Control of Gene Expression in Bacteria

Introduction

- Most control is exerted through the regulation of transcription initiation

By DNA binding proteins

- The classic model is the *lac* operon
- Lactose utilisation in *E. Coli* requires three genes
  - *lacZ* - BETA-galactosidase, which cleaves lactose to galactose and glucose
  - *lacY* - lactose permease
  - *lacA* - lactose transacetylase
- These three genes are grouped together and transcribed as a **polycistronic unit** from a **single promoter**
- The structure of the operon is Promoter (lacI) - lacI gene - Promoter (lacZYA) - lacZYA gene
- It constitutes an **operon**
- **The *lac* repressor**
  - The lacZYA transcription unit contains an **operator site (Olac)** at the end of the lacZYA promoter region
    - Palindromic
  - It binds a protein called the **lac repressor** which is a potent inhibition of translation when bound (it physically impairs the advance of the polymerase)
    - It is a tetramer, and has a very high affinity for the operator site
    - This symmetry matches the symmetry of the operator
  - This lac repressor is encoded by a separate gene lacI which is also part of the lactose operon
  - Curiously, it **increases** the binding of the polymerase to the promoter region by two orders of magnitude
  - So both are there at the same time
- Induction
  - When lactose is present, low basal levels of the permease (lacY) allow it to enter the cell
  - Low levels of BETA-galactosidase (lacZ) convert it to allolactose
  - This acts as an **inducer** and binds to the lac repressor
  - This causes a comformational change in the lac repressor, reducing its affinity for the lac operator
  - The removal of the repressor allows the already-bound polymerase to rapidly begin transcription
  - Removal of the lactose leads to an almost immediate inhibition
    - The repressor rapidly re-occupies the operator site
    - The lacZYA transcript is extremely unstable
- cAMP promoter
  - The lactose promoter is not very strong, and in its absence, levels of transcription are not very high
  - For a high level of transcription, activity of **cAMP receptor protein (CRP)** is needed
    - or **Catabolite regulator protein**
    - or **Catabolite activator protein**
    - Contains helix-turn-helix domain, commonly found in many DNA-binding proteins
    - This forms a V shaped protrusion which inserts intself into the major groove of the DNA
    - Most DNA binding proteins are dimers, and most recognition sites are palindromic (symmetry)
  - When glucose is present, this is not activated
  - The activity is mediated by CRP, which exists as a dimer which cannot bind to DNA on its own
    - In the absence of glucose, cAMP levels increase, and CRP binds to cAMP
    - The CRP-cAMP complex binds to the lactose promoter just upstream from the site for the RNA polymerase
    - This induces a 90 degree bend in DNA, which is believed to enhance RNA polymerase binding to the promoter
    - Enhances transcription by 50-fold
  - The cell uses glucose preferentially over lactose

Modulating promoter specificity of RNA polymerase

- Different sigma factors for different situations

Regulating transcription termination

- Attenuator structures are located prior to the structural genes they are intended to control
- Act to prevent transcription of those genes if need be
- Consists of
  - A sort upstream "leader sequence", rich in certain amino acids
  - An "attenuator" sequence, which can potentially form two different stem-loop secondary structures
- An example is in the *trp* operon, which codes for enzymes responsible for tryptophan synthesis
- The repressor region contains 4 regions, which can form two different structures
  - Either 1 with 2 and 3 with 4 (3+4 structure is a *terminator*)
  - Or 2 with 3 (**anti-terminator**)
- Just before these structure are two codons coding for tryptophan
- RNA polymerase pauses once it gets to the first of these sites, and waits for translation to reach these residues
- If tryptophan levels are **high**
  - The ribosome will whip past the two tryptophan residues
  - It is then occluding sequence 2
  - Therefore, as the polymerase approaches the terminator sequence, the 3:4 hairpin can form, and terminates
- If they're **low**
  - The ribosome pauses
  - This blocks sequence 1, and the 2:3 sequence is able to form
  - Transcription continues
- Commonly found in amino acid biosynthetic operons

Translational control

- Ribosomal proteins are organised into operons
- One protein from each operon binds near to the 5' end of their own polycistronic mRNA, inhibiting translation
- As the level of the ribosomal protein falls, less of the protein binds and the translational inhibition is relieved