GENE30002 Genes: Organisation & Function **[Notes]**

L1 Initiation of Transcription

RNA Polymerase: *E. coli* [Prokaryotes]

- Core enzyme catalyses RNA synthesis. Only one RNAP complex responsible for **all** transcription.
  - Essential: rpoA, rpoB & rpoC form α, β & β' subunits of complex [loss: lose RNAP activity].
  - Non-essential: rpoZ gene product: ω subunit aids assembly of complex [loss: complex still forms].
- σ factor recognises promoters by their consensus sequence. Most genes are transcribed by σ^{70}.
  - Essential: rpoD gene prod.: σ subunit initiates transcription by allowing RNAP to recognise & bind promoter.
  - Different σ-factors initiate transcription at different sets of promoters & allow global shift in gene expression under particular circumstances (i.e. environmental stress).

Phases of Transcription

- Recognition: RNAP holoenzyme (core + σ^{70}) associates & recognises -35 & -10 consensus. **Closed complex** formed.
- DNA melting: DNA melted & unwound at -10 promoter region (AT rich): **Open complex**: Allows RNAP to transcribe.
- Initiation: Short chains are synthesised & released as bound RNAP shutters until >10 bp length transcript made.
- Elongation: ssDNA is protected by RNAP as it moves through & double-helix is re-established. RNAP: DNA -> mRNA
- Termination: RNAP & RNA transcript are released.

RNAP can initiate without a pre-existing template (DNAP requires primer). Signals within the DNA seq allow RNAP to initiate promoter sequence. Investigate functionally important DNA seq: (a) comparative; (b) genetic; (c) biochemical.

(A) Comparative

- Comparing *E. coli* promoter sequences: Conserved sequences at -10 & -35; A **consensus** seq can be compiled.
- Region between -10 & -35: no strong consensus for any sequence, appear to separate -10 & -35 regions.
- Weak promoters tend to differ most from consensus while strong promoters tend to match consensus most.

(B) Genetic Approach

- Isolate mutants that affect promoter function to determine location of nt in sequence.
- 75% mutants in -10 & -35 regions while remainder are close to these regions.
- Occasional mutations ↑ promoter strength, generally ↑ the similarity to the consensus sequence e.g. lacI.
- Most mutations which ↓ promoter strength ↓ the similarity to the consensus sequence.

Prokaryotic promoters: Mostly contain -10 & -35 consensus sequences.

- May have an additional UP-element, found in genes that are highly expressed e.g. rRNA genes.
- May be w/o -35 but have extended 'discriminator' sequence downstream of -10. Allows RNAP to interact strongly.

(C) Biochemical Approaches to analysing protein-DNA interactions

In vitro experiments using cloned fragments of DNA (e.g. promoter region); & DNA-binding proteins (e.g. RNAP) in crude extracts or purified.

Electrophoresis gel mobility shift assays:

- Radioactively label dsDNA fragment, Mix with protein, Electrophorese.
- Mobility of DNA fragment through gel will be retarded if protein binds (protein capable of binding within DNA).
- Indicates protein binds DNA (**somewhere**: does NOT infer function).
- Absence of protein (-): labelled dsDNA fragment is run on gel & naked DNA will run to a point.
- Add protein (+): If protein can bind sequence within, complex runs slowly & unbound DNA will run to same point.
- Add more protein (++): If more binding can occur, complex will run even more slowly.

DNA Footprinting (DNA modification protection)

- Radioactively label dsDNA on one end of one strand; Mix with protein; Treat with DNase I to cleave DNA "random"
- Denature, electrophorese; Use label to visualise DNA fragments. Lac promoter (DNA) + RNAP (protein):
- IF binding site present, protein & DNA can bind together.
- (-) Protein: Fragments cut at every nt; band on gel represents 1nt fragment correspond to entire sequence.
- (+) Protein: Prevent Dnase I from cleaving DNA: missing band: locate bound protein (-10 & -35)
- Indicates region of DNA protected by protein & ~binding sites. Doesn't indicate **which** bases required for binding.

ChIP-seq: Genome-wide identification of binding sites

- Identify protein binding sites to its DNA target *in vivo* & expose cells to agent to cross-link proteins to DNA.
- Crosslink protein & DNA, lyse cells, fragment DNA
- Immunoprecipitate protein, reverse cross linking; Purify DNA & sequence.
- Align common sequences of bound protein: identify sequence within genome & adjacent genes.

POSSIBLE mechanisms that MIGHT regulate expression of a prokaryotic gene:

Regulating transcription initiation by regulatory proteins. Terms:

- POSITIVE vs. NEGATIVE: Refers to action of **regulatory protein**.
  - Negatively acting protein actively represses expression of a gene.
  - Positively acting protein actively promote expression of a gene.
- *CIS* & *TRANS*: Whether regulatory component is a **site in DNA** (cis) or a **diffusible gene product** (trans: acts at a cis site) e.g. RNAP is trans-acting diffusible protein that can bind cis-acting sites within DNA.
- REPRESSIBLE vs. INDUCIBLE: Refers to action of small MW **effector** molecule.

Complementation of a *lacI* mutant "in trans":

- lacI: repressor protein (trans-acting factor) that binds at operator to prevent expression of lac structural genes.