

BCH 4054 Chapter 31 Lecture Notes

Slide
1

Chapter 31

Transcription and
Regulation of Gene Expression

Slide
2

Messenger RNA

- Central Dogma (Francis Crick, 1958)
 - DNA → RNA → Protein (Fig 31.1)
- Jacob-Monod Hypothesis: Four properties
 - Base composition reflecting DNA
 - Heterogeneous in size
 - Can associate with ribosomes
 - High rate of turnover

Early labeling experiments showed ribosomes as the site of protein synthesis. An alternate hypothesis was that each protein had its own specific ribosome where it is made.

Slide
3

Other Forms of RNA

- Ribosomal RNA
 - Major RNA component of cell
- Transfer RNA
 - Small RNA molecules carrying the amino acids in protein synthesis
- Eukaryotic “small nuclear” RNA
 - mRNA processing in eukaryotes
- Viral RNA

All forms are made by copying a DNA template (excepting some RNA viruses), so the process of transcription is common to all.

Slide
4

Transcription, General Features

- A “template” DNA strand is copied using the Watson Crick base pairing rules
 - See Fig Page 1016 for nomenclature convention
- Only small segments of DNA copied at any one time.
 - Must be specific start and stop sites
 - Copying must be regulated
- Chemistry is similar to that of DNA polymerase
 - Nucleoside triphosphates add to 3' OH end with pyrophosphate as the product.

Slide
5

Transcription, General Features, con't.

- DNA separates and forms a “transcription bubble”
- No 3'-5' proofreading, so error rate is about 1 in 10^4
- Topoisomerases needed to introduce negative supercoiling in front, remove it in rear
 - Otherwise RNA would end up wrapped around the DNA (See Fig 31.6)

Slide
6

Transcription in Prokaryotes

- A single polymerase
 - $\alpha_2\beta\beta'\sigma$ structure
 - β' binds to DNA, β binds substrate NTP's
 - σ recognizes start site, called the **promoter**
 - Several different σ 's, recognizing different promoters
 - Not required for RNA synthesis; dissociates after transcription starts, leaving **core enzyme**

Slide
7

Stages of Transcription in Prokaryotes

- Binding of Polymerase at promoter site
- Initiation of polymerization
- Chain elongation
- Chain termination
 - (See Fig 31.2)

Slide
8

Transcription in Prokaryotes: Binding

- **Closed promoter** complex formed
 - K_d is 10^{-6} to 10^{-9} M
- Polymerase unwinds DNA to form **open promoter** complex
 - K_d is $\sim 10^{-14}$ M
- Promoter sites characterized by **DNA footprinting**
 - See Fig page 1018

Slide
9

Prokaryotic Promoter Sites

- About 40 bp region on the 5' side of the transcription start site
- Two consensus sequence elements
 - “-35 region”, consensus TTGACA
 - Binding site for σ subunit
 - Pribnow box near -10, consensus TATAAT
 - Easier to unwind DNA
- See Fig 31.3

Different σ units recognize different promoter sites. For instance, “heat shock” proteins are recognized by a specific σ unit. Not all variation in rate is due to σ factor difference, however. There are “strong promoters” and “weak promoters”, depending on how close they are to the consensus sequence. Rate of transcription therefore is dependent on the promoter structure.

Slide
10

Transcription in Prokaryotes: Initiation

- Two polymerase binding sites
 - Initiation site and elongation site
 - Initiation usually begins with a purine
- After 6-10 nucleotides added, σ subunit dissociates
- Rifamycin B blocks initiation by blocking NTP binding to β subunit
- Rifampicin blocks translocation
 - See Fig 31.4 for structures

Slide
11

Transcription in Prokaryotes: Elongation

- 20-50 bases /second in *E. coli*
 - Slower in GC rich regions
- Cordycepin (Fig 31.5) blocks elongation
- Actinomycin D, an intercalating agent, also inhibits (Page 370 not discussed in book)
- Topoisomerases necessarily involved

Slide
12

Transcription in Prokaryotes: Termination

- Two mechanisms
 - Rho-termination factor protein (Fig 31.8)
 - An ATP dependent Helicase that unwinds DNA-RNA hybrids
 - Recognizes C rich regions, overtakes polymerase
 - Specific sequence site termination (Fig 31.7)
 - Inverted repeat, rich in G-C, forms stem-loop
 - Followed by 6-8 U's

Slide
13

Transcription in Eukaryotes

- Three classes of RNA polymerase
 - I (or A), makes rRNA precursors
 - Not inhibited by α amanitin
 - II (or B), makes mRNA
 - K_i for α amanitin 10^8 M
 - III (or C), makes tRNA's and some small RNA's
 - K_i for α amanitin 10^6 M
- Large multimeric proteins
 - All have large subunits similar to β and β' of E coli.

Slide
14

RNA Polymerase II in Yeast

- 10 different peptides (Table 31.1)
- RPB1 and RPB2 homologous to β and β'
- RPB1 binds to DNA
 - Has C terminal domain PTSPSYS
 - Many OH groups for phosphorylation
 - Only unphosphorylated form can initiate synthesis
 - Elongation requires some phosphorylation

Slide
15

Eukaryotic Transcription Factors

- Interaction with promoters involves “transcription factors” which recognize and initiate transcription at specific promoter sequences
- Binding begins at promoter element called the **TATA box**
 - By **TBP**, the TATA Binding Protein
 - See Fig 31.11 and 31.12

Note that TBP binding to TATA sequence causes a large bend in the DNA

Slide
16

Transcription Factors, con't

- There are many general transcription factors (see Table 31.2),
- There are also additional **promoter elements**
 - For example, CAAT box, GC box
 - Fig 31.29 and Table 31.4
- **Enhancer sequences** located far away can also bind transcription factors and interact with polymerase by DNA looping
 - Fig 31.30

Regulation in Eukaryotes is much more detailed and complex than we can deal with in this course. We can see some of the principles of how protein-DNA interactions can activate or inhibit polymerase activity, though, by looking at some of the regulatory phenomena in prokaryotic transcription.

Slide
17

Regulation of Transcription in Prokaryotes

- Genes with common functions (i.e. biosynthetic or catabolic pathway genes) often clustered together.
- Transcription of total gene cluster produces a “polycistronic mRNA”
- This collection of genes is called an “operon”, and transcription of the total operon is under control of common regulatory elements.

“Cistron” is a unit of heredity defined by a “cis-trans” test. In a diploid condition, if mutant A on one chromosome can complement mutant B on another, they are on different cistrons. If not, they are on the same cistron. Hence a cistron produces an identifiable gene product—I.e. a polypeptide chain.

Slide
18

Induction and Repression

- Induction refers to the increase in production of a protein in response to a metabolite.
 - Enzymes of catabolic pathways are often induced by the presence of the metabolite.
- Repression refers to the decrease in production of a protein in response to a metabolite.
 - Enzymes of anabolic pathways are often repressed by the final product of the pathway.

Slide
19

Induction and Repression, con't.

- Both phenomena involve regulatory proteins that bind near the promoter region of the operon, called the “operator”.
 - All genes of the operon are coordinately regulated.
- The protein might **inhibit** (negative control) or **activate** (positive control) transcription of the operon.
- These regulatory proteins bind to small molecules (ligands), which either **increase** or **decrease** binding to the operator.

Slide
20

Four Combinations of Control (Fig 31.21)

- Inhibitory protein, ligand decreases DNA binding: ligand is an **inducer**
- Inhibitory protein, ligand increases DNA binding: ligand is a **repressor**
- Stimulatory protein, ligand decreases DNA binding: ligand is a **repressor**
- Stimulatory protein, ligand increases DNA binding: ligand is an **inducer**

The inhibitory protein is often called a repressor, and its ligand a **corepressor** or **coinducer**, while the stimulatory protein is often called an inducer, and its ligand a **coinducer** or **corepressor**.

Slide
21

The *lac* Operon

- Genes for lactose catabolism (Fig 31.16)
 - Beta-galactosidase (*lac Z*)
 - Lactose permease (*lac Y*)
 - Transacetylase (*lac A*) (unknown function)
- All proteins induced by lactose (or other beta-galactosides)
- Regulatory mechanism first determined by genetic analysis

Slide
22

The *lac* Operon, con't.

- Mutant types observed:
 - z^- , y^- or a^- , mutations in structural genes
 - o^- , constitutive, maps next to operon
 - o^-/o^+ partial diploids are constitutive, so the effect is **cis**—the o gene only affects its operon
 - i^- , constitutive, maps elsewhere
 - i^-/i^+ partial diploids are inducible, so the effect is **trans**—the I gene affects both operons

A constitutive mutant is one that is no longer under control. It makes the enzymes without the need for an inducer.

Slide
23

The *lac* Operon, con't.

- The i gene product is a **repressor** protein.
- The o gene is the binding site of the repressor.
- The repressor binds to DNA and inhibits transcription.
- The **inducer** (or **co-inducer**) binds to the repressor and inhibits its binding to DNA, thus removing inhibition.
 - This only occurs when there is lactose to be metabolized.

Slide
24

Catabolite Repression

- The *lac* operon and other catabolite operons are also under a control called **catabolite repression**.
- The promoters are weak, and require an activator protein (**CAP**) to be transcribed.
- The CAP binding site is also near the promoter.

CAP stands for **catabolite activator protein**.

Slide
25

Catabolite Repression, con't.

- Binding of the CAP protein to DNA is stimulated by 3',5'-cyclic AMP.
- Adenyl cyclase of E. coli is inhibited by the uptake of glucose.
- Hence CAP is inactive if there is glucose present, making glucose the preferred substrate.
- Cyclic-AMP has been called an **ancient hunger signal**.

CAP is also sometimes called **CRP**, which stands for **cyclic AMP receptor protein**.

Remember the conditions that lead to cyclic AMP production in the liver!

Slide
26

The *araBAD* Operon

- Also a catabolite pathway for L-arabinose metabolism.
- Stimulated by CAP-cyclicAMP
- The repressor (*araC*) also regulates its own synthesis, and looping of DNA is involved with interaction of two *araC* molecules.
 - See Fig 31.22, but don't bother with details.

Slide
27

The *trp* Operon

- Codes for enzymes of tryptophan biosynthesis. (See Fig 31.24)
- Also regulated by a repressor protein.
- Repressor binding to DNA requires a **corepressor**, which is tryptophan.
- Thus synthesis of pathway enzymes is regulated by the end product.

Slide
28

The *trp* Operon Attenuation

- Another regulatory mechanism in some operons.
- Translation of operon begins before transcription is completed.
- A stem-loop stop termination signal occurs early in the operon.
- An alternative stem-loop structure can prevent the termination loop from forming.
 - Fig 31.27

Slide
29

Attenuation, con't.

- Early message contains several Trp codons in a row.
- If Trp concentration is low, ribosome pauses here until Trp-tRNA can bind.
- Pausing allows alternative stem-loop structure to form, preventing termination structure from forming. (Fig 31.28)
- Leader peptides in several operons show codon sequences characteristic of attenuation.
 - Fig 31.26

Slide
30

Transcription Regulation in Eukaryotes

- Much more complicated
- Already mentioned “transcription factors”, promoter elements, and enhancer sequences
- Regulation through binding of proteins
 - Some can activate or speed up transcription
 - Some can inhibit or slow down transcription
- Binding proteins are being classified according to “binding motifs”
 - Helix-turn-helix (Fig 31.34)
 - Zinc finger (Fig 31.37 and 31.38)
 - Leucine Zipper (Fig 31.43 and 31.44)

The alpha helix fits into the major groove. Specific hydrogen bonding between side chain amino acids and the bases is one source of sequence recognition. Arginine can hydrogen specifically to guanine, and glutamine can specifically bind to adenine (See Fig 31.36).

Slide
31

RNA Processing

- Ribosomal RNA and transfer RNA from both prokaryotes and eukaryotes are cut from longer **precursor RNA** structures.
 - Seven ribosomal gene “operons” in *E. coli*. Spliced into the three rRNA’s and several tRNA’s (Fig 33.1)
- Some bases methylated in rRNA, many bases modified in tRNA. (Fig 11.26)
- CCA is added to eukaryotic tRNA, while part of transcribed structure in prokaryotes

Slide
32

RNA Processing, con’t.

- Big difference in mRNA processing in eukaryotes versus prokaryotes.
- Prokaryotes
 - Multi-gene operons are copied
 - Translation can begin before transcription ends
- Eukaryotes
 - Many genes are split (Fig 31.45) so mRNA transcript must be spliced
 - 5’ end is “capped”
 - 3’ end is extended with a poly A tail (all but histone mRNA)

Slide
33

Capping

- GTP is added to the 5’ end to form a 5’-5’ triphosphate bridge (Fig 31.47)
- This guanine and the first two ribose residues are methylated (Fig 31.48)
- CAP structure may protect against RNase degradation
- CAP structure also important in binding orientation on ribosome

Slide
34

3'-Polyadenylation

- Termination of transcription does not occur until polymerase has passed a consensus AAUAAA sequence - the poly(A)⁺ addition site.
- The transcript is cleaved by an endonuclease about 10-30 bases downstream of this consensus site.
- Poly(A) polymerase adds 100-200 adenine residues to the 3' hydroxyl end.
- Histone messages are not polyadenylated.
- Polyadenylation may help identify the molecules for splicing and export as well as protect against RNase.

The poly A tails provide an easy means to separate mRNA's from other RNA's by passing over a column containing poly dT. Base pairing causes the poly A tails to anneal with the column while other RNA's are washed off. Then the mRNA's are removed by raising the temperature or changing the salt conditions.

Slide
35

Splicing

- "Introns" removed; "Exons" joined
- Substrate is the capped, polyadenylated messenger transcript.
- The spliced-out introns appear as "lariat" structures. (Fig 31.50)
 - Note 5'-splice site, 3'-splice site, branch site
- Small nuclear RNA's are involved as components of a ribonucleoprotein particle called a **snRP**

Slide
36

Splicing, con't.

- **snRP** is called a "spliceosome"
- Five small nuclear RNA's
 - **U1**, 165 nt, recognizes 5' splice site (Fig 31.52)
 - **U2**, 189 nt, recognizes branch site
 - **U4**, 145 nt, **U5**, 115 nt, **U6**, 106
 - Complex set of interactions (Fig 31.53, 31.54)

Slide
37

Splicing, con't.

- Some transcripts can undergo alternative splicing, producing related polypeptides called **protein isoforms**.
- Some primitive “self-splicing” RNAs have been discovered.
 - Pre rRNA from *Tetrahymena thermophila*
 - mRNA of a fungal mitochondria

Thomas Cech received the 1989 Nobel Prize in Chemistry for his discovery of the self-splicing RNA. This was the first demonstrate that RNA can have catalytic activity.