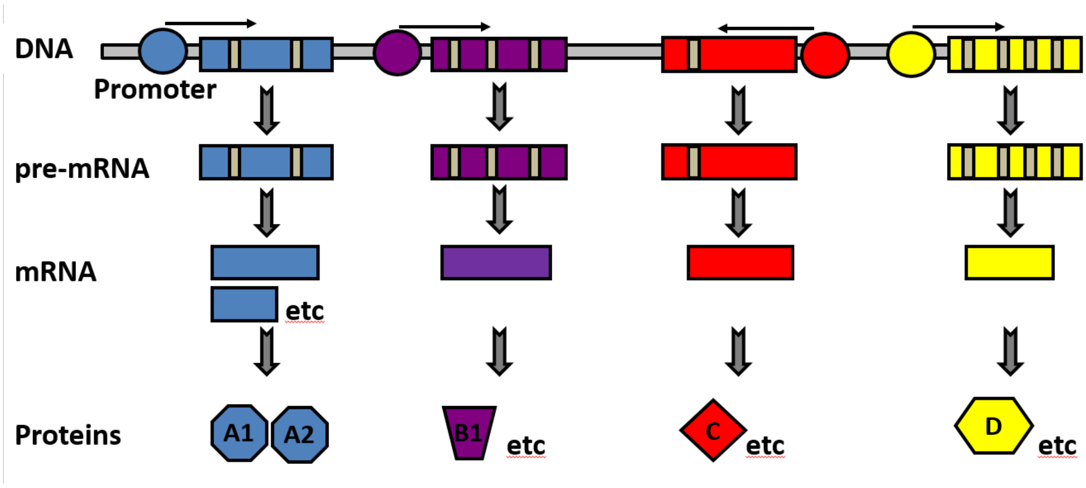
Gene Architecture in Eukaryotes: Single transcriptional units

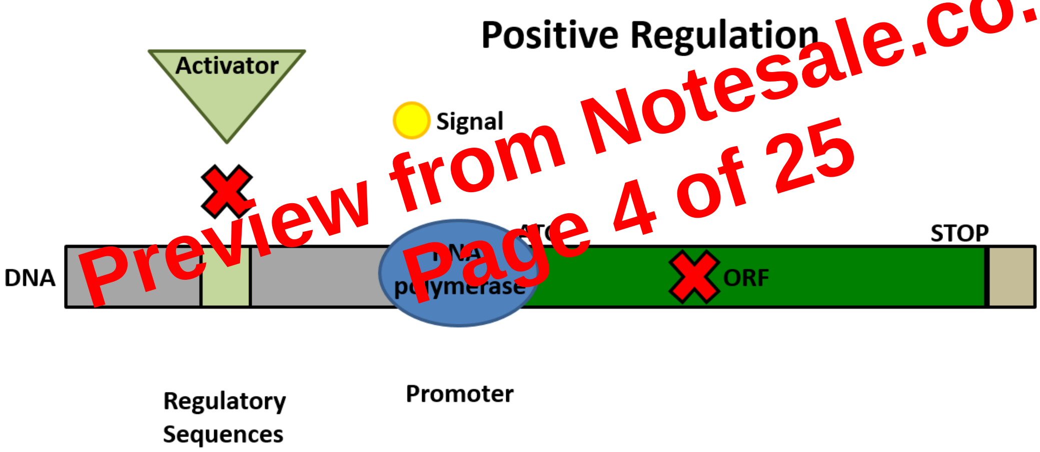
Eukaryotes – every single gene will have its own promoter (monocistronic)



- **Coding sequences are often interrupted by intervening sequences (introns)**

Alternative splicing can produce multiple proteins from the same mRNA by different order of combinations of exons

General principles of Gene Regulation



RNA polymerase can bind to the promoter of a gene, but the binding is not strong for the initiation to occur to make the mRNA

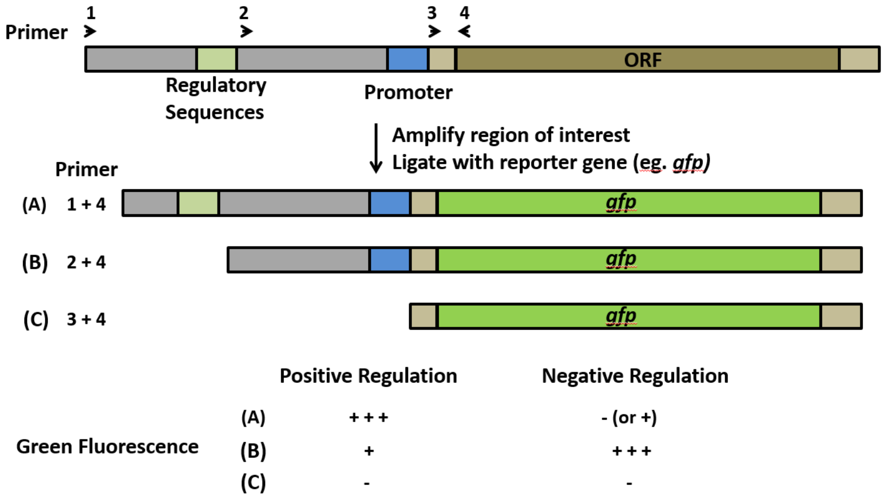
Therefore, we need an activator, it doesn't bind to the regulatory sequence until it gets activated by a signal in the form of molecules/light/phosphorylation etc

Signal binds and induce a conformational change which usually is a dimerization of the regulatory proteins binding together, then bind to the DNA and influence the strength of the binding of the polymerase. RNA polymerase notices the binding of the activator which may be a kink or directly in contact with the polymerase to allow a stronger binding with the promoter. Polymerase gets released and start transcription.

This is positive regulation – molecule comes in, bind to regulator molecule and up regulate transcription

How to identify *cis* elements

Reporter Gene Assay: Promoter mapping



Take the regulatory unit of the DNA and with PCR make different size versions of the regulatory DNA

Primer 1+4 gives the entire section of the regulatory unit, cis regulatory unit in green and promoter region in blue. If positively regulated, gene will be able to express and gfp is transcribed and translated. If negatively regulated promoter, we don't get green fluorescence, but this is not 100% negative as the repressor protein may sometime falls off (equilibrium between the bound and unbound state) and the polymerase can quickly bind to promoter and transcribe.

Primer 2+4, losing a positive regulatory site means the activator cannot bind to promote transcription and the polymerase would just hover around the promoter site without that push to activate transcription. Therefore, only little transcription. Losing a negative regulatory site, so the gfp gene cannot be switched off and lots of transcripts produced

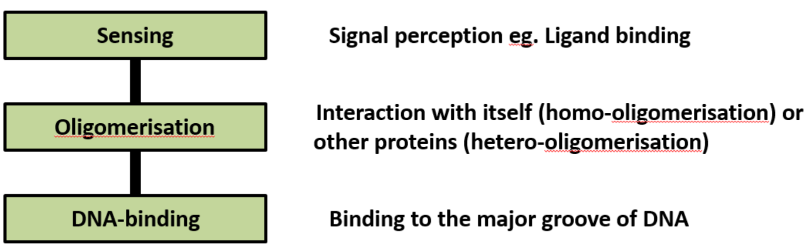
Primer 3+4 cuts both regulatory sequences and promoter, so no gfp produced.

By producing lots of cunctation, we slowly reduce the length of the regulatory parts of the gene and observe the outcomes to map where these regulatory units are.

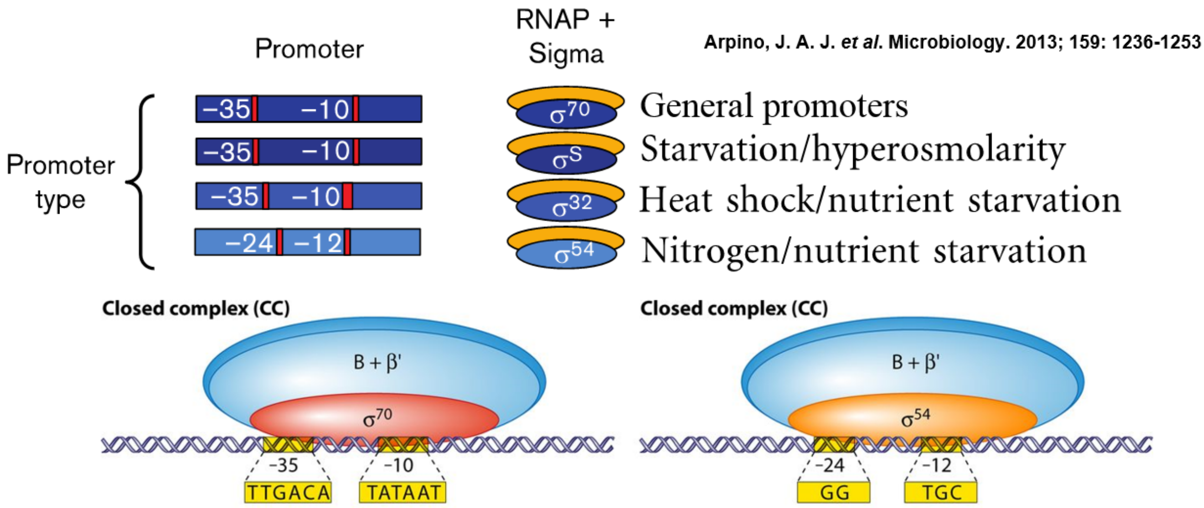
Trans elements

- Transcription factors (Proteins) that bind control regions (cis elements) to stimulate or inhibit gene expression
- They are diffusible factors that act on a different molecule of DNA from where they are encoded

Domain Architecture of *trans* elements



Sigma domain recognises -35 and -10 consensus sequences and recruit the apo-RANP to bind to the DNA.

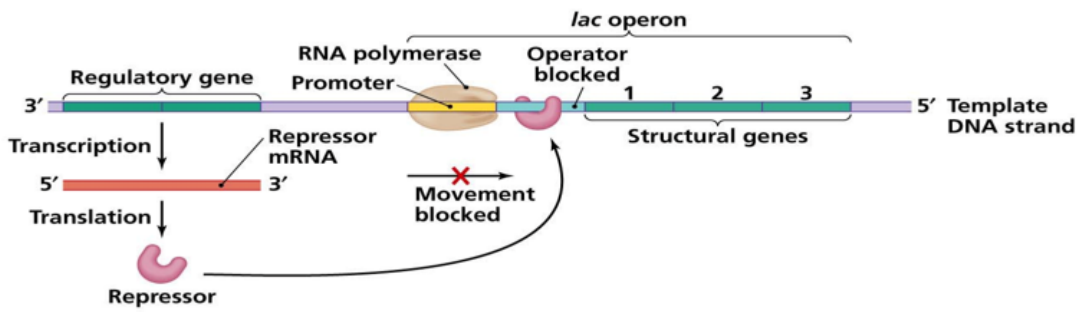


Different types of sigma units:

Initiation of transcription by the RNAP-σ^{70} (A) and RNAP-σ^{54} (B) holoenzymes. The σ^{70} factor directs the binding of polymerase to the consensus −10 (TATAAT) and −35 (TTGACA) sequences to form an energetically unfavorable closed complex (CC) that is readily converted into an open complex (OC) to initiate transcription. In contrast, the σ^{54} factor directs the binding of RNAP to conserved −12 (TGC) and −24 (GG) promoter elements that are part of the wider consensus sequence YTGGCACGrNNNTTGCW (where uppercase type indicates highly conserved residues, lowercase type indicates weakly conserved residues, N is nonconserved, Y is pyrimidines, R is purines, and W is A or T) (10). This forms an energetically favourable CC that rarely isomerizes into the OC. In order to form the transcription “bubble,” a specialized activator (a bacterial enhancer binding protein [bEBP]) must bind and use the energy from ATP hydrolysis to remodel the holoenzyme.

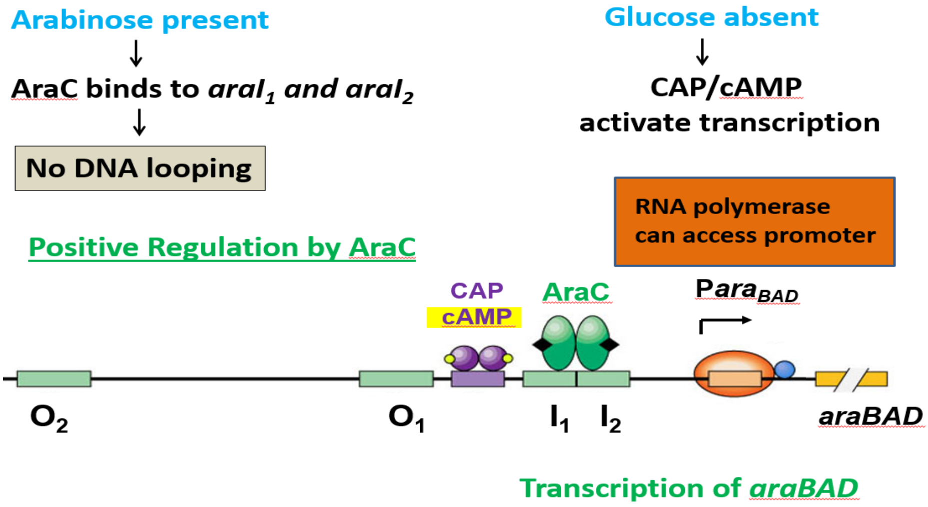
The *lactose* operon

Negative Regulation



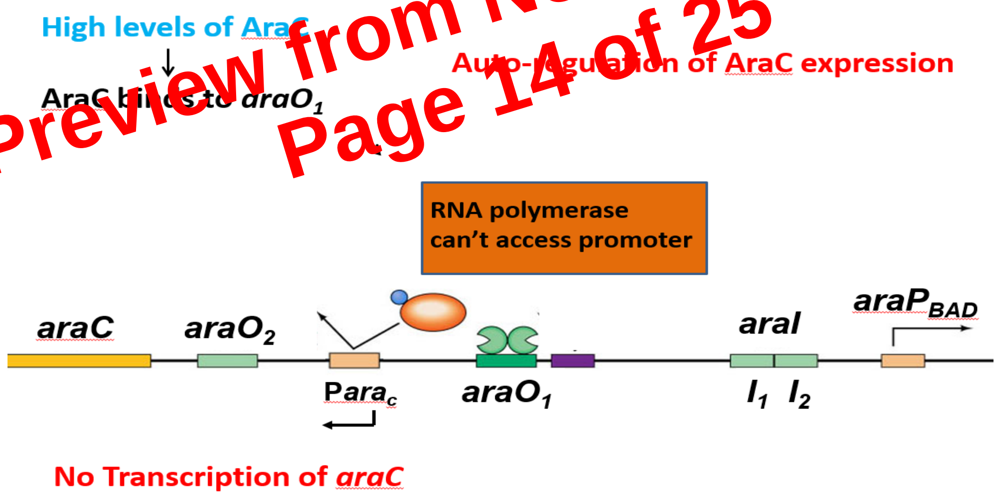
No lactose - Repressor protein binds to the operator site & blocks RNA polymerase

In the absent of arabinose, AraC binds to two sites, araO2 and araI1. I1 which is to the promoter gene and O2 upstream of the promoter form a DNA loop and blocks the promoter ParaBAD. There is therefore no space for the RNAP to bind. (Negative regulation by araC in the absent of arabinose)



In the present of arabinose, it binds to the araC protein and causes a conformational change in the protein. Instead of binding to I2 and O2, it now binds to I1 and I2. It loses the DNA loop and AraC acts as an activator. Simultaneously, the adenylate cyclase is activated in a low glucose level to form CAP cAMp that binds to the catabolite site. RNAP can access promoter, transcribing araBAD gene and now we can metabolise arabinose.

Alter regulation, araC acts as both negative and positive regulator. The level of araC should not be too high to get accidental activation of the system.



Self-regulation: AraC made can binds to the promoter that drives araC and repress the production of araC. At high level of araC, it binds to its own promoter and switches it off. Then level of araC in the system falls. When the araC level drops, the araC falls off the promoter site and level starts to rise again. Always at a constant level!

When no araC proteins are present, RNAP is free to transcribe the gene that codes for the production of araC proteins. When this protein is present, it binds to the O1 site. This physically prevents transcription of the araC gene. AraC protein is able to control its own transcription. During repression, a dimeric form of the araC protein binds I1 and O2 site forming a DNA loop. Present of