Supplemental Materials for

Studying Gene Expression: Database Searches and Promoter Fusions to Investigate Transcriptional Regulation in Bacteria

Betsy M. Martinez-Vaz, Irina Makarevitch, and Shane Stensland
Department of Biology, Hamline University, Saint Paul, MN 55104

Table of Contents
(Total pages 53)

Appendix I: Questions used to explore student’s knowledge of gene expression and regulation
Appendix II: Bioinformatics Worksheet
Appendix IIA: Bioinformatics Worksheet- Answer Key
Appendix III: Absorbance and Fluorescence Worksheet
Appendix IV: Literature Search and Experimental Design
Appendix V: Pre- Test and Post- Test
Appendix VA: Answer Key for Pre- and Post- Test
Appendix VI: Bioinformatics Search Exercise and Exam Questions
Appendix VII: Example of Student’s Laboratory Report
Appendix VIII: Rubric borrowed from LabWrite and modified to grade laboratory
Appendix IX: Laboratory Prep
Appendix X: Student’s Handout
Appendix XI: Instructions for Database Searching
Appendix XII: Map for Plasmid pUA66
Appendix I: Questions used to explore student’s knowledge of gene expression and regulation

1. Define gene expression

2. Describe in simple terms the function of the main players of transcriptional gene regulation (and the general mechanism of their action):
   a) promoter
   b) terminator
   c) activator (transcription factor)
   d) repressor (transcription factor)
   e) metabolic substrate as a co-regulator (transcription factor)

3. Which biological parameters are used to determine whether a gene is being expressed?
   a) mRNA synthesis
   b) Protein production

4. Mention one experiment commonly employed to investigate changes in gene expression
   a) Common answers include: northern Blots, Reverse transcriptase PCR (RT-PCR), western blots and DNA microarrays

5. Can you give an example of positive and negative regulation of gene expression? What about repression and induction of gene expression?
Appendix II: Bioinformatics Worksheet

NCBI Genome

A) Describe the genome region surrounding gltB and serA. Which genes are located near gltB? Near serA? Discuss the function of at least two of these genes.

Regulon DB

A) List the transcription factors that are known to control the gltB or serA unit.

B) List the binding sites determined for the transcription factors that control the gltBDF and the serA transcriptional units.

Ecocyc

A) Name two biochemical pathways involving the enzymes encoded by gltB and serA. You can name pathways involving one or both enzymes.

B) Go to the gltB and serA transcriptional unit link and answer the following questions:

i. How many genes are parts of the gltB and serA transcriptional units?

ii. Which proteins do they encode?

iii. Are all the genes in the serA and gltB transcriptional units under the control of the same promoter? Justify your answer.
Comparing Databases

A) Briefly describe the goal of Ecocyc and Regulon DB. What type of information do these databases provide?

B) Mention two similarities and two differences in the data presented in these databases.

   Differences:
   1.
   2.

   Similarities:
   1.
   2.

C) Which features of each database did you find more helpful? Give two examples. Discuss your answer.

D) How can the information presented in these databases be applied to modern research in Biochemistry?

E) Mention two other databases that can be accessed through Regulon DB and Ecocyc
Appendix IIa: Bioinformatics Worksheet- Answer Key

NCBI Genome

A) Describe the genome region surrounding gltB and serA. Which genes are located near gltB? Near serA? Discuss the function of at least two of these genes.

Two genes located near gltB are gltD and yhcC. The GltD protein is the small subunit of glutamate synthase. YhcC is a hypothetical protein predicted to be an iron sulfur reductase.

Two genes located in genome region containing serA are rpiA and pepP. These encode a ribose 5 phosphate isomerase and a proline aminopeptidase, respectively.

Most of the genes surrounding serA and gltB encode proteins that function in cellular metabolism.

Regulon DB

A) List the transcription factors that are known to control the gltB or serA units. Mention whether the transcription factors are activators or repressors.

gltB is regulated by the following transcription factors: ArgR (repressor), Lrp (activator), IHF (activator), GadE (activator), CRP (repressor) and FNR (repressor)

serA is regulated by the following transcription factors: Nac (repressor), Lrp (activator) and CRP (activator)

B) List the binding sites determined for the transcription factors that control the gltBDF and the serA transcriptional units.

gltBDF:
ArgR- aatcaaaattACCGAAATTTCATGCATAatcacataaa
ArgR-gcttttatcaACTGCATAATCAATCAAattaccgaaaa
CRP-tttaataaagAATTTTGCGCTAAAGCACATTtctgtaccaat
IHF-ttttaatttcAGTCATTTAATAAagaattttgc
IHF-tccattttaaTTTCAGTCATTtTataaagaatt
LRP-ctttatgacaGTACGAAAAATGCTgtttttgc
LRP-aattaaggcaGTATAAAATGTGCTggtttttgctg
LRP-ggcaaaaaccGTAGCATGAACGcattacca
LRP-leucine-ttttagccctAAAGATAAAATCCATTttaatttca
Nac, FNR and GadE- No binding sites reported in RegulonDB
SerA:
Nac-tcaaaaaataATATTGCGGCAATATgaacgtttgc
Nac-ctttcaacatATCAAAAAATAATATTgcggcaaata
CRP-cAMP-acatgtcaccAAATTTAATGAAGAGAATTTTTTtaacggggg
Lrp- No binding sites reported.

Ecocyc

A) Go to the gltB and serA transcriptional unit link and answer the following questions:

i. How many genes are parts of the gltB and serA transcriptional units?

**gltB- This gene has one transcriptional unit, it contains three genes (gltBDF).**

**serA- This gene has tow transcriptional units, both units have only one gene.**

ii. Which proteins do they encode?

**gltBDF** encodes the following proteins:
- GltB- glutamate synthase- large subunit
- GltD- glutamate synthase- small subunit
- GltF- periplasmic protein of unknown function

**serA encodes the following protein:**
- SerA- α-ketoglutarate reductase / D-3-phosphoglycerate dehydrogenase

iii. Are all the genes in the serA and gltB transcriptional units under the control of the same promoter (s)? Justify your answer.

**There are two transcriptional units for the serA gene; each transcriptional unit is under the control of a different promoter. In contrast, all the genes in the gltBDF transcriptional unit are under the control of the same promoter. This assumption is based on the observation that only one promoter is shown in RegulonDB and EcoCyc databases. In addition, a transcriptional unit is defined as states that “a set of one or more genes transcribed from a single promoter”, therefore that will always be only one promoter per transcriptional unit.**
Appendix IIa- continued

EcoCyc- continued

B) Go to the “search “ tab on the top right of the EcoCyc web page. Select pathways from the pull down menu. Under search for pathways name type: serine and then glutamate. Explore the pathways matching the names of these amino acids. Name the biochemical pathways involving the enzymes encoded by \textit{gltB} and \textit{serA}. You can name pathways involving one or both enzymes.

- Biochemical pathways involving \textit{GltB}- Glutamate biosynthesis, nitrogen metabolism and ammonia assimilation cycle

- Biochemical pathways involving \textit{SerA}- serine biosynthesis

Comparing Databases

A) Briefly describe the goal of Ecocyc and Regulon DB. What type of information do these databases provide?

The goal of both databases is to provide basic information on the genes, proteins and metabolic processes in the model organism, \textit{Escherichia coli}.

RegulonDB and EcoCyc contain information on transcriptional regulation, metabolic pathways, gene and protein sequences. In addition, these resources provide maps with the position of the genes in the \textit{E.coli} chromosome and search tools to compare the homology of gene and protein sequences across different microbial genomes

B) Mention two similarities and two differences in the data presented in these databases.

Differences:

1. RegulonDB focuses on gene regulation and contains many resources to learn about this topic. These resources include links to regulatory networks, gene expression and promoter analyses tools. RegulonDB also allows the users to do a search for growth conditions that might affect the expression of different genes and operons; EcoCyc does not have this capability.
2. Overall, EcoCyC provides more information on gene and proteins than RegulonDB. For example, EcoCyc contains several paragraphs describing the function of serA and its protein product. In addition, this database provides the chemical reactions catalyzed by the SerA and GltB enzymes with their kinetic parameters. These data is harder to find in RegulonDB because the links in this database are not properly labeled. Accessing protein data in RegulonDB requires clicking on a tab labeled as “notes” which does not direct the viewer to the type of information it contains.

Similarities:

1. Both database contain color-coded graphic representations of the components of each transcriptional unit.
2. Both databases provide easy access to protein and DNA sequences for the genes that are being searched.

C) Which features of each database did you find more helpful? Give two examples. Discuss your answer.

- Some of the most helpful features of RegulonDB include:
  A) The tables containing the transcription factors that regulate a particular gene and their DNA binding sites. These tables provide a quick and easy way to summarize the published data available on a particular transcriptional regulator.
  B) RegulonDB allows viewing the information available for a gene in one page; there is no need of going through separate links to see the graphic representation of transcriptional units or basic information such as GC content and genome position of a gene.

- Some of the most helpful features of EcoCyc include:
  A) The search feature of EcoCyc is very powerful and allows us to search for biochemical reactions and pathways containing your gene of interest.
  B) Overall, EcoCyc is better organized and easier to search than Regulon DB. In EcoCyc the user has to used several links to access a given set of data, these links have been labeled clearly and are very easy to navigate.
D) How can the information presented in these databases be applied to modern research in Biochemistry?

By using EcoCyc and RegulonDB scientists can learn about the regulation any given gene in *E.coli* without having to do extensive literature searches that are often necessary before designing and performing an experiment. Looking at information on RegulonDB or EcoCyc can allow scientists to predict how gene expression will change in response to a given genetic or environmental condition. In addition, these databases contain information on the location of any gene in the genome and provide easy access to DNA and protein sequences. These features are advantageous when planning an experiment because one can easily access protein and DNA sequences for primer design and homology searches.

E) Mention two other databases that can be accessed through Regulon DB and Ecocyc

- NCBI GeneBank
- Protein data bank (PDB)
**Appendix III: Absorbance and Fluorescence Worksheet**

**Fluorescence Measurements.** Please record the data for absorbance and fluorescence as x=average and (+/-) standard error.

<table>
<thead>
<tr>
<th>Strain / Sample Tube</th>
<th>Before/After Incubation</th>
<th>Growth Medium</th>
<th>Average Absorbance</th>
<th>Average Fluorescence</th>
<th>Normalized Fluorescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT (gltBp-GFP)</td>
<td>Before</td>
<td>LB (rich)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT (gltBp-GFP)</td>
<td>Before</td>
<td>M9 (minimal)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT (pUA66)</td>
<td>Before</td>
<td>LB (rich)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT (pUA66)</td>
<td>Before</td>
<td>M9 (minimal)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT (gltBp-GFP)</td>
<td>After</td>
<td>LB (rich)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT (gltBp-GFP)</td>
<td>After</td>
<td>M9 (minimal)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT (pUA66)</td>
<td>After</td>
<td>LB (rich)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT (pUA66)</td>
<td>After</td>
<td>M9 (minimal)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This is an example of the table used to record absorbance and fluorescence measurements during the laboratory exercise. Students were asked to record the average absorbance and fluorescence (after three or four replicas) and calculate an average normalized fluorescence. The second column of the table refers to the time at which the fluorescence measurements were taken.
Appendix IV: Literature Search and Experimental Design

This exercise consists of formulating a testable hypothesis. What type of questions can you ask about the factors that regulate gltB or serA expression using the microbial strains and materials available in the lab? In order to figure things out, you need to learn more about gltB and its regulation.

1. Literature Analysis

Do a literature search and find a source describing the physiological role of **TWO** of the following gltB or serA regulators (LRP, CRP, Nac or ArgG). Answer the following questions:

   a. What environmental conditions affect the activity of each transcription factor? Explain whether conditions activate or repress the activity of the transcription factor.

   b. How do these transcription factors allow gltB or serA to respond to different environmental stimuli? Describe whether the factor is an activator or a repressor. Explain the conditions that stimulate such activity.

Do an NCBI (medline) search to investigate the current knowledge about gltB and serA regulation. Select three abstracts that relate to a specific aspect of gltB or serA regulation that you would be interested in exploring. Start filling out the experimental design worksheet.

2. Experimental Design

Your lab group needs to design an experiment that would address a question about gltB or serA regulation. The experiment must involve the use of the gltBp-GFP or serAp-GFP reporter constructs and any of the materials and strains listed on the "supply list". As you work on your experimental design, make sure to clearly specify the experimental groups, controls and physical conditions (temperature, growth conditions, strains needed, etc). Once you have finished your experimental design, fill out the experimental design worksheet and turn it in to your instructor. After instructor’s approval, start preparing media that you will need for the following week.
Appendix IV: Experimental Design - continued

1. List the three medline references that you are using to design your project. Include a brief statement of a major point of each of these three papers.
   a.
   b.
   c.

2. State the main question that you want to address in your experiment?

3. What is your hypothesis?

4. Experimental Plan

Strain or strains to be used:

Growth conditions (media, temperature, etc):

Controls:

Supplies:
List what you would need to have prepared for the next lab to run this experiment (solutions, media, etc…)

Measurements and outcomes: Indicate how you would be assessing changes in gene expression. What is the expected outcome of the proposed research?
Appendix V: Pre-Test and Post-Test

1. In presence of a transcriptional activator, the activity of a promoter should:
   a. Increase
   b. Decrease
   c. Remain constant
   d. Change only in response to cellular cues
   e. None of the above

2. In the presence of a transcriptional repressor, the activity of a promoter should:
   a. Increase
   b. Decrease
   c. Change only in response to cellular cues
   d. Remain constant
   e. None of the above

3. A repressor inhibits the synthesis of given mRNA by
   a. Mutating RNA polymerase
   b. Cleaving the mRNA as soon as it is transcribed
   c. Inhibiting the binding of RNA polymerase to the promoter region
   d. Producing early termination of transcription
   e. Recruiting RNA degrading enzymes into the cell

4. Cellular metabolites often influence the activity of transcription factors due to:
   a. Their role as allosteric effectors of RNA Polymerase
   b. Their ability to bind DNA and block transcription
   c. Their ability to bind transcription factors and act as co-activators and co-repressors of transcription
   d. Their role as inhibitors of protein synthesis

5. The activity of metabolic enzymes is controlled:
   a. Allosterically
   b. At the level of transcription
   c. By covalent modifications
   d. By proteolytic cleavage
   e. All the above
Appendix V: Pre- Test and Post- Test- continued

6. Which of the following is true about transcription factors
   a. They can have one or more binding sites in DNA of their target genes
   b. They usually have no DNA binding specificity
   c. They are often unique to a particular bacterial specie
   d. They are produced in large quantities and stored in the nucleoid until needed
   e. None of the above

7. In cells, transcription factors function to …
   a. Regulate one gene at the time
   b. Coordinate the expression of hundreds of genes in response to a given condition
   c. Control of the activity of RNA polymerase
   d. Modulate the expression of operons
   e. All the above

8. The direct binding of a cellular metabolite to a transcription factor (TF) most likely would:
   a. Antagonize the activity of TF
   b. Enhance the activity of the TF
   c. Have no effect in TF activity
   d. Produce a mutation RNA Polymerase
   e. Both, enhance or antagonize TF activity

9. Scientists use gene reporter assays when
   a. They need to quantify or monitor a cellular process that is not easily measurable
   b. DNA mutations prove to be unsuccessful to study gene function
   c. RNA levels in cells are easy to measure
   d. Cells lack membrane stability
   e. The organism of interest is hard to isolate and culture

10. Gene reporters are useful to measure gene expression because:
    a. They are expressed in cells at high quantities
    b. They produce stable proteins
    c. They encode protein products that are easy to measure
    d. They can be used in different types of cells
    e. All the above
Appendix V: Pre-Test and Post-Test-continued

11. List three examples of reporter genes
   a. ________
   b. ________
   c. ________

12. A reporter construct effective to monitor transcriptional regulation should have
   a. The promoter region of the gene of interest attached to DNA encoding the reporter gene
   b. The promoter region of the reporter gene attached to the DNA sequence encoding the gene of interest
   c. The protein encoded by the reporter attached to RNA polymerase
   d. The RNA polymerase gene replaced with the reporter gene.
   e. All the above

13. You have been given the task to investigate the regulation of a newly identified anticancer protein named Hampiper. The DNA sequence of the gene encoding this protein has been determined and is available in a database known as GeneBase. You decide that your very first experiment to investigate the regulation of Hampiper should be:

   a. To create a mutant version of the anticancer protein
   b. To measure the level of hampiper mRNA at different temperatures
   c. To measure the level of the anticancer protein in normal cells as they progress through the cell cycle
   d. To fuse the promoter to a reporter gene, transfet the construct into normal and cancerous cells and assay for activity under different conditions.
   e. None of the above

14. You have decided to investigate the function and regulation of hamP1, an E.coli gene encoding a cold resistant enzyme. The first part of the project is bioinformatics; you decide to start by doing a search in NCBI genomes. What kind of information are you expecting to obtain?

   a. Gene sequence only
   b. A link to biochemical pathways related to hamP1 function
   c. Gene sequence and position of hamP1 in the E.coli genome
   d. A list of transcription factors that regulate hamP1
   e. None of the above
Appendix V: Pre-Test and Post-Test—continued

15. Your boss has asked you to investigate whether an *E.coli* gene is under the control of the catabolic repressor protein (CRP). The gene encodes a hypothetical protein of unknown function and cellular phenotype. As an efficient scientist you decide that the first research task should be:

a. Mutate *crp* and compare the level of mRNA of every gene in the *E.coli* genome in wild type and mutant cells
b. Treat genomic DNA with different amounts of CRP and determine the parts of a DNA sequence where CRP binds
c. Search the literature for CRP binding sites; then analyze the promoter region of the gene of interest to determine the presence of CRP binding sites
d. Given that the function and cellular phenotype of the protein encoded by the gene are not known, is impossible to study its regulation
e. Both, a and c are reasonable approaches to learn about gene regulation

16. Which of the following databases provide detailed information about transcription factors and the genes that they regulate

a. Protein data bank
b. GeneBank
c. Ecocyc
d. Regulon DB
e. I do not know of any databases that provide this information

17. You want to know how many transcriptional regulators are known control the activity of *lacZ* in *E.coli*. Your advisor suggest that you do this computationally, which of these databases is most likely to provide the information needed for your research:

a. Blast
b. GeneBank
c. Regulon DB
d. NCBI genomes
e. Pubmed
18. Which of the following methods will allow you to successfully measure transcriptional regulation. Circle as many as you think are correct
   a. Microarrays
   b. Protein quantification
   c. Measurements of mRNA abundance
   d. Gene Reporter Assays
   e. RegulonDB

19. Databases containing information on transcriptional regulation present data obtained in:
   a. Computer simulations
   b. Experimental analysis
   c. Literature searches
   d. Genome sequences
   e. All the above

20) Ecocyc provides information about
   a. Gene regulation and metabolism in E.coli
   b. Different types of E.coli strains
   c. Gene regulation and transporters
   d. Gene regulation, transporters and metabolism in E.coli
   e. Gene sequences and transcription factors only

21) Choose the best definition for the term “reporter gene”
   a. A gene that researchers attach to another gene of interest in cell culture, animals or plants
   b. A gene encoding a protein that lacks stability
   c. Any gene that “reports” changes in the cell physiology
   d. A gene that is used to follow changes in mutation rates
   e. A reporter gene is one that codes for a product that is easy to measure and whose phenotypic expression is easy to monitor

22) Which of the following properties are often monitored as part of reporter assays:
   a. Beta galactosidase activity
   b. Fluorescence
   c. Luminescence
   d. Colorimetric changes (optical density)
   e. All the above
23) Your Biology professor created a library of reporter constructs in which different pieces of DNA were cloned upstream of a promoter-less beta galactosidase gene. What are these “reporter” constructs meant to report?

a. The ability of the cloned DNA sequences to drive transcription of the beta galactosidase gene
b. The stability of the beta galactosidase protein
c. The time needed to induce the expression of the beta galactosidase gene
d. The location of betagalactosidase in the cell
Appendix Va: Answer Key for Pre-Test/Post-test

1) a  
2) b  
3) c  
4) c  
5) e  
6) a  
7) e  
8) e  
9) a  
10) e  
11) lacZ, gfp, cat, gus  
12) a  
13) d  
14) c  
15) c  
16) c and d  
17) c  
18) a,b,c,d  
19) e  
20) a  
21) e  
22) e  
23) a

Database search questions- Answers

a. How many transcriptional regulators are known to control the activity of ompF?

   Seven

b. Mention two transcriptional activators and two transcriptional inhibitors of ompF

   Lrp and CRP- activators  
   IHF and OmpR- repressors

c. Write the sequence of a binding site for an inhibitor. Specify the inhibitor chosen

   OmpR- cgtaatgttGCAATGTAATTTCCGTAACaggatgatcg
d. Write the sequence of a binding site for an activator. Specify the activator chosen

Lrp- atttggttcAAAAAGAAAAAAAtatgtaacca

e. Provide a citation for a literature article confirming the activity of transcriptional activator or repressor mentioned above.


f. Provide a URL and indicate the bioinformatics resource that was used to answer the questions

http://regulondb.ccg.unam.mx

Answer Key for final exam questions

1) a
2) c
3) c
4) a
5) a
6) d
7) e
8) Controls missing: 1) Cells containing promoter less GFP to account for background fluorescence, 2) sample with honey and growth media only to investigate if the presence of honey impairs or affects fluorescence measurements.
Appendix VI: Bioinformatics Search Exercise and Exam Questions

Database search questions

a. How many transcriptional regulators are known to control the activity of *ompF*?

b. Mention two transcriptional activators and two transcriptional inhibitors of *ompF*.

c. Write the sequence of a binding site for an inhibitor. Specify the inhibitor chosen.

d. Write the sequence of a binding site for an activator. Specify the activator chosen.

e. Provide a citation for a literature article confirming the activity of transcriptional activator or repressor mentioned above.

f. Provide a URL and indicate the bioinformatics resource that was used to answer the questions.

Exam Questions

1. Normalized fluorescence is a better measurement of gene expression because
   a. It accounts for differences in growth between the cultures that are being compared.
   b. Correct for errors due to instrument malfunctions.
   c. Correct for mutations in the GFP gene.
   d. It is equivalent to the measurement used to calibrate the fluorometer.
2. A student is investigating the role of CRP in the regulation of a gene involved in serine biosynthesis. He constructed a GFP promoter fusion to be used in a reporter assay. According to his hypothesis, the gene is under the control of CRP and should be expressed only in the absence of glucose. The data obtained in shown in the table below

<table>
<thead>
<tr>
<th>Sample</th>
<th>Normalized Fluorescence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E.coli (serAp-GFP)</em> M9 +glucose</td>
<td>5997</td>
</tr>
<tr>
<td><em>E.coli (serAp-GFP)</em> M9 +lactose</td>
<td>48000</td>
</tr>
<tr>
<td><em>E.coli:crp- (serAp-GFP)</em> M9 +glucose</td>
<td>5500</td>
</tr>
<tr>
<td><em>E.coli:crp- (serAp-GFP)</em> M9 +lactose</td>
<td>5300</td>
</tr>
<tr>
<td><em>E.coli (promoterless GFP)</em></td>
<td>5000</td>
</tr>
</tbody>
</table>

According to the data shown, the gene under study is

a. Not regulated by CRP, its expression does not change in response to alternative carbon sources
b. Not active under the growth conditions used for the experiment
c. The gene is regulated by CRP, that is why there is higher fluorescence in the presence of lactose compared to glucose
d. Activated by glucose and repressed by lactose
3. A graduate student wants to investigate the role of SorA in the regulation of a gene encoding the Ham protein. She hypothesizes that SorA is an activator of *ham* expression. She constructs a *ham* promoter GFP fusion to investigate *ham* expression in response to SorA. What should her first experiment be?
   a. Test the *hamp*-GFP promoter fusion in the presence of different carbon sources
   b. Test the expression of the *hamp*-GFP promoter fusion in different cell types
   c. Test the *hamp*-GFP promoter fusion in wild type cells and mutants lacking SorA
   d. Test the *hamp*-GFP promoter fusion at different phases of the bacterial growth cycle

4. The data from the initial SorA/*ham* experiment is shown in the table below

<table>
<thead>
<tr>
<th>Sample</th>
<th>Normalized Fluorescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT <em>E.coli</em> (<em>hamp</em>-GFP) in M9</td>
<td>65000</td>
</tr>
<tr>
<td><em>E.coli</em>: sorA- (<em>hamp</em>-GFP) in M9</td>
<td>15000</td>
</tr>
<tr>
<td><em>E.coli</em> (promoterless GFP) in M9</td>
<td>7000</td>
</tr>
<tr>
<td>M9 only</td>
<td>900</td>
</tr>
</tbody>
</table>

According to the data presented above:
   a. SorA is an activator of *ham*
   b. Sor A is a repressor of *ham*
   c. The data is inconclusive, the role of SorA in *ham* regulation cannot be determined
   d. The *ham* promoter is not active under the conditions used for this experiment
5. You have discovered a novel transcription factor, Hampiper, that regulates the expression of genes involved in lysine biosynthesis. As part of your investigative efforts, you decide to test the effects of other amino acids on Hampiper activity. In order to accomplish your goal, you construct a klausp-GFP promoter fusion. Klaus is a protein, which is activated by Hampiper. What experiment would you do next?

  a. Test the activity of the klausp-GFP fusion on WT cells containing Hampiper and several concentration of different amino acids, including lysine
  
  b. Test the klausp-GFP promoter fusion at different phases of the bacterial growth cycle
  
  c. Test the klausp-GFP promoter fusion at different in cells lacking Hampiper grown in media with different amino acids
  
  d. Test the activity of klausp-GFP in different types of bacteria

6. The leucine responsive protein is a transcription factor that responds to changes in the levels of several metabolites, such as amino acids and sugars. Your classmate performed an experiment to investigate what amino acids influence Lrp activity. He used a gltB promoter GFP fusion given that gltB is known to be activated by Lrp. His data is shown below

<table>
<thead>
<tr>
<th>Normalized fluorescence</th>
<th>M9 + glucose+ Arg</th>
<th>M9+glucose+ Leu</th>
<th>M9+ glucose+ alanine</th>
<th>M9 + glucose</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT E.coli (gltBp-GFP)</td>
<td>85000</td>
<td>15000</td>
<td>45000</td>
<td>45000</td>
</tr>
<tr>
<td>WT E.coli (promoterless GFP)</td>
<td>7000</td>
<td>7000</td>
<td>7000</td>
<td>7000</td>
</tr>
</tbody>
</table>

According to the data;

  a. Leucine antagonizes the activity of Lrp
  b. Alanine enhances Lrp activity
  c. Arginine antagonizes Lrp activity
  d. Arginine enhances Lrp activity and leucine antagonizes it
  e. None of the above
7. You have decided to investigate how many transcription factors possibly regulate catA, a gene involved in a carrot juice degradation pathway in *E.coli*. Which of these databases are mostly likely to be useful in your quest for potential catA regulators?

   a. GeneBank  
   b. Regulon DB  
   c. Ecocyc  
   d. Pudbmed  
   e. B and c will be most helpful

8. Your classmate conducted an experiment to investigate whether honey activated the transcription of a flagella regulator, *flhDC*. He used an *flhDCp-GFP* fusion for his experiment. The data is shown below:

<table>
<thead>
<tr>
<th>Sample</th>
<th>Normalized fluorescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT. <em>E.coli</em> (<em>flhDCp</em>-GFP) M9 + honey</td>
<td>65000</td>
</tr>
<tr>
<td>WT. <em>E.coli</em> (<em>flhDCp</em>-GFP) M9 only</td>
<td>30000</td>
</tr>
</tbody>
</table>

He claims that honey acts as an activator of *flhDC* transcription. You tell him that his experiment is lacking several controls. Mention two important controls that are missing and briefly explain their importance.
Appendix VII: Example of Student’s Laboratory Report

Effect of Nitrogen Levels on Glutamate Synthase (gltB) Expression in *Escherichia coli*

**Introduction**

This experiment was conducted to investigate the transcriptional regulation of genes involved in amino acid metabolism. The gene that will be scrutinized is *gltB*, which encodes the large subunit of glutamate synthase, an enzyme important for L-glutamamine synthesis. This gene is regulated by multiple regulons including Nitrogen Assimilation Control (Nac) which is a global regulator that responds to nitrogen concentrations. Current literature has shown that Nac acts as a repressor which binds to the *gltB* promoter under nitrogen limiting conditions thus halting *gltB* transcription (see Figure 1) (1). Using a single point mutation, researchers produced Nac knock out strains in *K. aerogenes*. These strains transcribed *gltB* regardless of their endogenous nitrogen levels whereas *gltB* transcription was abrogated in control strains under nitrogen limiting conditions. Therefore, it was concluded that nac allows the cell to respond to nitrogen concentrations (2).

![Figure 1](image-url)

Figure 1. Nac binds to the *gltB* promoter (P*gltB*) which is upstream from the glutamate synthase operon.

This experiment seeks to further understand Nac’s control over *gltB* expression in *E. coli* by varying the nitrogen concentrations in their growth media. Under nitrogen limiting conditions, resources are not available for cells to produce amino acids. Therefore, *gltB* expression is expected to decrease when nitrogen is scarce due to the repressive action of Nac.

**Methods**

Four strains of *E.coli* were incubated in M9 media supplemented with 2.5 g, 5.0g, or 10.0g NH₄Cl. The strains included WT *gltBp*-GFP (inserted GFP gene and *gltB* promoter), WT pUA66 (GFP gene with no promoter), Nac⁻*gltBp*-GFP(Nac gene knockout; GFP gene and *gltB* promoter), and Nac⁺ pUA66 (Nac gene knockout, GFP gene with no promoter). Previously cultured bacteria were given fresh M9 media (supplemented with the appropriate amount of NH₄Cl) and plated in 96-round bottom wells in triplicates and stored at 4°C. During this time,
Appendix VII: Example of Student’s Laboratory Report-continued

a second set of samples were prepared and incubated for 2 hours at 37°C after which they were plated. Media with no bacteria was also plated in order to measure the baseline fluorescence. Absorbance and fluorescence measured at 600nm and 490nm, respectively, using Modulus Fluorometer Microplate reader. *gltB* expression was compared between wild-type and Nac knockout strains.

**Data**

Table 1.

<table>
<thead>
<tr>
<th>Media Type and Incubation</th>
<th>Normalized Fluorescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild Type</td>
<td>Nac^{-/-}</td>
</tr>
<tr>
<td>Low (2.5g NH4Cl) – No Incubation</td>
<td>95137.9</td>
</tr>
<tr>
<td>Low (2.5g NH4Cl) – 2hr Incubation</td>
<td>123311.5</td>
</tr>
<tr>
<td>Medium (5.0g NH4Cl) – No Incubation</td>
<td>155661.2</td>
</tr>
<tr>
<td>Medium (5.0g NH4Cl) – 2hr Incubation</td>
<td>133988.3</td>
</tr>
<tr>
<td>High (10.0g NH4Cl) – No Incubation</td>
<td>90466.4</td>
</tr>
<tr>
<td>High (10.0g NH4Cl) – 2hr Incubation</td>
<td>87094.9</td>
</tr>
</tbody>
</table>

Table 1. Average normalized fluorescence levels (fluorescence at 490nm/absorbance at 600nm) in wild type and Nac^{-/-} *E.coli* cells before and after incubation.

**Results**

For all cell types, absorbance increased between 1.5x-3.0x after a two hour incubation at 37°C. The WT *gltB*-GFP cells in each media type increased in cell count as a measure of absorbance about 2x. The other cell types (WT pUA66, Nac^{-/-} *gltB*-GFP, and Nac^{-/-} pUA66) had variable increases between each media type, as evident in Figure 1.

In the low nitrogen concentration (2.5g NH4Cl) M9 media, *gltB* expression was higher in the Nac^{-/-} than in the WT cells in both the pre- and post-incubation groups. The difference decreased but was still present after incubation (Figure 2 and 3). In the medium nitrogen concentration (5.0g NH4Cl) media, *gltB* expression was higher in the WT than in the Nac^{-/-} cells. The difference is greater before incubation, and while still statistically significant after incubation, it may not be biologically relevant. In the high nitrogen concentration (10.0g NH4Cl) media, *gltB* expression was greater in the Nac^{-/-} cells than in WT cells both pre- and post-incubation. Post-incubation, the difference is much greater.
**Figure 1.** Graph depicting growth of four cell types (WT gltB-GFP, WT pUA66, Nac -/- gltB-GFP, and Nac -/- pUA66) after two hour incubation at 37°C. Y-axis represents ratio of post-incubation absorbance to pre-incubation absorbance (post abs/pre abs). Absorbance levels are a measure of cell growth and cell population. The greater the absorbance, the greater the cell count.

**Figure 2.** Graph of average normalized fluorescence as a measure of gltB expression in wild type (blue) and Nac -/- (red) cells before incubation at 37°C in M9 media supplemented with either low (2.5g NH₄Cl), medium (5.0g NH₄Cl), or high (10.0g NH₄Cl) nitrogen concentrations.
Figure 3. 

**gltB Expression in M9 Media with Varying Nitrogen Concentration After Incubation**

**Figure 3.** Graph of average normalized fluorescence as a measure of *gltB* expression in wild type (blue) and Nac−/− (red) cells after two hour incubation at 37°C in M9 media supplemented with either low (2.5g NH₄Cl), medium (5.0g NH₄Cl), or high (10.0g NH₄Cl) nitrogen concentrations.

**Discussion**

The goal of this experiment was to examine the effect of varying nitrogen concentrations on *gltB* expression in *E.Coli*. It was hypothesized that under nitrogen-limiting conditions, the nitrogen assimilation control (nac) regulator would inhibit *gltB* expression. The findings of this experiment partially support this hypothesis. Under nitrogen limiting conditions, WT *gltB* expression was reduced when compared to their counter parts grown in medium nitrogen concentration media. *gltB* expression in the Nac knockout cells is also higher under low nitrogen conditions when compared to the wild type cells. This is to be expected if nac represses *gltB* expression under nitrogen limiting conditions.

There were some surprising findings with the cells grown in high nitrogen concentration M9 media. It was expected that *gltB* expression in wild type cells would continue to increase in response to increasing nitrogen concentrations because, the cell would have the necessary resources for amino acid synthesis when there are ample amounts of nitrogen present. The data, however, shows a decrease in *gltB* expression in wild type cells under high nitrogen concentrations when compared to the other concentrations. This patter runs contrary to the Nac knockout cells, which show an increase in *gltB* expression under the high nitrogen condition.
One possible explanation for these findings is due to the cells’ nitrogen assimilation needs. Under nitrogen-limiting conditions, cells may utilize glutamine synthase to scavenge nitrogen from the environment. When nitrogen is no longer limiting, high levels of glutamine synthase are no longer needed for both scavenging and amino acid synthesis pathways. A 2004 study by Hua et. al. also found increased intracellular glutamate synthase levels in *E.coli* under ammonia limitation (3). The researchers also confirmed the up-regulation of nac under nitrogen-limiting conditions.

Nac is activated by nitrogen limiting conditions and it in turn inhibits the transcription of *gltB*. The findings of this experiment support the inhibitory action of nac on *gltB*, however the extent to which nac responds to nitrogen concentration is unclear and the mechanism by which it controls *gltB* expression remains to be elucidated. The affect incubation has on nac’s control of *gltB* also warrants further research.

References

## Appendix VIII: Rubric borrowed from LabWrite and modified to grade laboratory reports

<table>
<thead>
<tr>
<th>Section</th>
<th>Points</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
<td>2</td>
<td>Describes lab content concisely, adequately, appropriately</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td>3</td>
<td>Effectively defines the research problem and states the research question</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Successfully establishes the scientific concept of the lab</td>
</tr>
<tr>
<td></td>
<td></td>
<td>States hypothesis and provides logical reasoning for it</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>3</td>
<td>Gives enough details to allow for replication of procedure</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td>6</td>
<td>Opens with effective statement of overall findings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Presents visuals clearly and accurately</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Presents verbal findings clearly and with sufficient support</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Successfully integrates verbal and visual representations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use of scientific units appropriately</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Provides legends for graphs and figures</td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td>6</td>
<td>Opens with effective statement of support of hypothesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Backs up statement with reference to appropriate findings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Provides sufficient and logical explanation for the statement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gives answer to the research question and solution for unknowns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effectively links answer of research question to solution of problem</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sufficiently addresses other issues pertinent to lab</td>
</tr>
<tr>
<td><strong>Conclusion</strong></td>
<td>3</td>
<td>Convincingly describes what has been learned in the lab</td>
</tr>
<tr>
<td><strong>References</strong></td>
<td>3</td>
<td>All appropriate sources in the report are listed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Citations and references adhere to proper format</td>
</tr>
<tr>
<td><strong>Presentation</strong></td>
<td>4</td>
<td>Format of tables and figures is correct</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Report is written in scientific style: clear and to the point</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grammar and spelling are correct</td>
</tr>
<tr>
<td><strong>Overall aims of the report: the student...</strong></td>
<td>5</td>
<td>Has successfully learned what the lab is designed to teach</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Demonstrates clear and thoughtful scientific inquiry</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accurately measures and analyzes data for lab findings</td>
</tr>
</tbody>
</table>

### Points Earned

<table>
<thead>
<tr>
<th>Total Possible</th>
<th>Points</th>
<th>Percentage</th>
</tr>
</thead>
</table>
Appendix IX: Laboratory Prep

Gene Reporters Lab- Week #1

Part I. Bioinformatics

Equipment and Materials per Lab Session (20 students):

- Computers with internet access

Part II. Measuring Fluorescence to Monitor Gene Expression

Equipment and Materials per Lab Session (20 students):

- LB broth (1 liter)
- M9 broth + 0.4% glucose and 0.025% yeast extract (1 liter)
- Fluorometer 96 well microtiter plate reader
- 37 C Incubator shaker
- Transparent 96 well microtiter plates (6 plates)- Fisher Brand catalog # 12565501
- 14 ml Sterile culture tubes (30 tubes total)
- 50 ml of overnight cultures of the following *E.coli* strains grown on LB+ 25 micrograms per ml of Kanamycin:
  - *E.coli* MG1655 (*gltB* p-GFP)
  - *E.coli* MG1655 (pUA66)
- 50 ml of overnight cultures of the following *E.coli* strains grown on M9+0.4% glucose + 0.025% yeast extract +25 micrograms per ml of Kanamycin:
  - *E.coli* MG1655 (*gltB* p-GFP)
  - *E.coli* MG1655 (pUA66)

Equipment and Materials per student group (3 students per group)

- Computer (laptop or desktop) with Internet access; one computer per student is ideal but one computer per group works well too.
- Transparent 96 well microtiter plate (1)- Fisher Brand catalog # 12565501
- (6) sterile culture tubes (14 ml falcon tubes with round bottom)
- Tube rack that can hold 14ml tubes
- Pipettors (1000 ul and 200 ul)
- Sterile tips (1000 ul and 200 ul)
- Sterile tubes containing 3 ml of overnight cultures of *E.coli* MG1655 (*gltB* p-GFP) and *E.coli* MG1655 (pUA66) grown on LB+ Kanamycin
• Sterile tubes containing 3 ml of overnight culture of *E.coli* MG1655 (*gltBp-GFP*) and *E.coli* MG1655 (pUA66) grown on M9+0.4% glucose + 0.025% yeast extract +25 micrograms per ml of Kanamycin.

**Gene Reporters Lab-Week #2**

**Literature Search and Experimental Design**

**Equipment and Materials per Lab Session (20 students):**

• Laptop or desktop computers with internet access

**Equipment and Materials per student group (3 students per group)**

• Computer (laptop or desktop) with Internet access; one computer per student is ideal but one computer per group works well too.

**Gene Reporters Lab- Week#3**

**Implementing the Gene Expression Experiments Designed by your Group**

**Equipment and Materials per Lab Session (20 students):**

• Fluorometer microtiter plate reader
• Incubator shaker at 37°C
• Transparent 96 well microtiter plates- Fisher Brand catalog # 12565501
• 50 ml of 20% glucose- filtered sterilized
• 50 ml of 20% glycerol –filtered sterilized
• 50 ml of 20% lactose- filtered sterilized
• 500 ml of LB broth
• 500 ml M9 +0.4% glucose +0.025% yeast extract
• 500 ml M9 +0.4% lactose+0.025% yeast extract
• 500 ml M9+0.4% glycerol +0.025% yeast extract
• 500 ml M9 (low ammonia concentration) +0.4% glucose +0.025% yeast extract
• 500 ml M9 (high ammonia concentration) +0.4% glucose +0.025% yeast extract

**The media needed for this part of the laboratory exercise might vary according to the experiments designed by the students; you might need to prepare one or all versions of the media listed.**

• 50 ml of 100mM solutions of the 20 amino acids made in sterile water

**The students might no need all the amino acids, wait to get the request sheet before you start making solutions. Common requests usually include arginine, leucine, lysine and serine.**
Gene Reporters Lab- Week#3- continued

Equipment and Materials per Lab Session (20 students)- continued:

- 50 ml of overnight cultures of the following *E.coli* strains:
  - *E.coli* MG1655 (*gltBp*-GFP); *E.coli* MG1655 Δ*lrp* (*gltBp*-GFP), *E.coli* MG1655 (*serAp*-GFP); *E.coli* MG1655 Δ*lrp* (*serAp*-GFP), *E.coli* MG1655 Δ*argG* (*gltBp*-GFP), *E.coli* MG1655 Δ*nac* (*gltBp*-GFP) and *E.coli* MG1655 (pUA66)
  - The overnight cultures should be grown in the microbial medium that will be used for the gene expression experiments. For example if testing gene expression in M9 (low ammonia concentration) grow the overnight cultures in this medium

** The strains needed for this part of the exercise might vary according to the experiments designed by the students; you might need to grow one or all the strains listed.

Equipment and Materials per student group (3 students per group)

- Transparent 96 well microtiter plate (1)
- (10) sterile culture tubes (14 ml falcon tubes with round bottom)
- Tube rack that can hold 14ml tubes
- Pipettors (1000 ul and 200 ul)
- Sterile tips (1000 ul and 200 ul)
- Sterile tubes containing 3 ml of overnight cultures of the *E.coli* strains requested by the students
- Sterile flasks containing 40 ml of the microbial media requested by the students
- 1 box of sterile 10 ml pipets or 6 sterile 10 ml pipets per group
Recipes for Microbial Media
All sterile media can be stored at room temperature for several weeks; if antibiotics are added, the media should be stored at 4C for up to one month.

LB Broth

Recipe for 1 liter of LB Broth
To 950 ml of water add:
10 grams of tryptone
5 grams of yeast extract
10 grams of NaCl
Mix and complete the volume to 1000 ml. Autoclave for 15 minutes to sterilize it.

M9 Broth + 0.4% glucose and 0.025% yeast extract (Recipe for 1 liter)
To 750 ml of sterile DI water add:
200 ml of 5X M9 salts
2 ml of 1M MgSO4
40 ml of 20% glucose (filter sterilize, DO NOT autoclave)
0.1 ml of 1 M CaCl2
3 ml of 10% yeast extract (10 grams in 100 ml of water)

5X M9 salts- Recipe for 1 liter
Dissolve the following salts in deionized water to a final volume of 1 liter. The salt solution can be divided into 200 ml aliquots and sterilized by autoclaving.

NaCl – 2.5 grams
NH₄Cl- 5.0 grams
KH₂PO₄ - 15.0 grams
Na₂HPO₄ - 64.0 grams

Prepare the M9 salts, yeast extract, MgSO₄ and CaCl₂ solutions; sterilize by autoclaving. Add 750 ml of DI water to a glass bottle and autoclave to sterilize the water. After autoclaving, let the solutions cool down. Add the M9 salts, yeast extract, MgSO₄ and CaCl₂ to the sterile water. Add the filtered sterilized glucose to the medium. Mix gently. To keep things sterile, make all your addition in the laminar flow hood.

To make M9 medium with high concentration of ammonia, the M9 salts mix was modified by adding 10 grams of NH₄Cl per liter; 2.5 grams of ammonium chloride per liter were added to the M9 salts to prepare M9 with a low concentration of ammonia.
The carbon sources used to supplement the M9 medium (glucose, lactose and glycerol) were prepared as 20% (w/v) solutions in DI distilled water. The sugars were sterilized by filtration using a 0.22 um filter system.
Cells contain hundreds of enzymes that work together to facilitate growth and metabolism. Enzymes must be regulated to adjust their levels and activity to cellular needs. Enzyme regulation normally involves two common mechanisms: controlling the synthesis of the enzyme (genetic control) or controlling the activity of the enzyme. Generally, the activity of an enzyme is controlled by binding of small molecules (allosteric regulation), proteolytic cleavage or covalent modifications (e.g. phosphorylation and dephosphorylation) (2). In contrast, genetic control of enzyme activity is accomplished by transcriptional regulation (2). This mechanism is effective because it decreases the number of enzyme molecules that cells produce at a given time and avoids spending resources in protein synthesis and post-translational modifications.

Transcriptional regulation is facilitated by a group of proteins known as transcription factors (TF). These proteins bind to either the promoter or enhancer region of a gene and thereby control the synthesis of mRNA. Transcription factors use several mechanisms to mediate transcriptional regulation, these include: stabilizing or blocking the binding of RNA polymerase to DNA, recruitment of proteins (co-regulators) to enhance the activity of TF-DNA complexes and chromatin modification. The regulation of gene expression is linked to changes in environmental and cellular conditions. Transcriptional regulators can activate or repress transcription in response to different signals, including the presence of simple metabolites, such as sugars and amino acids. Binding of these molecules to transcription factors can either enhance or block their activity. For example, the LacI regulatory protein is called a repressor because it keeps RNA polymerase from transcribing the structural genes needed for the catabolism of lactose in *E. coli*. Thus the LacI repressor inhibits transcription of the *lac* operon (2). The effect of the LacI repressor on the *lac* genes is referred to as negative regulation (Figure 1a).

**Figure 1.** The LacI repressor negatively regulates the expression of the genes necessary for lactose catabolism in *E. coli* (*lacZ, lacY and lacA*). When LacI is bound to the operator region of the promoter, transcription of lacZYA is inhibited. Lactose binding to LacI blocks the activity of the repressor by preventing DNA binding; RNA polymerase can then transcribe the lac genes. The figures were obtained from: [http://www.blc.arizona.edu/Marty/411/Lectures/Figures/Lac_Operon_rep.GIF](http://www.blc.arizona.edu/Marty/411/Lectures/Figures/Lac_Operon_rep.GIF).
When lactose binds to a regulatory (allosteric) site on the repressor protein, it causes a conformational change. As a result of this change, the repressor’s ability to bind DNA is blocked and it can no longer bind to the operator region of the gene. RNA polymerase can then bind to the promoter and transcribe the *lac* genes (figures 1b). In contrast, positive regulation of transcription involves a regulator protein (activator) that activates the binding of RNA polymerase to DNA. An excellent example of positive regulation is the catabolism of the disaccharide sugar maltose in *Escherichia coli*. The enzymes for maltose catabolism in *E. coli* are synthesized only after the addition of maltose to the medium (6). However, control of the synthesis of maltose-degrading enzymes is not under negative control as in the *lac* operon, but under positive control; transcription requires the binding of the activator protein (MalT) to the DNA (2). In this case, binding of maltose to MalT produces a conformational change that allows this regulator to bind DNA and recruit RNA polymerase to the promoter to initiate transcription (Figure 2).

Transcription factors are essential for the regulation of gene expression and are found in all living organisms. Recent advances in the field of genomics have provided the complete genome sequence of hundreds of organisms. Genome sequence analysis suggests that the number of transcription factors found within an organism increases with genome size, and larger genomes tend to have more transcription factors per gene. The availability of genome sequences has allowed scientists to perform computational analyses and predict the position and number of binding sites for a particular transcription factor in a given genome. This information have been compiled and organized in public biological databases. For example, the Ecocyc and RegulonDB databases provide valuable data on gene structure, transcription factors and regulation of gene expression in the model organism, *Escherichia coli* (2,4). By searching Ecocyc or RegulonDB, scientists can learn about the structure and function of any gene in the *E. coli* genome and the transcription factors that control its expression. Furthermore, these databases contain data on the consensus DNA sequences recognized by transcription factors and the number and position of their binding sites in the genome. This information is useful when investigating...
transcriptional regulation because it allows researchers to make predictions about the factors and conditions that control the expression of a given gene.

Studying transcriptional regulation is essential to understand the function of a gene and its role in the cell. Scientists employ a variety of methods to study transcriptional regulation; these techniques are based on quantification of the mRNA of interest or the protein encoded by it. Quantification of proteins is normally carried out by Western Blot analyses with specific antibodies. In contrasts, techniques such as Northern Blots, DNA microarrays and reverse transcriptase polymerase chain reaction (RT-PCR), are utilized to detect and quantify mRNA. While these technologies are accurate and effective for studying gene expression, they often require expensive reagents and laboratory equipment. An alternative method for studying transcriptional regulation is the use of reporter assays. These assays consists of using a “reporter gene” to study investigate a cellular process that cannot be easily studied under normal conditions. Compared to other technologies used to study gene expression, reporter assays are cost effective, can be performed quickly and do not require complex technical manipulations.

A reporter gene is one that encodes a product that is easy to measure. It could be an easily visible protein (like green fluorescent protein from jellyfish, GFP) or an enzyme that could be assayed by a color reaction (like beta-galactosidase). When reporters are used in studying gene regulation at transcriptional level, the reporter gene would typically be attached to the regulatory region (promoter and operator, in bacteria), using recombinant DNA methods, making a gene (or operon, in bacteria) fusion (1). Here instead of a normal gene, a reporter gene (for example, GFP) is controlled by regulatory sequences of the gene of interest.

Figure 3. Schematic representation of a transcriptional fusion. The promoter region of the gene of interest has been cloned upstream of the gene encoding the green fluorescent protein (GFP). Any factors or conditions influencing the expression of the natural gene would affect the expression of GFP.
In this arrangement, anything that ordinarily affects the expression of the natural gene would also affect the expression of the reporter gene. For example, in a particular strain of *E. coli*, gene A is expressed only when bacteria are grown in the presence of arabinose. If a GFP gene is controlled by gene’s A promoter, cells will look green (as on figure 4 to the right) only when grown on medium containing arabinose.

Figure 4. Petri dishes containing *E. coli* (gene A*) in regular medium (left) and medium supplemented with arabinose (right).

Reporter genes offer a big advantage, because one doesn’t need a separate assay for each regulatory region being studied (1). Also, for many reporter genes in unicellular organisms, simple indicator plate assays have been developed. These plate assays allow one to determine levels of gene expression by the color of a bacterial or yeast colony. For instance, a bacterial strain making substantial amounts of beta-galactosidase will form a blue colony on a plate with X-gal; this allows one to identify cells with different levels of beta-galactosidase, and to use X-gal plates to study any regulatory region driving expression of beta-galactosidase. The following table provides examples of commonly used reporter genes/proteins and their activity and detection method (1).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT (chloramphenicol acetyltransferase)</td>
<td></td>
</tr>
<tr>
<td>Transfers radioactive acetyl groups to chloramphenicol; detection by thin layer chromatography and autoradiography.</td>
<td></td>
</tr>
<tr>
<td>GAL (b-galactosidase)</td>
<td>Hydrolyzes colorless galactosides to yield colored products.</td>
</tr>
<tr>
<td>GUS (b-glucuronidase)</td>
<td>Hydrolyzes colorless glucuronides to yield colored products.</td>
</tr>
<tr>
<td>LUC (luciferase)</td>
<td>Oxidizes luciferin, emitting photons, visible.</td>
</tr>
<tr>
<td>GFP (green fluorescence protein)</td>
<td>fluoresces upon excitation with UV light.</td>
</tr>
</tbody>
</table>

Bacteria are excellent model systems to study transcriptional regulation because these organisms can quickly adapt to changes in environmental conditions including shifts in temperature and nutrients. *Escherichia coli* regulates the synthesis of the enzymes necessary for sugar or amino acid metabolism in response to availability of these compounds. In this laboratory exercise, you will investigate the transcriptional regulation of genes involved in amino acid biosynthesis. Specifically, we will be studying two genes: *gltB* which encodes the large subunit of glutamate synthase, and *serA*, a gene encoding D-
3-phosphoglycerate dehydrogenase (4,7). gltB is involved in amino acid metabolism in E. coli cells and catalyzes the following reactions:

\[
\text{L-glutamate} + \text{H}_2\text{O} + \text{NADP}^+ &\rightarrow & 2\text{-ketoglutarate} + \text{ammonia} + \text{NADPH} \\
2 \text{L-glutamate} + \text{NADP}^+ &\leftrightarrow & \text{L-glutamine} + 2\text{-ketoglutarate} + \text{NADPH}
\]

Glutamate synthase uses ammonia as a nitrogen source for the generation of L-glutamate from alpha-ketoglutarate. In addition, glutamate synthase catalyzes the single-step conversion of L-glutamine and alpha-ketoglutarate into two molecules of L-glutamate. In doing so, it simultaneously operates as the major source of L-glutamate for the cell and as a key step in ammonia assimilation during nitrogen-limited growth.

SerA catalyzes the first committed step in the biosynthesis of L-serine from glucose.

\[
\text{3-phosphoglycerate} + \text{NAD}^+ &\leftrightarrow & 3\text{-phospho-hydroxypyruvate} + \text{NADH} + \text{H}^+
\]

Transcriptional regulation of gltB and serA in response to a variety of cellular conditions is mediated by several transcription factors; these include but are not limited to the Leucine Responsive Protein (Lrp), the Catabolic Repressor Protein (CRP), and the Arginine repressor (ArgR). These transcription factors control the gene expression in response to changes in the levels of different cellular nutrients. For example, Lrp controls the expression of genes involved in amino acid biosynthesis and catabolism. It does so by responding to the levels of cellular nutrients, such as amino acids, nitrogen compounds, and sugars. In this laboratory exercise, you will study the activity of a gltB and serA GFP promoter fusion in E.coli cells grown on rich and defined microbial media. LB medium is rich in amino acids and carbon sources while M9 medium is considered “defined” medium containing only one carbon source and the essential amount of nitrogen and trace elements necessary for growth. Where would you expect to observe more gltB or serA promoter activity? Why?

**Lab Objective**

The main goal of this laboratory exercise is to design an experiment to address the following questions: What factors affect the expression of the gltB and serA operons? Do conditions that affect carbon and nitrogen metabolism influence gltB and serA expression? How is the effect of cellular nutrients on gltB and serA regulation mediated? In order to address these questions, we will use gene reporters, specifically transcriptional fusions in which the serA and gltB promoters have been cloned upstream a promoter-less GFP. Plasmids containing these transcriptional fusions have been introduced into E.coli strains with wild type and
mutant backgrounds. Any factor that affects the activity of the \textit{gltB} or \textit{serA} promoter will produce a change in the level of GFP activity. Measuring fluorescence in a special instrument known as a “fluorometer”, will allow you determine the activity of your reporter construct.

References


Experimental Approach

Part I.
To understand the experimental technique used in this Lab and to determine the role of growth media in regulation of \( \text{gltB} \), you will perform fluorescence measurements of cells containing a \( \text{gltBp-GFP} \) and promoterless GFP and grown in the rich and minimal media. By analyzing the results of this experiment, you should be able to answer the following question: Is there a difference in the level of expression of \( \text{gltB} \) when cells are grown in rich and minimal medium?

1) Take one tube of each of the following four cell cultures labeled as WT (\( \text{gltBp-GFP} \)) M9, WT (\( \text{pUA66} \)) M9 and WT (\( \text{gltBp-GFP} \)) LB and WT (\( \text{pUA66} \)). The strain labeled as WT pUA66 contains a plasmid with a promoterless GFP gene. This is the control strain that will allow us to correct for background fluorescence artifacts produced by the cells and the growth media. Add 0.5 ml of each culture to individual tubes containing 2.5 ml of the appropriate media, for example add 0.5 ml of the overnight culture grown on M9 to 2.5 ml of fresh M9 medium.

2) Take 0.150 ml of each diluted culture and place it in a microtiter plate; make sure to have three replicas of each sample (12 samples total). Take initial absorbance and fluorescence measurements of the samples immediately. Incubate the diluted cultures (in the culture tubes, not the plate!!!) by shaking at 225 rpm and 37°C for 2 hours.

3) After incubation (approximately 2 hours), take 0.150 ml of each cell culture to new wells of the same microtiter plate you used before. Make sure to have three replicas of each culture on the microtiter plate.

4) Measure fluorescence and absorbance by using the Fluorometer Microplate reader (you will have 24 wells total). This instrument measures fluorescence by using an optical kit that contains a light source, excitation (490 nm) and emission (510-570 nm) emission filters; absorbance is measured at 600nm. Fluorescence measurements are an indicator of promoter activity, while absorbance measurements indicate microbial cell growth. The data files will be posted on blackboard later during the day. You will analyze the data the following week!

5) During the second week, calculate normalized fluorescence by dividing the raw fluorescence by the absorbance of your sample. Normalized fluorescence is a better measurement of gene expression because it accounts for differences in cellular growth among microbial strains. Be prepared to share and discuss your data during the third
6) Calculate promoter activity by subtracting the normalized fluorescence of the control (promoterless, GFP in pUA66) from the normalized fluorescence of the WT (gltBp-GFP).

7) Record your data by using the table provided in the worksheet.

Part II (To be done while cells are incubating).

The second part of this exercise consists of investigating the gltB and serA transcriptional units. We will use a bioinformatics approach to learn about the function and regulation of these genes. First, choose your favorite promoter, gltB or serA. You will start by looking at NCBI genomes and other two databases, Regulon DB and Ecocyc. These publicly available resources have specific information about E.coli genes and their regulation. In addition, familiarize yourself with the pathways and the regulatory strategies involved in the control of gltB or serA expression, to do this, visit ecocyc (http://www.ecocyc.org/), an electronic encyclopedia of microbial genes and pathways. Once in these databases, type gltB or serA in query box and answer the questions provided in the worksheet.

Procedure

1. Go to NCBI Genomes. Under genomes select bacteria and search for the E.coli MG1655 genome.
2. Once on the E.coli MG1655 genome page, type gltB or serA in the gene query page. Observe the genome region surrounding your gene of choice. Answer the questions in the worksheet.
3. Go to Regulon DB by using the following URL: regulondb.ccg.unam.mx/. Type gltB or serA in query box. Answer the questions in the worksheet.
4. Go to Ecocyc DB by using the following URL: http://www.ecocyc.org. Type gltB or serA in query box. Answer the questions in the worksheet.
Part III – The Second Week

The third part of this exercise consists of formulating a testable hypothesis. What type of questions can you ask about the factors that regulate \textit{gltB} or \textit{serA} expression using the microbial strains and materials available in the lab? In order to figure things out, you need to learn more about \textit{gltB} and its regulation. To do this, you will complete a literature search to examine recent research reports on \textit{gltB} and \textit{serA}. After examining the current literature, you must come up with a research question that could be investigated using the strains and reagents available in the lab. Please complete the literature search and experimental design worksheet. Both these worksheets must be turned in by the end of today’s laboratory period.

Part IV- The Third Week

You will perform the experiments you outlined during the last lab. When you are done, you need to write a paper discussing your results (Value 35 points). The paper should contain Introduction that explains the rationale for your proposed scientific question and your hypothesis in light of the current literature. This section should end with the statement of your research question and hypothesis. In Materials and Methods, you should describe all the methods used in the experiment. Remember to write this section in a passive voice and be as concise as possible. This section should provide information sufficient to repeat your experiment by a professional. Results section of your paper should start with a brief re-statement of your experimental goals and have explanation of all of your figures and tables. Do not forget about figure legends. Figure legends should not repeat the text word by word, but should be understandable without reading the text. Finally, conclude with the Discussion section comparing the results of your experiment to the current literature and pointing out all inconsistencies. You should also include some future directions. The paper should end with re-statement of major conclusions that could be drawn from your experiment.
Appendix XI: Instructions for Database Searching

EcoCyc

1. Type the following URL: http://ecocyc.org/ in the search window of any internet browser.
2. A window containing the home page of the EcoCyc database will appear.
3. Look at the left side of the page to locate the informational section; it is labeled as “About EcoCyc”. Click on “guided tour” or “instructional videos” for detailed examples and information on how to use the database.
4. At any point during your EcoCyc search use the green boxes with questions marks to get information about the content of a particular section, labeling schemes and figure descriptions.
5. Look at the top right corner of the page and locate the “search” tab of the database. Notice the options next to search tab: Quick search and Gene search.
6. Start your search by typing the name of the gene of interest.
7. As an example, type lacZ in the search tab.
8. Click on quick search.
9. A page containing information on the genes, proteins and transcriptional units matching the term lacZ should appear. Read the description provided for each term (genes, proteins and transcription units) to become familiar with the information provided by these links.
10. Next, click on transcriptional units. Notice that lacZ has four transcriptional units. A transcriptional unit is defined as stretch of DNA transcribed into an RNA molecule and encoding at least one gene.
11. Click on the each of the transcriptional units and observe the color boxes. Each box represents a transcription factors that regulates lacZ. Transcriptional activators are represented with green boxes; repressors are shown in magenta boxes.
12. Move the mouse over any of the transcription factor’s (TF) boxes. Read the information provided regarding the distance of the TF’s binding site from the transcription start site and experimental evidence to support these data.
13. Roll the mouse to the right if the diagram to explore the purple arrows representing the genes that are part of the lacZ transcriptional units.
14. Scroll down the page until you get to the area showing the promoter sequences. Examine the promoter region and pay attention to the DNA sequences where the transcription factors are known to bind.
15. Roll the mouse to right of the page until you find the flask and computer icons. Click on this icons to learn about the evidence supporting the data on promoter sequences and binding sites for transcription factors. The flask icon indicates experimental data while the computer icon refers inferences made based on computational analyses.
16. Scroll down the page until you reach the section containing the transcriptional terminators. Notice the DNA sequences their position relative to the transcription start site.
17. Return to the main EcoCyc page. Click on the tab labeled as search, select Gene/Proteins/ RNA from the pull down menu.
18. On the search by protein name tab type: LacZ (doing the search using the gene name tab will take you to the same page).
19. In the next window, click on product name.
20. Scroll down the page to get to genome browser.
21. Click on genome browsers to observe the position of the lacZ gene in the E.coli Genome.
22. Click on the align multi-Genome browser to search for genes homologous to lacZ in other organisms.
23. Select different E.coli strains for genome comparison by checking the tabs next to the organism names.
24. Click submit query.
25. Look at the genome maps of the organisms containing the lacZ homologs.
26. Continue to scroll down the page until you find the enzymatic reaction section.
27. Click on the reaction to view the substrates and products of the reaction catalyzed by LacZ.
28. Return to the main EcoCyc page. Click on the tab labeled as search, select pathways from the pull down menu.
29. On the search tab, type: lactose. How many pathways does this search retrieve? How many involve LacZ?

**Regulon DB**

1. Type the following URL: [http://regulondb.ccg.unam.mx/](http://regulondb.ccg.unam.mx/) in the search window of any internet browser.
2. Roll the mouse to the tab labeled as “Using Regulon DB”. Click on this tab and select “Demos” from the pull down menu.
3. Watch the Demo video before you begin your Regulon DB search.
4. Move to the search tab on the right of the page, select “gene” from the pull down menu.
5. On the search window type “lacZ”, then click on GO.
6. Look at the diagram showing the genome region containing the lacZ gene. Identify the transcriptional regulators that control the gene. The repressors are in red boxes while the activators are shown in blue. Roll the mouse to the left of the page and click on the tab labeled “Object color and code”.
Read the information regarding the colors and symbols used to represent the different components of a transcriptional unit.

7. Scroll down the page to the “external database links” section. What databases can you access using this Regulon DB link?

8. Continue to scroll down the page until you get to the “Product” section. This part of the database contains information on the protein encoded by LacZ. Click on “notes” to access general information on the function of beta galactosidase.

9. Roll the mouse to the section labeled as “external database links”. What protein databases can you access using this Regulon DB link?

10. Move to the bottom part of the page and look for the section labeled as “transcriptional regulation”. Click on “display regulation”, a table with a list of transcription factors and their binding sites will appear. Examine the table to learn about the sequence, genome position and function of the regulators that control the activity of lacZ.

11. Continue to scroll down the page to observe the individual transcriptional units associated with the lacZ in the E.coli genome.

**NCBI Genomes**

1. Start the search by typing the following URL in any internet browser: http://www.ncbi.nlm.nih.gov/sites/genome/

2. Type: coli, MG1655 on the search tab.

3. Scroll down the page and click on the link labeled as NC_000913

4. A page containing information on the features of the E.coli genome will appear. Notice the size of genome, the number of genes and proteins, topology and GC content.

5. Locate the tab labeled as “search gene, GeneID or locus_tag”; type lacZ and press find gene.

6. Look at the diagram illustrating the genome region containing lacZ. Which genes are located near lacZ? Click on lacZ to access information on the protein sequence and function of LacZ.
Note for instructors: These searches can be modified to meet the laboratory exercise goals. Searches can be focused on the transcriptional units or protein sequences encoded by any given gene. Alternatively, the students can utilize EcoCyc to learn about the position of a gene in a genome and the presences of orthologs (genes in different species that evolved from a common ancestral gene by speciation). Orthologs retain the same function in the course of evolution. Identification of orthologs is critical for reliable prediction of gene function in newly sequenced genomes in different species). RegulonDb provides information on regulatory networks, regulons and growth conditions. These parameters can be incorporated into the bioinformatics part of the laboratory exercise. For example, the students can be asked to select “growth condition” from the search menu and the type “amino acids”. This search will provide a list of genes which activity is affected by the presence of amino acids in *E.coli* cells.
Appendix XII: Map for Plasmid pUA66

LOCUS pMS201 4249 bp DNA circular 13-FEB-2000
DEFINITION low copy vector for cloning promoters, contains pLtet01, derived from pZS21-luc.
KEYWORDS surette.
FEATURES Location/Qualifiers
CDS 3514..4230
   /vntifkey="4"
   /label=GFPmut2
   /note="GFP mut 2 coding sequence"
BASE COUNT 1103 a 974 c 909 g 1263 t

BclI  BglII
ctgtcttttgcagatcttttgccctgccccctgcagatctcttgaccggaagaagctaccatcactttgacttcgagggc
^3290  ^3300  ^3310  ^3320  ^3330  ^3340  ^3350
NarI  PvuII  AatII
tcccaacccctacgagggccgccccagcttgccacgtgcataagaaaccattattatcatgacattaccttataa
^3370  ^3380  ^3390  ^3400  ^3410  ^3420  ^3430
AvaI  XbaI  NdeI
XhoI  BamHI  M  S  K
aataggcgatcagagggccctttggtctccacgaggggacctctagttaagaaggagatacatATGTCTA
^3450  ^3460  ^3470  ^3480  ^3490  ^3500  ^3510
One Site
AatII gacgt|c: 3405.
BamHI g|gatcc: 3483.
BclI tigatca: 3290.
BglII a|gatct: 3295.
BssHII g|cgcgc: 2734.
EcoRV gat|atc: 554.
HincII gty|rac: 4003.
HpaI gtt|aac: 4003.
SacI ga|cgt|c: 2449.
SpeI a|ctagt: 1650.
XhoI c|tcgag: 3476.

Multiple Sites
AvaI c|ycgrg: 1962, 3476.
DraI tt|aaa: 466, 990, 1034, 3904.
NaeI gcc|gcc: 200, 2633.
NarI gg|cqcc: 2089, 3133, 3381.
NcoI c|catgg: 2699, 3680, 4204.
NdeI ca|tattg: 532, 3513, 3744.
PstI ctgca|g: 3086, 4238.
PvuII cag|ctg: 3029, 3389.
SphI gcatg|c: 2734, 4244.
XbaI t|ctag: 1, 3489.

No Sites
AccIII tc|ccgga: none.
ApaI gggcc|c: none.
ClaI at|cgat: none.
EcoRI g|aatto: none.
HinIII a|agctt: none.
KpnI ggtac|c: none.
MluI a|ccgct: none.
NheI g|ctagc: none.
NotI gc|gcggcc: none.
NruI cgat|gcga: none.
PvuI cgat|gcg: none.
SacII cccg|gg: none.
SalI gjtcgac: none.
ScaI ag|tact: none.
SmaI ccc|ggg: none.
Stul ag|ct: none.
VspI at|taat: none.
XmaI c|ccggg: none.