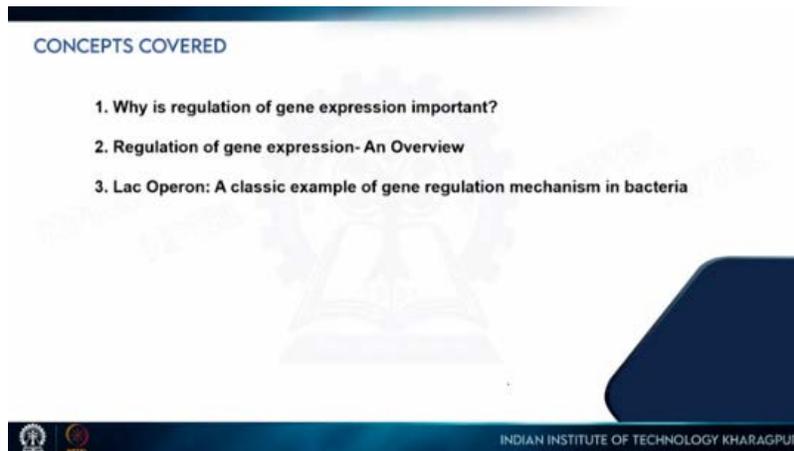


Introduction to Complex Biological Systems
Professor Dibyendu Samanta and Professor Soumya De
Department of Bioscience and Biotechnology
Indian Institute of Technology, Kharagpur
Lecture 10
Regulation of gene expression

Welcome to the NPTEL online course on Introduction to Complex Biological Systems. Now, we are on module 2. We are discussing transcription and translation, making products from the stored information in DNA. So, this is lecture 10, and we will be discussing the regulation of gene expression.

Here, mostly, I will cover a few important points. Why is the regulation of gene expression important? That is the first point, followed by an overview of the regulation of gene expression. And finally, I will give you one classic example of a gene regulation mechanism in bacteria, which is also known as the Lac Operon system. So, first, why is the regulation of gene expression important?



If you see, each cell contains all the genetic material for growth and development. But the thing is, some of these genes are expressed all the time, and other genes are not expressed all the time. They are switched on and off as needed. For example, if I now talk about the human body, every cell has the same information, whatever is written in DNA.

Why is regulation of gene expression important?

- Each cell contains all the genetic material for growth and development
- Some of these genes are expressed all the time
- Other genes are not expressed all the time. They are switched on and off at need

But now, if I compare a nerve cell and a liver cell, those two cells are completely different in how they look, as well as their functions. So, why do these different behaviors and different functions actually occur? All these things actually lie in how the genes are getting expressed in these two different types of cells, and that determines the fate of the cells. So, if I give some examples, like here, one protein, if I say GAPDH, this protein is present in every cell in our body. The full form of this protein is glyceraldehydes, glyceraldehyde 3-phosphate dehydrogenase.

So, GAPDH, this helps in glucose metabolism. So as we know, for living cells, each cell needs to metabolize glucose and produce ATP. So this enzyme is present everywhere. And I would say this gene is also being expressed.

That means it is making a product from this gene all the time. Like some genes might be expressed, might be making a product all the time, but some genes will only make a product when needed. And apart from that, we can also classify these kinds of genes in two defined ways. One is called constitutive gene expression. That means those genes are constantly expressed or they are available in every cell.

Like genes are present in every cell, but they are being expressed. They are making a product in every cell. For example, GAPDH. Insulin, this insulin gene is a tissue-specific gene because insulin is produced in the pancreas only. So as a result of that, although the insulin gene is present in the DNA, present in every cell, it is making the product, it is being expressed inside the pancreas so that the pancreas can make insulin for us.

So this is called tissue-specific gene expression. So the summary of this slide is that gene expression is very much controlled. Although we have the same set of information in every cell, because of gene expression and its regulation, we see different types of complexities and different types of functions. This is not only in eukaryotes, even this is true for bacterial systems, for example, a bacterium when present in nature, and that bacteria might be pathogenic also, but when it enters inside a host, then for survival, it can make different types of products.

Why is regulation of gene expression important?

- Each cell contains all the genetic material for growth and development
- Some of these genes are expressed all the time
- Other genes are not expressed all the time. They are switched on an off at need

$GAPDH \rightarrow \text{Glyceraldehyde 3-phosphate dehydrogenase}$
Insulin \rightarrow Insulin

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

Different sets of genes will be expressed so that they can actually evade the host immune system. So, as a result of that, this gene expression and its regulation are true everywhere in prokaryotes as well as in eukaryotic cells. Now, I will just give you an overview of gene expression. So, gene expression means to make a product from the gene, right? So, if I say it again, just like the central dogma, DNA goes to RNA.

So, that is transcription. And here, particularly, I am writing messenger RNA so that it is the direct template for translation, and it will make a product that is our protein. This is a translation. Now, major regulation is actually restricted to this part, the transcription and translation. Like, if we can control the regulation of transcription and translation, then the product formation will differ.

Regulation of gene expression- An Overview

DNA $\xrightarrow{\text{Transcription}}$ mRNA $\xrightarrow{\text{Translation}}$ Protein



INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR



So, for example, transcriptional control, like, if I give some examples, RNA polymerase transcribes DNA or a particular gene to make RNA. Now, if the promoter sequence, which is present just before the gene sequence, if the promoter is very strong, that means it can recruit RNA polymerase very efficiently, then the amount of RNA or the transcription will be higher. So, as a result of that, more mRNA will be present inside the cell. As a result of that, if more templates for protein production are there, naturally, we might get more protein, but that is not always true because there are so many other types of regulation mechanisms that exist. Now, this mRNA can form protein, which I already mentioned, but there is a specific mechanism that is called the degradation of mRNA because mRNA exists all the time inside the cell, then it will constantly make protein, but I just told you that sometimes we do not need that protein.

So, as a result of that, mRNA should be degraded also. So, that is the degradation of mRNA. And then, protein is produced after translation. But this protein, there are many fates of this protein. Although protein is present, this protein, maybe active or inactive also because finally, proteins will perform functions for us. But now, even though we have proteins, if those proteins are inactive, we are not getting the function. So, there are very well known mechanisms present in our body as well as in bacteria. For example, phosphorylation and dephosphorylation of proteins are very well known mechanisms that regulate protein activity. So, even though proteins are present, it depends on whether they are active or inactive.

Inactive versions of proteins can also be degraded by well known mechanisms, meaning proteins can be degraded as well. So, this is the degradation of proteins. Now, this scheme, whatever I mentioned, is true for bacteria as well as eukaryotic systems, but there are many complexities present in eukaryotic systems, and I will highlight a few of them. So, particularly in the case of eukaryotes, we have a nuclear membrane. So, if I say this is the nuclear membrane and this is a nuclear pore. So, this is the cytoplasm, and this is the nucleus. Now, we have DNA here inside the nucleus. DNA first goes to mRNA transcription. The transcript, before processing in eukaryotic systems, should be processed properly to make a mature mRNA, which is the template for translation. After processing, you will get mRNA.

This processing step is followed by the transport of mRNA into the cytoplasm, where it is readily available for translation. It will then undergo translation and produce proteins. Specifically, I would like to mention this processing step here. During processing, I mentioned that processing means at the 5' end of mRNA, the addition of a 7-methylguanosine cap, and at the 3' end, a poly-A tail should be added. Also, very importantly, there are many introns present in eukaryotic genes, but those introns should be removed, and finally, we will have our mature mRNA without introns. So, this processing, this splicing of mRNA, is also a very controlled process and can also affect gene expression.

So, I will give you just one example here of how the splicing of mRNA actually helps in gene regulation or that can control gene regulation. So if I mention here, this is 5', 3' and this is an mRNA, and this mRNA here is exon 1, exon 2, and exon 3, and this is intron I1 and I2, this is intron 1 and intron 2. So now, splicing means that introns should be removed. So, as a result of that, what I should get is 5' to 3', this is mature mRNA and you have E1, E2, and E3.

So now, this is your mature mRNA, and it will make some protein. After translation, it will make a protein. Now, sometimes for specific genes, alternate splicing can also happen. In that case, what we will get because of the alternate splicing, we might get 5' to 3' this mRNA, but here you have just E1 and E2, but the E3 portion is not present in this mRNA.

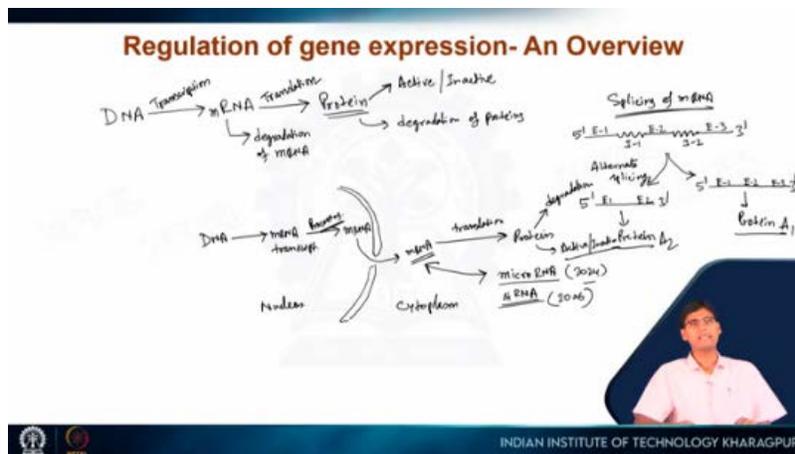
So, as a result of that, it will also make a protein. Now, there should be some difference in these two proteins. For example, I would say if in this protein, for example, I am saying that protein A1 and protein A2. In protein A1, if I say this is a membrane-bound protein and the A3 portion actually helps this protein to be anchored on the membrane of a cell. Now, protein A2 doesn't have this E3 portion, right? E3, that means the exon 3 portion, so as a result of that, when it will be translated, that portion which will be required for anchoring on the surface of the cell membrane should not be there, so as a result of that, it will not be present on the cell surface; it can be secreted outside the cell.

So, this is a kind of complex mechanism of gene regulation also, and this kind of different type of regulation or gene expression is present in our system, and this is just one example that I mentioned, particularly this is for the eukaryotic system, and also this mRNA degradation mechanism. It is also very robust. There are many different types of mRNA degradation mechanisms present. For example, I just mentioned in previous classes that there are three different types of RNA. Particularly messenger RNA, ribosomal RNA, and transfer RNA.

Those are the very important classic RNAs, which are very basic things. But apart from that, we have some small RNAs which regulate gene expression. For example, microRNA. So, that means this microRNA can also regulate the degradation of mRNA.

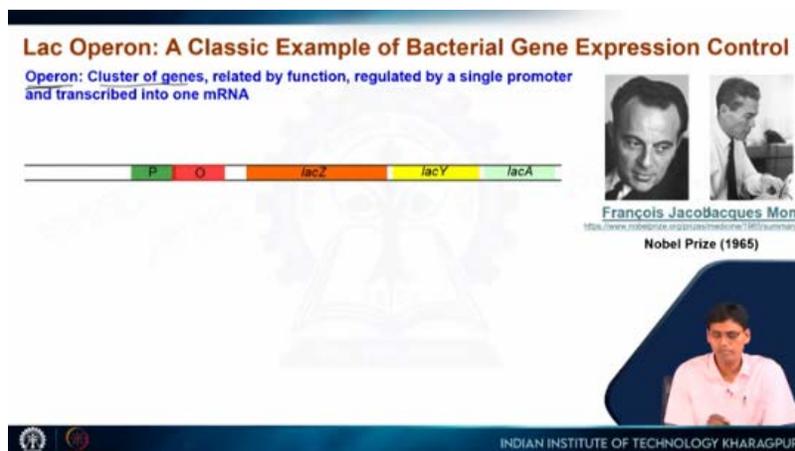
So, this microRNA is a small RNA, but it can control the mRNA and determine how long it will be present inside the cell. Similarly, siRNA So, siRNA is also very important and it also regulates the mRNA function. So, these are microRNA and siRNA. Very recently, in 2024, Victor Ambros and Gary Ruvkun received the Nobel Prize in Physiology and Medicine for their work on microRNA.

Similarly, in 2006, Andrew Fire and Craig Mello also received the Nobel Prize for their work on siRNA, so you can understand how important they are in gene expression, the microRNA and siRNA. This protein degradation, whatever I told you about degradation, and then the protein, whether it will be active or inactive. They are also very important in gene expression. So, now, specifically, I would like to mention a case study, for example, this is in a bacterial system, the Lac Operon.



This is a classic example of bacterial gene expression control. Whatever I discussed in the previous, during the last few minutes, that is an overall understanding of gene expression, but in this case, I would be specifically mentioning a set of genes and how they are regulated in order to metabolize, for example, lactose in this case, the Lac Operon. So, an operon means a cluster of genes related by function and regulated by a single promoter and transcribed into one mRNA. This is an operon.

So, this operon actually, you know, this word I believe came from the Latin word opera. So, that means to perform or work. So, something like that. So now, if you see, this is a cluster of genes, as I just mentioned, a cluster of genes. So you can see here we are mentioning three genes: lacZ, lacY, and lacA.



So, this gene is lacZ. A gene means that some sequence is present in DNA and it will make some product after transcription and translation. So lacZ will make a protein, particularly an enzyme called beta-galactosidase. This is the product of lacZ. Similarly, lacA will

produce another enzyme called lactose permease, and the third one, lacA, will make some enzyme called thiogalactoside transacetylase. This actually helps in cell detoxification, but this kind of complex mechanism which we are not going to discuss but still, it should be fine. So, this is Galactoside transacetylase, which helps in cell detoxification. And beta-galactosides, many of you know already that it is lactose. So this Lac Operon, we are discussing to understand how bacteria regulate or they actually metabolize lactose. So, lactose should be metabolized into glucose and galactose. Lactose should be broken down into glucose and galactose because lactose is a disaccharide, so it should be broken down into glucose and galactose for, you know, for production of energy. So, lactose will make glucose plus galactose. So, this conversion is done by this enzyme called beta-galactosidase.

This is the enzyme that will break lactose into glucose and galactose and lactose permease will help to bring or it will help to bring lactose from outside the cell to inside the cell. So that's why the name is lactose permease. So, it will bring lactose molecules inside the cell. So, we are talking about bacterial cells here, particularly you may consider it's an E. coli cell.

So, import of the lactose. So, now naturally, what will happen? So, if this is the gene and in the case of the Lac Operon, all these three genes, lacZ, lacY, and lacA, are controlled by a single promoter. So, a cluster of genes related by function, how are they related?

All these three products of these three genes help in lactose metabolism. That is why they are related by function; they are also regulated by a single promoter and transcribed into one mRNA. So, what will happen? So, as you can see here, you have the promoter here.

So, then, as I already mentioned before during the transcription lecture, the promoter will recruit this enzyme called RNA polymerase. So, then RNA polymerase will move in this direction and it will make mRNA, right? So, it will make mRNA, and then, when RNA polymerase moves in the reaction, you will get transcription. So, transcription, and you will get mRNA 5' to 3' and this mRNA will be translated into protein because, in this mRNA, you have the information to make the lacZ product, the lacY product, and the lacA product. That means beta-galactosidase, lactose permease, and thiogalactoside transacetylase. So,

now the thing is, although you can see here just one mRNA, those three proteins will not be fused together. There will be three separate proteins because those genes are separated by their own start codon and stop codon.

So after the translation we will get three different proteins that are beta-galactosidase, lactose permease, and thiogalactoside transacetylase. Now the thing is this is the normal scenario when the RNA polymerase binds to the promoter site, and then it will make the mRNA and it will make all the products. But the question here is that if lactose is not present there in the environment then where the bacterium is growing. Then why would the bacteria make lactose?

It is not required to make these lactose metabolizing gene products. If lactose is not there, there is no use of beta-galactosidase or lactose permease. So as a result of that, the genes should be turned off. Right? So I will discuss this part.

Lac Operon: A Classic Example of Bacterial Gene Expression Control

Operon: Cluster of genes, related by function, regulated by a single promoter and transcribed into one mRNA

<i>lacZ</i>	β -galactosidase	Breaks lactose into galactose and glucose.
<i>lacY</i>	lactose permease	Imports lactose into the bacterial cell.
<i>lacA</i>	thiogalactoside transacetylase	Cell detoxification.

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

So whatever I mentioned in the previous slide that is the outcome of the lac operon. If everything gets transcribed, translated, then all these three products will be formed and they might help in lactose metabolism. Now, if lactose is not present, then bacteria will not make this product, those are unnecessary. So, that means, this operon is somehow regulated so that without requirement it will not make the product. How?

So, I am going to explain that although I mentioned *lacZ*, *lacY*, *lacA* and the promoter site, I did not mention the operator site. So, you can see here it is mentioned O. So, this means

the operator site present on DNA in this lac operon operator site. So, this operator site in this place, one protein will be present and it will bind here.

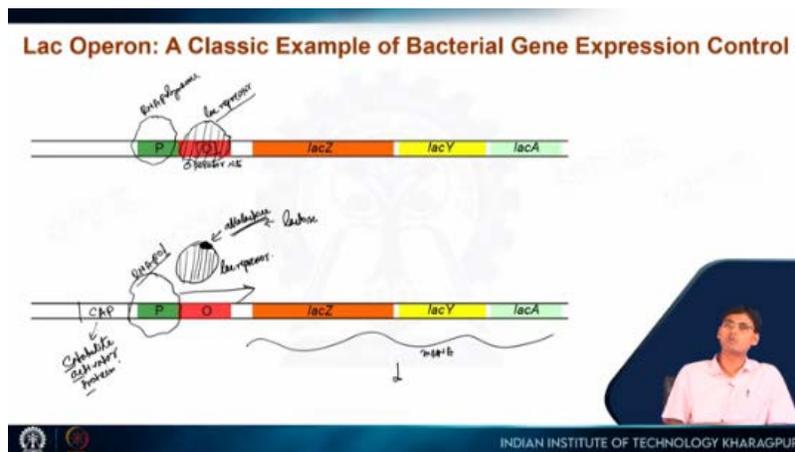
So, I would say this is a protein here, which is called the lac repressor, the lac repressor. It is a protein that will bind to the operator site. This is, I would say, its default state. This lac repressor is always bound to the operator site. So, unnecessarily, this product will not be produced. So, now if the lac repressor is bound here, then this is our RNA polymerase, even if it gets recruited to the promoter site, cannot transcribe these genes because the lac repressor is actually inhibiting the progression of RNA polymerase. So, as a result of that, there is no expression of these genes. So, now in this case, I am telling that there is no lactose here but if lactose is present, then what will happen? If lactose is present, then see, in this case, I am saying that lactose is present. So, if lactose is present, now from this lactose one isomer called allolactose should be there inside the cell in very low amounts. Allolactose is just an isomer of lactose. Now, this allolactose will bind to this lac repressor. This is the lac repressor. So, now this allolactose, if I say this is my allolactose, will bind to the lac repressor and change the conformation of the lac repressor, so that it cannot bind to the operator site anymore. So, as a result of that, this operator site will be free. Now, the RNA polymerase, so here this is again RNA polymerase.

So now it can move in this direction because the lac repressor is no longer bound to the operator site. As the allolactose binds to the lac repressor and changes its conformation so it cannot recognize the operator site anymore. So, as a result of that, transcription will happen here and you will get mRNA and finally, you will get the protein required for lactose metabolism. So, I just mentioned how it behaves in the presence and absence of lactose.

Now, this is a very simplistic view of the lac operon, but there are some more complexities there. I would like to mention that. So now, bacteria, if lactose is present and lactose is absent, I told. But what about if lactose and glucose both are, like both are present? Then, they evolve in such a way that they will prefer glucose.

Because they don't need to metabolize lactose anymore because a ready source, like glucose, is available in the media, so they will uptake, they will utilize glucose only, not

the lactose. So as a result of that, again, during that time, they should turn off their lac operon, right? So that kind of mechanism is also there. Because I'm not going into much more details that are not required at this level, but the thing is here you have another site that is called the cap binding site. This cap binding site means catabolite activator protein. CAP, catabolite activator protein binding site. So now the thing is here, one protein called catabolite activator protein will bind here, and it will actually, I would say, this is some kind of glucose sensor that measures how much glucose is present inside the cell, depending on that, it will bind. And it will regulate this gene expression. Very briefly, if I mention, if a huge amount of glucose is present inside the cell, then the amount of cyclic AMP should be low. So if cyclic AMP is low, then cyclic AMP will not bind to catabolite, activator protein. So as a result of that, it will not actually turn on this system.

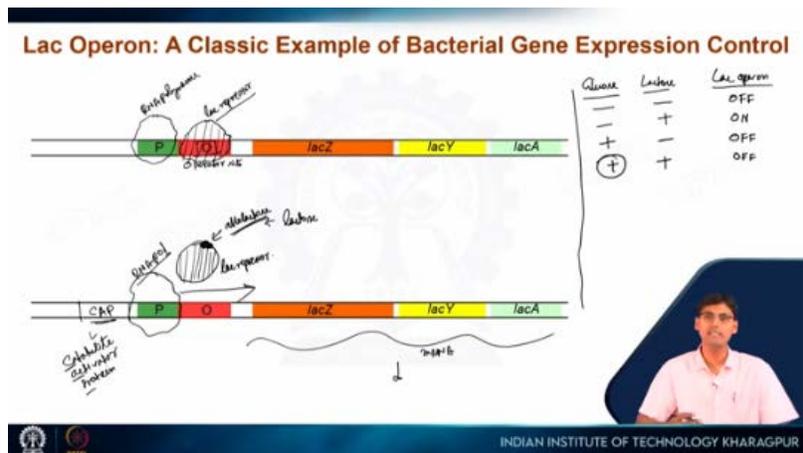


But if the amount of glucose is low inside the cell, then because of the metabolism, then we will have more cyclic AMP. So, the cyclic AMP actually binds to this catabolite activator protein, and after that, it will actually help the expression of this set of genes. So, as a result of that, glucose and lactose both decide the fate of this lac operon, not just lactos. But finally, they will use, I would say, there are four different scenarios if I mention here now. Here Glucose, Lactose and Lac Operon whether it is on and off that I will mention. For example, glucose is not there, lactose is not there. So, I am just putting minus in both the cases. Then why does Lac Operon need to make this product? No lactose is required.

So, it should be turned off, this whole system. And now you don't have glucose, but you have lactose. Then bacteria should depend on lactose for an energy source, right? So, as a

result of that, they have to make this product from the Lac Operon so that they can metabolize lactose and they can survive. I mentioned that if lactose is present, then how they regulate the gene expression here. So, no glucose, and then lactose is present, so this Lac Operon should be turned on. Now, you have glucose, but you do not have lactose. Then there is no question of Lac Operon product because no lactose is present here. They can readily use glucose for their energy source.

So, this should be again but both present, then this is a little bit complicated case that Lac Operon should not make a product because they have glucose, and they do not need to metabolize this lactose for their energy source. They can readily use glucose for this purpose. But the explanation relies on this position that this glucose sensor, catabolite activator protein. And because of high and low amounts of glucose present inside the cell, that will dictate the expression of this Lac Operon also. And that way, these genes regulate this, and they help in lactose metabolism. And because of this beautiful work,



Jacob and Monod received the Nobel Prize. I just mentioned it in the previous slide; you can look into it. So now, this is just one example, and this is some kind of overview of gene expression. That is all for today, and you can look into any textbook; any standard textbook will be fine for further reading.

REFERENCES

1. RNA: Life's Indispensable Molecule by James Darnell
2. Biochemistry by Lubert Stryer (4th Edition)
3. Lehninger Principles of Biochemistry by Nelson and Cox (8th Edition)
4. Essential Cell Biology by Alberts et al., (4th Edition)

