formed; 1,000 different transformant cells will be recovered, each carrying one of the original 1,000 pieces of genetic information. Such a collection is called a DNA library. If the original extract came from human cells, the library is a human library.

Individual DNAs of interest can be fished out of such a library by screening the library with an appropriate probe.

Genetic Engineering

The manipulation of an organism's genetic material (DNA) by introducing or eliminating specific genes.

Gene Regulation	Gene expression in all organisms is carefully regulated to allow for differing conditions and to prevent wasteful overproduction of unneeded proteins. The genes involved in the transport and breakdown of food are good examples of highly regulated genes. For example, the simple sugar, arabinose, can be used as a source of energy and carbon by bacteria. The bacterial enzymes that are needed to breakdown or digest arabinose for food are not expressed in the absence of arabinose but are expressed when arabinose is present in the environment. In other words when arabinose is around the genes for these digestive enzymes are turned on. When arabinose runs out these genes are turned back off. See Appendix D for a more detailed explanation of the role that arabinose plays in the regulation and expression of the green fluorescent protein gene.
Green Fluorescent Protein	Green Fluorescent Protein (GFP) was originally isolated from the bioluminescent jellyfish, *Aequorea victoria*. The gene for GFP has recently been cloned. The unique three-dimensional conformation of GFP causes it to resonate when exposed to ultraviolet and give off energy in the form of visible green light.
Lysozyme	Enzyme needed to lyse, or break open bacteria cell walls. The enzyme occurs naturally in human tears, acting as a bactericidal agent to help prevent bacterial eye infections. Lysozyme gets its name from its ability to lyse bacteria.
Pellet	In centrifugation, the heavier particles such as bacteria or the cellular membranes and other debris of lysed bacteria are found at the bottom of a microfuge tube in a pellet.
Plasmid	A circular DNA molecule, capable of autonomous replication, carrying one or more genes for antibiotic resistance proteins.
pGLO	Plasmid containing the GFP sequence and ampicillin resistance gene which codes for Beta-lactamase.
Recombinant DNA Technology	The process of cutting and recombining DNA fragments as a means to isolate genes or to alter their structure and function.
Screening	Process of identifying wanted bacteria from a bacterial library.
Sterile Technique	Minimizing the possibility of outside bacterial contamination during an experiment through observance of cleanliness and using careful laboratory techniques. See Appendix C.
Streaking	Process of passing an inoculating loop with bacteria on it across an agar plate.

Supernatant	Liquid containing cellular debris that are lighter than the debris in the pellet formed after centrifugation.
Vector	An autonomously replicating DNA molecule into which foreign DNA fragments are inserted and then propagated in a host cell (*i.e.* plasmid).

Appendix B

Basic Molecular Biology Concepts and Terminology

A study of the living world reveals that all living organisms organize themselves in some unique fashion. A detailed blueprint of this organization is passed on to offspring.

Cells are the smallest functional units capable of independent reproduction. Many bacteria, for instance, can survive as single cells. The chemical molecules within each cell are organized to perform in concert.

Cells can be grown in culture and harvested

Cells can be gathered from their natural locations and grown inside laboratory containers. Appropriate food and environment must be provided for the cells to grow. Bacteria and yeast are very easy to grow in culture. Cells taken from plants, insects, and animals can also be grown, but are more difficult to care for.

After growth is complete, cells in culture can be harvested and studied.

Cloning

When a population of cells is prepared by growth from a single cell, all the cells in the population will be genetically identical. Such a population is called "clonal". The process of creating a clonal population is called "cloning". The purpose of streaking bacteria on agar is to generate single colonies, each arising from a single cell.

Looking inside cells

The molecules inside a cell each perform a given function. For instance, DNA molecules store information (like the hard drive in a computer). Proteins are the workhorses of the cell.

To study these molecules we prepare a clonal population from a cell type of interest, break open the cells, and sort the contents. For instance, it is fairly easy to separate all the proteins from all the DNA molecules.

Purifying a single species of protein out of the mixture of proteins found inside a cell type is also possible. Each type of protein has unique physical and chemical properties. These properties allow the separation of protein species *i.e.* based on size, charge, or hydrophobicity, for instance.

Special molecules, specialized functions

We will take a close look at three very special kinds of molecules found inside cells: DNA, RNA and proteins. Each of these molecules performs a different function. DNA molecules are like file cabinets in which information is stored. RNA helps to retrieve and execute the instructions which are stored in DNA. Proteins are designed to perform chemical chores inside (and often outside) the cell.

DNA: The universal template for biological information

The master "script" for each organism is encoded within its deoxyribonucleic acid (DNA). The information within the DNA molecule/s of each cell is sufficient to initiate every function that cell will perform.

DNA molecules are very long chains composed of repeating subunits. Each subunit ("nucleotide") contains one of four possible bases protruding from its side:

ADENINE ("A")	CYTOSINE ("C")
THYMINE ("T")	GUANINE ("G")

Since nucleotides are joined head-to-tail, a long strand of DNA essentially consists of a chemical backbone with bases protruding along its side. The information carried by this molecule is encoded in the sequence of As, Gs, Cs and Ts along its length.

Some further points to note about DNA structure:

1. Because the subunits of DNA chains are joined head-to-tail, the sequence is directional *i.e.* "AACTG" is **not** the same as "GTCAA". By convention, we write DNA sequence from the free 5' end of the backbone and work our way towards the other free end (3').

   i.e. 5'...AACTG...3'

2. The protruding bases along the chain are free form spontaneous bonds with available bases on other DNA strands according to the following rules:

   (i) "A" pairs with "T"
   (ii) "C" pairs with "G"

   Because of these rules, "A" and "T" are said to be complementary bases; "G" and "C" are also complementary.

   (iii) For two DNA strands to pair up, they must be complementary and run in opposite directions.

   i.e. (5'...AGGTC...3') can pair with (5'...GACCT...3'). These two strands have complementary sequences. The double-stranded pair is written as follows:

   5'...AGGTC...3'
   3'...TCCAG...5'

   The above molecule contains five base pairs. Indeed, in nature, DNA almost always occurs in double-stranded form, the two strands containing complementary sequences.

3. DNA molecules are typically thousands, sometimes millions of base pairs long!! Sometimes the two ends of a DNA molecule are joined to form circular DNA.

4. Double-stranded DNA, in its native form, occurs as a coiled spring, or helix. Because it is two-stranded, it is often referred to as a double helix.

The architecture of DNA allows for a very simple strategy during reproduction: The two strands of each DNA molecule unwind and "unzip"; then, each strand allows a new complementary copy of itself to be made by an enzyme called "DNA polymerase". This results in two daughter molecules, each double-stranded, and each identical to the parent molecule.

Proteins and RNA are the workhorses of the cell

The biochemistry of life requires hundreds of very specific and efficient chemical interactions, all happening simultaneously. The major players in these interactions are short-lived protein and RNA molecules which can work together or independently to serve a variety of functions. Like DNA, RNA and proteins are also long chains of repeating units.

RNA

RNA (ribonucleic acid), like DNA, consists of four types of building blocks strung together in a chain. It differs from DNA in the following respects:

The four bases in RNA are “A”, “G”, “C” and “U” (uracil); the pairing rules are the same as for DNA except that “A” pairs with “U”. Although RNA can pair with complementary RNA or DNA, in cells RNA is usually single-stranded. The sugar in the RNA backbone is ribose, not deoxyribose. RNA molecules are generally short, compared to DNA molecules; this is because each RNA is itself a copy of a short segment from a DNA molecule. The process of copying segments of DNA into RNA is called transcription, and is performed by a protein called RNA polymerase.

Proteins

Proteins (more precisely, polypeptides) are also long, chain-like molecules but more structurally diverse than either DNA or RNA. This is because the subunits of proteins called amino acids, come in twenty different types. The exact sequence of amino acids along a polypeptide chain determines how that chain will “fold” into a compact structure. The precise 3-dimensional features of this structure, in turn, determine its function.

What a protein will **do** depends on the exact **sequence** of its amino acids.

In most cases, a protein will perform a single function. Very diverse functions can be performed by proteins: Some proteins, called “enzymes”, act as catalysts in chemical reactions; some carry signals from one part of a cell to another - or, in the case of “hormones”, from one cell to another; some proteins (“antibodies”) have the task of fighting intruders; many become integral parts of the various physical structures inside cells; and still others (“regulatory proteins”) police various activities within cells so as to keep them within legal limits.

Linear code, three-dimensional consequences

DNA is the primary depot for information in living systems. As mentioned, this information is linear *i.e.* encoded in the sequence of “A”, “G”, “C”, “T” building blocks along the DNA molecule. This linear code can be passed on to offspring — because DNA can make exact copies of itself.

Short segments of each DNA molecule are chosen for transcription at any given time. These segments are called “genes”. The enzyme RNA polymerase copies the entire segment, base by base, assembling an RNA molecule which contains a sequence of “A”, “G”, “C” and “U” exactly complementary to the DNA sequence of the transcribed gene.

In addition to providing a master template for copying RNAs, DNA also contains sequence information which tells the RNA polymerase where to start transcribing a gene (“promoter”) and where to stop; how many copies it should make and when; and it can even embed certain information within the RNA sequence to determine the longevity and productivity of that RNA!

There are three major classes of RNAs copied off DNA templates: messenger RNAs, or mRNAs, which relay the sequence information required for assembling proteins; transfer RNAs, or tRNAs, which work in the assembly line for proteins; and RNAs which perform structural functions. For example, ribosomal RNAs, or rRNAs, help build the scaffolding for ribosomes, the factories where proteins are assembled.

mRNAs carry the sequence information for making proteins. Ribosomes read this sequence of nucleotides, by a process called “translation” into a sequence of amino acids. How is this accomplished? There are only four kinds of nucleotides, but twenty kinds of amino acids!

During translation, the ribosome reads 3 nucleotides at a time and assigns an amino acid to each successive “triplet”. [NOTE: Triplets are often referred to as “codons”.] Each amino acid is then attached to the end of the growing protein chain. There are 64 possible triplets, or codons. Thus, the linear information residing in DNA is used to assemble a linear sequence

of amino acids in a protein. This sequence, in turn, **will determine the way that protein will fold into a precise shape** with characteristic chemical properties.

In summary, the primary transfer of information within cells follows the order:

DNA >>> RNA>>> PROTEIN>>>>TRAIT

Although the information itself is linear, the implications are 3-dimensional. A fundamental assumption of recombinant DNA technology is that permanent and desirable changes in the functioning of living cells can be accomplished by changing the linear sequence of their DNA.

Genes are discrete files of DNA information

A gene is a segment within a DNA molecule singled out for copying into RNA. Directly or indirectly, this RNA will perform a function. It is convenient to think of a gene, therefore, as a unit of function.

Many traits, such as bacterial resistance to an antibiotic, are governed by single genes. Most traits—such as the color of a rose, or the shape of a nose—are governed by several genes acting in concert.

Genes can vary in length: Some are only a few hundred base pairs long; some can be tens of thousands of base pairs long. A DNA molecule may carry from a handful to thousands of genes. A cell, in turn, may contain one or several DNA molecules (chromosomes). Thus the number of genes in a cell can vary greatly. *E.coli*, a bacterium, contains one DNA molecule with about five thousand genes on it. A human cell contains 46 DNA molecules carrying a total of about 100,000 genes.

All genes in a given cell are **not** copied into RNA (*i.e.* "expressed") at the same time or at the same rate. Thus, when speaking of gene function, one refers to its "expression level". This rate can be controlled by the cell, according to predetermined rules which are themselves written into the DNA.

An example: The cells in our bodies (all 100 trillion of them) each contain identical DNA molecules. Yet liver cells, for example, express only those genes required for liver function, whereas skin cells express a quite different subset of genes!

DNA can be cut into pieces with restriction enzymes

Restriction enzymes are proteins made by bacteria as a defense against foreign, invading DNA (for example, viral DNA). Each restriction enzyme recognizes a unique sequence of typically 4–6 base pairs, and will cut any DNA whenever that sequence occurs.

For example, the restriction enzyme BamHI recognizes the sequence (5'..GGATCC..3') and cuts the DNA strand between the two G nucleotides in that sequence.

Restriction enzymes will cut DNA from any source, provided the recognition sequence is present. It does not matter if the DNA is of bacterial, plant or human origin.

Pieces of DNA can be joined by DNA ligase

DNA ligase is an enzyme that "glues" pieces of DNA together, provided the ends are compatible.

Thus, a piece of human or frog or tomato DNA cut with BamHI can be easily joined to a piece of bacterial DNA also cut with BamHI. This allows the creation of recombinant DNAs *i.e.* hybrids, created by joining pieces of DNA from two different sources.

Genes can be cut out of human DNA, or plant DNA, and placed inside bacteria. For example, the human gene for the hormone insulin can be put into bacteria. Under the right conditions, these bacteria can make authentic human insulin!

Plasmids are small circular pieces of DNA

Plasmids are small circular DNAs found inside some bacterial cells. They replicate their own DNA by "borrowing" the cells' polymerases. Thus they can persist indefinitely inside cells without doing very much work of their own!

Because of their small size, plasmid DNAs are easy to extract and purify from bacterial cells. When cut with a restriction enzyme, they can be joined to foreign DNAs - from any source - which have been cut with the same enzyme.

The resulting hybrid DNAs can be re-introduced into bacterial cells by a procedure called transformation. Now the hybrid plasmids can perpetuate themselves in the bacteria just as before except that the foreign DNA which was joined to it is also being perpetuated. The foreign DNA gets a free ride, so to speak.

Every hybrid plasmid now contains a perfect copy of the piece of foreign DNA originally joined to it. We say that foreign piece of DNA has been "cloned"; the plasmid which carried the foreign DNA is called a cloning vehicle or vector.

In addition to their usefulness for cloning foreign genes, plasmids sometimes carry genes of their own. Bacteria die when exposed to antibiotics. However, antibiotic-resistance genes allow bacteria to grow in the presence of an antibiotic such as ampicillin. Such genes are often found on plasmids. When foreign DNA is inserted into such plasmids, and the hybrids introduced into bacterial cells by transformation, it is easy to select those bacteria that have received the plasmid—because they have acquired the ability to grow in the presence of the antibiotic, whereas all other bacterial cells are killed.

DNA libraries

When DNA is extracted from a given cell type, it can be cut into pieces and the pieces can be cloned en masse into a population of plasmids. This process produces a population of hybrid (*i.e.* recombinant) DNAs. After introducing these hybrids back into cells, each transformed cell will have received and propagated one unique hybrid. Every hybrid will contain the same vector DNA but a different "insert" DNA.

If there are 1,000 different DNA molecules in the original mixture, 1,000 different hybrids will be formed; 1,000 different transformant cells will be recovered, each carrying one of the original 1,000 pieces of genetic information. Such a collection is called a DNA library. If the original extract came from human cells, the library is a human library.

Individual DNAs of interest can be fished out of such a library by screening the library with an appropriate probe.

Appendix C

Sterile Technique

With any type of microbiology technique (*i.e.* working with and culturing bacteria), it is important to not introduce contaminating bacteria into the experiment. Because contaminating bacteria are ubiquitous and are found on fingertips, bench tops, etc., it is important to minimize contact with these contaminating surfaces. When students are working with the inoculation loops, pipetters, and agar plates, you should stress that the round circle at the end of the loop, the tip of the pipetter, and the surface of the agar plate should not be touched or placed onto contaminating surfaces. While some contamination will not likely ruin the experiment, students would benefit from an introduction to the idea of sterile technique. Using sterile technique is also an issue of human cleanliness and safety.

Working with *E. coli*

The host organism, an *E. coli* K-12 strain, and the vector containing the recombinant GFP protein and the subsequent transformants created by their combination, are not pathogenic organisms like the *E. coli* O157:H7 strain that has been in the news. However, handling of the *E. coli* K-12 entities of the GFP Purification Kit requires the use of Standard Microbiological Practices. These practices include but are not limited to the following: Work surfaces are decontaminated with 10% bleach once a day and after any spill of viable material. All contaminated liquid or solid wastes are decontaminated before disposal. Persons wash their hands: (i) After they handle materials involving organisms containing recombinant DNA molecules, and (ii) before exiting the laboratory. All procedures are performed carefully to minimize the creation of aerosols. Mechanical pipetting devices are used; mouth pipetting is prohibited. Eating, drinking, smoking, and applying cosmetics are not permitted in the work area.

Decontamination and Disposal

If an autoclave is not available, all solutions and components (loops and pipettes) that have come in contact with bacteria can be placed in a fresh 10% bleach solution for at least 20 minutes for sterilization. A shallow pan of this solution should be placed at every lab station. No matter what you choose, all used loops and pipettes should be collected for sterilization. Sterilize Petri dishes by covering the agar with 10% bleach solution. Let it stand for 1 hour or more. Once sterilized, the agar plates can be double bagged and treated as normal trash. Safety glasses are recommended when using bleach solutions.

Appendix D

Gene Regulation. One Gene: One Protein.

Our bodies contain thousands of different proteins which perform many different jobs. Digestive enzymes are proteins; some of the hormone signals that run through our bodies and the antibodies protecting us from disease are proteins. The information for assembling a protein is carried in our DNA. The section of DNA which contains the code for making a protein is called a gene. There are over 100,000 genes in the human genome. Each gene codes for a unique protein: one gene-one protein. The gene which makes a digestive enzyme in your mouth is different from one which makes an antibody or the pigments that color your eyes.

Organisms regulate expression of their genes and ultimately the amounts and kinds of proteins present within their cells for a myriad of reasons including developmental, cellular specialization and adaptation to the environment. Gene regulation not only allows for adaptation to differing conditions, but also prevents wasteful overproduction of unneeded proteins which would put the organism at a competitive disadvantage. The genes involved in the transport and breakdown (catabolism) of food are good examples of highly regulated genes. For example, the sugar arabinose is both a source of energy and a source of carbon. *E. coli* bacteria produce three enzymes (proteins) needed to digest arabinose as a food source. The genes which code for these enzymes are not expressed when arabinose is absent, but they are expressed when arabinose is present in their environment. How is this so?

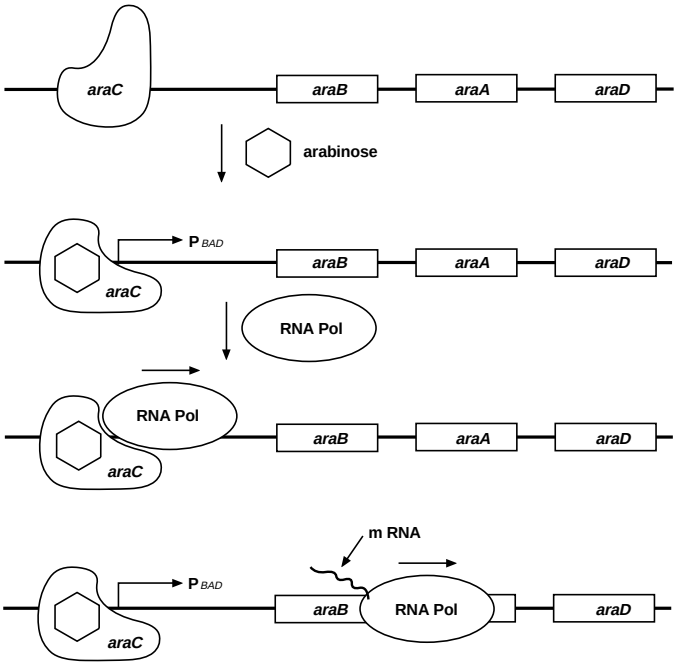
Regulation of the expression of proteins often occurs at the level of transcription from DNA into RNA. This regulation takes place at a very specific location on the DNA template, called a promoter, where RNA polymerase sits down on the DNA and begins transcription of the gene. In bacteria, groups of related genes are often clustered together and transcribed into RNA from one promoter. These clusters of genes controlled by a single promoter are called operons.

The three genes (*ara*B, *ara*A and *ara*D) that code for three digestive enzymes involved in the breakdown of arabinose are clustered together in what is known as the arabinose operon. These three proteins are dependent on initiation of transcription from a single promoter, (P_{BAD}). Transcription of these three genes requires the simultaneous presence of the DNA template (promoter and operon), RNA polymerase, a DNA binding protein called AraC, and arabinose. AraC binds to the DNA at the binding site for the RNA polymerase (the beginning of the arabinose operon). When arabinose is present in the environment, bacteria take it up. Once inside, the arabinose interacts directly with AraC which is bound to the DNA. The interaction causes AraC to change its shape which in turn promotes (actually helps) the binding of RNA polymerase and the three genes B, A, and D, are transcribed. Three enzymes are produced, they do their job, and eventually the arabinose runs out. In the absence of arabinose, the AraC returns to its original shape and transcription is shut off.

The DNA code of the pGLO plasmid has been engineered to incorporate aspects of the arabinose operon. Both the promoter (P_{BAD}) and the *ara*C gene are present. However, the genes which code for arabinose catabolism, *ara*B, A and D, have been replaced by the single gene which codes for the Green Fluorescent Protein (GFP). Therefore, in the presence of arabinose, AraC protein promotes the binding of RNA polymerase and GFP is produced. Cells fluoresce a brilliant green color as they produce more and more protein. In the absence of arabinose, AraC no longer facilitates the binding of RNA polymerase and the GFP gene is not transcribed. When the GFP protein is not made, bacteria colonies will appear to have a wild type (natural) phenotype - of white colonies with no florescence.

This is an excellent example of the central dogma of molecular biology in action; DNA>RNA>PROTEIN>TRAIT.

The Arabinose Operon



Expression of Green Fluorescent Protein