Journal of Genetic Mutation Disorders
Volume 1 | Issue 1

Review article Open Access

E. Coli Bacteria and It's Transcription Regulation

Bhattacharya PK*

Retired ex Professor and Head of Pathology Department, Calcutta School of Tropical Medicine, Kolkata, India

***Corresponding author:** Bhattacharya PK, MD (Pathology Calcutta University), FIC path, WBMES, Retired ex Professor and Head of Pathology Department, Calcutta School of Tropical Medicine- Kolkata 73, West Bengal, India. 700073. Tel: 9231510435, Email: profpkb@yahoo.co.in

Received Date: June 08, 2025 **Accepted Date:** June 26, 2025 **Published Date**: June 28, 2025

Citation: Bhattacharya PK (2025) E. Coli Bacteria and It's Transcription Regulation. J Gene Mutat 1(1): 106.

Abstract

Escherichia coli is the commonest Bacteria for urinary tract Infection do regulate its transcription predominantly at the stage of initiation by repressors and/or activators.

Keywords: E-Coli, Transcription, Bacteria

Introduction

Bacteria always do live in an ever-changing environment. They have to adapt again and again to their environment changes. Escherichia coli, for example, which lives usually in the gut of humans, has to adapt to the constantly changing, with food human beings eat each time. It is thus a necessity for E. coli to regulate the synthesis of its enzymes according to its needs to survive in the gut of humans. Enzymes that break down such as lactose are only needed when these compounds are present.

Historical Development of the Field

The first systematic analysis of such adaptation was detected probably in 1930s by the Finnish biochemist Henning Karstro¨m [10]. His work was then repeated in the 1940s by the French geneticist Jacques Monod [8,19]. Monod critically analysed the growth of E. coli in presence of mixtures of glucose and other sugars, such as lactose, arabinose or maltose. Monod confirmed Karstro¨ m's observation that glucose is always used first by E Coli bacteria. This is the reasons that uncontrolled diabetic patients who have glycosuria suffer from recurrent E coli infections regularly and UTI most repeatedly produces biofilm in the urethra or in the bladder neck of diabetic patients [8,19].

In the 1940s it thus became clear that certain sugars, like lactose, arabinose or maltose, induce the synthesis of enzymes that break these sugars down to glucose that can be used directly by E Coli bacteria. On other hands, certain amino acids, such as tryptophan, methionine or histidine, repress the synthesis of enzymes that synthesize them. It thus became necessary to decide which system should be analysed extensively. The answer was unambiguous. It was the" lac system " of E. coli, for which the best tests had been developed in the following way. o-Nitrophenyl-b-D-galactoside (ONPG) was a perfect substrate to test for the presence of b-galactosidase. ONPG is colourless but when hydrolysed by b-galactosidase its product, nitrophenol, turns into yellow. Moreover, isopropyl-1-thio-b-D-galactoside (IPTG), an analogue of lactose, was synthesized. It induces the synthesis of b-galactosidase but is not a substrate of b-galactosidase. With these two substances it can be demonstrated that b-galactosidase can be induced about 1000-fold. Moreover, it can be also demonstrated that the synthesis of b-galactosidase is a de novo synthesis. This destroyed the then prevailing instruction hypothesis that the inducer forced an already synthesized precursor protein to fold correctly into active b-galactosidase, like antibodies which were thought to fold in the presence of antigens in their final conformation. It was not just excellent biochemistry that led to the solution of the problem, for genetic mutant analysis also helped enormously.

Three types of E. coli mutants are usually found. One mutant type does not grow on lactose. At that time, three types of such lac2 mutants were found which did not grow on lactose. There were mutations that destroyed activity of b-galactosidase (Z2) and others that left b-galactosidase activity intact. It was known that these latter mutants were defective in lactose permease (Y2), the pump system that pumps lactose from the outside into the cell of E Coli.

There was also a third type of mutant in which both b galactosidase and lactose permease activity were absent. Genetic mapping indicated that the Z and Y mutations were closely linked. The two genes are indeed part of one transcription unit, of one operon.

The other type of mutant was constitutive, i.e., it synthesized b-galactosidase in large amounts even in the absence of an inducer. These mutants were called I2. Again, these mutations were closely linked to the lac Z and lacY genes.

The question was now, how does regulation of b galactosidase and Lac permease then work in E. coli? Does it work in a positive manner? Does the product of the lac I gene activate transcription of the lacZ and lacY genes in the presence of an inducer or a suitable mutation? Or does it regulate in a negative manner, i.e., is the synthesis of lac ZY message inhibited by the presence, but decontrolled in the absence, of the product of the lacI gene?

While these questions were raised in the late 1950s and early 1960s, the existence of messenger ribonucleic acid (mRNA) was demonstrated. It then seemed to make sense to assume that the product of the lacI gene inhibits the initiation of lac mRNA synthesis. The evidence for the model which Jacob and Monod then presented in 1960 and 1961 was the following [8,19]. They presented the two types of lac constitutive mutants that were predicted for negative control and were incompatible with positive control:

1. They found constitutive recessive mutations mapping in the lac I gene. In other words, the presence of an I+ gene on an episome or a chromosome is dominant over an I2 mutation on the chromosome, or vice versa, on an episome. This is what one would expect in the case of negative control. The absence of repressor leads to overproduction of b-galactosidase and Lac permease. Its presence is dominant over its absence.

2. Lac repressor should, according to the model, interact with the operator, a region close to the promoter, the starting site of transcription. Jacob and Monod found weak, constitutive Oc mutations which were cis dominant. They thus result in constitutive synthesis of the genes which are covalently linked to the region which carries the Oc mutation. Genes in trans are not turned on. This is exactly what one would expect from negative control.

The repressor–operator model was so convincing because it was not only presented by Monod for the lac system but also, by Jacob, for the machinery of phage lambda (l) determining lysis or lysogeny of E. coli. Phage l, discovered by Esther Lederberg, has two ways to develop after entering E. coli. In the lytic pathway it may multiply and lyse the E. coli host cell. In the lysogenic pathway it may incorporate its deoxyribonucleic acid (DNA) at a particular site in E. coli DNA and stay there forever. Phage l makes turbid plaques on E. coli not carrying l gene. The turbid centres of the plaques consist of bacteria which have become lysogenic for phage l. CI repressor of Lambda prevents phage l from making plaques on E. coli lysogenic for l. There are mutants of l which follow the lytic path exclusively. Some of them occur in the CI repressor gene. They make clear plaques on E. coli. There are operator mutants of l that are capable of making plaques on E. coli lysogenic for l. Here the operator site with which l repressor interacts has been damaged. It was the combination of the lac and l systems which made the Jacob–Monod model of negative control of transcription so convincing [8,19]. Thus, most molecular biologists at the time believed that all genetic control is negative. It took years until the first cases of positive control in E. coli (ara, mal) were generally accepted. And it took a decade or longer until it was accepted that in eukaryotes almost all genetic control is positive.

Steps Involved in Transcription of Multicistronic Operons

In E. coli most DNA is effectively naked. In contrast, in eukaryotes DNA is packed in chromatin of simple or higher order. Because of the nucleosomes, RNA polymerase cannot easily recognize TATAAT sequence in eukaryotes to start transcription. However, a DNA sequence may be easily recognized by a protein in E. coli DNA. In E. coli a 5'TATAAT sequence at the 210 position and a 5'TTGACA sequence at the 235 positions (+1 is the start of transcription) suffice to bind RNA polymerase effectively and specifically. In E. coli, a particular subunit of RNA polymerase, s (sigma), recognizes these sequences and guarantees effective, specific DNA binding. Recently it has been shown that a subunit of RNA polymerase may bind to an AT-rich region around 250, the UP element. DNA is in its B form when recognized and so forms the closed complex. The DNA then opens up from the 210 regions to the start position and forms the open complex. In the presence of the four trinucleotides UTP, ATP, GTP and CTP (uridine, adenosine, guanosine and cytosine triphosphate) the start of transcription begins, optimally every second. Transcription proceeds then at 378C at the rate of about 70 bases per second.

In E. coli, the DNA that is transcribed often contains several genes. Such an assembly of several genes that are transcribed from one promoter is called an operon.

It should be noted that E. coli mRNA is, in general, not spliced, in contrast to eukaryotic mRNA. It is also not edited, reverse transcribed or modified at its 5' and 3' ends. It is used for translation as it is. It should be noted that most mRNA of E. coli is rather

unstable. In general, it has a half-life of about 2 min. Thus, in order to be used efficiently it has to be translated into protein while being synthesized. There are special DNA sequences where transcription stops. Some of these signals do not need the help of a protein. Others do. Regulatory proteins may inhibit the stop of transcription. Transcription may in principle be regulated at all stages between start, elongation and stop. Yet it is evident that most effective negative or positive regulation occurs by inhibiting or enhancing formation of the closed complex.

Negative Control of Transcription Through a Repressor: The lac Operon

The lac system was in 1960, and is today, one of the best understood regulatory systems of E. coli. Lac repressor was the first repressor ever isolated [6] and sequenced. It is a tetramer. Each monomer consists of 360 residues. Its X-ray structure has been solved [1,2]. It has a modular structure that has been analysed genetically in great detail [18,14]. Residues 1–60 code for the head piece, which binds to operator DNA. To do so it uses a helix, the recognition helix, the second helix of a helix-turn-helix motif, to bind in the major groove of one-half of the operator. Residues 61–330 of Lac repressor code for the core, which carries one inducer binding site and one dimerization site. Residues 330–360 code for two heptad repeats, which guarantee that two dimers form a tetramer.

Thus, there can be mutations in the head piece which destroy the capacity of Lac repressor to bind to lac operator and to repress. They retain the capacity to bind inducers and to form tetramers. Such mutants are negative dominant constitutives (I2d). There are mutations in the core which destroy the capacity of Lac repressor to bind inducers but keep the capacity of dimer formation and operator binding. Such mutants are dominant negatives (Is). The lac operator shows imperfect dyadic symmetry. Each half is bound by one recognition helix. The DNA sequences which are recognized by RNA polymerase, the promoter, and by Lac repressor, the operator, overlap. The centre of symmetry of lac operon is 11 bp downstream of the start of transcription. That promoter and operator overlap can be demonstrated by DNAase protection experiments. Thus, RNA polymerase and Lac repressor cannot bind to their binding sites at the same time. The stronger the Lac repressor binding, the less RNA polymerase is bound and the less mRNA of the lacZ gene is produced.

When the DNA of the lac operon was sequenced, two additional sequences which looked like lac operator were found. One of them, O2, is situated 401 bp downstream from the centre of symmetry of Lac O1 (the author will now call it lac operator O1). The other sequence is situated 92 bp upstream of the centre of symmetry of O1. O2 binds Lac repressor less well than O1 by a factor of 5. O3 binds Lac repressor less well than O1 by a factor of 300.

No constitutive mutations were ever found in these O2 or O3 sequences. Therefore, for a long time they were regarded as nonfunctional leftovers from evolution. This was all wrong. It was indeed shown that when just O2 or O3 is destroyed, repression decreases by a factor of 2–3. However, when O2 and O3 are both destroyed, repression decreases by a factor of 70! How can this be explained? It was proposed that tetrameric Lac repressor forms loops either between O1 and O2 or between O1 and O3. If formation of one of the loops is made impossible by the destruction of either O2 or O3, then the other loop can still form. But if both O2 and O3 are destroyed no more loops can be formed. Repression has to occur exclusively from O1.

But why does a loop between O1 and O2 or O3 lead to stronger repression than binding of repressor at O1 alone?
There is a simple, reasonable explanation: if tetrameric Lac repressor is bound with two subunits to O2 or O3, then the two subunits which are free have come rather close to O1. Their local concentration in regard to O1 is thus strongly increased. It is increased more so from O3 than from O2. This increases the binding capacity of these Lac repressor dimers to O1. This then leads to less of the RNA polymerase binding to the promoter, i.e., more repression. In the distance of more than 600 bp the local concentration of a repressor becomes unimportant: it reaches the level of repressor floating freely around in the cell, i.e.,1028M [18].

Two such independent, parallel devices of any functionare often called redundant. This term is misleading. The author thus prefers to call O2 and O3 auxiliary or supplementary instead of redundant, pseudo (i.e., lying) operators. Only complexes that are unique

in E. coli allow the actions of repressors which are further away than
TAACACTCGCCTAT TGTTAA '5
5' AATTGTGAGCGGATAA CAA T TT 600 bp from the promoter. Such complexes do not exist in the lac system but they exist in the genome of phage l. If two dimers of l repressor bind to two properly spaced l operators (to OR1 and OR2 and to OL1 and OL2), then the tetramers formed on DNA interact with each other and form an octamer [14, 5]. The loop is now 2.4 kbp long.

Positive Control of Transcription Through an Activator: The lac Operon

When negative control of the lac operon was proposed in 1960–1961 by Jacob and Monod, that the lac operon is also positively regulated was completely overlooked. A Detailed genetic analysis indicated that mutations in two loci outside the lac operon lead to a 50-fold decrease of b-galactosidase synthesis [8,19]. It was found that one of these genes codes for cyclic adenosine monophosphate (cAMP) synthetase, whereas the other codes for a protein which was then called catabolite activator protein (CAP) or catabolite repressor protein (CRP). It was shown that CAP/CRP binds in the presence of cAMP to specific sequences of DNA and bends the DNA strongly. Is it the bending itself that activates RNA polymerase? I think it is not. The bending, of course, is an integral part of the complex form of CAP and RNA polymerase on the DNA. The centre of symmetry of one such CAP-binding sequence was found 61.5 bp upstream of the transcriptional start of the lac operon.

How does CAP increase the efficiency of the lac promoter 50-fold? It binds weakly to RNA polymerase. It has been shown in detail which regions of CAP bind to which regions of a subunit of RNA polymerase (a CTD) [4]. In the case of the lac system, it is the downstream subunit which does the interaction. Thus, the simplest explanation of the 50-fold activation of transcription is that the presence of CAP at the CAP site increase the local concentration of RNA polymerase at an otherwise inefficient lac promoter with an inefficient 210 region. Therefore, the concentration of closed complexes has increased 50-fold. This then leads to correspondingly more open complex and to more starts of transcription. This Interpretation is in line with the recently determined X-ray structure of the CAP-a CTD-DNA complex. The direct protein–protein interactions work without conformational changes [3].

In the case of a repressor loop between O2 and O1, CAP can bind to its binding site; RNA polymerase cannot bind to its promoter. In the case of the loop between O3 and O1, CAP protein cannot bind to its proper site. It can possibly bind when it displaces Lac repressor from O3 to a sequence b5 bp upstream. However, this is not clear.

The level of cAM depends to some extent on the level of glucose. In E. coli CAP is thus used to turn on or to turn off (it then acts as a repressor!) more than a dozen systems. There exists another, similar protein, FNR, with similar regulatory powers, whose activity is regulated by the availability of oxygen.

How General is the Regulation of the lac Operon? The Lac Repressor Family

The E. coli chromosome codes for 16 proteins which are homologous to Lac repressor. Some, but not all of them, have been analysed in detail. What is the bottom line? Every system differs from every other system [11]. The" Gal system", for example, produces a dimeric Gal repressor. Yet the system has two operators separated by 114 bp. The two operators are brought closer to each other by the protein HU, which forms a complex with the DNA binding between the two operators. So, the two dimeric Gal repressors therefore somehow interact with each other. The CAP site is positioned 41.5 bp upstream of the promoter. This then has the consequence that the upstream subunit of CAP interacts with RNA polymerase. In the deo system, Cyt repressor is also dimeric. It does not inhibit RNA polymerase binding to the promoter. It Bonds between two CAPs bound at 293.5 and 240.5 and prevents them from binding with their active sites to the active sites of RNA polymerase. So CytR repressor inhibits the functioning of, but not the binding of, two CAPs to their sites on DNA.

Are there systems where a repressor may form a 'road block' that stops progression of RNA polymerase? Pur repressor is reported to repress the pur B gene by a road block 242 bp downstream of the start of transcription by a factor of 2–3. I do not know of any natural system with a larger effect. It seems that every system of the lac family has its own way of functioning. The same can be said of transcriptional regulation in general. So, if we compare Lac repression with l repression, we note that both of them do so by competing directly with RNA polymerase. Here the similarity is remarkable. Lac and l repressors both use a helix to bind specifically to the major groove of DNA. This is true for most repressors or activators including CAP. But there are a few repressors, such as methionine repressor and some phage repressors, that use b sheets to recognize the major groove of DNA. And finally, we recall that l repressor is able to activate its own transcription.

Other Systems

l Repressor was the other example supporting the Jacoband Monod theory of gene regulation. It inhibits the formation of the closed complex by RNA polymerase at promoters PLand PR (Ptashne, 1987). In the other systems the trp operon – the operator which binds to Trp repressor in the presence of the corepressor tryptophan – has its centre of symmetry within the 210 boxes of the Trp promoter. It Thus seems that direct competition of the repressor with RNA polymerase for binding at its promoter, or competition with CAP or another activator for its binding sites, is a major mode of repression. Instead of an inducer or a corepressor, covalent phosphorylation may be used to change the capacity of a repressor or an activator. Such is the case with OmpR, which binds better to its DNA sites when phosphorylated [7].

There are other methods of regulation of mRNA synthesis which we have not so far encountered; for example, termination and antitermination may be used for transcriptional control. The N protein of phage l acts as an anti-terminator. It is thus essential for the lytic phase of this phage. The trp operon uses antitermination to regulate the amount of trp mRNA. In the presence of tryptophan, a short peptide which is coded by DNA directly in front of a termination signal is translated. This leads to termination of transcription. However, when tryptophan is lacking, the ribosome stops on the two trp (UGG) codons which code for this amino acid. This leads then to antitermination. A similar device is found in the operon. Thus, both start and stop of transcription are used for regulation of transcription. Finally, there are very few cases in E. coli where start of transcription of an operon is inhibited by the formation of a protein–DNA complex that encompasses a large region of the entire operon. Such is the case of the bgl operon. However, it is unclear which proteins bind to which sequences. The major shortcoming of the Jacob–Monod model was the exclusion of the possibility of positive control. Indeed, a detailed analysis indicated that l repressor activates the synthesis of its own mRNA. And then there are systems like the arabinose or the maltose system. In both cases activators, AraC and MalT, are necessary for transcription of these operons in the presence of arabinose or maltose. Both activators are bound directly upstream of the relevant promoters. Presumably they act by increasing the local concentration of RNA polymerase over their weak promoters.

The problems that repressors and activators have to solve in order to work optimally are quite different. repressor has to bind specifically and tightly to its operator. When it occupies 99% of its operator, it represses 100-fold. In contrast, the activator need not bind too tightly to its site. If it saturates only half its site, it activates to 50% of its capacity. The activator, not the repressor, can bind weakly but specifically to RNA polymerase. It needs only an accurate position on the DNA.

Conclusion

When E. coli enters an environment where no more growth is possible the rpoS gene is turned on. It produces a sigma factor, ss, which allows RNA polymerase to use a new set of promoters that so far had been unavailable. In fact, bacteria spend most of their life in the stationary phase.

The corresponding sigma factor gene of Salmonella typhimurium is a major virulence determinant. Genes involved in degradation of organic compounds such as methyl benzoate and m-toluate are controlled by the corresponding gene in Pseudomonas putida. It is a great pleasure for today's researchers to analyse these wonderful constructs of evolution.

References

1. Bell CE, Lewis M (2000) A closer view of the conformation of the Lac repressor bound to operator. Nature Structural Biology 7:209-214.

2. Bell CE, Lewis M (2001) The Lac repressor: a second generation of structural and functional studies. Current Opinion in Structural Biology 11:19-25.

3. Benoff B, Yang H, Lawson CL et al. (2002) Structural basis of transcription activation: the cap a CTD-DNA complex. Science 297:1562-1566.

4. Busby S, Ebright RH (1994) Promoter structure, promoter recognition, and transcription activation in prokaryotes. Cell 79:743-746.

5. Dodd IB, Perkins AJ, Tsemitsidis D, Egan JB (2001) Octamerisation of l CI repressor is needed for effective repression of PRM and efficient switching from lysogeny. Genes and Development 15:3013- 3022.

6.Gilbert W, Mu¨ ller-Hill B (1966) Isolation of the lac repressor. Proceedings of the National Academy of Sciences of the USA 56:1891-1898.

7. Head CG, Tardy A, Kenney LJ (1998) Relative binding affinities of OmpRand OmpR-phosphate at the ompF and ompCregulatory sites. Journal of Molecular Biology 281:857-870.

8. Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. Journal of Molecular Biology 3:318-356.

9. Karstro¨m H (1938) Enzymatische Adaption bei Mikroorganismen. Ergebnisse der Enzymforschung 7:350-376.

10. Mu¨ ller-Hill B (1998) The function of auxiliary operators. Molecular Microbiology 29:13-18.

11. Mu¨ ller-Hill B (2002) What can be learned from the lacR family of Escherichia coli? In: Hodgson DA and Thomas CM (eds) Sixty-first Symposium of the Society of General Microbiology, Cambridge: Cambridge University Press. 143-154.

12. Pace HC, Kercher MA, Lu P et al. (1997) TITEL. Trends in Biochemical Sciences 22:334-339.

13. Ptashne M (1987) A Genetic Switch. Phage l and Higher Organisms, Cambridge, MA: Blackwell Scientific.

14. Re´vet B, von Wilken-Bergmann B, Bessert H, Barker A and Mu¨ ller HillB (1999) Four dimers of l repressor bound to two suitably spaced pairs of l operators form octamers and DNA loops over large distances. Current Biology 9:151-154.

15. Suckow J, Markiewicz P, Kleina LG et al. (1996) Genetic studies of theLac repressor XV:4000 single amino acid substitutions and analysis ofthe resulting phenotypes on the basis of the protein structure. Journal of Molecular Biology 261:509-523.

16. Lwoff A, Ullmann A (eds) (1978) Selected Papers in Molecular Biology by Jacques Monod. London: Academic Press.

17.Miller JH, Reznikoff WS (eds) (1978) The Operon. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

18. Muller-Hill B (1998) The lac Operon. A Short History of a Genetic Paradigm.Berlin:de Gruyter.Bacterial Transcription Regulation

19. Ptashne M (1987) A Genetic Switch. Phage l and Higher Organisms. Cambridge, MA: Blackwell Scientific.

20.Ptashne M, Gann A (2002) Genes & Signals. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

21. Steitz TA (1993) Structural Studies of Protein-Nucleic Acid Interaction, The Source of Sequence-Specific Binding Cambridge: Cambridge University Press.

22.Wagner R (2000) Transcription Regulation in Prokaryotes. Oxford: OxfordUniversity Press.Bacterial Transcription Regulation