Gene Regulation:
Positive and Negative Control

[redacted]

INTRODUCTION:

A gene can be regulated by a variety of different mechanisms. The two most understood mechanisms are those involving a negative mode of control and a positive mode of control. In a negative mode of control, the removal of a restraining element permits expression from a promoter, whereas in a positive mode of control, a stimulatory element is needed to induce the expression of the gene (Fig. 1). Both these modes of control perform the same function, so why are some genes regulated by a negative mode of control while others or control by a positive mode? In a multi-compartment model, I will explore this question and determine the condition for the selection of one mode of gene control over the other. The model keeps track of the population of the wild type organism in comparison with different mutant populations to see under what condition the wild type population is overcome by the mutant population, which signals the loss of that mode of control. (Savageau, 1974)

The model is patterned after the life cycle of *Escherichia coli* and more specifically, *E. coli*'s lactose and maltose operons. The lactose operon is controlled by the negative mode of control. To induce the expression of β-galatosidase from the lactose operon, the repressor protein, called the lac repressor, must bind to the inducer, in this case, lactose. Once this occurs, the lac repressor changes it's conformation and no longer can suppress the initiation of transcription. This chain of events induces the production of β -galactosidase, the *E. coli* enzyme that breaks down lactose for its use. (Savageau, 1998, II)

In contrast, the maltose operon is controlled by a positive mode of control. The inducer, maltose, must bind to the activator protein, malT. This allows the complex,

malT-maltose, to interact with RNA polymerase and facilitate initiation of transcription of enzymes necessary for *E. coli's* utilization of maltose. (Schwartz, 1987)

In this model, a complete life cycle is defined as the entry into a host, colonization of that host, and finally the exit from the host and the entry of a new host. The host for an *E. coli* is the mammalian digestive system. The digestive tract is divided into two parts: the proximal portion, which is the stomach to the beginning of the small intestines, and the distal portion, which includes the end of the small intestine to the colon. The host digests lactose at the end of the proximal portion of the stomach and maltase at the end of the distal portion of the stomach (Fig. 2). To exploit the host's nutrients, *E. coli* uses the lactose as its energy source when it is in the proximal portion of the digestive track and maltose when it is in the distal region of the digestive system. Thus, *E. coli* utilizes these carbon sources before the host is able to digest and absorb the lactose and maltose for itself. (Savageau, 1974)

β-galactosidase, a gene product of the lactose operon, is in high demand in the proximal portion since there is a high concentration of lactose (inducer) in this region. However, the lac genes are in low demand in the distal portion of the digestive system, where there is little if any lactose. It takes 3 hrs for *E. coli* to pass through the high demand environment for lactose utilization. The demand for the products of the maltose operon is the opposite, low demand in the proximal portion and high demand in the distal portion, corresponding to the location *E. coli* starts utilizes maltose as an energy source. It takes about 6 hours to pass through the low demand environment for the utilization of maltose. (Savageau, 1998, II)

Along with understanding the maltose and lactose operons and the alternation between the high and low demand environments in *E. coli's* life cycle, the different phenotypes the mutants exhibit are also important in understanding the model. There are many mutations that can obstruct the production of the needed enzyme but this model will only address mutations in the promoter site, the modulator site, or in both sites of the operon. Each mutation will exhibit a different phenotype, depending on the mode of control and the environment the organism is in (Savageau, 1998, I). The phenotypes will in turn affect the growth rate of that type of mutation if it causes constituitive expression or no expression in both environments. This is further discussed later in each model.

In negative regulation (lactose operon), a high level promoter is needed for full expression of the necessary enzymes when the repressor is removed, while a functional modulator is needed for the inhibition of the expression when the repressor is present. Thus, promoter and double mutants will have no expression with or without the inducer, since both has a mutation which no longer gives a high level promoter. In contrast, the modulator mutant can produce β-galactosidase when the inducer is present but is unable to shut down synthesis when the inducer is gone since it doesn't have a functional modulator. (Fig 3. A & B) Both these defects will affect the growth rate of these mutants.

In positive regulation (maltose operon), a low level promoter is necessary for expression to be turned off upon removal of the activator, while a functional modulator is needed for the full expression in the presence of the activator-inducer complex. Therefore, the promoter and double mutants are constituitively expressing the maltose enzymes. In a high demand environment, the modulator mutant is unable to produce

these enzymes although it should and in a low demand, it also isn't producing the enzymes since it has a functional promoter that enables the shut down of the operon in the presence of the inducer-activator complex. A normal (wild type) *E.coli* should have expression in the high demand environment and inhibit expression in the low demand environment. (Fig. 3. C & D)

The muti-compartment model will take into account the environment (high or low demand), the growth rate, and the mutation rates of each population to monitor the changes in the population sizes. The population sizes will in turn determine if the wild type population will be selected for and thus the preservation of that regulatory mechanism. Otherwise, if one of the mutant populations overcomes that of the wild type population, the regulatory mechanism is considered lost. The purpose of this model is to determine if the mode of regulation is arbitrarily picked or if there are specific conditions where a positive regulation or a negative regulation is selected.

BASICS OF MODEL:

The model will be based on the following population dynamic equations:

$$d(Xw) / dt = Aww * Xw$$
$$d(Xp) / dt = (Apw * Xw) + (App * Xp)$$
$$d(Xm) / dt = (Amw * Xw) + (Amm * Xm)$$
$$d(Xd) / dt = (Adm * Xm) + (Adp * Xp) + (Adm * Xd)$$

Subscripts should all be defined.

"X" represents the population of the respective type, indicated by the lower case letter after the X. For example, Xw is the population size of the wild type. "A" represents the ratio of one population that mutated into another population. For example, Apw is the ratio of wild type population, which mutated to become a promoter mutant. (Savageau, 1998, I)

So Aww means what + neg?

"A" values are dependent on the mutation rates and growth rates:

Aww = [1- (Mpw + Mmw)] * Gw
Apw = Mpw * Gw
App = (1-Mdp) * Gp
Amw = Mmw * Gw
Amm = (1 – Mdm) * Gm
Adm = Mdm * Gm
Add = Gd

At first glance this does not seem consistent with saying "A" represents the ratio of 2 diff. mutated populations.

The mutation rate (M) determines the rate at which a member of one population mutates into a member of another population (Fig. 4). The two letters after the M will indicate what the organism has become and what it was before the mutation. For example, Mpw means the mutation rate from wild type organism to a promoter mutant. The growth rate (G) also has a letter after it, which indicates the population it represents. Thus, Aww is equal to the Gw times what's left of the wild population after subtracting the ones, which mutated into promoter (Mpw) and modulator (Mmw) mutants. The other A equations follow the same logic. (Savageau, 1998, I)

The growth rates (G) and mutation rates (M) are all constants; however, G and M changes between two values depending if the organism is in a high or low demand environment. Thus, a switch was created to alternate between the two values as the environment switches from a high demand to a low demand and vice versa. The switch is defined as:

$$d / dt\ (hswitchl) = pulse\ (1, 0, C) - pulse\ (1, H, C)$$

where C is equal to the time it takes for a gene to complete an on and off cycle and H is equal to the period in which the gene is turned on. H is in turn dependent on D which is the fraction of the cycle time (C) when the gene is on. The hswitchl reservoir is filled to one whenever the environment is a high demand environment and empties itself once

high demand (the period H) has passed. The hswitchl remains equal to zero until a new cycle C begins. By making the M values and G values dependent on the hswitchl value, it enables the mutation rates and growth rates to switch to the appropriate value automatically as the environment alternates between a high demand and a low demand environment. (Graph 1)

Lastly, the model consists of the equations for the ratio of mutant population to that of the wild type. The ratios are used to analyze whether the regulation mechanism of the wild type population is selected or lost.

MODEL 1: Negative Regulation-Lactose Operon

Since the time period in the high demand environment (H) is equal to 3 hours in the lactose operon as stated in the introduction, C can be derived:

$$H = D * C = 3$$
$$C = 3 / D$$

We also know that the first portions of the digestive system is a high demand environment, while the second half is a low demand environment for the lactose operon. (Savageau, 1998, II) Thus, M and G are equal to:

M = IF hswitchl > 0 THEN high demand value
ELSE low demand value
G = IF hswitchl > 0 THEN high demand value
ELSE low demand value

Thus, the environment switches from first a high demand to a low demand environment.

The curve that determines whether the negative regulation mechanism is kept is that of the ratios of the mutant to the wild type population. When D is set at 0.5, the curve can be seen to alternate between two slope. This is seen for all three ratios

(Xp/Xw, Xm/Xp, and Xd/Xw). The curve for the ratios of Xp/Xw and Xd/Xw are identical since from Fig. 3, we see these two mutants exhibit the same phenotype in the low and high demand environment. The first slope corresponds to the high demand environment and the second slope corresponds to the low demand environment.

As seen from Graph 2, the curves for Xp and Xd ratios have a negative first slope, while the second slope is zero. This occurs because in a high demand environment, the two mutants are unable to express the desired genes since they both have a defective promoter. Thus, the wild type population, which is able to make the appropriate lactose enzymes will eventually forces the promoter and double mutant populations into extinction. The second slope is zero because the mutants exhibit the same phenotype as the wild type in the low demand environment.

Graph 3 is the curve of the modulator mutant ratio at D = 0.5. The first slope in this case is zero while the second slope is a negative slope as predicted by the phenotypes the modulator mutant expresses in the two different environment. In the high demand environment, the modulator mutant is able to produce the necessary enzymes like the wild type so grow at the same rate as the wild type population. However, in the low demand environment, the modulator mutant is unable to shut down expression of the now no longer needed enzymes. Thus, this waste of its energy puts it at a disadvantage to the wild type organism, causing the slope to change to a negative slope.

Although Graph 2 and 3 proves that the model does work, we have not determine what conditions cause the selection of the negative mode of control for the lactose operon. To do this, I chose one mutant ratio, which had the highest overall slope and thus, the best chance of overcoming the wild type population. When all three ratios are

graphed on the same plot, we see the modulator mutant has the highest slope (Graph 4). This ratio was used to determine the condition for the selection of the wild type mechanism.

Through parameter runs, I determined the most sensitive parameter was the demand. Graph 5 shows how the Xm/Xp curve changes with high or low demand. The overall slope decreases as the demand decreases from 0.9 to 0.1. The more positive this curve is indicates the mutant population is rising quicker than the wild type population and with time the wild type regulation mechanism will be lost. However, if this curve has an overall negative slope, the wild type population is increasing quicker than the mutant population and thus the negative regulation is selected. Therefore, from this graph, we can see the condition for selection of the negative regulation is low demand.

MODEL 2: Positive Regulation-Maltose Operon

In the maltose operon, as stated in the introduction, it takes 6 hours to pass through the low demand environment. Thus, C is different from that in the lactose operon model.

$$L = (1-D) * C = 6$$
$$C = 6 / (1-D)$$

The value of M and G above are also different for the maltose operon since in this case, the first portion of the digestive track is a low demand environment and the second half is a high demand environment. (Savageau, 1998, II)

M = IF hswitch > 0 THEN low demand value
ELSE high demand value
G = IF hswitch > 0 THEN low demand value
ELSE high demand value

Thus, the positive regulation model alternates between first a low demand and then a high demand environment.

Again, the ratios of mutant to wild type populations were graphed. Like in the negative regulation, the Xp and Xd ratios gave the same graph since both mutants constituitively expressed the mal genes in both the low and high environment. From Graph 6, we can see the first slope is negative while the second slope is positive. This indicates in a low demand environment, the mutants are at a disadvantage since they are wasting energy to make enzymes that are not needed. Thus, the wild type population rises quicker then both mutant populations in a low demand environment. The second slope describes the fight between the mutants and the wild type in a high demand environment. The positive slope indicates the promoter and mutant populations are increasing quicker than the wild type population in the high demand environment. However, since the overall curve is decreasing with time, the increase in slope in the high demand environment is not enough to compensate for the low growth in the low demand environment. When a parametric run was done on the parmaeter D, we see the overall slope decreases as the value of D increases (Graph 7). This reveals that positive regulation is selected for in high demand environments.

Unlike the other two mutants, the modulator mutant has an overall increasing slope (Graph 8), and thus at demand 0.5, the positive mode of regulation is lost. The first slope of the curve is zero because the phenotype of the modulator mutant resembles that of a normal wild type organism. The second slope is positive indicating the positive regulation is lost at a demand of 0.5. By varying the value of D between 0.9 to 0.1, the curve of the Xm/Xp curve decreases with the increasing value of demand (Graph 9). The

curve at 0.9 is close to being linear, thus only at high demand is the positive mode of regulation not lost. This result agrees with the results of Graph 7.

CONCLUSIONS:

From the negative regulation (lactose operon) model and the positive regulation (maltose operon) model, the condition for the selection of one control mode over another was determined. A negative mode of control is selected when the demand for the genes of the operon is in low demand, while positive regulation is selected when the demand is high. This shows a mode of regulation is chosen depending on the environment the organism is in and thus, this decision is not arbitrarily made. The model can be used to predict the demand in the environment for other operons when the nature of regulation is known or vice versa. These predictions proved to be very accurate as seen in Savageau's article, "Design of molecular control mechanisms and the demands of gene expression".

If time had permitted, this model could have been expanded through more research. By finding out the criteria (exact slope of the ratio curve) for the selection of the wild type control mechanism, the model can then be used to determine the exact range of demand where the selection for the positive and negative mode of control is realizable. Once the range of demand is determined, we can solve for C and find the time it takes to complete an organism's life cycle. With more work, the two models can produce more specific results and actual numbers such as the range of demand where selection is realizable and the period of time to complete the life cycle of an organism.

Figures, Tables and Graphs:

Figure 1: Schematic representation of negative-controlled and positive-controlled operons. (A) In the negative control operon, the repressor prevents expression of the gene in the absence of the inducer but when the inducer is present the repressor can no longer bind to the modulator site so expression of the gene is seen. (B) In the positive control operon, the activator is unable to bind to the modulator without the inducer so no transcription occurs without the inducer present.

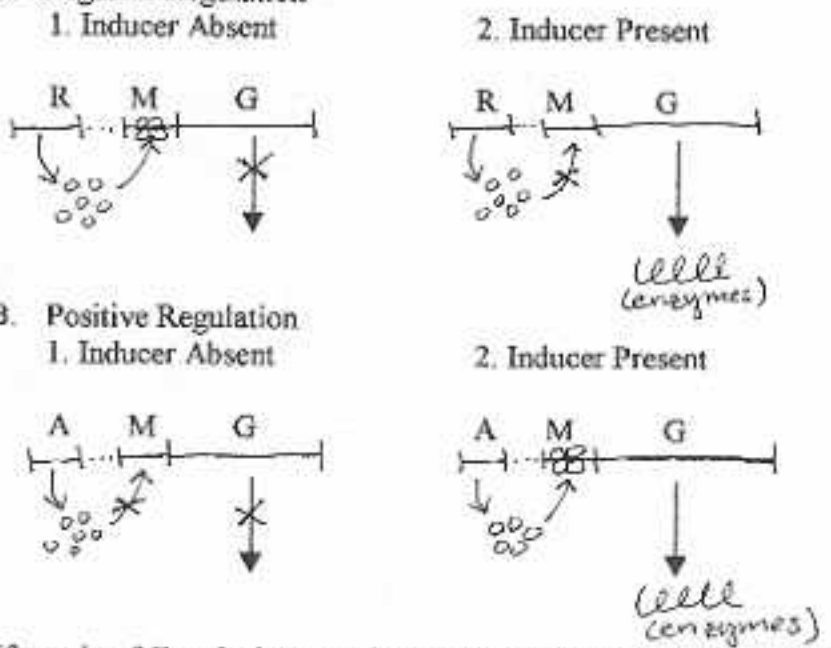


Figure 2. The life cycle of *E. coli* alternates between two different environments: (A) the proximal portions of the digestive tract (stomach to the beginning of the small intestines) where there is a high demand for the lac genes and low demand for the mal genes and (B) the distal portions of the digestive tract (small intestines to the colon) where there is low demand for the lac genes and high demand for the mal genes.

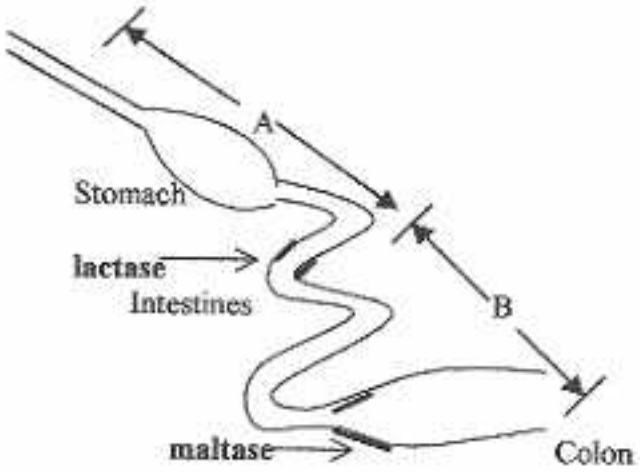


Figure 3: Expression of gene depends on the mode of control and the environment. In the negative mode of regulation, the figure shows the phenotypes of the wild type and mutants in a high demand environment (A) and a low demand environment (B) using symbols R = repressor, M = modulator, P = promoter, and G = desired genes. In the positive mode of control, the low demand environment (C) and the high demand environment (D) are also depicted using the symbol A = activator along with the other symbols above. The arrow symbolizes the expression of the desired genes.

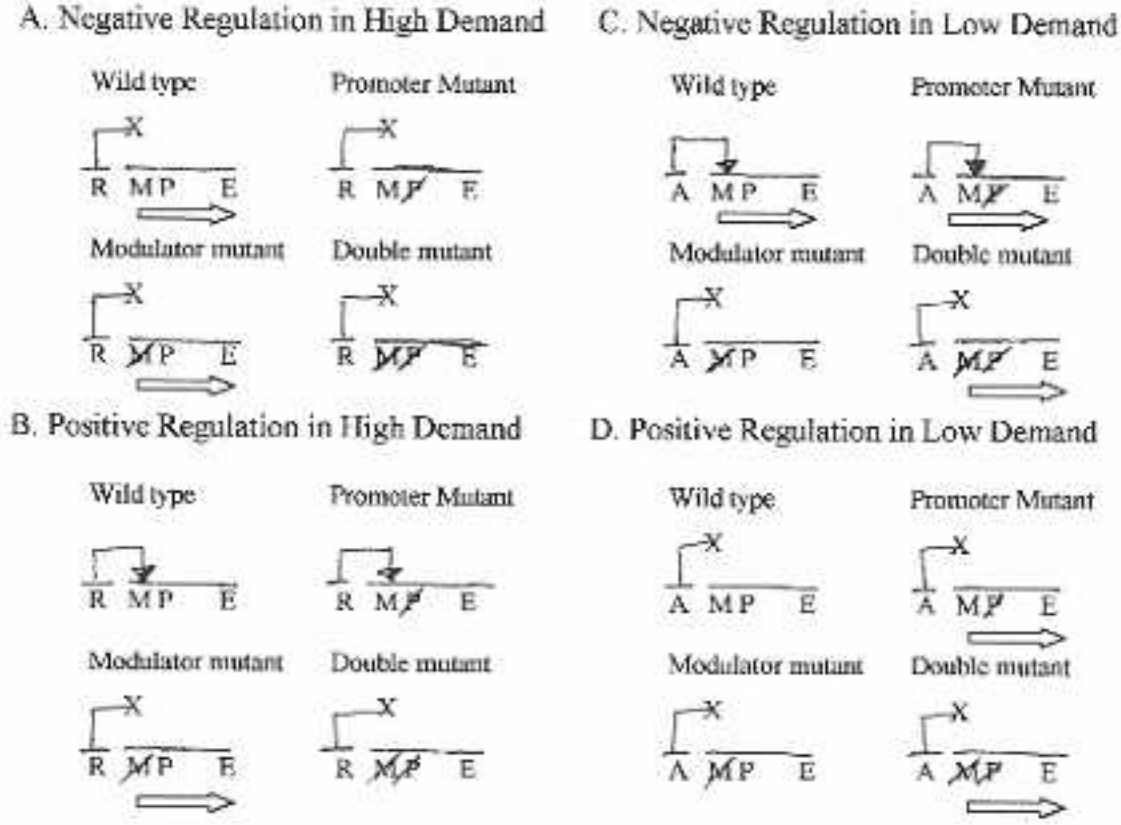


Figure 4: Schematic diagram showing how the population sizes are affected by the mutation rates and growth rates.

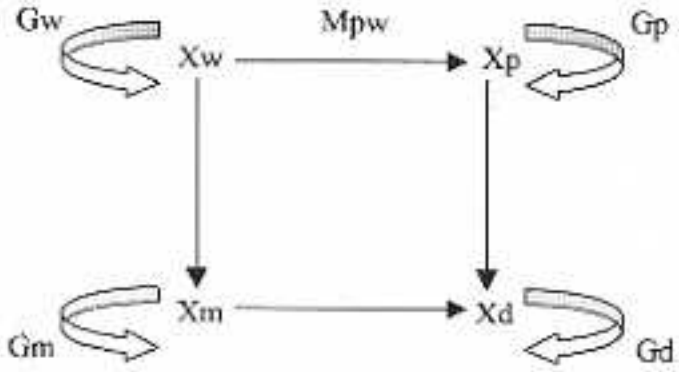


Figure 5: The diagram of the four population reservoirs, which shows the relationship of the parameters and variables.

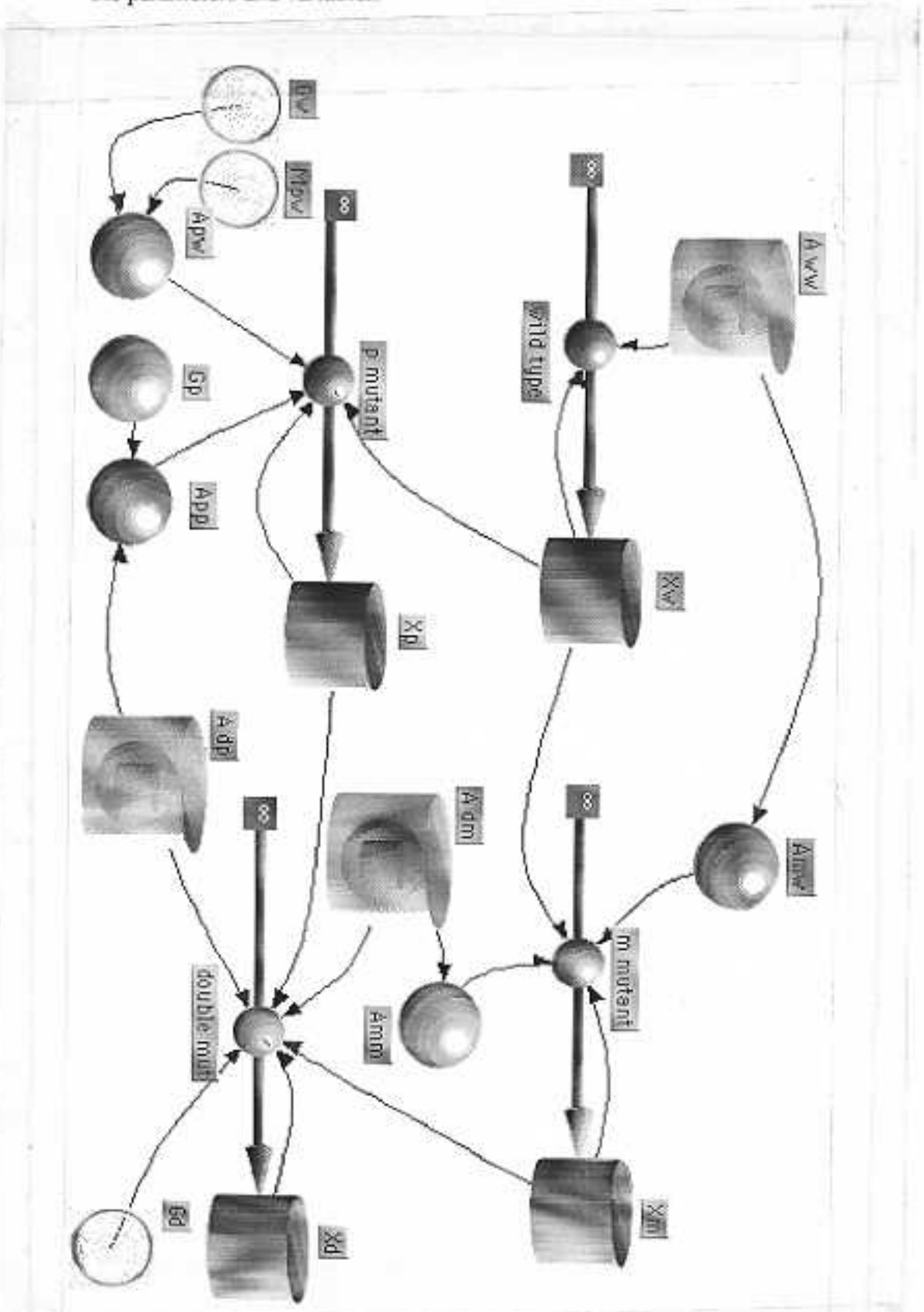


Table 1: Relationships and Parameters for Negative Regulation-Lactose Operon (G = generation/hr, M = base-1 gen-1, C = hrs).

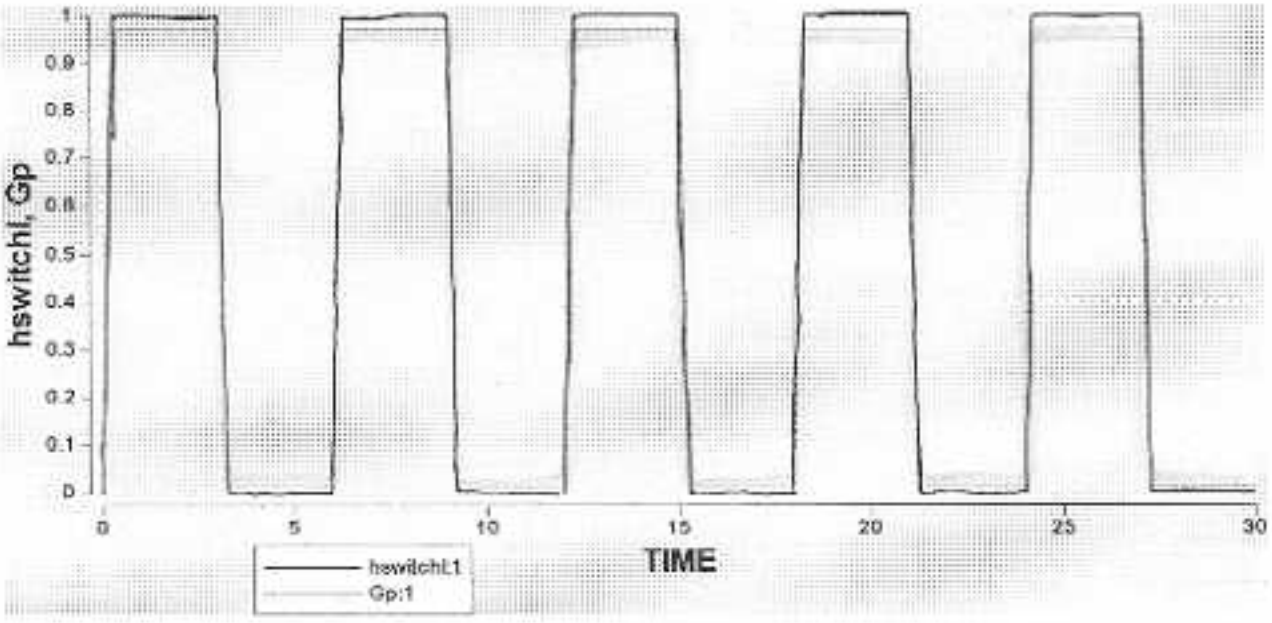
```
d/dt (hswitchl) = flowin - flowout
  INIT hswitchl = 0
flowin = pulse (1, 0, C)
flowout = pulse (1, H, C)
H = D*C
L = (1-D)*C
D = .5
C = 3/D

Mdm = 6e-9
Mdp = 4.8e-8
Mmw = 4.8e-8
Mpw = IF hswitchl > 0 THEN 6e-9 ELSE 6e-10
Gm = IF hswitchl > 0 THEN 1 ELSE 0.0124875
Gd = IF hswitchl >0 THEN 0.97 ELSE 0.0125
Gw = IF hswitchl > 0 THEN 1 ELSE 0.0125
Gp = IF hswitchl > 0 THEN 0.97 ELSE 0.0125

App = (1-Mdp)*Gp
Amw = Mmw*Gw
Amm = (1-Mdm)*Gm
Adm = Mdm*Gm
Adp = Mdp*Gd
Apw = Mpw*Gw
Aww = (1-(Mpw+Mmw))*Gw

d/dt (Xw) = +wild_type
  INIT Xw = 10
wild_type = Aww*Xw
d/dt (Xp) = +p_mutant
  INIT Xp = 1
p_mutant = (Apw*Xw)+(App*Xp)
d/dt (Xm) = +m_mutant
  INIT Xm = 1
m_mutant = (Amw*Xw)+(Amm*Xm)
d/dt (Xd) = +double_mut
  INIT Xd = 1
double_mut = (Adm*Xm)+(Adp*Xp)+(Gd*Xd)

ratiop_w = Xp / Xw
ratiom_w = Xm / Xw
ratiod_w = Xd / Xw
ratiow_pmd = Xw / (Xp + Xm + Xd)
```

Table 2: The Relationships and Parameters for the Positive Regulation is like the negative control except for:

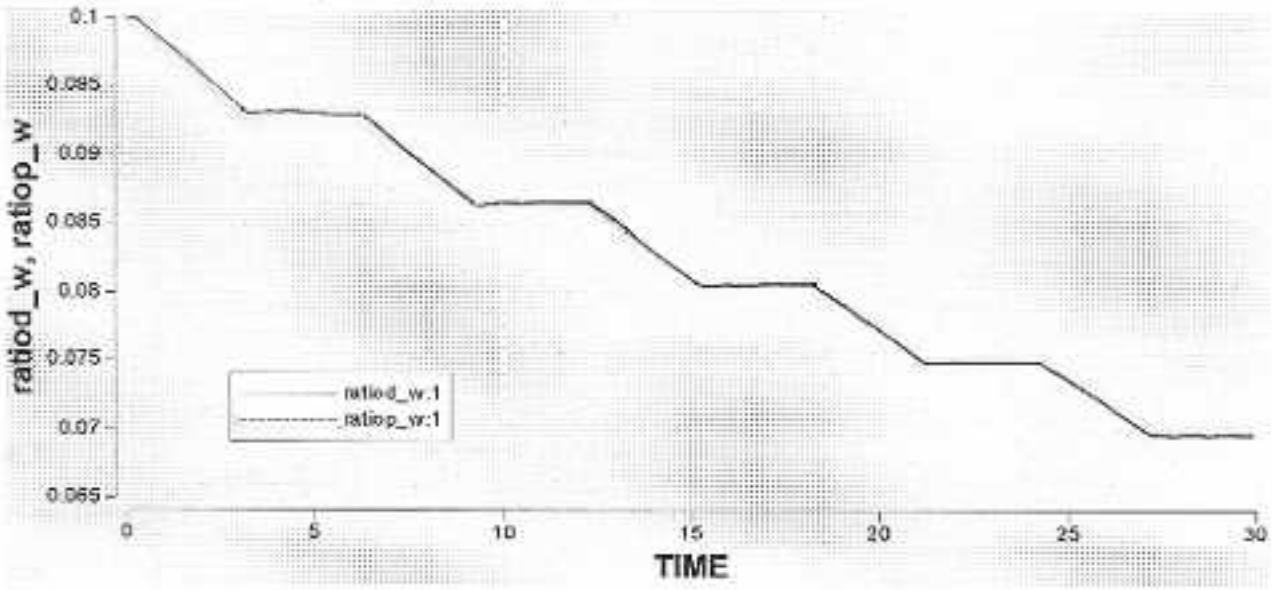
d/dt (hswitchl) = flowin - flowout
INIT hswitchl = 0
flowin = pulse (1, 0, C)
flowout = pulse (1, H, C)
H = D*C
L = (1-D)*C
D = .999
C = 6/(1-D)

Mdm = 6e-9
Mdp = 4.8e-8
Mmw = 4.8e-8
Mpw = 6e-10
Gm = IF hswitchl > 0 THEN 1 ELSE 0.0125
Gd = IF hswitchl >0 THEN 0.999 ELSE 0.0125
Gw = IF hswitchl > 0 THEN 1 ELSE 0.012125
Gp = IF hswitchl > 0 THEN 0.999 ELSE 0.0125

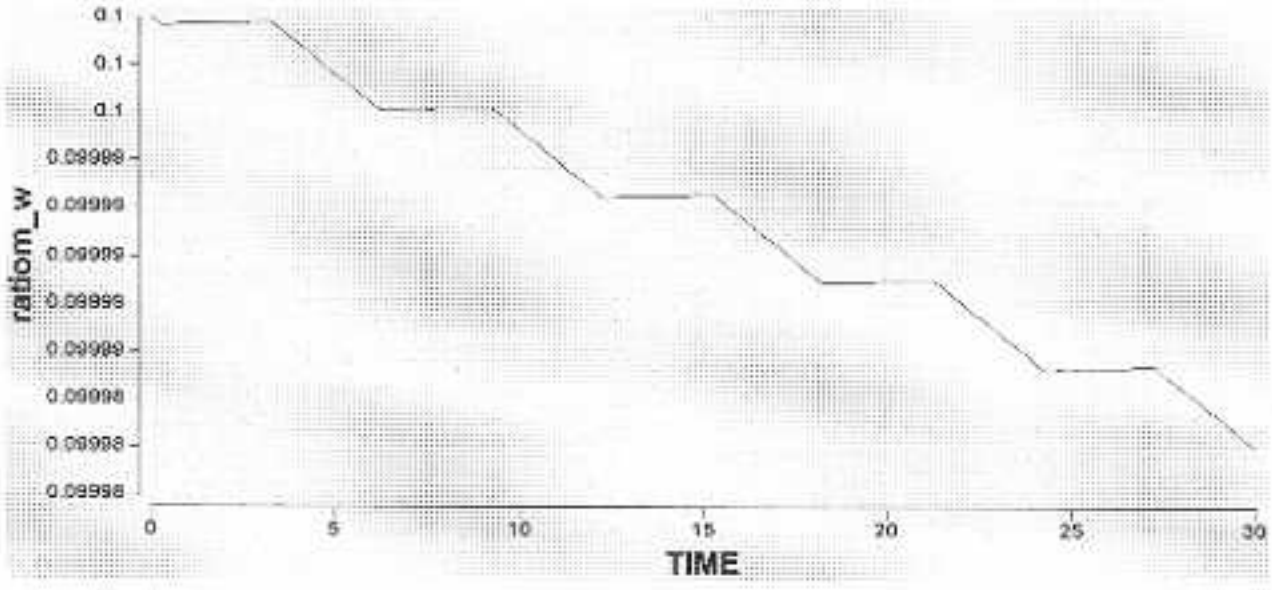
Graph 1: An example of how the fluctuation of the hswitchl curve is used to alternate between the two values of a parameter. As can be seen from the curve from the lactose operon model, when the time is within that of the high demand period, Gp = 0.97 and when the time is within the low demand period, Gp = 0.0125.



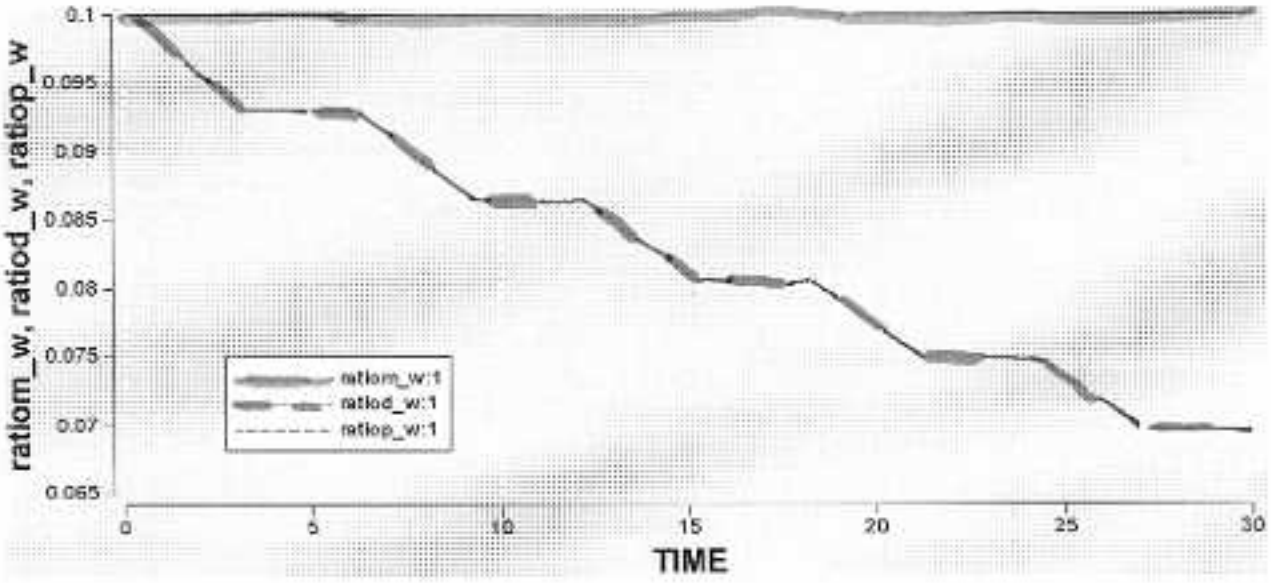
Graph 2: The graph of Xp/Xw and Xd/Xw vs. Time. The two curves are identical because the two mutants exhibit the same phenotypes in the high demand environment and low demand environment.



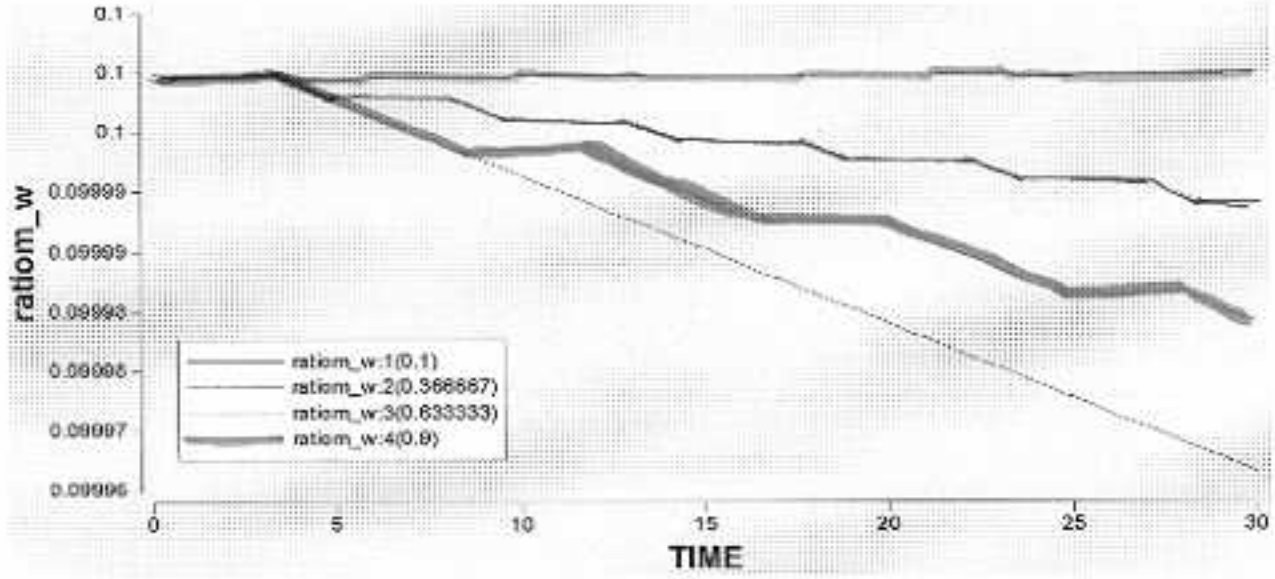
Graph 3: The graph of Xm/Xw vs. Time. This curve shows the slope is zero, when the phenotype is the same as the wild type. However, the slope turns negative when the mutant continuously expresses the lac genes although they are no longer needed in the low demand environment.



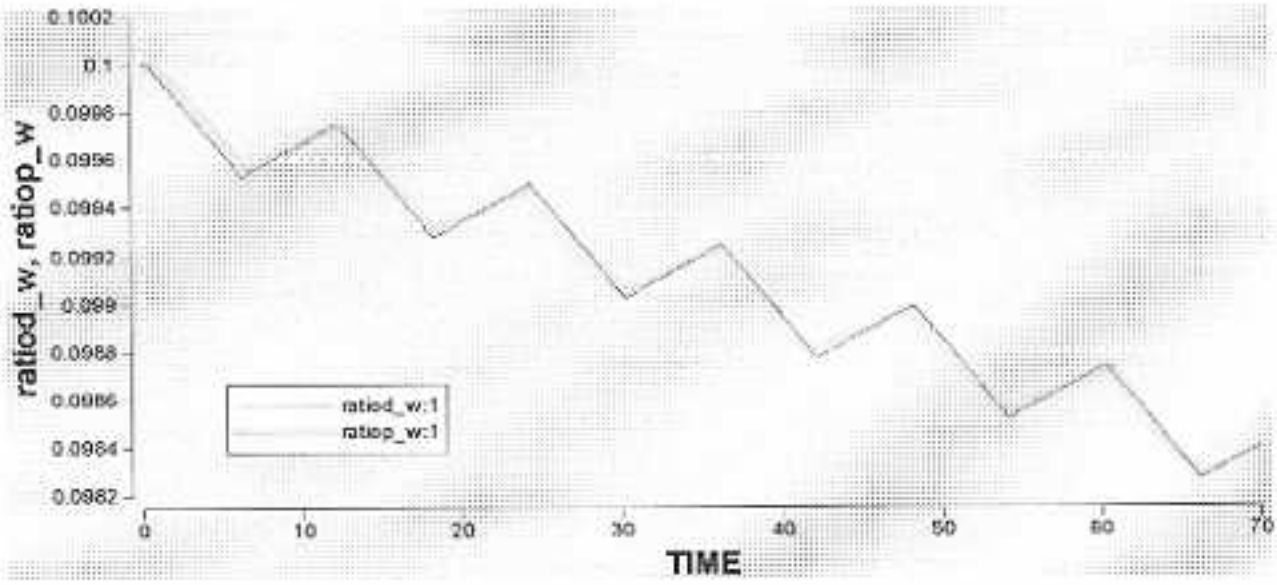
Graph 4. The graph of all the ratios to determine which ratio is most likely to overcome the wild type population. As seen from the graph, the modulator mutant has the highest overall slope and thus is the one which has the greatest chance of becoming selected over the wild type between the three mutants.



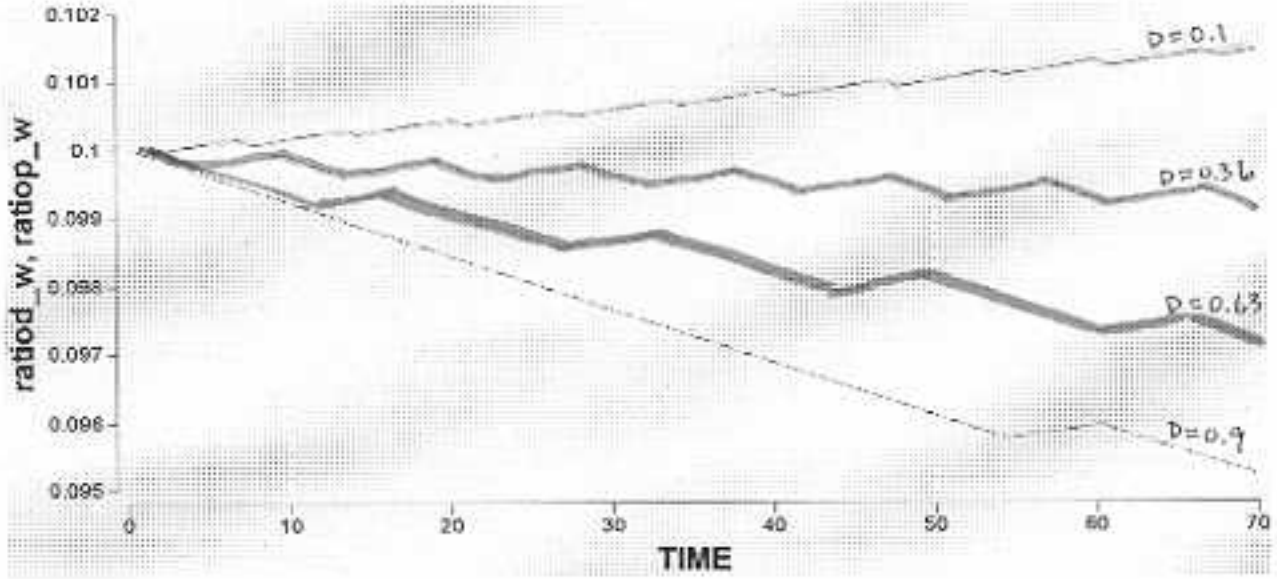
Graph 5: The ratio Xm/Xp curves as the demand (D) values changes between 0.1 to 0.9. In the legend, the demand value of each curve is in parenthesis. From the graph, we see that as the demand decreases the overall slope decreases. Thus, the negative regulation is selected when the demand is low.



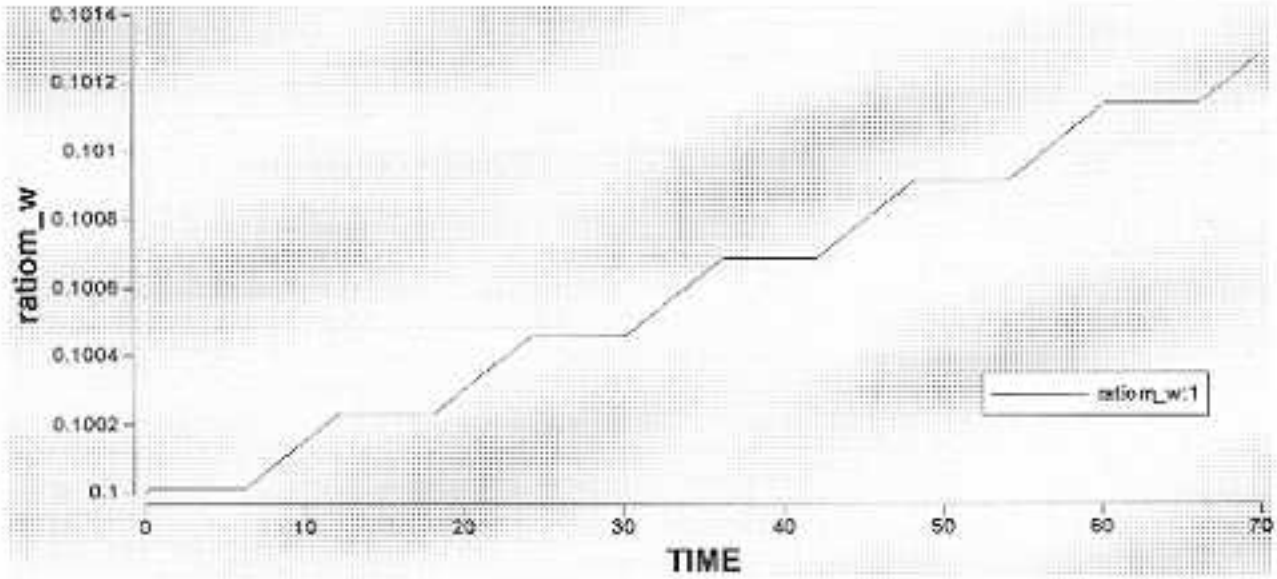
Graph 6: The graph of the promoter and double mutants' population ratios vs. time at D = 0.5. Both mutants constituitively expresses the mal genes and thus, they have a disadvantage in the low demand environment, shown by the negative first slope.



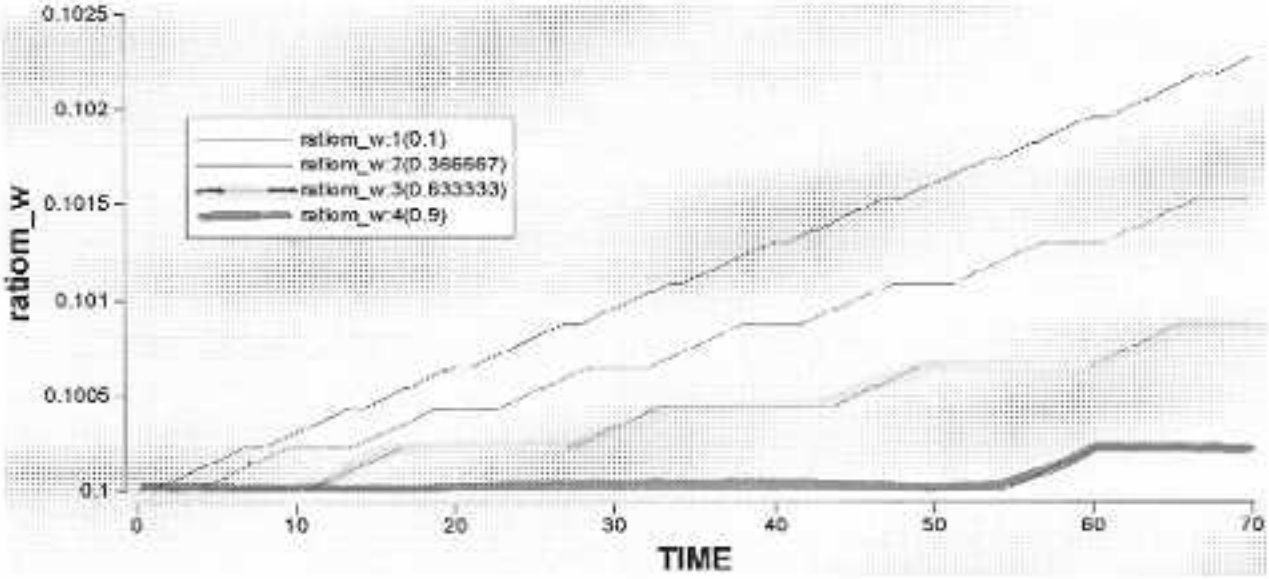
Graph 7: The changes in the overall slope of the Xp/Xw and Xd/Xw curves as the demand decreases from 0.9 to 0.1. As the demand increases, the overall slope decreases and the positive mode of control is conserved.



Graph 8: The graph of the modulator mutant vs. time when D = 0.5. The modulator mutant is the only mutant with an overall positive slope within the three mutants.



Graph 9: As demand increases from 0.1 to 0.9, the curve of Xm/Xw becomes more linear and thus, the positive mechanism of control is not lost. The D value corresponding to each curve is in parenthesis.



Bilbliography

Savageau, M. A., 1974 Genetic regulatory mechanisms and the ecological niche of Escherichia coli. Proc. Natl. Acad. Sci. USA 71: 2453-2455.

Savageau, M. A., 1977 Design of molecular control mechanisms and the demand for gene expression. Proc. Natl. Acad. Sci. USA 74: 5647-5651.

Savageau, M. A., 1998 Demand theory of gene regulation. I. Quantitative Development of the theory. Genetics 149: 1665-1676.

Savageau, M. A. 1998 Demand theory of gene regulation. II. Quantitatinve application to the lactose and maltose operons of *Escherichia coli*. Genetics 149: 1677-1691.

Schwartz, M., 1987 The maltose regulon, pp. 1482-1502 in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, edited by F. C. Neidhardt. American Society for Microbiology, Washington, DC.