Discovering Prokaryotic Gene Regulation with Simulations of the \textit{trp} Operon

by

Audrey Crowther, Heather Bergan-Roller, Nicholas Galt, Christine Booth, Joseph Dauer, and Tomáš Helikar

University of Nebraska-Lincoln

Learning Objectives

1. Perturb and interpret a simulation of the \textit{trp} operon
2. Define how simulation results relate to cellular events
3. Describe the biological role of the \textit{trp} operon
4. Describe cellular mechanisms regulating the \textit{trp} operon
5. Explain mechanistically how changes in the extracellular environment affect the \textit{trp} operon
6. Define the impact of mutations on \textit{trp} operon expression and regulation

The \textit{E. coli} genome contains approximately 4,300 genes that code for metabolic enzymes needed for cellular respiration, transport proteins essential for acquiring nutrients, regulatory proteins needed to control the production of other proteins, and many others. Because protein synthesis requires a tremendous expenditure of energy (ATP), only a subset of the available genes are actively being expressed (turned ON) at any given time (Figure 1). The expression of many of these genes are influenced by external and internal conditions. Natural selection has favored \textit{E. coli} and other prokaryotes that are able to regulate the expression of genes so that they are only expressed when they need to be expressed.

In the activities that follow, you will be exploring the genetic control mechanisms that regulate the production of the amino acid tryptophan, which is needed for protein synthesis, in prokaryotes to exemplify gene expression and regulation present in all organisms.

The Operon Model

In prokaryotes, genes that share a similar function are often clustered together on the chromosome and their expression is coordinately controlled (i.e., if one gene is going to be expressed, all of the genes will be expressed) by a single \textbf{promoter} and \textbf{operator}. This form of gene regulation differs from eukaryotes in that eukaryotic genes are regulated individually. Collectively, the promoter, operator, and protein-coding genes are called an \textbf{operon} (Figure 2).

![Figure 2 Operon Model](image)

The regulatory or control region of the operon consists of the promoter, operator, and binding sites for regulatory proteins (transcription factors) called \textbf{activators} that function to activate transcription. The promoter is a sequence of DNA to which RNA polymerase binds to initiate transcription. The operator is a short sequence of DNA that recognizes transcriptional regulators and is analogous to an ON/OFF switch. Operators typically bind \textbf{repressor proteins} that shut off transcription by blocking RNA polymerase from binding to the promoter. Repressor proteins, along with \textbf{co-repressors}, will be described in more detail in the following activities.

When the operon is turned ON, the genes within the operon are transcribed by RNA polymerase to produce a single mRNA. The mRNA is then translated into individual polypeptides (proteins).
The trp Operon

The trp operon is a cluster of genes that function together to produce the amino acid tryptophan (trp) and is one of the most basic examples of gene regulation in response to changes in both the external and internal environment. Most bacteria obtain tryptophan by either synthesizing it from precursor molecules within the cell or by transporting it into the cell from the environment. For example, E. coli cells living in the gut of an omnivore such as a grizzly bear will experience fluctuations in environmental tryptophan depending on the food sources available to the bear. When the bear is feeding on a rich protein source such as salmon, tryptophan would be readily available to the E. coli cells. However, when the bear is primarily feeding on berries (low protein), environmental tryptophan would be low and the E. coli cells would synthesize their own tryptophan. To do so, the cell must express genes of the trp operon.

Regulation of the trp Operon

The activity of the trp operon is controlled by a regulatory protein called the trp repressor and by intracellular levels of tryptophan (an amino acid), which acts as a corepressor. The trp repressor is produced from the regulatory gene called trpR that is located upstream of the trp operon. Cells can respond rapidly to changes in cellular tryptophan levels because the trpR gene is continuously expressed. When cellular tryptophan levels are low, RNA polymerase is able to bind the DNA at the trp promoter and transcribes the protein-coding genes of the trp operon (Figure 3). The five genes are transcribed into a single mRNA that is translated into five individual proteins. TrpE and trpD (enzyme 1), trpC (enzyme 2) and trpB and trpA (enzyme 3) form three enzymes that produce tryptophan from precursor molecules within the cell.

When tryptophan levels are high, the tryptophan binds to and activates the trp repressor. Activation causes the trp repressor to change shape and allows it to bind to the DNA in the trp operator. This binding shuts off the trp operon by blocking the RNA polymerase (Figure 4). The trp operon is considered a repressible operon and is an example of negative gene regulation because it is actively being transcribed (trp operon is considered ON) in the absence of its regulatory protein, the trp repressor. The trp repressor is required to stop transcription of the trp operon (trp operon is considered OFF).

Further, the trp operon demonstrates feedback inhibition at the level of gene expression. When tryptophan levels are low, the operon is ON which leads to the production of tryptophan. When tryptophan accumulates to a sufficient level, it activates the trp repressor which then inhibits further production of tryptophan. As the cell translates other proteins, it will use up its store of tryptophan. When tryptophan levels drop, the trp repressor will no longer bind tryptophan and no longer block the RNA polymerase from transcribing the trp operon. This cycle will continue indefinitely unless the cell is able to acquire an adequate amount of tryptophan from the environment, which would then inhibit the trp operon.

Throughout this lesson on the trp operon, you will be asked to provide mechanistic explanations to justify your predictions/conclusions. What is a mechanistic explanation? A mechanistic explanation explains HOW something happens by including the components involved and how those components behave or interact. For example, when tryptophan levels are high, the trp operon is OFF; but HOW does that happen? Here is an example of an appropriate mechanistic explanation: tryptophan binds to and activates the trp repressor; the trp repressor then binds to the promoter region of the trp operon and blocks transcription of the trp operon by RNA polymerase.
Homework Questions

Complete the following questions:

1. What is tryptophan?

2. What cellular signal does the *trp* operon detect?

3. Collectively, what is the function of the proteins encoded in the *trp* operon?

4. Use your own words to define a repressible operon.
**trp Operon Simulation**

**Dynamic Simulation**
In this activity you will be using a computational modeling and simulation software called the Cell Collective to explore the components and dynamics of the trp operon. Figure 5 represents a simplified computational model of the trp operon that was built using the Cell Collective. Before running the simulation, it is important that you understand the components and properties represented in the model (Figure 5). When building models in the Cell Collective, arrows represent positive regulation (activation) and blunted lines represent negative regulation (inhibition). This convention allows scientists to build models that incorporate different types of regulation. In a biological context, however, the interactions between components of a system are diverse and complex. Therefore, arrows connecting components in models like shown in Figure 5 often represent diverse interactions and relationships. To begin, complete Table 1 by describing/defining what each arrow and blunted line in the model represents.

![Figure 5 Diagram of trp Operon Computational Model](image)

| Table 1 Describe/define what each interaction represents in Figure 5. |
|---------------------------------------------------------------|
| a. The trp operon is transcribed by RNA polymerase to produce trp_mRNA. |
| b. The trp_mRNA is ________ into proteins, which assemble to form the _________. |
| c. The trp_enzymes are ________ for tryptophan synthesis from precursor molecules within the cell. |
| d. |
| e. |
| f. |
**trp Operon Simulation**

In this activity you will be simulating how the external environment and intracellular (inside the cell) conditions influence the expression of the *trp* operon.

**Part 1: Access to the Dynamic Model of Cellular Respiration**

a. Go to learn.cellcollective.org
b. Register and login to Cell Collective.
c. From the Home page, select “Prokaryotic Gene Regulation: the trp operon” (Figure 6).
d. Click “Model” on the top middle menu bar.
e. Within the “Model” tab, click “Simulation” to enter the simulation workspace (Figure 7). In this workspace you will be able to observe dynamic behavior of the *trp* operon.
f. You are now ready to simulate the model.

**Part 2: Simulation Setup**

a. To begin, adjust the “Sliding Window Size” to 1 in the Simulation Control panel (Red box, Figure 7) and in the Internal Components panel (Blue box, Figure 7), check the box under the “ ” icon next to “trp_operon”. Do NOT adjust External Components activity levels at this time.
b. Click the play ( ) button under Simulation Control to start the simulation.
c. You can now evaluate the activity (shown on the y-axis) of the *trp* operon when no tryptophan is present in the external environment.
d. To see the activity of other components in the model just check the “ ” icon box next to its name. By checking the components in the order in which they are activated you will be able to observe the dynamics of the *trp* operon.
e. Reset your simulation by clicking the stop ( ) button under Simulation Control.
f. Continue this activity by completing the questions on the next page.
Investigation 1:

A. Prediction. How does the presence of tryptophan in the environment influence the expression of the genes in the \( trp \) operon? (For each prediction select one of the options below)

1. The \( trp \) operon will be active (ON) inactive (OFF) alternating between active and inactive
2. The \( trp \) repressor will be active inactive alternating between active and inactive

B. Defend your prediction. To support your prediction, use the information provided in Figure 3-4 (Pg2) and Figure 5 (Pg3) to complete a mechanistic description in the space below (i.e., circle either active or inactive, and fill in the blanks). This mechanism describes \textbf{HOW} the components involved would interact when environmental tryptophan is available to a cell.

The tryptophan (trp) would enter the cell, bind to and \textbf{activate/inactivate} the __________. The trp repressor would then bind to the __________. As a result, the trp operon would become \textbf{active/inactive} because the trp repressor would block the __________ from transcribing the genes in the operon.

C. Test your prediction

1. In the \textbf{External Components} panel, adjust the “environmental_trypophan” slider to 100.
2. Monitor the activity levels of the \( trp \) repressor and \( trp \) operon by checking the box under the “○” icon next to “\textbf{trp_repressor}” and “\textbf{trp_operon}”.
3. Start the simulation by clicking the play (▶) button.

D. Record the results. Record your simulation results by selecting one of the options for each component you observed.

1. The \( trp \) operon is active (ON) inactive (OFF) alternating between active and inactive
2. The \( trp \) repressor is active inactive alternating between active and inactive

E. Evaluate your prediction. Do your simulation results match your prediction? (circle one) Yes No

If your prediction was not correct, continue to play with the simulation to understand the following:
1) how simulation results translate to events inside the cell
2) how environmental tryptophan affects the \( trp \) operon

F. Describe the mechanism correctly. Describe what your results indicate is occurring in the cell (use your results to support a mechanistic explanation). Specifically, describe \textbf{HOW} the components of the \( trp \) operon system are affected by adding environmental tryptophan.
Investigation 2:

A. **Prediction.** Predict the activity of the *trp* operon if tryptophan was no longer supplied to a cell by the environment. (For each prediction select one of the options below)

   1. The *trp* operon will be active (ON) inactive (OFF) alternating between active and inactive
   2. The *trp* repressor will be active inactive alternating between active and inactive

B. **Defend your prediction.** To support your prediction, use the information provided in Figure 3-4 (Pg2) and Figure 5 (Pg3) to complete a mechanistic description in the space below. This mechanism should describe **HOW** the components involved would interact when environmental tryptophan is no longer available to a cell.

C. **Test your prediction.**

   1. In the **External Components** panel, adjust the “environment_tryptophan” slider to 0.
   2. Monitor the activity levels of the “trp_repressor” and “trp_operon”.
   3. Start the simulation.

D. **Record the results.** Record your simulation results by selecting one of the options for each component you observed.

   1. The *trp* operon is active (ON) inactive (OFF) alternating between active and inactive
   2. The *trp* repressor is active inactive alternating between active and inactive

E. **Evaluate your prediction.** Do your simulation results match your prediction? (circle one)  Yes  No

   If your prediction was not correct, continue to play with the simulation to understand the following:

   1) how simulation results translate to events inside the cell
   2) how environmental tryptophan affects the *trp* operon

F. **Describe the mechanism correctly.** Describe what your results indicate is occurring in the cell (use your results to support a mechanistic explanation). Specifically, describe **HOW** the components of the *trp* operon system are affected by removing environmental tryptophan.
Investigation 3: The trpR gene, which is upstream of the trp operon, codes for the trp repressor and is continuously expressed to produce the trp repressor protein.

A. Prediction. Predict the activity of the trp operon and trp repressor if a cell acquires a mutation in its trpR gene that no longer allows its gene product, the trp repressor protein, to bind to the trp operator.

1. The trp operon will be active (ON) inactive (OFF) alternating between active and inactive
2. The trp repressor will be active inactive alternating between active and inactive

B. Defend your prediction. To support your prediction, use the information provided in Figure 3-4 (Pg2) and Figure 5 (Pg3) to complete a mechanistic description in the space below. This mechanism should describe HOW the components involved interact.

C. Test your prediction.

1. In the External Components panel, adjust the “trp_repressor_mutation” slider to 100
2. Monitor the activity levels of the “trp_operon” and “trp_repressor”
3. Start the simulation.

D. Record the results. Record your simulation results by selecting one of the options for each component you observed.

1. The trp operon is active (ON) inactive (OFF) alternating between active and inactive
2. The trp repressor is active inactive alternating between active and inactive

E. Evaluate your prediction. Do your simulation results match your prediction? (circle one) Yes No

If your prediction was not correct, continue to play with the simulation to understand the following:
1) how simulation results translate to events inside the cell
2) how a mutation in the trpR gene affects the trp operon

F. Describe the mechanism correctly. Describe what your results indicate is occurring in the cell (use your results to support a mechanistic explanation). Specifically, describe HOW the components of the trp operon system are affected by the mutation.
Investigation 4: The trpC gene is part of the trp operon and codes for an enzyme important for the synthesis of tryptophan.

A. Prediction. In a situation where tryptophan is NOT supplied to a cell by its environment, predict the activity of the trp operon if that cell had acquired a mutation that made the trpC gene product (the enzyme) nonfunctional.

1. The trp operon will be active (ON) inactive (OFF) alternating between active and inactive
2. The trp repressor will be active inactive alternating between active and inactive
3. Tryptophan synthesis will be active inactive alternating between active and inactive
4. The trp mRNA will be present absent alternating between present and absent
5. Tryptophan will be present absent alternating between present and absent

B. Defend your prediction. To support your prediction, use the information provided in Figure 3-4 (Pg2) and Figure 5 (Pg3) to complete a mechanistic description in the space below. This mechanism should describe HOW the components involved interact.

C. Test your prediction.

1. In the External Components panel, adjust the “trpC_mutation” slider to 100
2. Monitor the activity levels of the “trp_operon”, “trp_mRNA”, “trp_enzymes”, “trp_synthesis”, “tryptophan” and “trp_repressor”
3. Start the simulation.

D. Record the results. Record your simulation results by selecting one of the options for each component you observed.

1. The trp operon is active (ON) inactive (OFF) alternating between active and inactive
2. The trp repressor is active inactive alternating between active and inactive
3. Tryptophan synthesis is active inactive alternating between active and inactive
4. The trp mRNA is present absent alternating between present and absent
5. Tryptophan is present absent alternating between present and absent

E. Evaluate your prediction. Do your simulation results match your prediction? (circle one) Yes No

If your prediction was not correct, continue to play with the simulation to understand the following:
1) how simulation results translate to events inside the cell
2) how a mutation in the trpC gene affects the trp operon

F. Describe the mechanism correctly. Describe what your results indicate is occurring in the cell (use your results to support a mechanistic explanation). Specifically, describe HOW the components of the trp operon system are affected by the mutation.
Insight: Attenuation

At first, scientists believed that regulation of the *trp* operon was only controlled by the repressor-operator interaction. However, later experimental findings indicated a secondary mechanism of control. This secondary mechanism is now known as **transcription attenuation**, or attenuation for short.

During the process of attenuation, transcription is halted before RNA polymerase reaches the protein-coding genes of the *trp* operon. This is possible because transcription and translation happen simultaneously in prokaryotes. Before transcription of the *trp* operon genes, RNA polymerase must first transcribe a segment of DNA called the **leader region**, located in between the operator and the *trpE* gene. The leader mRNA contains self-complementary regions (colored green, purple, yellow, and blue segments in Figure 8) that fold into different hairpin structures depending on the availability of tryptophan in the cell. Depending on which hairpin forms, transcription will either continue or terminate before reaching the *trpE* gene. When tryptophan levels are high, the ribosome translates the leader mRNA quickly, leading to the formation of the **terminator** hairpin (Figure 8A). This terminator hairpin causes RNA polymerase and the mRNA chain to be released from the *trp* operon, terminating transcription. When tryptophan levels are low, the ribosome stalls while translating the leader mRNA. This encourages the formation of the **antiterminator** structure, allowing for transcription of the *trp* operon to continue (Figure 8B).

**Data Analysis**

You are a researcher in the late 1960’s examining the *trp* operon. Your colleagues have found that the *trp* operon is controlled by a repressor-operator interaction; however, these findings do not fully explain the phenomena of regulation in the *trp* operon. Your hypothesis is that *trp* expression is controlled by an additional regulatory mechanism (regulation of the *trp* operon by attenuation has not been discovered yet).

You have created two mutant strains that contain mutations in the leader region of the *trp* operon (Mutant 1 and Mutant 2) with the goal to demonstrate that this region is important for regulation of the operon. You decide to compare the production of *trp* enzyme anthranilate synthase, which is dictated by the levels of *trpE* expression, in these mutant strains to wild-type (“normal”) cells. You produce the following table (Table 2) summarizing your results. Use Table 2 to answer the questions on the next page.

| % activity of *trp* operon enzyme | Anthranilate Synthase (*trpE*) |
|-----------------------------------|--------------------------------|
| Low Tryptophan                    |                                |
| Wild type                         | 100%                           |
| Mutant 1                          | 21-44%                         |
| Mutant 2                          | 31-46%                         |
| High Tryptophan                   |                                |
| Wild type                         | 4-10%                          |
| Mutant 1                          | 5-10%                          |
| Mutant 2                          | 5-10%                          |

Table 2. Enzyme activity from wild-type cells in a low-tryptophan environment were set as 100% activity. 2ug/ml of tryptophan was added in the low tryptophan environment; 50ug/ml of tryptophan was added in the high tryptophan environment. Modified from Hiraga et al., 1967.
1. You know that the \textit{trp} operon is regulated by the repressor-operator interaction. Compare the activity of wild type under low tryptophan and high tryptophan. Do your data reflect this type of regulation on the \textit{trp} operon? Explain why below.

2. When looking at your data, under which conditions do you notice the biggest difference in enzyme expression between the wild-type and mutant strains? (Circle one)

| Condition          | Wild-type | Mutant | Difference |
|--------------------|-----------|--------|------------|
| Low tryptophan     |           |        |            |
| High tryptophan    |           |        |            |

3. Should the \textit{trp} repressor be active or inactive under the environmental conditions you selected above? Does this indicate that the \textit{trp} repressor is involved in decreasing Mutant 1 and Mutant 2’s enzyme activity in comparison to wild type? Why or why not?

4. Do your data support the hypothesis that there is a secondary mechanism regulating \textit{trp} operon expression? How do the data either support or discredit this hypothesis?

5. With your current knowledge of attenuation, how do you think the mutations in Mutant 1 and Mutant 2 impact the leader region of the \textit{trp} operon to generate the data in Table 2?

6. Another lab has also published findings involving mutations of the \textit{trp} operon leader region. This lab, however, has created a mutant strain that expressed an increase in \textit{trp} operon activity under low tryptophan conditions compared to wild type. Explain how a mutation in the leader region can lead to this increase in \textit{trp} expression.

**Discussion question**

We now know that the leader region encodes for a peptide (a chain of 2-10 more amino acids) that contains two tryptophan amino acids. You also know that the formation of the antiterminator vs. terminator hairpin is dependent on the ribosome stalling during translation of the leader region, but what causes the ribosome to stall? With these two pieces of information, describe the mechanism of how tryptophan levels regulates transcription of the \textit{trp} operon by means of transcription attenuation.

**References**

Hiraga S, Ito K, Hamada K, Yura T. 1967. A new regulatory gene for the tryptophan operon of Escherichia coli., Biochem Biophys Res Commun, 26:522-7.