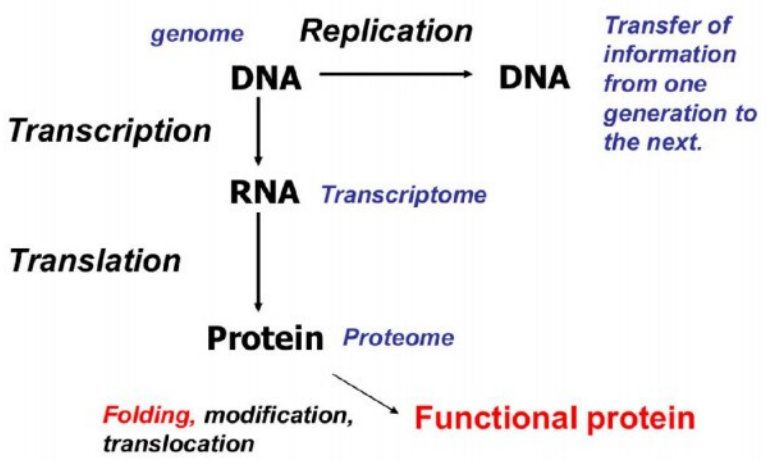
MBLG1001 Notes

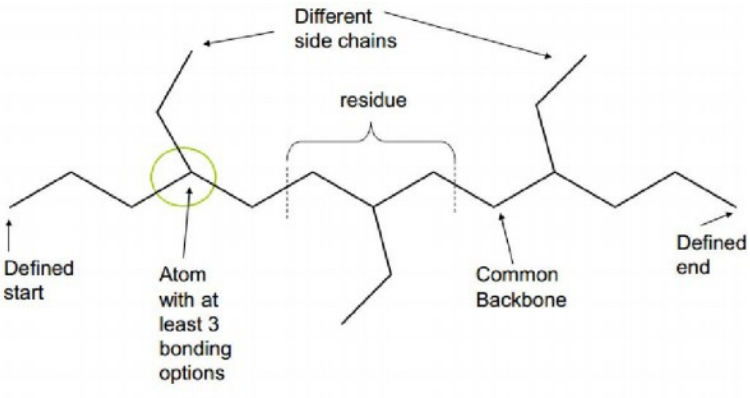
Molecules of Life: Carbon

- The genome (DNA)
- The transcriptome (RNA)
- The proteome (proteins)
- The DNA in almost every cell in the body contains the same information. There is usually one copy of each gene. However, cells in different parts of the body look different because different proteins are expressed. The **regulation of gene expression** (switching genes on/off) allows cells to differentiate themselves; explains cancers, defects, aging, etc.



- Major biopolymers not considered in detail:
  - Fats (lipids): $(-CH_2-)_n$
  - Carbohydrates: $(H\text{-}C\text{-}OH)_n$
    - Polymer: polysaccharide
    - Monomer: saccharide (or "sugar")
    - Sugars with 5-6 carbons readily cyclise forming ring structures (e.g. glucose).
    - Sugars (components of nucleic acids) are hydrophilic/water soluble due to the OH groups: impacts on properties of DNA/RNA.

» General Biopolymer Properties



- All linear biopolymers have a **defined beginning** (N terminal) and **end** (C terminal).
- Biopolymer synthesis is an **anabolic process** (requires energy input). Energy comes from high energy bonds of monomers.
- All biopolymers are synthesised in **one direction** only (sequentially adding next monomer to growing chain).
- Some of the monomer is lost in polymerisation, leaving a **residue** incorporated in the growing chain (e.g. amino acid residues joined by peptide bonds to form a polypeptide).
- The monomer must be **activated** before polymerisation.
- There are **3 phases to biopolymer formation** *in vivo* (in the cell): initiation, elongation and termination.

» Unique Properties of Carbon

- All major biopolymers have a substantially carbon backbone.
- However, carbon is ***not*** the most abundant element on the earth's crust (in order of abundance, O > Si > Al > Fe…)

» Initiation

- *E. Coli* has **one defined site: oriC** (origin of replication on the chromosome or replicator).
- There is a **specific sequence**, 200 – 300 bp long, which is recognised by the **initiator protein** DnaA.
- 20 – 30 DnaA monomers bind to oriC in a coordinated ATP-dependent (energy-requiring) fashion and recruit DnaB, a helicase.

Prokaryotes

- One site
- Specific sequence
- Specific proteins bind in a specific order

Eukaryotes

- Eukaryotes have multiple chromosomes and multiple oriC sites.

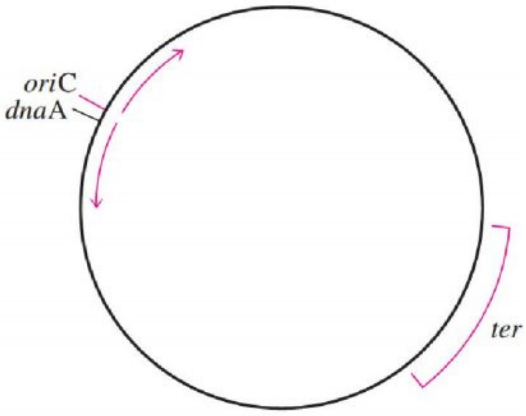
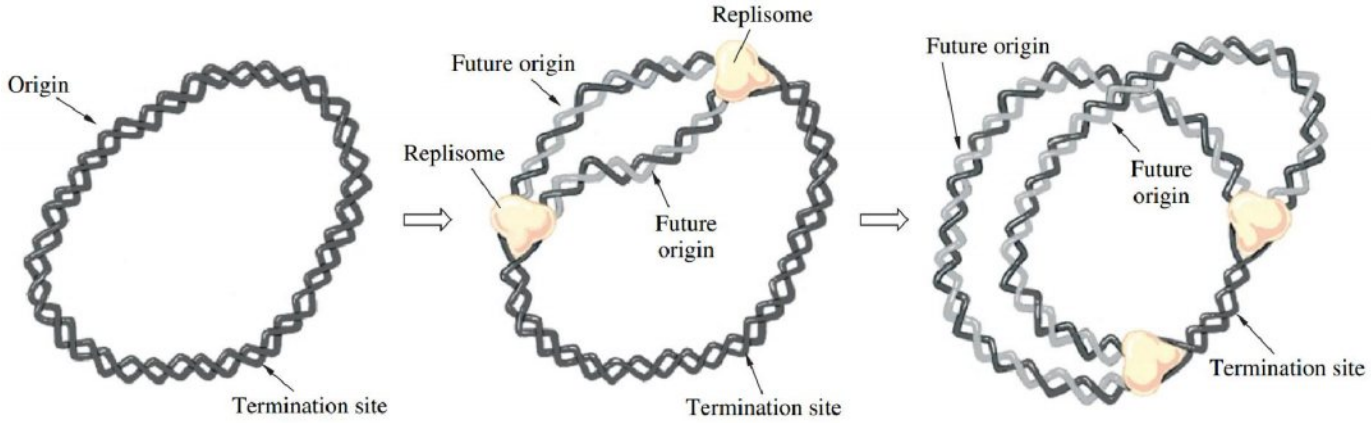


Figure 20.16 ▲
Location of the origin (*ori*C) and terminus (*ter*) of DNA replication in *E. coli*. *dna*A is the gene for the protein DnaA, which is required to initiate replication. The distance between *ori*C and *dna*A is about 40 kb. The red arrows indicate the direction of movement of the replication forks.

» Replication Forks

- The replication fork starts at the oriC and moves in opposite directions ('**bidirectional**'). It meets up with itself at the termination site.
- The actual enzyme, DNA polymerase, only works in one direction.
- DNA replication must occur with both strands at the fork.



» Proteins in Replication

The players:

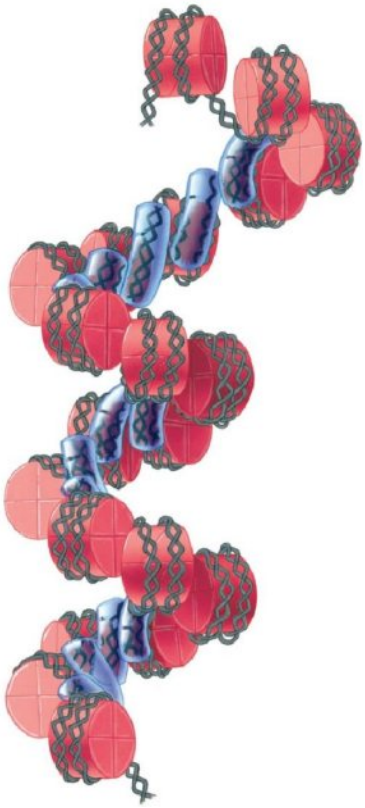
- **Helicase or DnaB**: unwinds the DNA double helix at the replication fork by disrupting hydrogen bonding between base pairs.
- **Single stranded binding protein (SSBP)**: binds to ssDNA and keeps the strand apart.
- **Topoisomerase**: untangles and relaxes the DNA from its supercoiled nature.
- **Primase**: lays down short RNA primers.
- **DNA polymerase III**: the main enzyme for synthesising a new DNA strand by adding nucleotides in the 5' to 3' direction.
- **DNA polymerase I**: fills in the gaps in the lagging strand.
- **Ligase**: seals the sugar-phosphate backbone; joins Okazaki fragments of the lagging strand.

» Steps in Replication

Unwinding the DNA

- **Helicase** (DnaB) first arrives in the initation complex and is present throughout replication.
- Replication of DNA requires access to the middle of the double helix, where the information is stored.

    - Unpacking of 3D structure needed for gene expression.



- **DNA lives in the nucleus**.
  - mRNA has to be transported out for translation.
  - Spatial and temporal separation of mRNA synthesis (*in nucleus*) from protein synthesis (*in cytoplasm*).
    - In prokaryotes, mRNA transcripts are translated directly (translation often initiates before transcription is complete).
    - In eukaryotes, transcription occurs in the nucleus and translation occurs in the cytoplasm.
  - Ribosomes attached to endoplasmic reticulum.
- **DNA generally very big**.
  - Lots of genes (>30,000 in humans)
- **Organelles have some of their own DNA**.
  - e.g. mitochondria, chloroplasts can make their own proteins.

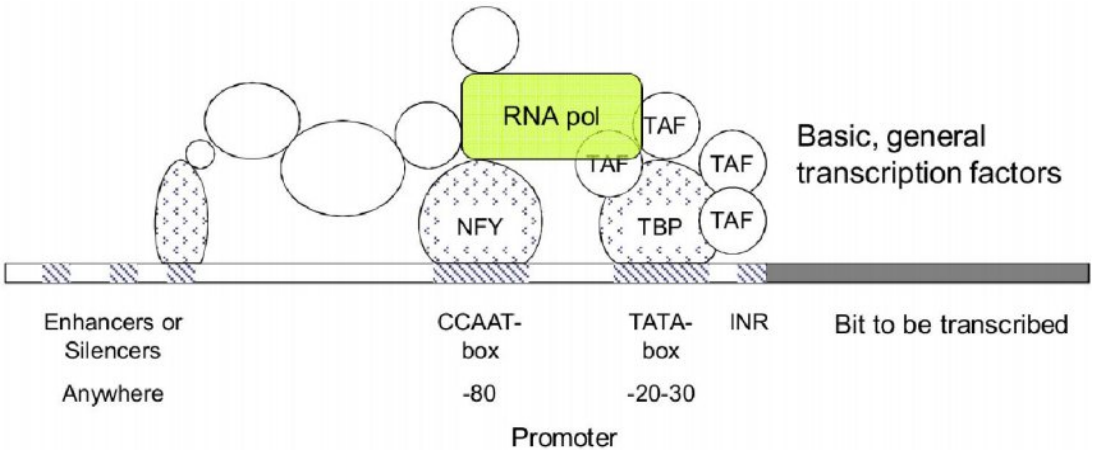
» Eukaryotic Transcription

Occurs similarly in eukaryotes and prokaryotes

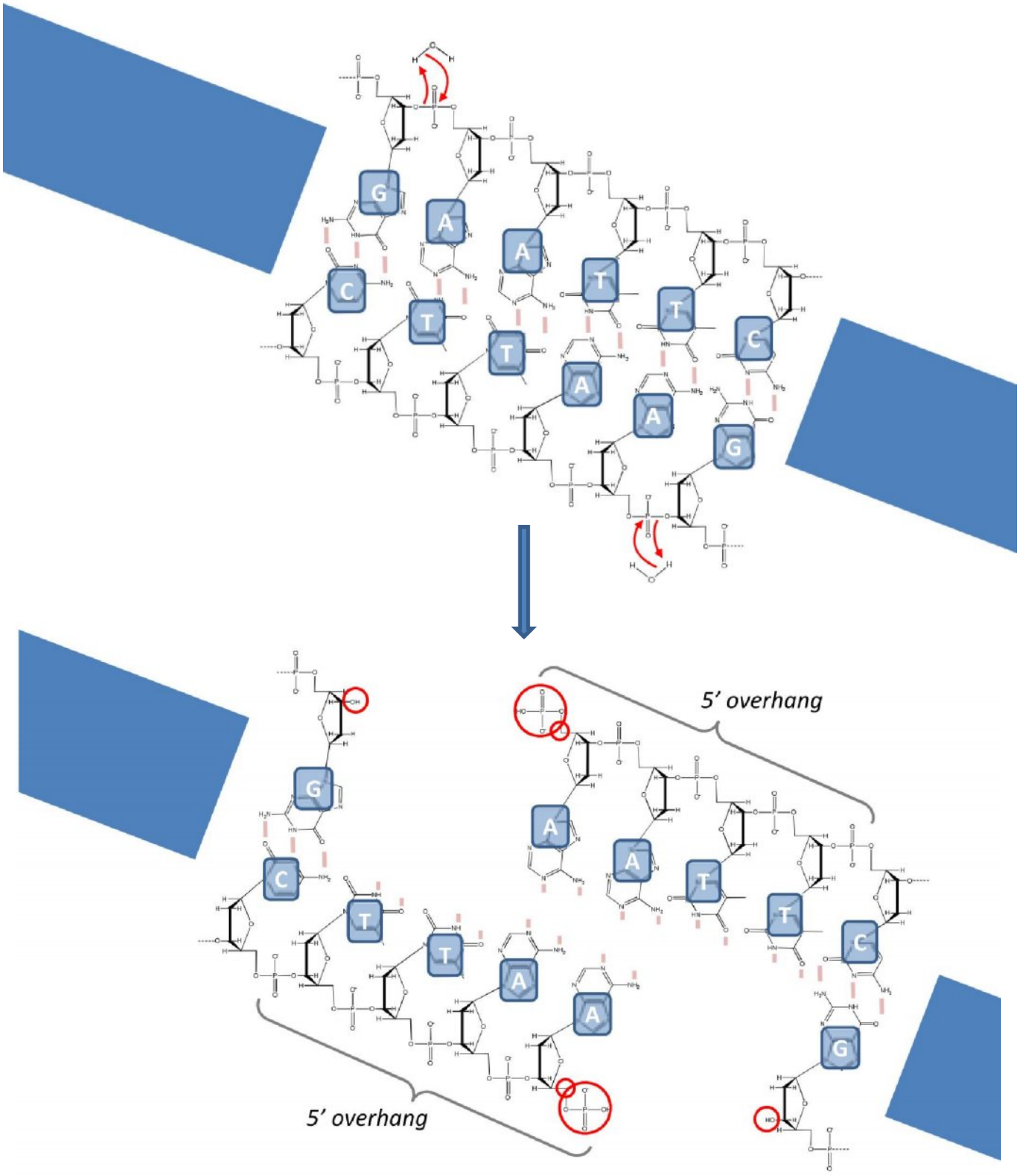
- 3 types of RNA polymerase used:
  - I for rRNA
  - II for mRNA ← *need to know this one*
    - **RNA pol II involved in eukaryotic transcription**
  - III for tRNA and rRNA
- All 3 RNA polymerases are **complex, multi-subunit (10 subunits) enzymes**.
  - *Alpha-amanitin inhibits RNA production: causes death over time*.
- **No equivalent of sigma**
  - Initiation of transcription requires a more complex way of recognising promoters (using RNA pol).
  - Need to open up the DNA: transcribing areas more prone to DNase digestion.

Eukaryotic Promoters

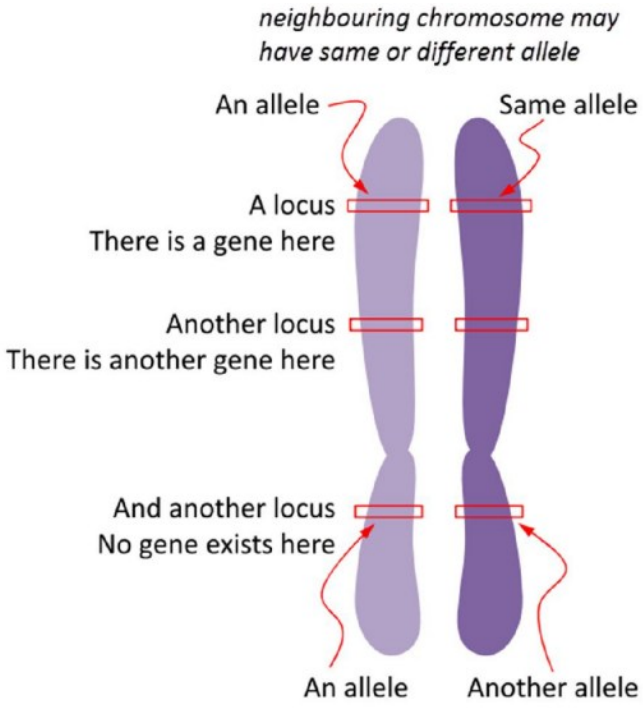
- **Eukaryotic transcription requires transcription factors**: **proteins that bind to specific DNA sequences in the promoter** and **affect the rate of gene transcription**.
  - Mixture of protein-protein interactions and DNA-protein interactions.
  - Activity of all the proteins can be modified.
- Transcription initiation requires supercoiled DNA to be unwound, histones dissociated and DNA exposed so that transcription complexes can be built at the promoter.



- Many restriction enzymes bind to DNA as **homodimers** (protein composed of 2 identical polypeptide chains).
  - They therefore recognise and cleave **palindromic sequences**.



- If the **5' overhangs** are complementary base sequences, then they are compatible.
  - 5' overhangs known as **sticky ends** as there are unfulfilled hydrogen bonds.
- Different restriction enzymes have different recognition and cleavage sites.
  - *Note: XmaI and SmaI recognise the same sequence but cut at different locations.*
- **Blunt ends** have no exposed bases for hydrogen bonding.
- Base pairing between overhangs is a random process. Sticky ends (with exposed hydrogen bonds) have higher likelihood of finding each other than blunt ends (cannot hydrogen bond together).

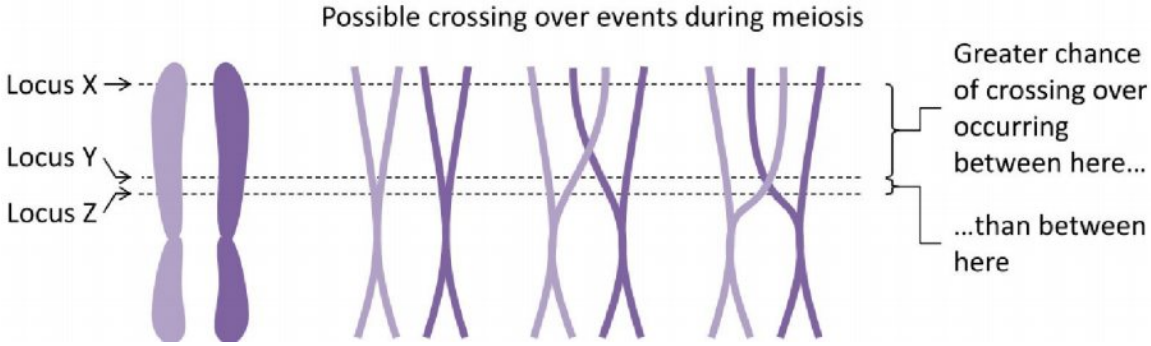


Genetic Markers

- Markers are loci that act as indicators or signposts: specific DNA sequences at known locations on a chromosome.
- Markers can have a **detectable phenotype** or **no detectable phenotype** (a **'molecular' marker**).
  - *Most markers are molecular markers*.
- Markers are useful if there are different alleles (forms) of them.

Linkage

- Alleles at loci that are **close together** on a chromosome tend to be **inherited together**, as they are less likely to experience a crossing-over event between them than far-apart loci.
- Crossing over: *exchange of genetic material between homologous chromosomes resulting in recombinant chromosomes*.
  - Crossing over is the inverse of inherited together.



Linkage Mapping

- **Linkage mapping** involves finding a **known molecular marker** that is closely linked with a locus of interest.
- This is typically done by analysing breeding experiments or large family trees.
  - Example: Through breeding experiments, the Ob gene was found to be **linked** to the D6Rck13 molecular marker (i.e. **crossing over rarely occurred** between D6Rck13 and the Ob loci).

Positional Cloning

- **Positional cloning** is used to locate the **gene sequence** of interest, starting from the molecular marker (found by linkage mapping).
- Genomic libraries are screened to get closer and closer to the gene sequence.
  - This process is called **chromosome walking** (*moving along the chromosome*).