PROKARYOTE GENE EXPRESSION

- negatively regulated - gene repressor protein that binds to an operator region
- positively regulated - gene activator proteins that binds to proper regions
- ligands binds repressors and activators and modulate their activity

LAC OPERON HAS THREE OPERATOR REGIONS

- Lac repressor is a single tetrameric protein that binds two operators simultaneously

1) O1 is the main operator region, first identifiers repressor and binds at position +11 from transcriptional start site
2) and 3) are secondary operators and bind +412 and -82 respectively

ACTIVATORS AND REPRESSOR PROTEINS

- dependent on location of binding sequence
- e.g. bacteriophage λ repressor

σ FACTORS

- binds to promoter sequence and associate with RNA polymerase
- most common is σ^{70} - initiation factor

E.COLI σ^{54}

- σ^{54} containing RNA polymerase transcribes genes that are controlled by activators that bind the DNA 80-160bp upstream in enhancer regions
- e.g. NtrC is a σ^{54} activator

TWO COMPONENT REGULATORY SYSTEM

1) sensor protein (NtrB and PhoR) - transmitter domain
   - is a transmembrane protein
   - regulated by a unique periplasmic domain that senses the environment
   - shows specificity for the receiver domain of the response regulatory protein they phosphorylate
2) response regulatory protein (NtrC and PhoB) - receiver domain
   - phosphorylation of receiver domain regulates the activity of a second functional domain → increase transcription

PhoR ACTIVARES PhoB

1) low phosphate in environment → low phosphate in periplasmic space
2) phosphate dissociates from PhoR periplasmic domain
3) conformational change in cytosolic domain → activating protein kinase
4) phosphate from ATP transferred to histamine residue on cytosolic domain and then transferred to PhoB
5) PhoB activated → induces transcription of several genes

NtrB ACTIVATES NtRC

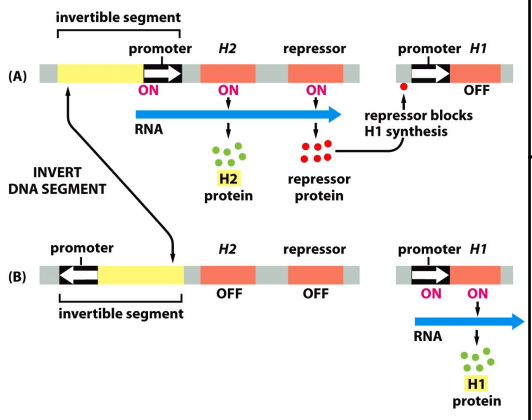
- NrtB phosphorylates NtrC in response to low glutamine levels leading to NtrC binding as enhancer

PHASE VARIATION

- switching gene expression by DNA arrangements
- e.g. salmonella - genes encoding flagellin

- inversion changes orientation of the promoter

1) genes encoding H2 flagellin and a repressor protein are transcribed, binding to H1 promoter repressing flagellin H1 gene
2) H1 gene is transcribed because H2 and repressor genes are not transcribed



TRASCRIPTION FACTORS

- gene regulator proteins
- have two domains:

1) DNA binding domain binds specific nucleotides sequences
2) activation of repression domain which interacts with other proteins to regulate transcription

- identified based on type of DNA binding domain they contain

TF BIND AT MAJOR GROOVE OF DNA

- edges of base pairs exposed giving a distinct pattern of features for each bp
- binding includes H bonds, ionic bonds and hydrophobic interactions
- minor groove do not have distinct pattern (GC and CG are identical)

HELIX-TURN-HELIX MOTIF

- two α-helices connected by short aa chain held at a fixed angle (turn)
- recognition helix - C-terminal helix at major groove
- structural helix - N-terminal helix positions recognition helix

HELIX-TURN-HELIX MOTIF CONTAINING PROTEINS

- proteins allow variations in presentation of the recognition helix ⇒ increases versatility of the motif
- bind DNA as symmetric dimers

HELIX-LOOP-HELIX MOTIF

- a non helical loop connects two α-helical regions
- dimerisation: homodimers or heterodimers
- C-terminal in both monomers form a coiled coil structure
- N-terminal regions of the α helices contain aa with basic side chains interact with DNA

ZINC FINGERS

- one or more Zn2+ ions are part of the motif
- regions of protein fold around the Zn2+ ions
- cys-cys-his-his zinc finger motif
  - 23-26 aa
  - Zn2+ ions hold α-helix and β-motifs together
  - α-helix contacts major groove
- cys-cys-cys-cys zinc finger motif
  - 55-56 aa including 4 conserved Cys residues
  - two α helices and two β sheets
  - Zn2+ ions stabilise the DNA recognition α-helix and a loop involved in dimer (homer or hetero dimer) formation

LEUCINE ZIPPER MOTIF

- leucine residue every 7th position ⇒ amphipathic helix formed (hydrophobic aa on one side)
- coiled-coil strucutre forms due to hydrophobic interactions of aa side chains
- bind DNA as dimers
  - each monomer is composed of an α-helix, N-terminus binds to -vely charged phosphates in backbone
  - dimer contacts two adjacent major grooves

TWO-STRANDED B-SHEET MOTIF

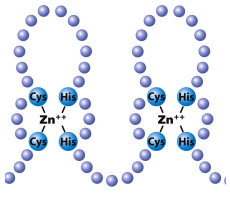
- aa side chains extend from β-sheet toward target DNA major groove

LOOP REGIONS e.g. p53

- some regulatory proteins have loop regions which recognise major and minor grooves

HETERODIMERISATION

- each monomer recognises the same DNA sequence - allows different transcriptional responses as different activation mains are brought together
- each monomer recognises different DNA sequences - increases combinatorial possibilities that a TF family can bind to



EUKARYOTIC GENE CONTROL REGIONS

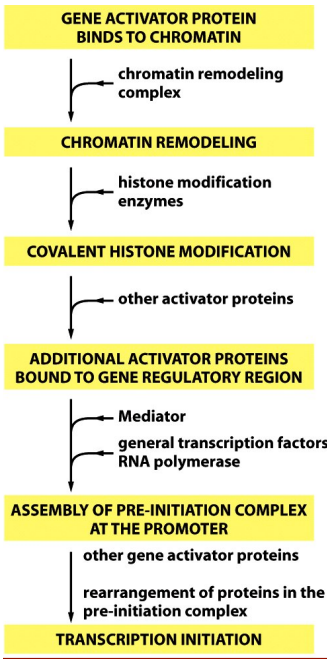
- promoter
- TATA box - promoter rich in As and Ts
- general transcription factors (GTFs) - proteins required for transcription initiation of all RNA pol III transcribed genes
  - TFIID - distorts DNA at promoter
  - TFIIH - helicase activity → unwinds DNA for RNA pol II access
- regulatory sequences - where gene regulatory proteins bind to control rate of assembly at promoter
- looping of DNA allows interaction of gene regulatory proteins
- proximal promoter regions - within 100-200 bp from TSS

MEDIATOR

- transcriptional complex -30 subunits
- binds RNA pol II
- one subunit has histone acetylene activities → promotes hyperacetylated state favouring transcription

DELETION ANALYSIS

- isolate upstream region on gene and remove using restriction enzymes or exonuclease to make a series of upstream deletions
- insert a vector containing reporting gene ⇒ identifies upstream region responsible for gene expression



REPRESSOR PROTEIN MECHANISM

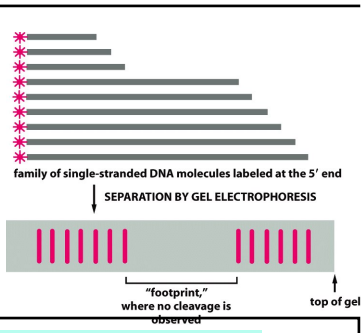
A. activator and repressor proteins compete for same regulatory DNA sequence
B. activator and repressor proteins bind the DNA, but repressor proteins binds activation domain of activator
C. repressor protein prevents assembly of GTFs or RNA pol II release from GTFs
D. repressor protein recruits chromatin remodelling complex that returns chromatin to pre-transcription state
E. repressor protein recruits histone deacetylase to promoter
   - histone acetylation can stimulate transcription initiation ∴ repressor reverses this modification
F. histone methylation typically leads to recruitment of other DNA-binding proteins and gene silencing

IDENTIFICATION OF GENE REGULATORY REGIONS: DNA FOOTPRINTING

- isolate upstream region of gene and label 5-end using polynucleotide kinase and labelled ATP
- nuclease digestion except where regulatory protein is bound

IDENTIFICATION OF GENE REGULATORY REGIONS: EMSA

- electrophoretic mobility shift assay
- isolate and label regulatory region and add cell extract
- DNA bound proteins in cell extract show different mobility (fragments bound to proteins are retarded



TRANSCRIPTION OF HUMAN MITOCHONDRIAL DNA

- genome is circular, double-stranded molecule
  - heavy strand (H) = encodes most genes
  - light strand (L)
  - strands transcribe in opposite directions
  - each strand transcribed as a single transcript from a single promoter
- mitochondrial RNA polymerase (POLRMT)
  - single subunit protein
  - POLRMT gene encoded and transcribed in nucleus, mRNA translated in cytoplasm and protein imported into mitochondrial matrix
- two transcription initiation factors (TFB2M, TFAM) involved in placing POLRMT at promoters

COORDINATED EXPRESSION OF MULTIPLE EU GENES

- single gene regulatory protein can coordinate the expression of several genes
- a gene may respond to many gene regulatory proteins
- some gene regulatory proteins are cell specific but most are active in a variety of cells
- e.g. glucocorticoid receptor
  - with bound glucocorticoid hormone ⇒ 3 genes expressed
  - each with an activator protein bound ⇒ low level of gene expression
- repressor or activity is dependent on other gene regulatory proteins bound at control centre

EUKARYOTIC NUCLEAR RNA POL

- each eukaryotes contain 3 nuclear pol
- have 5 core subunits homologous to E.coli pol ⇒ early evolutionary origin
- additional 4 small subunits
- enzyme specific submits

EUKARYOTIC POLYMERASE I

- located in the nucleus
- RNA pol I positions at TSS
- RNA pol I-specific regulatory sequences
  - core element at TSS and upstream element
- initiation transcription
  - binding of multimeric upstream activating factor (UAF)
  - binding of trimeric core factor (CF)

EUKARYOTIC POLYMERASE II

- largest subunit = RPB1
- essential carboxyl-terminal domain (CTD)
  - unique to RNA pol II, 7 conserved aa
- 5 potential phosphorylation sites
  - unphosphorylation at start of transcription
  - phosphorylation state changes during transcription elongation ⇒ allows proteins to interact

EUKARYOTIC POLYMERASE III

- RNA pol III-specific regulatory sequences and GTFs
  - promoter regions within the transcript sequences - A and B box (C box in 5s-rRNA gene)
  - 3 GTFs - required to initiation transcription (TFIIIA, TFIIIB, TFIIIC)

TRANSCRIPTION OF HIGHER PLANT CHLOROPLAST DNA

- genome is circular and double stranded
  - stands transcribe in opposite directions
  - multiple promoters, genes transcribed in groups
- two types of RNA polymerases involved

1) nuclear-encoded polymerase (NEP)
   - single subunit protein
   - transcribed in nucleus, translated in cytoplasm and imported into chloroplast stroma
   - transcribes subunits for PEP
2) chloroplast-encoded polymerase (PEP)
   - multi subunit protein
   - associated with σ^{70} factors

POSTTRANSCRIPTIONAL CONTROL OF EU GENE EXPRESSION

- (regulation of transcription initiation is the principle mechanisms for gene expression control)
- mRNA never occur as free RNA molecules - always have associated proteins; hnRNA (pre-mRNA) → mature mRNA → cap and poly(A) tail and introns removed → leave nucleus as cytoplasmic mRNPs