**Components of expression vector**

Following are the components of an expression vector:

**Promoter**

* A promoter ensures a reliable transcription of the gene of interest. Also, strong promoters are also necessary for an efficient mRNA synthesis with RNA polymerase.
* Regulation of the promoter is another critical aspect which should always be kept in mind while constructing an expression vector.
* The strongest promoters are those found in bacteriophages T5 and T7.

In *E. coli*, the promoter is regulated in two ways:

**Induction**: the addition of chemical switches on the transcription of the gene.

**Repression**: addition of chemical switches off the transcription of the gene.

The most commonly used promoters in *E. coli* expression system are:

1. The lac promoter:

* It regulates the transcription of the lac Z gene. The lac Z gene is responsible for the production of β- galactosidase.
* The lac Z gene can be induced by IPTG, isopropylthiogalactosidase.
* The lac promoter sequences can be fused to the target gene. It will, then, result in lactose- dependent expression of the target gene.
* The lac promoter has its drawbacks. It is quite weak and cannot be utilized for the high levels of production of the desired protein. In addition to this, the lac genes carry out the basal level of transcription even in the absence of induction (inducer molecule).

2. The trp promoter:

* It is responsible for the regulation of a cluster of genes which are involved in tryptophan biosynthesis.
* Tryptophan acts as its repressor molecule, and it is induced by 3-β-indoleacrylic acid.

3. The tac promoter:

* It is formed by hybridization of the lac and trp promoter. However, it is stronger than either of them.
* The tac promoter is induced by IPTG, isopropylthiogalactosidase.

4.  λPL:

* It is a strong promoter and is responsible for transcription of λDNA in *E. coli*
* The product of λcI gene acts as its repressor. It is called λ repressor.
* The expression construct with the λPLpromoter is used in combination with the *E. coli* mutant host. It is responsible for the production of a temperature sensitive form of λ repressor.
* At low temperatures, the repressor protein represses the transcription whereas the transcription of the cloned gene occurs at high temperatures; because the repressor is inactivated at high temperature.

**Constitutive Promoter**

A constitutive promoter is a kind of promoter which is unregulated and allows continual transcription of its associated gene.

Example of a constitutive promoter: GAP promoter of the gene encoding glyceraldehyde-3-phosphate dehydrogenase.

**Inducible promoter**

An inducible promoter is the one which works in a regulated manner and the expression of genes associated with them can be switched on or off at a particular stage of development or at a certain point of time.

Examples of inducible promoters: AOX1, GAL1, GAL10, nmt1, nmt42, and nmt81.

The **AOX1 promoter** of the gene encoding alcohol oxidase. It is induced by methanol and is best-suited for expression of the protein in *Pichia pastoris.*

The **GAL1** and **GAL10** promoters are other examples. They are induced by galactose and are suitable for protein expression in *Saccharomyces cerevisiae.*

The **nmt1, nmt42,** and **nmt81** promoters which are induced by thiamine for protein expression in *Schizosaccharomyces pombe.*

**Reporter Gene**

* The reporter gene is responsible for the production of the protein which can be detected and quantified with the help of a simple assay.
* They serve as a tool to measure the efficiency of the gene expression and also to detect the intracellular localization of the protein.
* The rate of expression of the structural gene is dependent upon the regulatory sequences which are located upstream to it.
* The rate of expression of the gene can be measured by replacement of its protein-encoding portion. Also, it can be fused to another gene which expresses another protein. The presence of this another protein can be easily identified.
* Reporter genes are useful in the identification of promoters, enhancers, and other proteins or regulatory elements which bind to them.

The most commonly utilized reporter genes are:

1. lac Z gene of *E. coli*

* It acts as a reporter in the presence of X- gal.
* Its levels are easily detected by the intensity of colour which is produced. The intensity of the blue colour produced is quantified.

2. CAT (chloramphenicol acetyltransferase) encoding gene of *E. coli*

* The CAT gene encodes chloramphenicol acetyltransferase.
* The transferase enzyme is responsible for the transfer of acetyl groups from acetyl CoA to the recipient antibiotic, chloramphenicol

3. Luciferase encoding gene of firefly, *Photinus pyralis*

* Luciferase is accountable for the oxidation of luciferin.
* The oxidation of luciferin results in the emission of yellow-green light. The emission of light is easily detected irrespective of the low levels.

4. Green fluorescent protein (GFP) encoding gene of jellyfish, *Aequorea victoria*

* It is an auto fluorescent protein with 238 amino acid residues produced by the bioluminescent jellyfish *Aequorea victoria.*
* In GFP, β-barrel is formed by eleven β strands. An α- helix runs through the center. The chromophore is located in the middle of the barrel. The amino acid residues from 65 to 67 with sequence Ser-Tyr-Gly form the chromophore, p- hydroxybenzylideneimidazolinone, which is fluorescent. The chromophore fluoresces at a peak wavelength of 508 nm (green light) when it is irradiated with UV or blue light (400 nm).
* GFP serves as a tool for determining protein localization.
* It serves as a tag whereby it is fused with a protein whose expression is to be monitored. Basically, the subcellular localization of the protein is investigated.
* Genetic engineering techniques help in the production of vectors which contain the coding sequence of the unidentified protein, X, cloned in the coding sequence of the GFP.
* This fusion product of GFP-X can now be transfected into target cells and the expression, as well as the subcellular location of the X protein, can easily be monitored and detected.

**Ribosome Binding Site and Translation Initiation Site**

* The ribosomal binding site (RBS) follows the promoter. It is responsible for the efficient translation of the cloned gene.
* The translation initiation site in case of prokaryotes is known as the Shine Dalgarno sequence. This sequence is enclosed within the RBS only.
* The consensus sequence of the translation initiation site includes a set of 8 base pairs present upstream the AUG start codon.
* The translation in eukaryotes is initiated at a particular sequence called Kozak sequence.
* The ribosomal machinery for the translation of mRNA is assembled on this site.

**Polylinkers**

* Each vector contains particular recognition sites for restriction enzymes. It is at the restriction site that the vector is excised to clone the foreign gene of interest.
* These sites often lie close together and, hence, are called polylinkers or multiple cloning sites (MCS).
* These regions are 50 to 100 base pair in length and may have a cluster of up to 25 restriction sites.

**Poly-A (polyadenylation) Tail**

* The poly-A tail present, at the end of the mRNA formed, protects the mRNA from degradation by the exonucleases or endonucleases.
* It is extremely critical for the stability of the mRNA.
* It is also responsible for the termination of transcription and translation and stabilizes the mRNA production.
* A nucleolytic enzyme complex and a poly-A-polymerase are prerequisites for the addition of poly-A tail at the end of the mRNA.