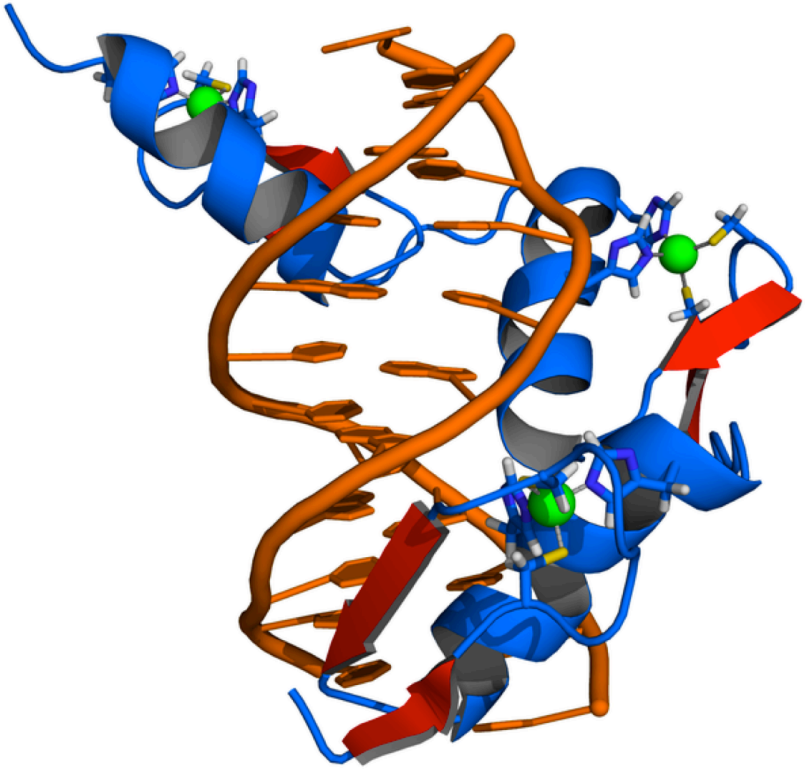


Cellular & Molecular Science

Dr Mick Jones
Infectious Diseases & Immunity
&
MRC Clinical Sciences Centre
Genomics Laboratory
Hammersmith Hospital Campus

Imperial College London

MRC Clinical Sciences Centre

MMBBS/BSc (6 year degree)	MBBS (Graduate Entry)
Year 1	Year 1 (accelerated programme)
Year 2	
Year 3	Year 2
Year 4 (BSc)	Exempt
Year 5	Year 3
Year 6	Year 4

Molecules Cells and Disease (MCD)

Cellular and Molecular Science (CMS)

49 Lectures from MCD not covered in lectures

Revision Lectures

Lead academics for the strands of the CMS Theme are as follows;

Proteins, Nucleic Acids and Gene Expression - Dr Mick Jones
Cellular Organisation of Tissues - Dr Anabel Varela Carver
Metabolism – Dr Katie Wynne and Professor Gary Frost
Cancer and the Cell Cycle - Professor Gerry Thomas and Dr Andy Porter

Depending upon which 'lectures' are to be delivered as SDL material (Self-Directed Learning) and which are to be given as traditional lectures/discussion sessions, then appropriate Q&A sessions will be timetabled to answer specific points raised by the students.

For these Q&A sessions to work productively will require students to email specific questions to the session academic well in advance of the session. The sessions are directed discussions to tackle specific areas in which the students are uncertain and want clarification and understanding.

Dr Jones will give an introductory lecture about the CMS Theme, and at the end of the theme, Exam Revision sessions, to assist in revision and studying practice exam questions, are planned.

Proteins, Nucleic Acids and Gene Expression - Dr Mick Jones

Cellular Organisation of Tissues - Dr Anabel Varela Carver

27/10/2011	Dr Mick Jones	Protein, Nucleic Acids and Gene Expression
28/10/2011	Dr Anabel Varela Carver	Tissues
03/11/2011	Dr Anabel Varela Carver	Cell Behaviour

Metabolism – Dr Katie Wynne and Professor Gary Frost

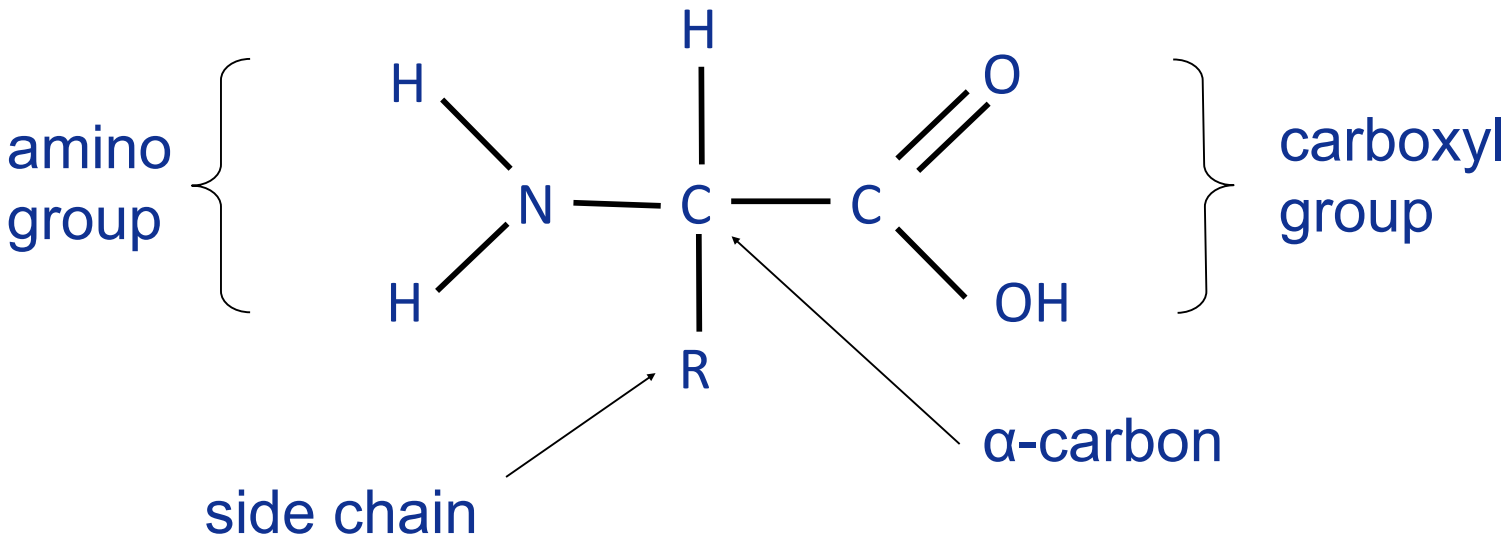
Date		Topic
02/12/2011	Lecture 1 Topic 1 Topic 2	Integrative Metabolism Cellular Energetics Metabolic pathways and ATP production (I)
08/12/2011	Topic 3 Topic 4	Metabolic pathways and ATP production (II) Metabolic pathways and ATP production (III)
12/12/2011	Topic 5 Topic 6	Mitochondria and Oxidative Phosphorylation Lipids and Membranes
13/12/2011	Topic 7 Topic 8	Membranes and Vesicles Cholesterol and Lipid Transport
16/12/2011	Topic 9 Lecture 2 Lecture 3	Human Metabolism and Energy Regulation Diabetes: The Archetypal Metabolic Disease Inborn Errors of Metabolism

Cancer and the Cell Cycle

Professor Gerry Thomas and Dr Andy Porter

Dr Andy Porter	Cell Cycle and its Regulation
Dr Andy Porter	Signalling Mechanisms
Dr Andy Porter	DNA Damage and Repair
Dr Andy Porter	Oncogenes and Tumour Suppressors
Prof Gerry Thomas	Cellular Pathology of Cancer
Prof Gerry Thomas	Cancer as a disease – Colorectal and Breast Cancer
Prof Gerry Thomas	Cancer as a disease – Skin Cancer
Prof Gerry Thomas	Cancer as a disease – Leukaemia

Anatomy of an Amino Acid



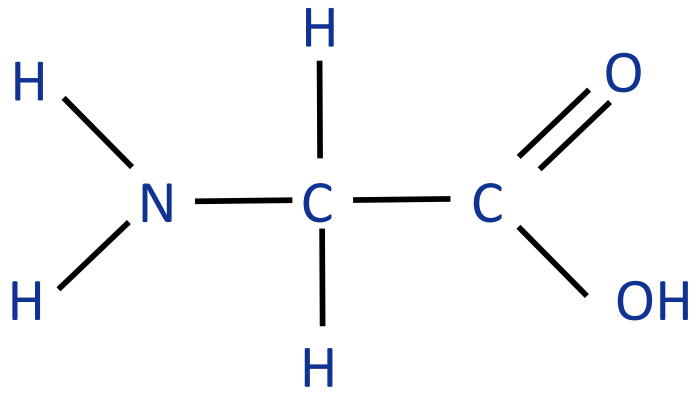
Substitutions at the R position or side chain, give rise to the 20 different amino acids e.g. R=CH_3 in alanine.

The whole of the amino acid minus the side chains is known as the backbone.

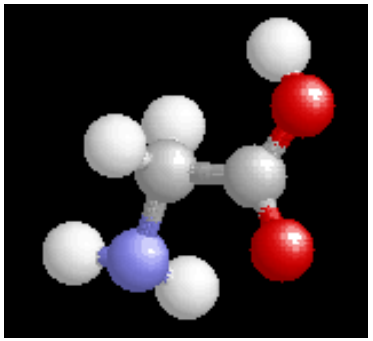
Glycine, the Simplest Amino Acid

In *all* organisms, from bacteria to humans, only 20 different amino acids go to make up any particular protein.

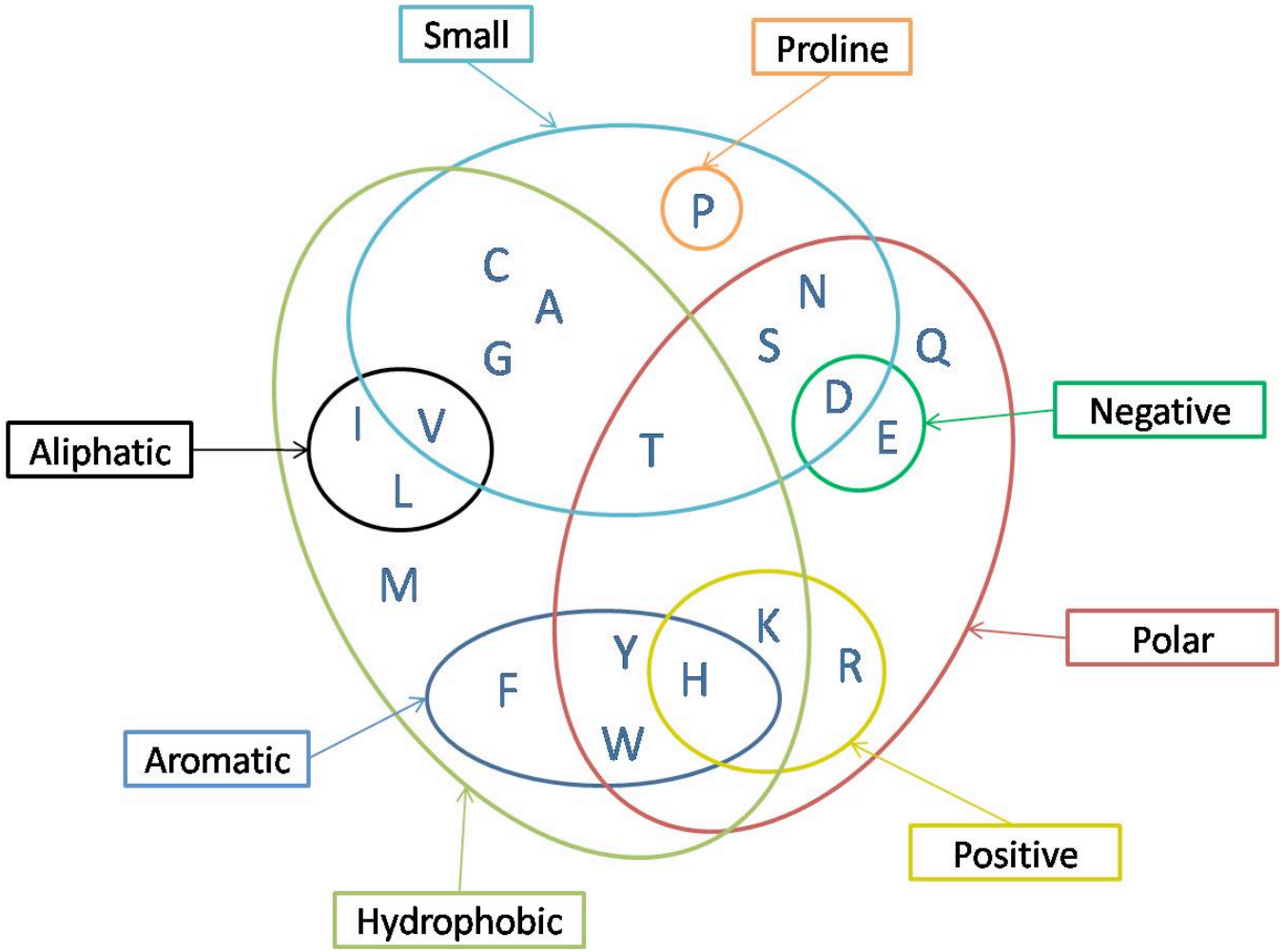
This gives us a protein alphabet that is at least 2 billion years old.



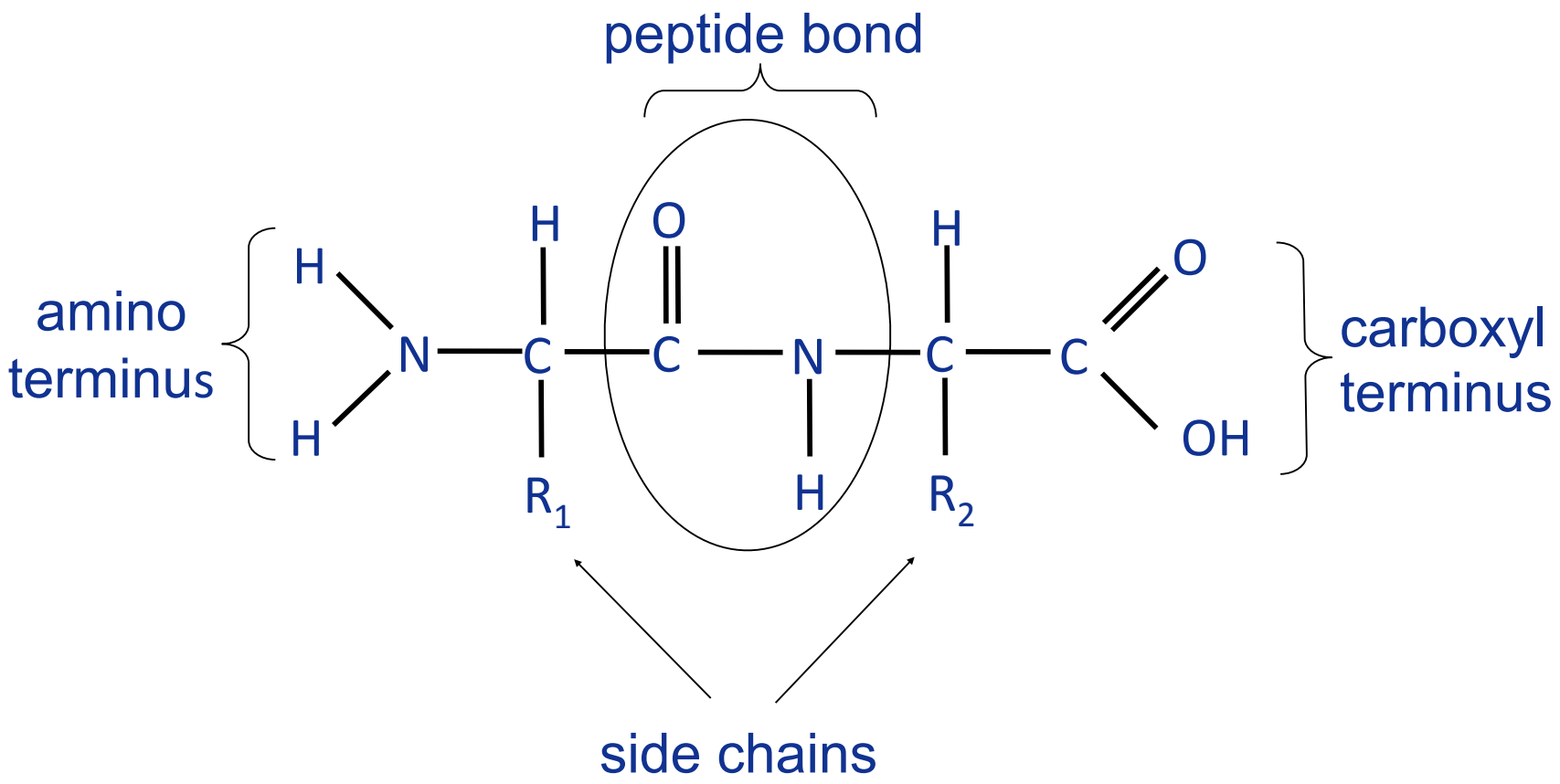
Glycine: G, gly



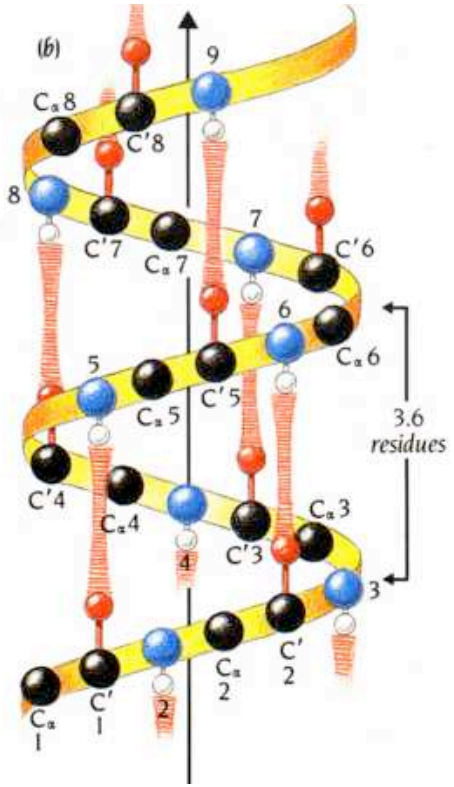
Ball & Stick Model


Small
Proline
P
C
A
G
N
S
Q
D
E
Negative
Aliphatic
I
V
L
T
M
K
Y
H
R
F
W
Polar
Aromatic
Positive
Hydrophobic

Anatomy of a Peptide



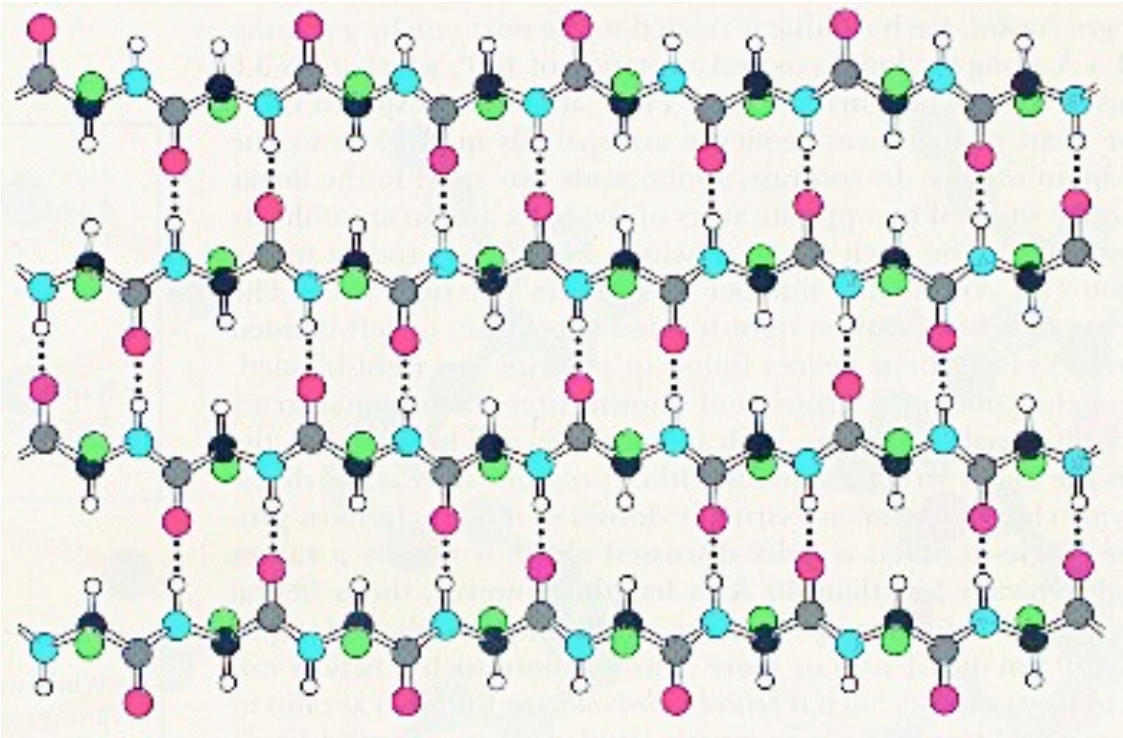
Helices are stabilised by Hydrogen Bonds



Hydrogen Bonds between the C=O of one residue and the N-H of another residue, 4 amino acids along the helix, stabilise the entire structure.

From Branden & Tooze,
'Introduction to Protein Structure'

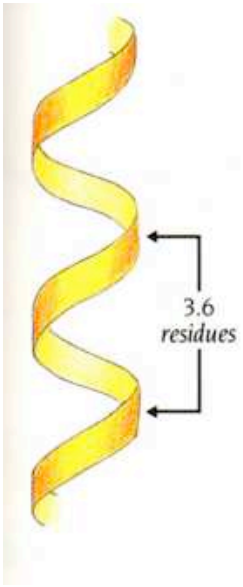
The β-pleated sheets are stabilised by Hydrogen Bonds



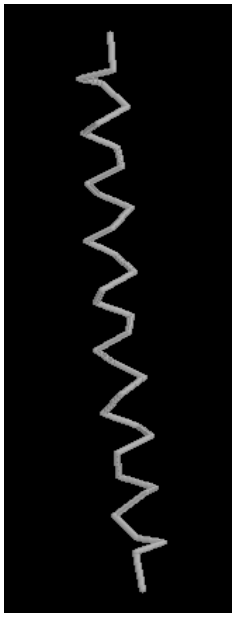
As with the alpha helix, hydrogen bonds between the N-H and C=O groups of two or more β-strands hold the β-*pleated sheet* sheet together.

The α-helix

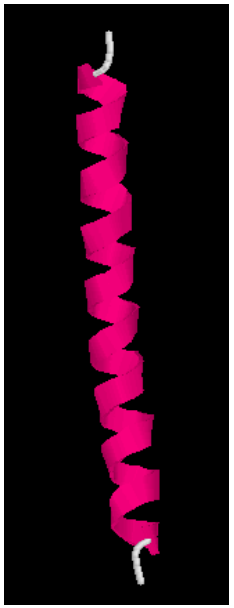
First described in 1951 by Nobel laureate, Linus Pauling,
The α-helix is a major structural element in proteins.



Right-handed α-helix
3.6 residues/turn



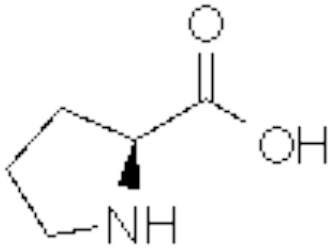
Backbone
representation
of an α-helix



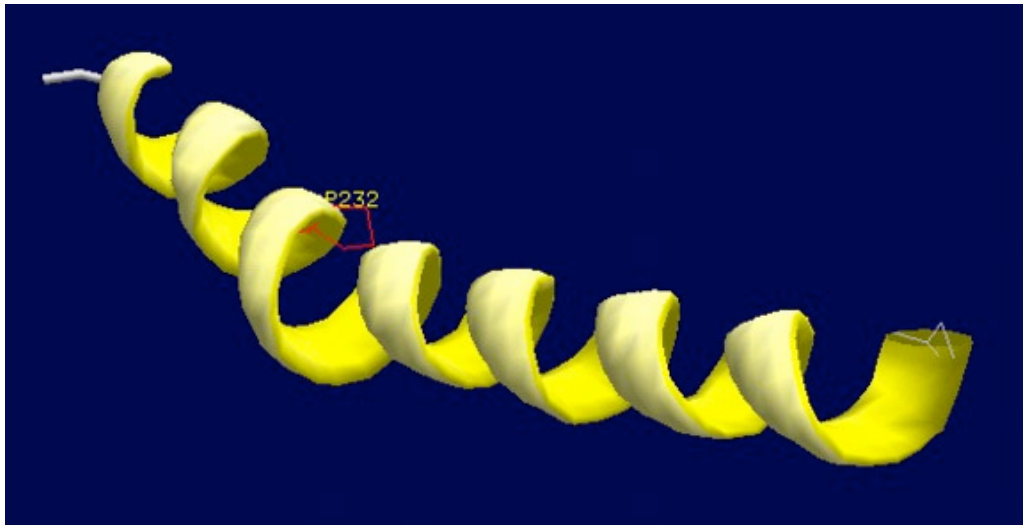
'ribbon'
representation
of an α-helix

Proline is a Kinky Amino Acid

In proline, the last atom of the side chain is bonded to the main chain N atom.

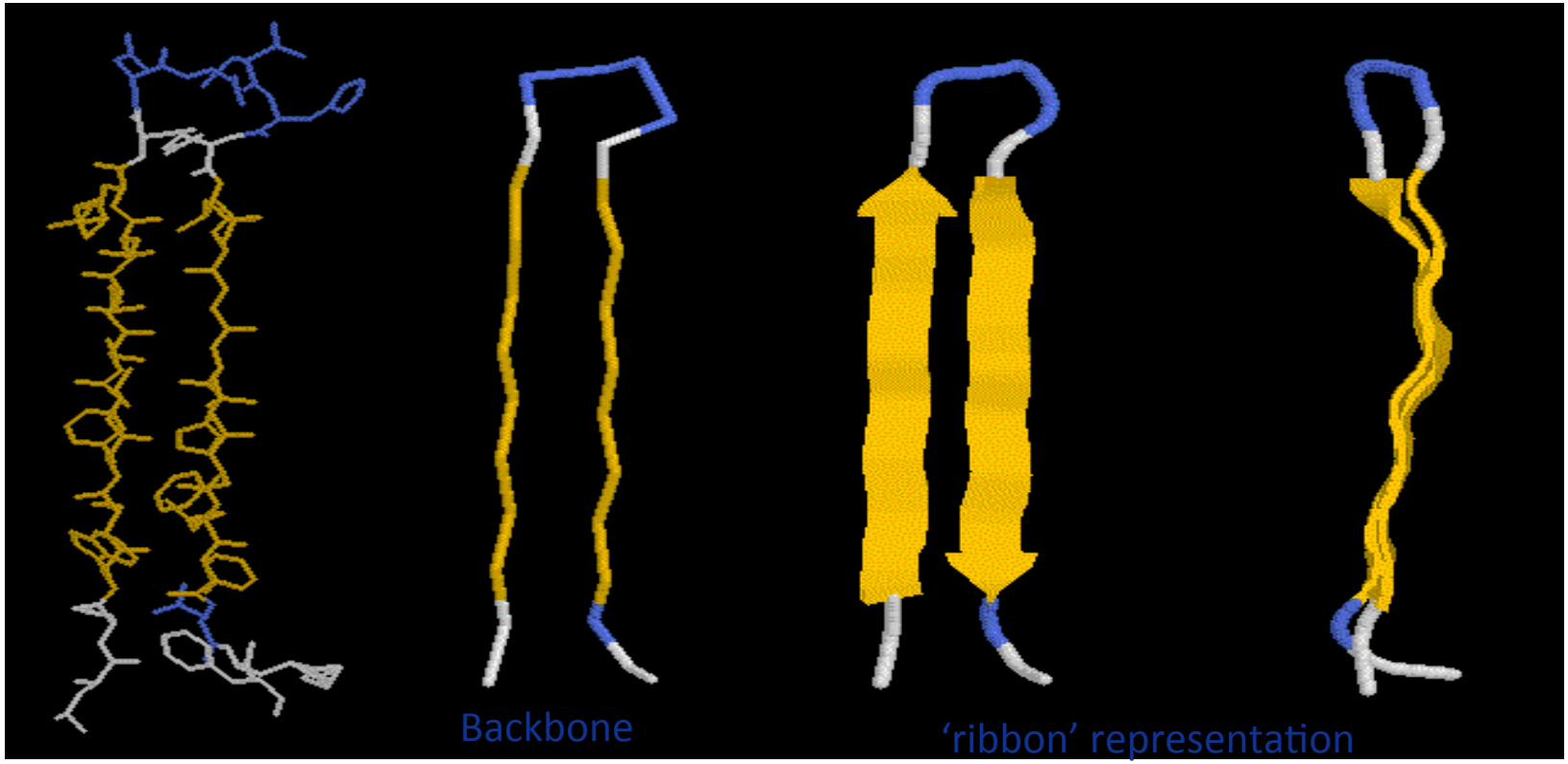


proline



This prevents the N atom from hydrogen bonding with the C=O groups of another residue within the helix, thereby distorting the helical conformation, putting a ‘kink’ into it.

The β-pleated sheet



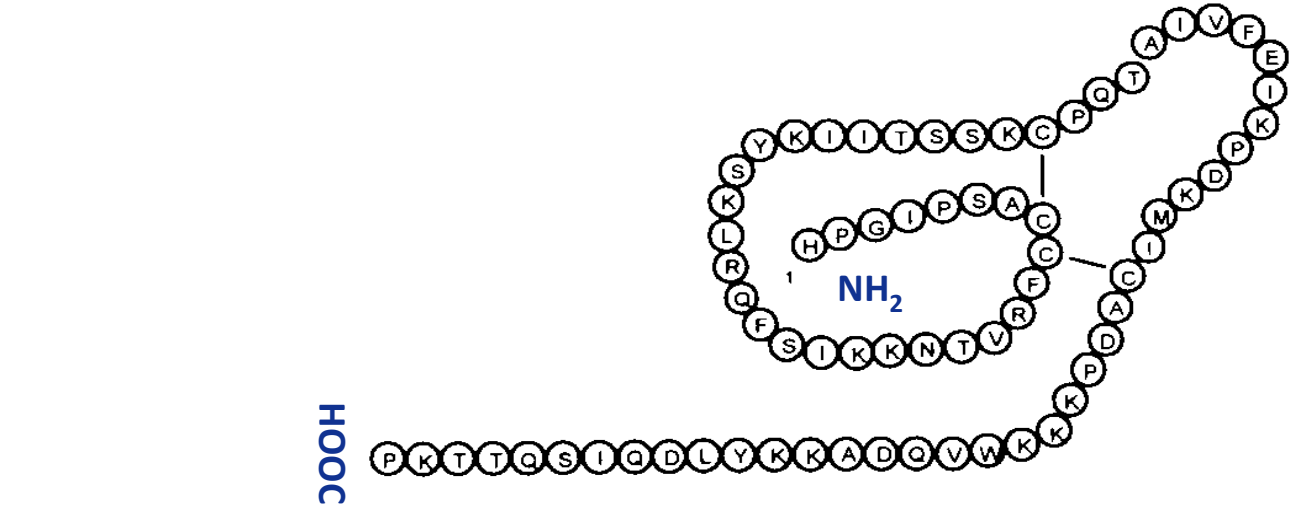
In the β-pleated sheet, the NH and C=O groups point out at right angles to the line of the backbone. This almost two dimensional sheet is pleated, like the bellows of an accordion.

Structural Levels of Proteins

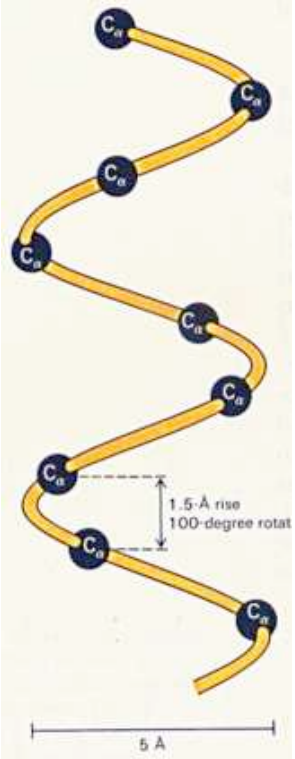
Primary structure

Simply the linear sequence of amino acids that make up the protein.

Standard nomenclature dictates that we write a protein sequence from the amino terminus to the carboxyl terminus.



Structural Levels of Proteins



α-helix

Secondary structure

Defined as local structural motifs within a protein, e.g. α-helices and β-pleated sheets.

Their existence within a protein is dictated by the primary structure or amino acid sequence.

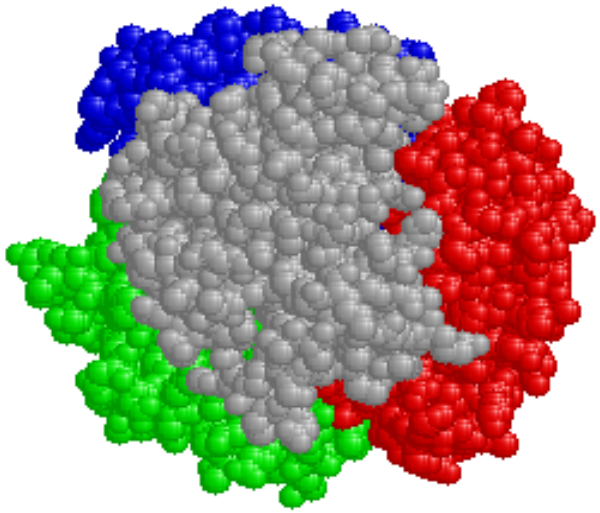
Structural Levels of Proteins

The Green Fluorescent Protein from the Pacific jellyfish, *Aequoria victoria*

Tertiary Structure

Defined as the arrangement of the secondary structure motifs into compact globular structures called domains.

Structural Levels of Proteins



β-Globin from
Homo sapiens

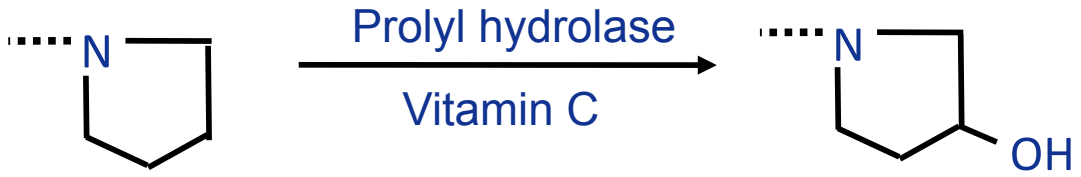
Quaternary Structure

Defined as the three dimensional structure of a multimeric protein composed of several subunits.

Post Translational Modification of Proteins

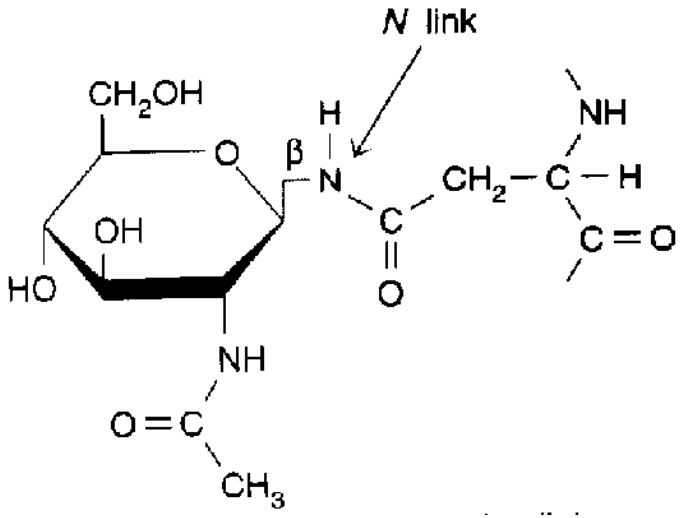
Even after synthesis, (post translation), the starting set of 20 amino acids can be modified to create novel amino acids, enhancing the capabilities of the protein.

Proline can be modified to produce hydroxyproline e.g. collagen fibres, a major constituent of skin, cartilage, teeth & bones.



These additional hydroxyl groups help to stabilise the fibres. A deficiency in vitamin C leads to the disease scurvy.

Post Translational Modification of Proteins



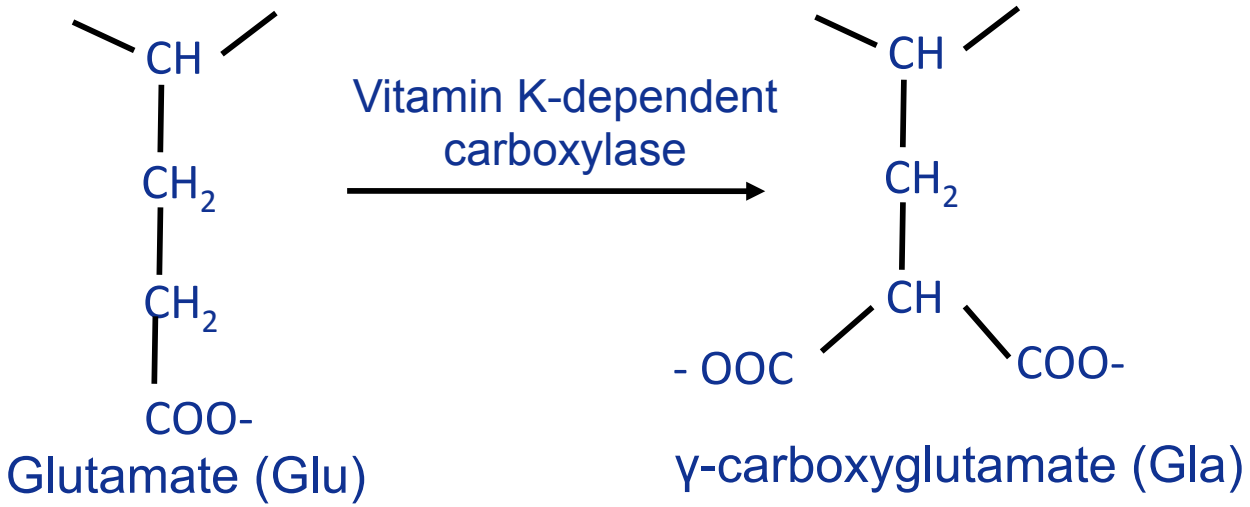
The addition of sugar residues to the asparagine residues of proteins (N-linked glycosylation)

This increases their solubility and also protects them from enzymatic degradation. Motif required in the primary structure is N-X-S/T (Asparagine-Any amino acid-Serine/Threonine).

Deficiency of N-linked sugar chain transfer is detected in congenital carbohydrate-deficient glycoprotein (CGD) syndrome which affects multiple tissues and has life threatening complications.

Post Translational Modification of Proteins

Similarly, γ-carboxyglutamate is produced by the carboxylation of glutamate.



Glutamate (Glu)

γ-carboxyglutamate (Gla)

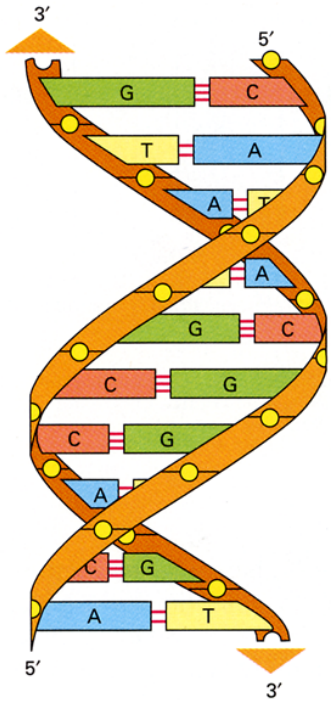
The formation of γ-carboxyglutamate residues within several proteins of the blood clotting cascade (e.g. factor IX) is critical for their normal function by increasing their calcium binding capabilities.

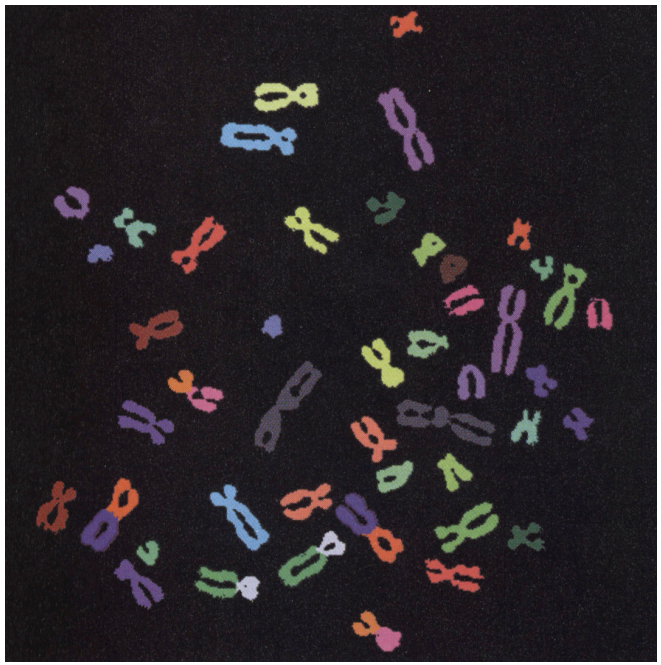
The anticoagulant warfarin works by inhibiting the carboxylation reaction.

Nucleic Acids and Chromosomes

Overview

- Chemical composition of DNA
- Structure of DNA
- How cells package DNA



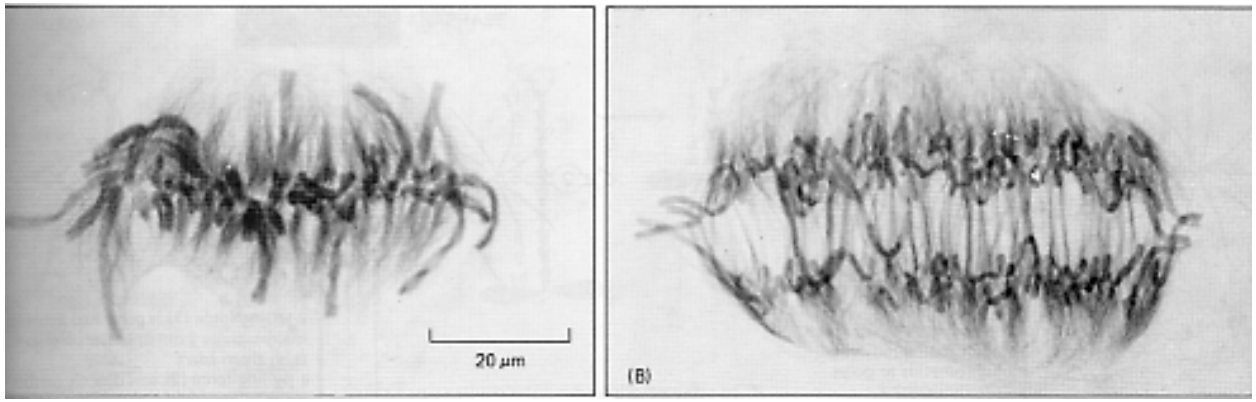


Nucleic Acids

DNA and RNA: molecules of heredity

DNA is the genetic material. The genes of all cells are made of DNA.

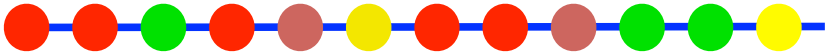
1900: Chromosomes segregate as the cell divides. Chromosomes consist of proteins and DNA. Proteins have 20 different amino acids. DNA has four different bases.



1944: Avery, MacLeod and Mc Carthy prove that DNA carries the genetic information.

Nucleic Acids

Nucleic acids are linear polymers of nucleotides.



DNA and RNA each contain 4 different types of nucleotides that are arranged in different sequences.

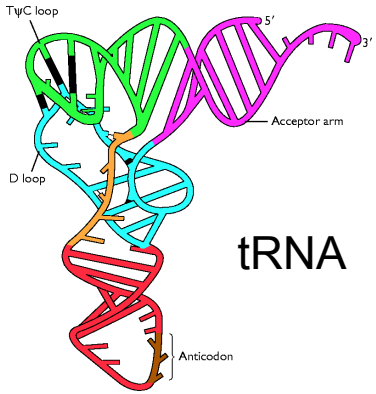
DNA A C G T

RNA A C G U

The 3D structure of DNA is a double helix made up of two chains wrapping around each other.

RNA can assume a variety of shapes.

DNA



tRNA

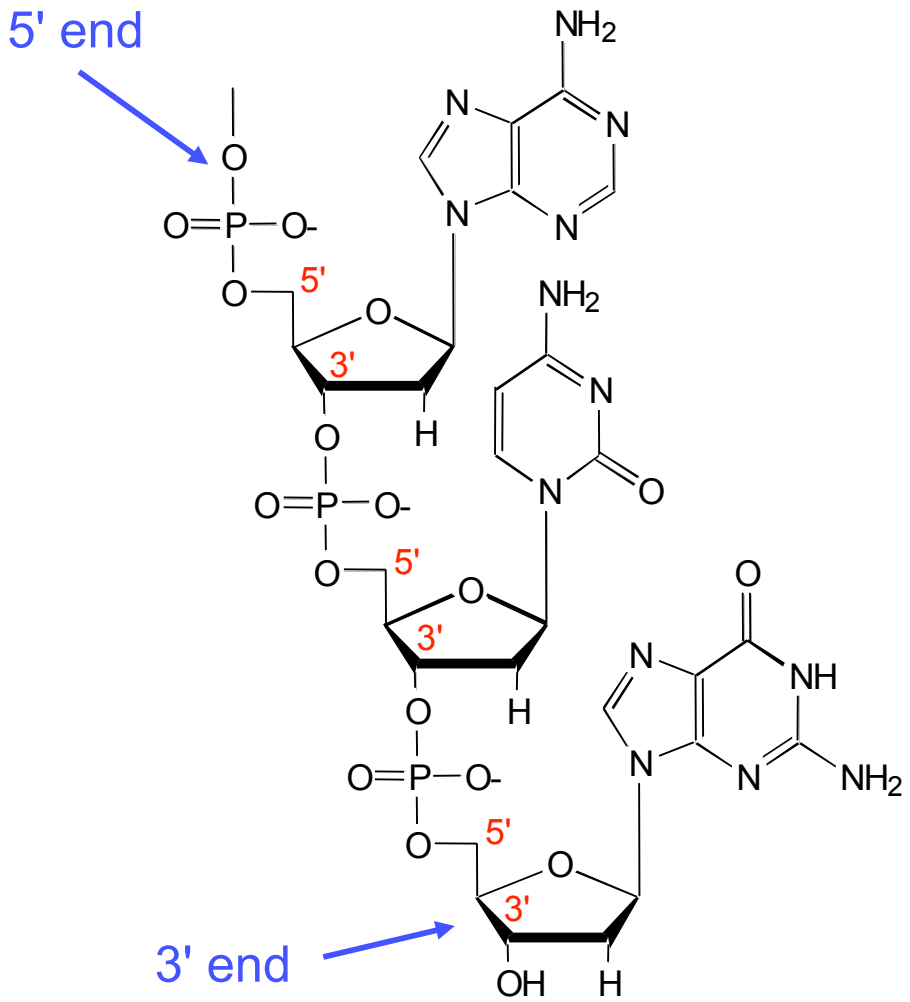
DNA - polymer of deoxyribonucleotide units

A long chain of deoxyribonucleotide units linked by phosphodiester links.

The 3’-OH of sugar of one nucleotide is linked to phosphate group, which in turn is joined to 5’-OH of adjacent sugar.

On each deoxyribose there is a base.

The chain has two ends, the 5’ end and the 3’ end. It is not symmetrical.

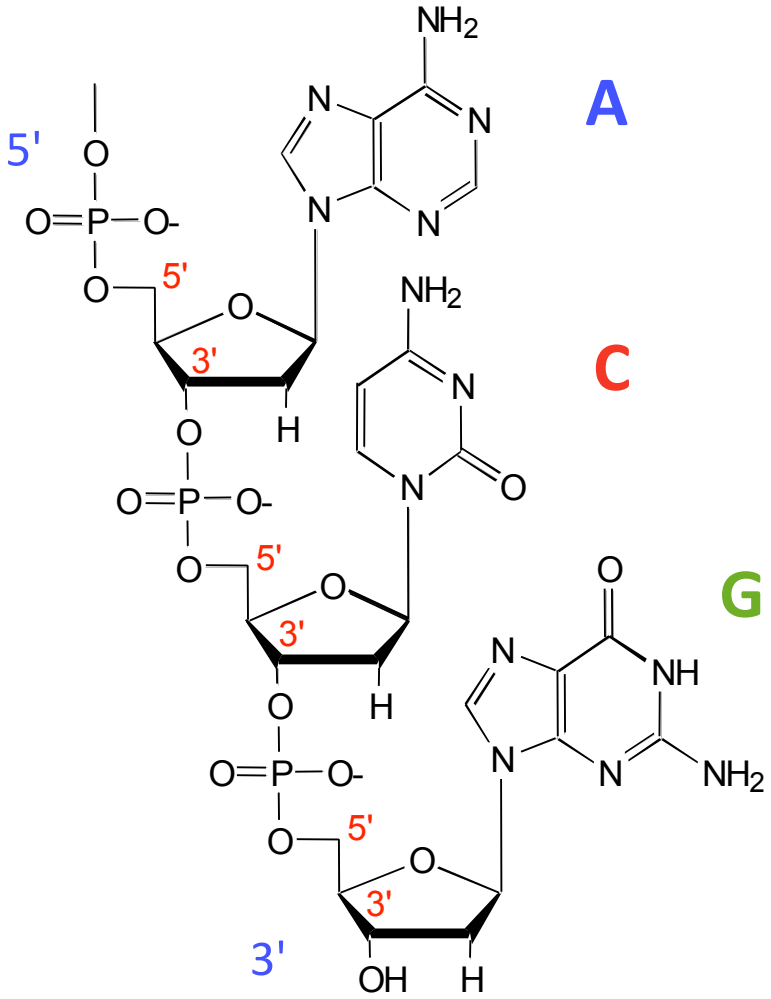


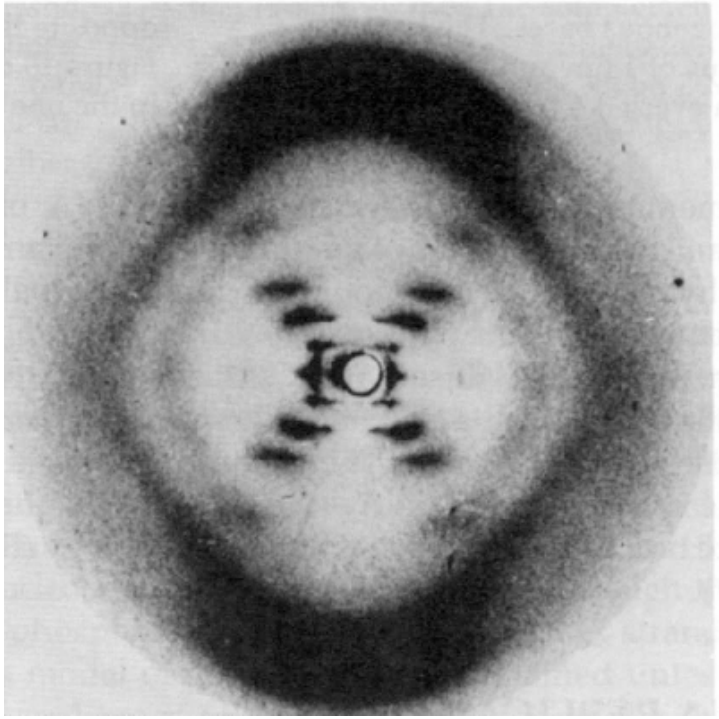
DNA primary sequence

The primary structure of a nucleic acid chain is the sugar-phosphate backbone with the bases, defined by the linear sequence of the bases. The bases carry genetic information, the sugar and phosphates perform a structural role.

By convention, the nucleotide sequence is specified from the 5’ to the 3’ end. The first nucleotide in a sequence has a free 5’ phosphate, and the last one a free 3’ OH.

Here: ACG (or 5’-ACG-3’)





X-ray Diffraction Pattern of DNA

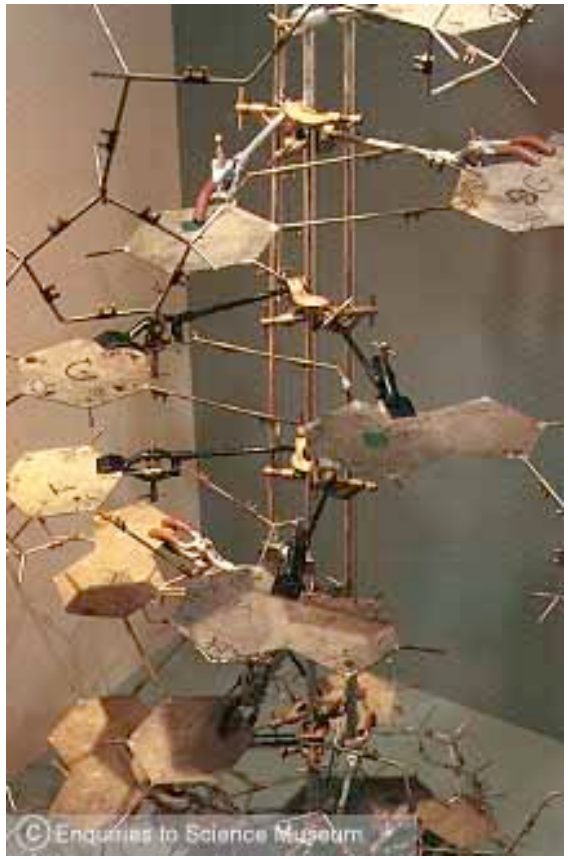
Francis Crick

James Watson

Maurice Wilkins

Rosalind Franklin

The Actual Models



The Double Helix

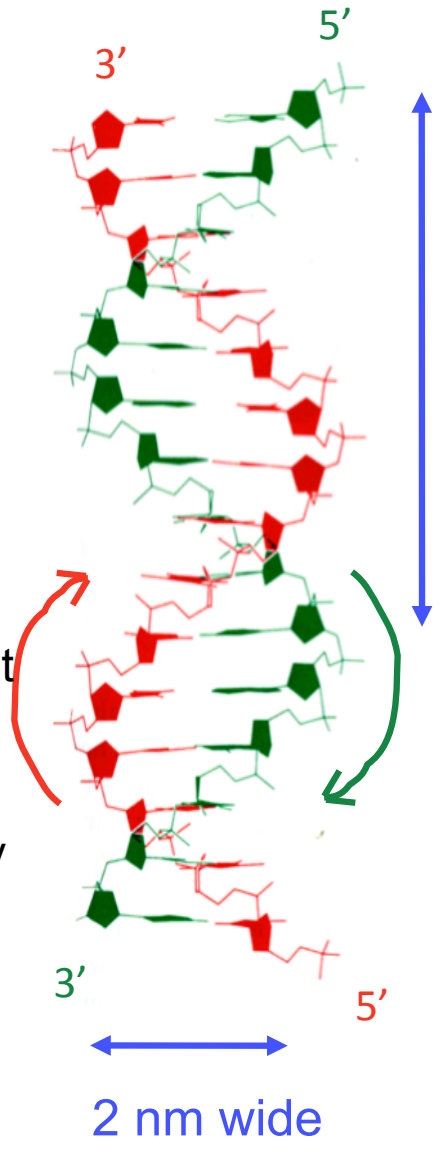
Two strands of DNA run alongside each other forming a right-handed double helix.

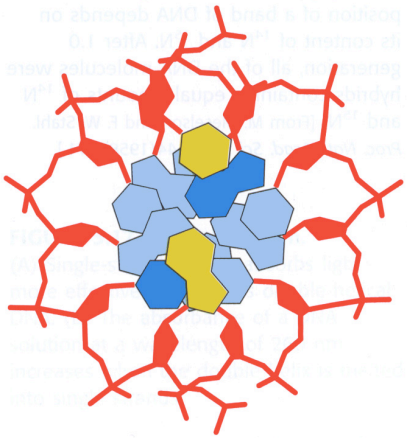
The chains are antiparallel - they run in opposite directions.

The deoxyribose and phosphate groups run along the outside with the negative charges outside.

The bases point inwards and the flat planes are perpendicular to the helix.

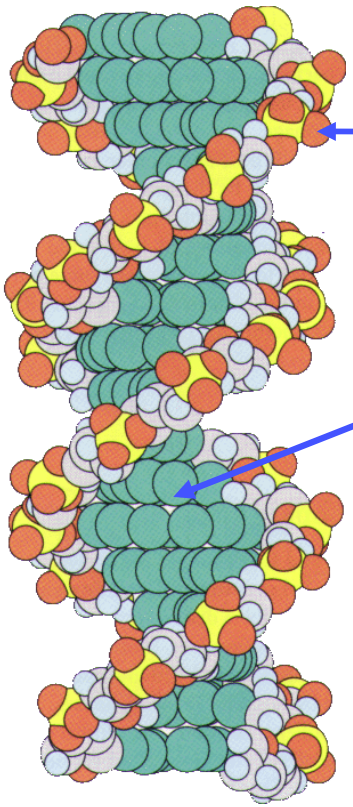
The two chains are held together by hydrogen bonds between the bases.





Top view

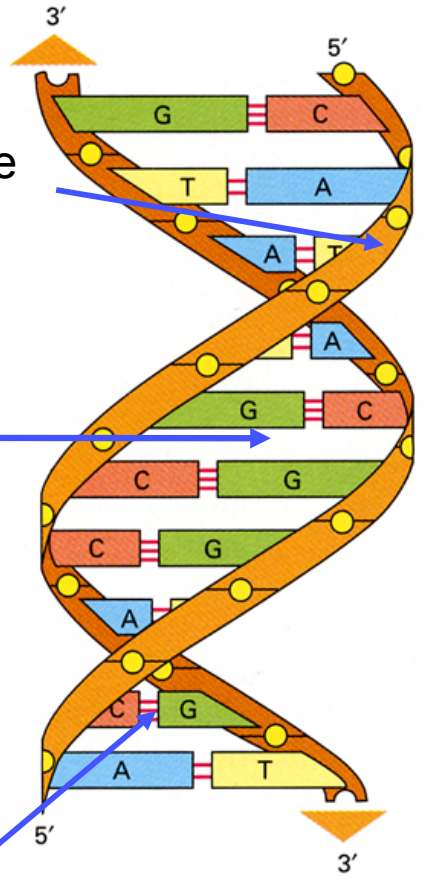
The Double Helix



phosphate backbone on the outside

DNA bases buried on the inside

Hydrogen bonds between the bases stabilise the structure. Pairing of specific bases (Watson - Crick base pairs).

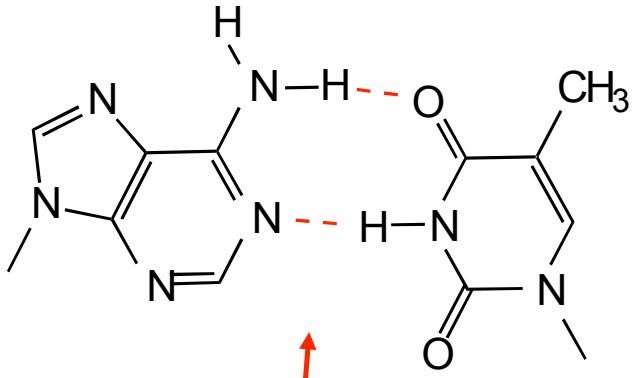


minor groove

major groove

Hydrogen Bonding

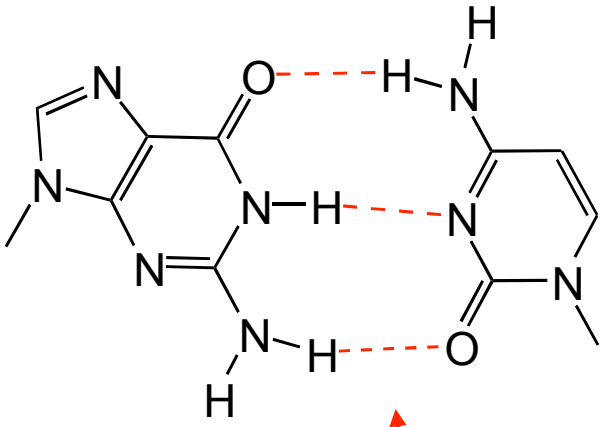
Watson-Crick base pairs



Adenine
(A)

Thymine
(T)

2 hydrogen bond
less stable



Guanine
(G)

Cytosine
(C)

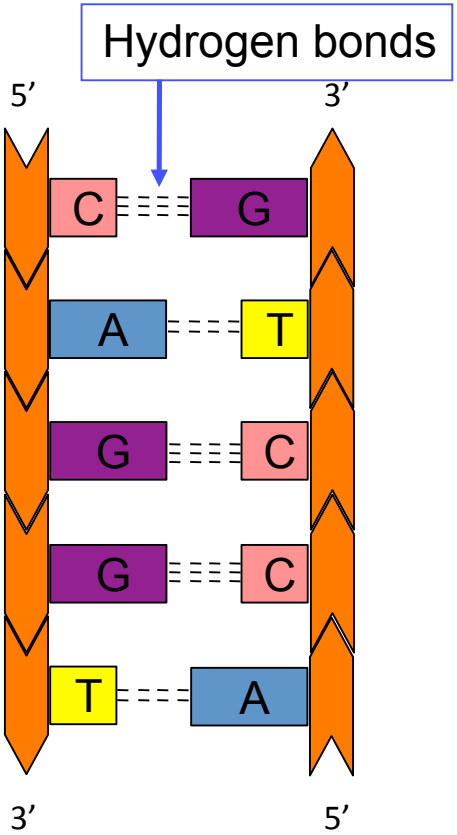
3 hydrogen bonds
more stable

Sequence complementarity of DNA strands

Most important aspect of DNA double helix: Specificity of pairing of bases. Adenine always pairs with Thymine. Guanine always pairs with Cytosine.

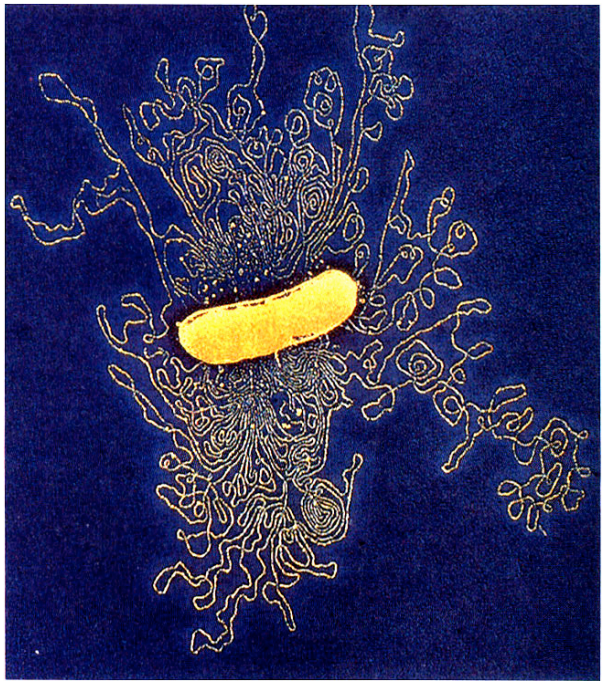
The two strands are complementary in their sequence. If the sequence of one strand is known, the sequence of the other strand follows.

```
5′  AATCGAACACCGTACCGT  3′
    ||||||||||||||||||
3′  TTAGCTTGTGGCATGGCA  5′
```



The *E.coli* genome

The entire DNA coding for an organism constitutes its genome. *E.coli* has 4.6 x 10^6 base pairs in a single circular double-stranded molecule. The length of *E.coli* DNA is 1.4 mm.



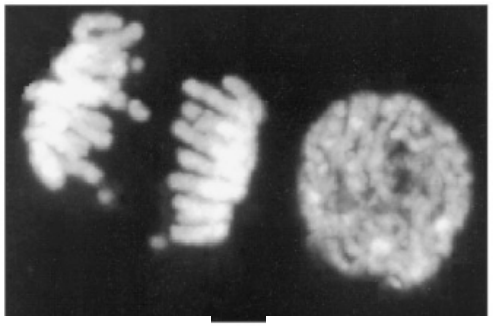
Electron micrograph of part of the *E.coli* genome

The Human Genome

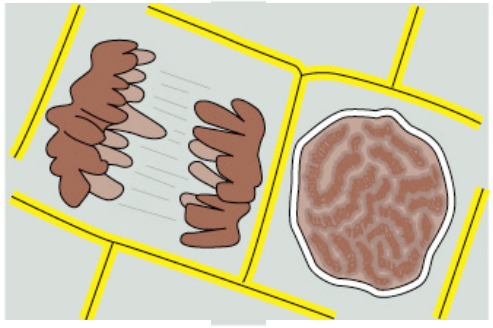
~ 3 x 10^9 base pairs of DNA divided into chromosomes that each contain a single, linear double-helical DNA molecule of ~ 200 x 10^6 base pairs.

The sequence of the human genome is now completed

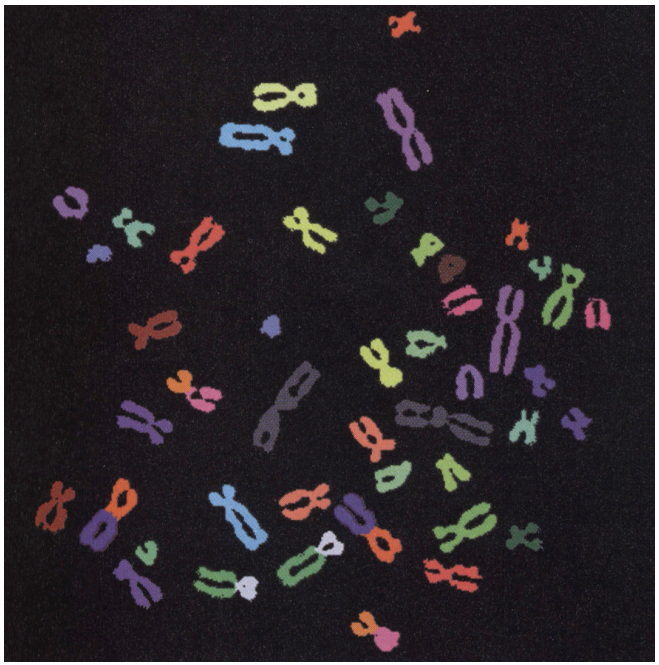
Human chromosomes; visible only just before cells divide but not in non-dividing cells



dividing non-dividing



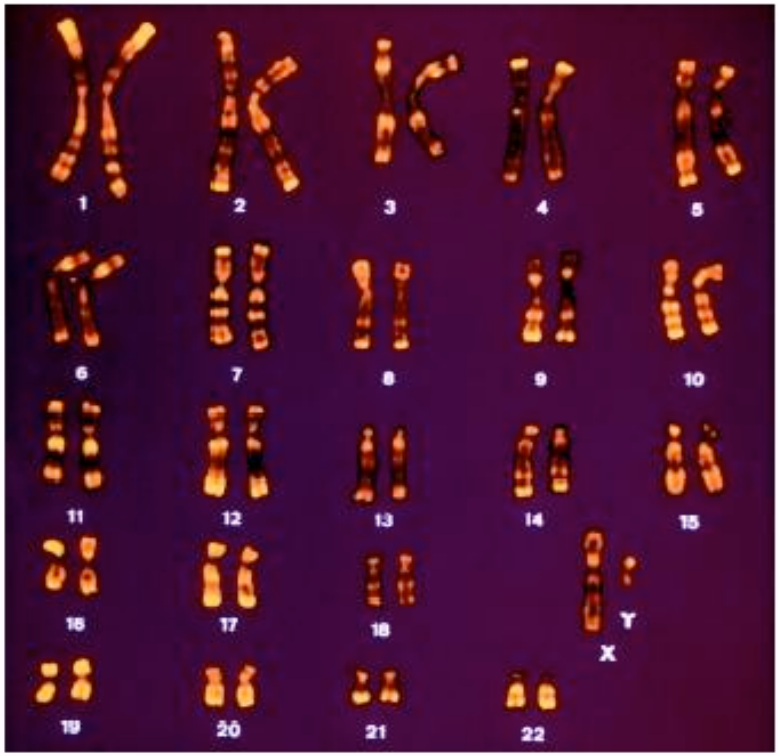
dividing non-dividing



Human chromosomes stained with fluorescent dyes

The Karyotype

A karyotype is an organised profile of someone's chromosomes.
A diploid human cell has 46 chromosomes
    22 pairs of 'normal' chromosomes (autosomes) and
    2 sex chromosomes (X and Y)
Sex chromosomes: XX for female; XY for male



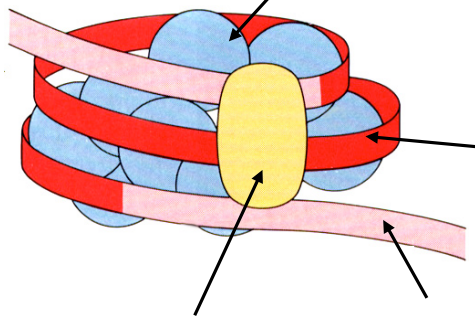
Two homologues of each chromosome

G-banded karyotype of a normal male cell

Structure of the Nucleosome

8 histones; 2 each of 2A, 2B, 3 and 4

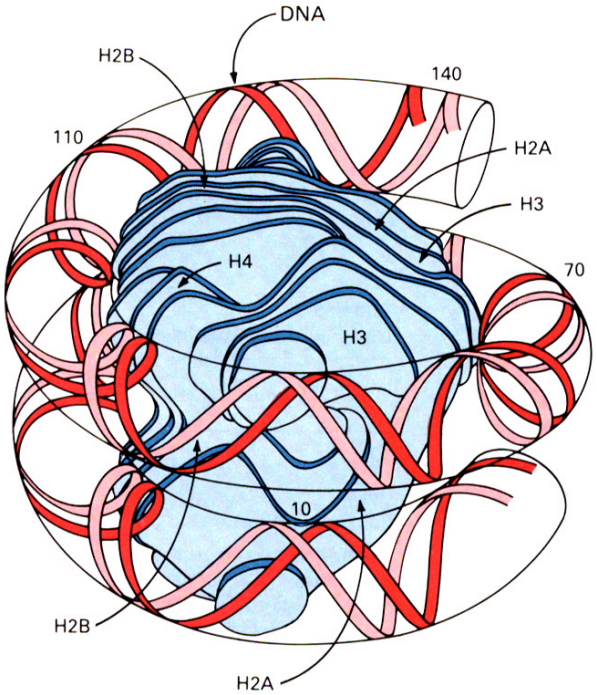
Histones: positively charged; interact with negatively charged sugar-phosphate backbone of DNA



About 150 base pairs of DNA wrapped around the histones (core DNA)

Histone 1 between the nucleosomes

Linker DNA



Model of nucelosome core showing DNA wound in left-handed superhelix around the histone octamer

Chromosomes

DNA double helix

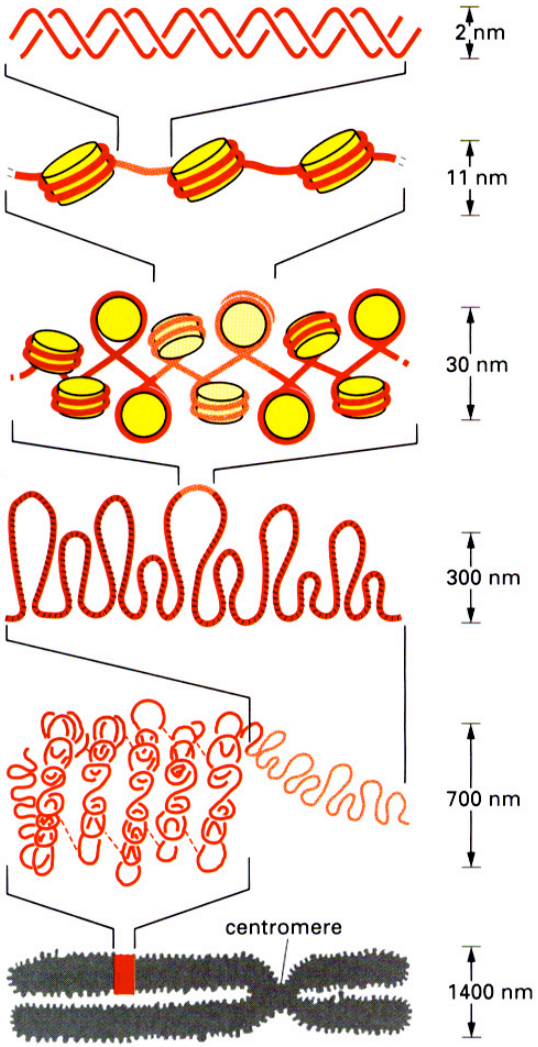
Beads on a string form of chromatin

30 nm chromatin fiber of packed nucleosomes

Section of chromosome in extended scaffold (interphase chromosome)

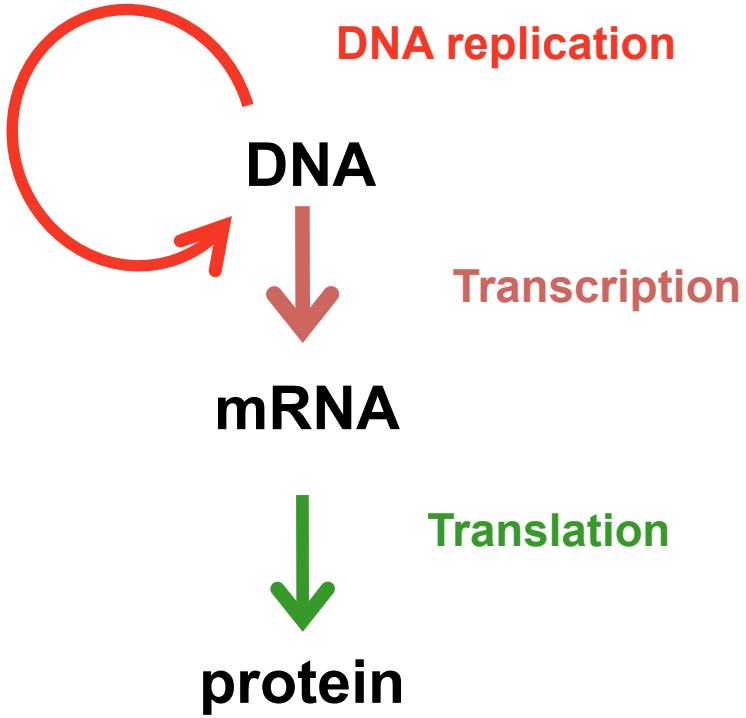
Condensed section of chromosome

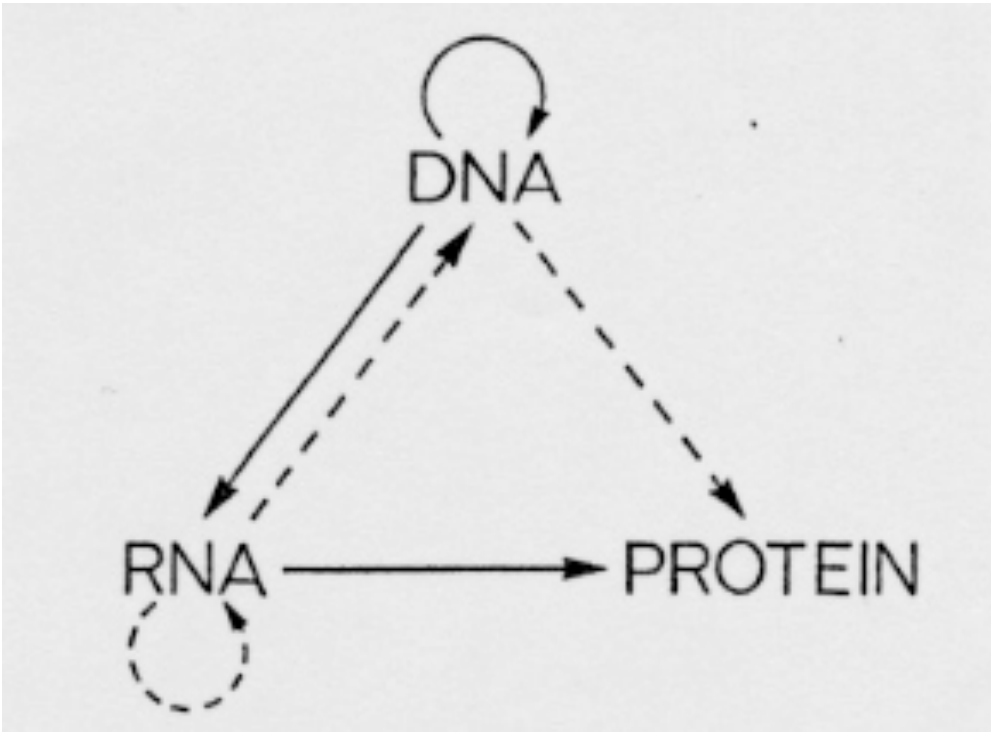
Entire mitotic chromosome



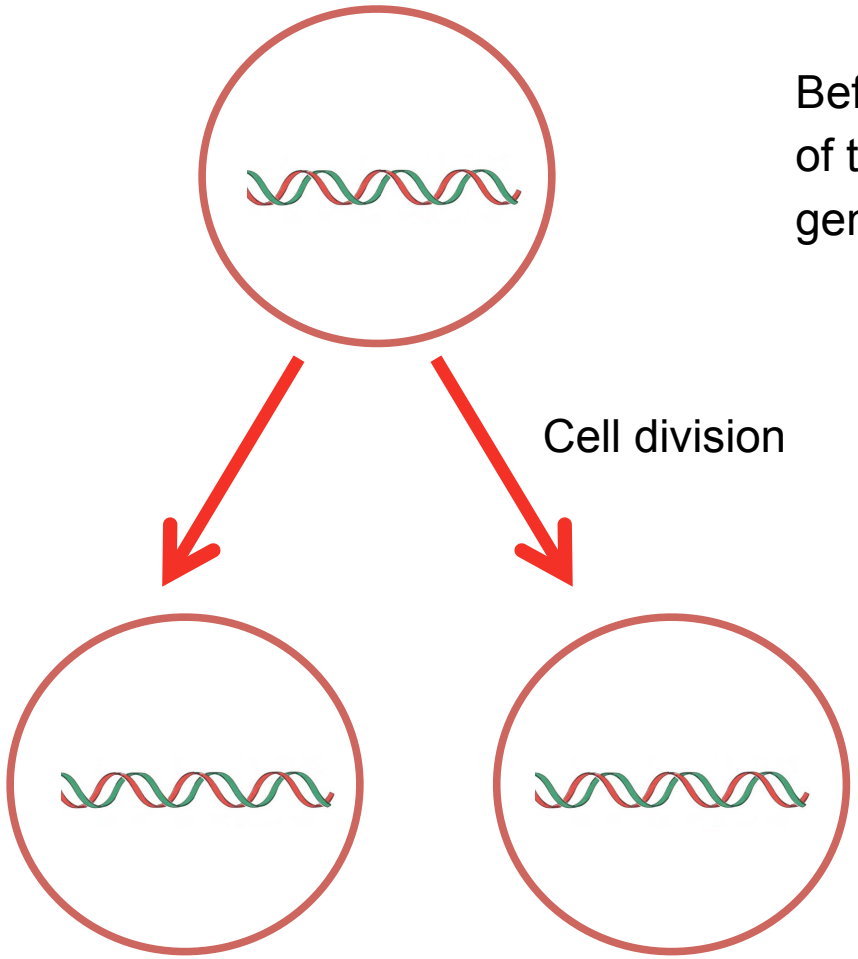
Net result: each DNA molecule has been packaged into a mitotic chromosome that is 10,000 fold shorter than its extended length.

Flow of Genetic Information





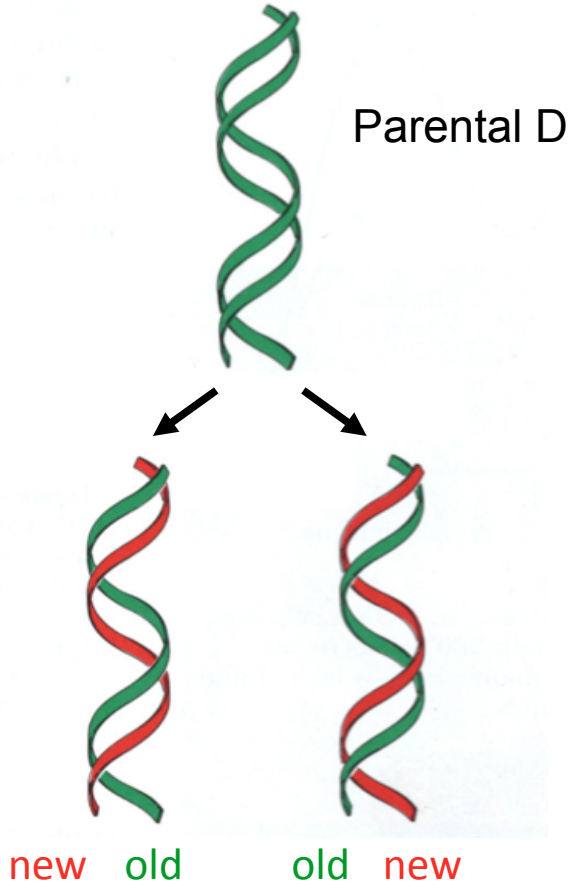
DNA Replication



Before each cell division, a faithful copy of the DNA needs to be made to ensure genetic continuity in the daughter cell.

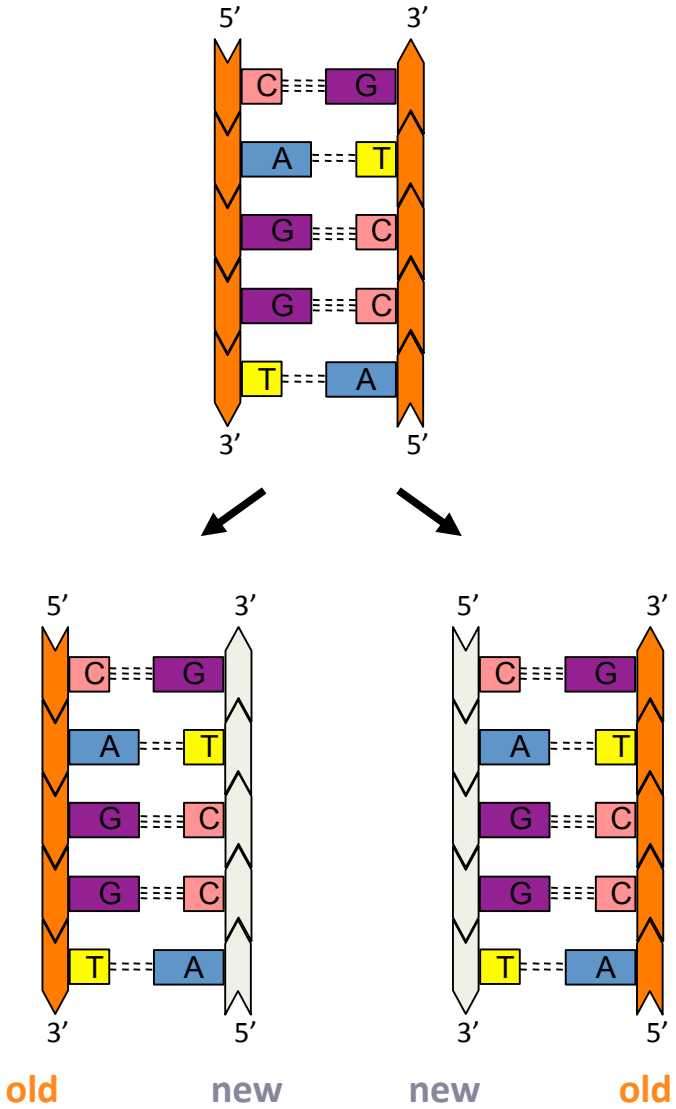
Identical copies of the DNA code

Semi-conservative Replication



DNA replication is semi-conservative. Each daughter cell inherits one old and one new strand.

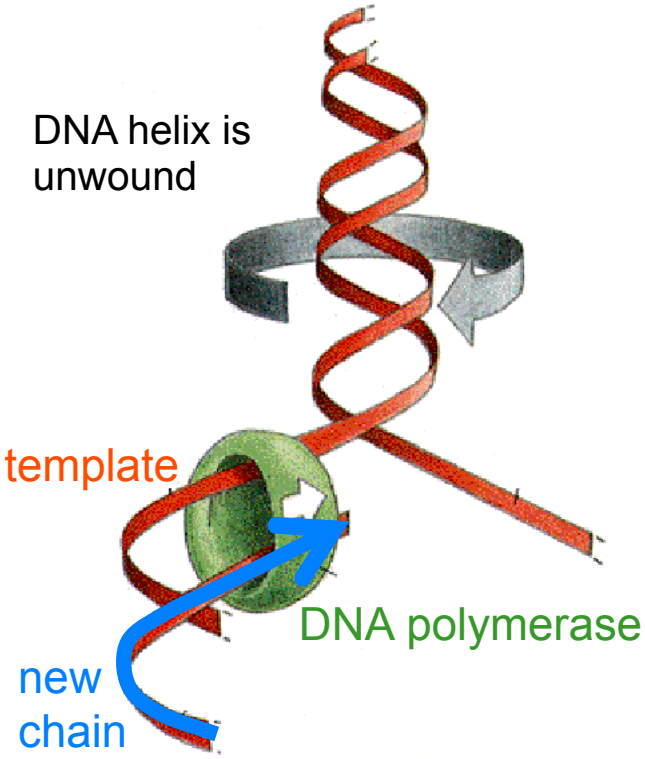
DNA Replication



The two strands are complementary to each other so each strand serves as a template for the synthesis of the other strand. This generates two identical copies.

Opening of DNA Helix

The DNA helix is very stable and has to be unwound before replication can occur. This is done by a DNA helicase, an anzyme that uses ATP as source of energy to break hydrogen bonds between base pairs.



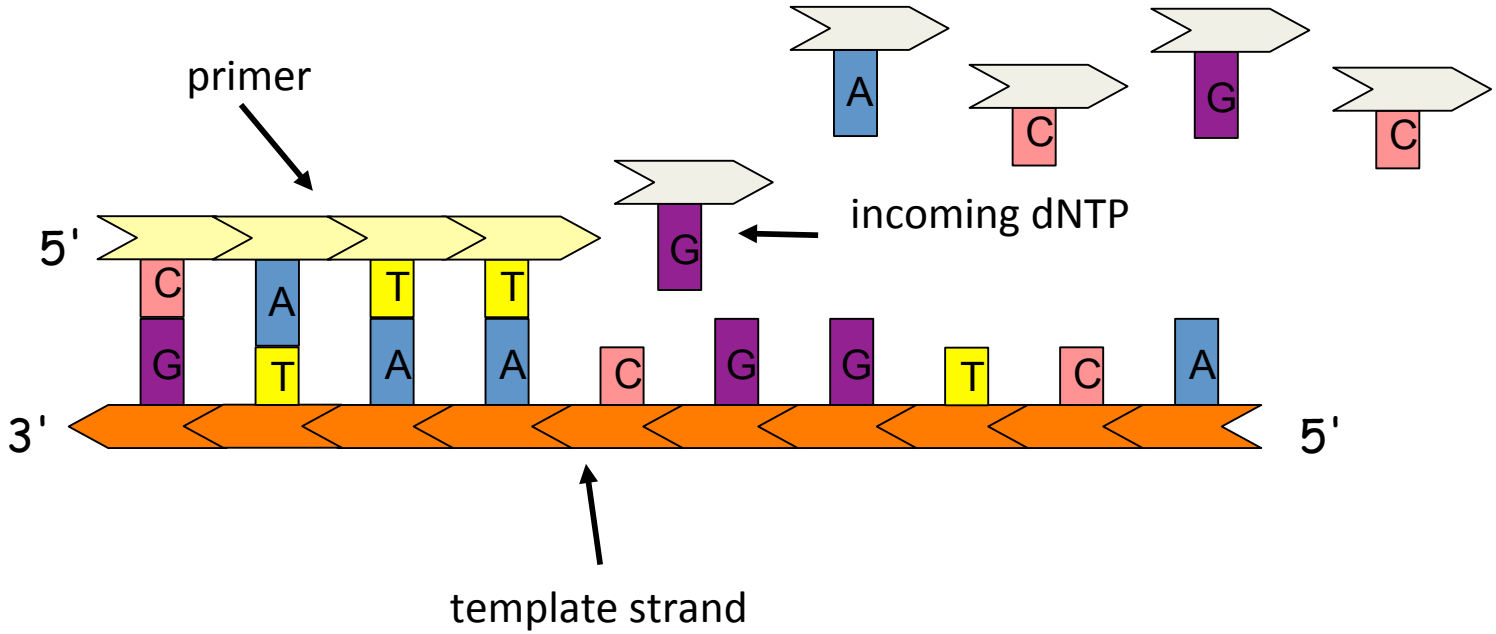
New DNA is synthesised by enzymes called DNA polymerases. DNA polymerases add nucleotides to the 3' end of a growing chain.

DNA Synthesis

DNA polymerases require:

- a template strand
- an oligonucleotide primer
- a supply of deoxynucleotide triphosphates (dNTPs)

DNA polymerases cannot start a new chain from scratch!

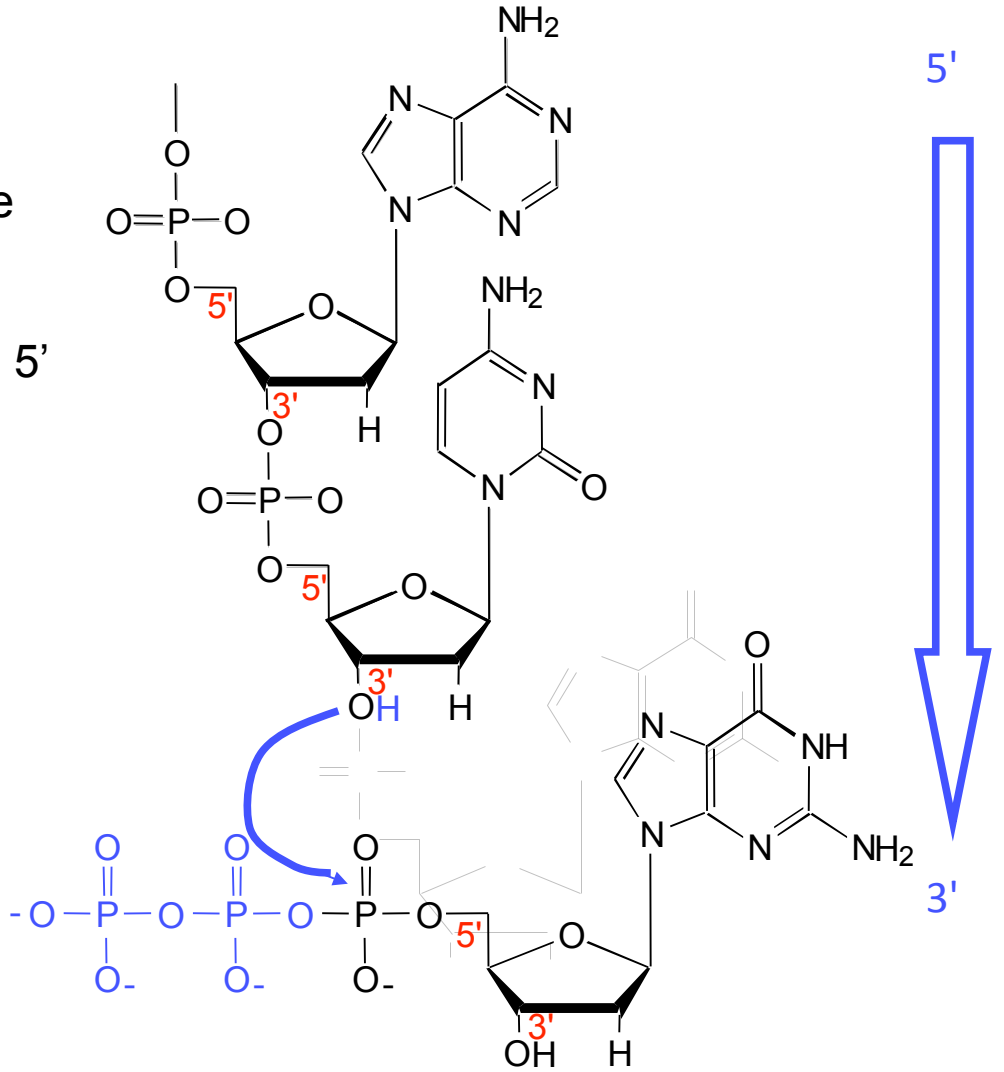


Enzyme Reaction

DNA polymerases add dNTPs to the 3’ end of a DNA molecule.

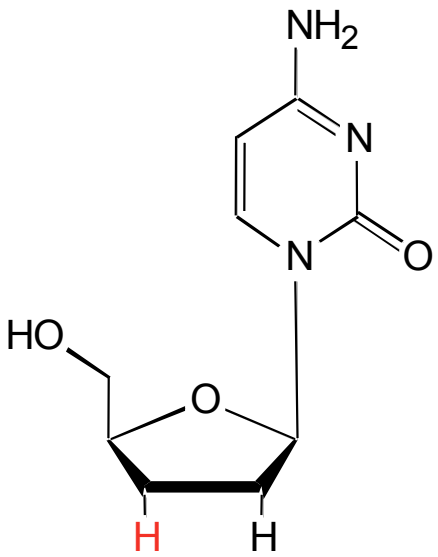
DNA (and RNA) synthesis occurs in 5’ to 3’ direction.

Energy is released by hydrolysis of the triphosphate. This drives the reaction.



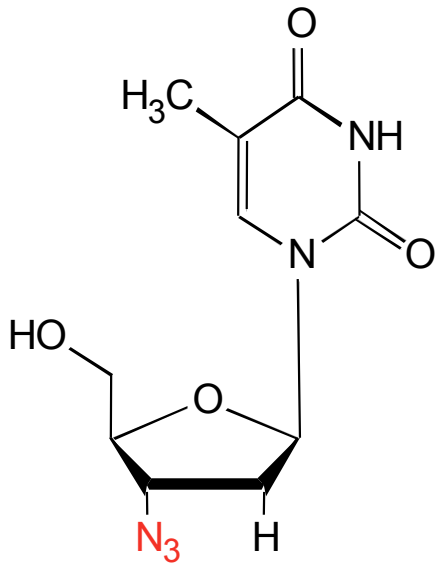
Drugs used as chain terminators I

Nucleoside analogues



dideoxycytosine (ddC)

Drug for HIV (zalcitabine)



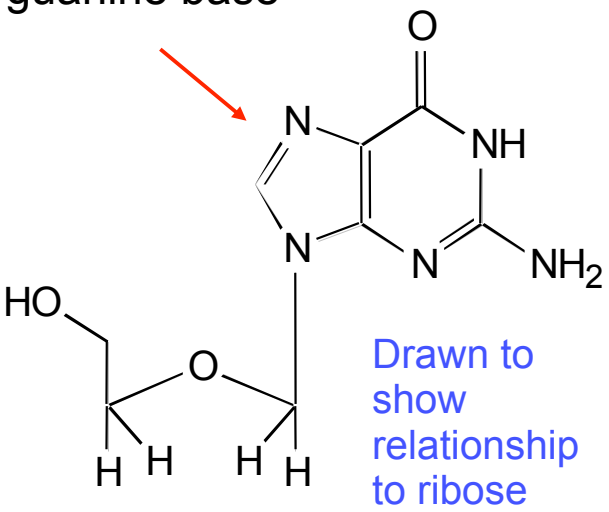
azidothymidine (AZT)

Drug for HIV (Zidovudine)

Drugs used as chain terminators II

Nucleoside analogues

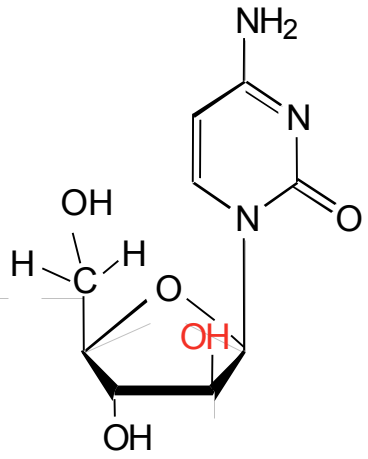
guanine base



Drawn to show relationship to ribose

Acyclovir

Drug for Herpes



Cytosine arabinose

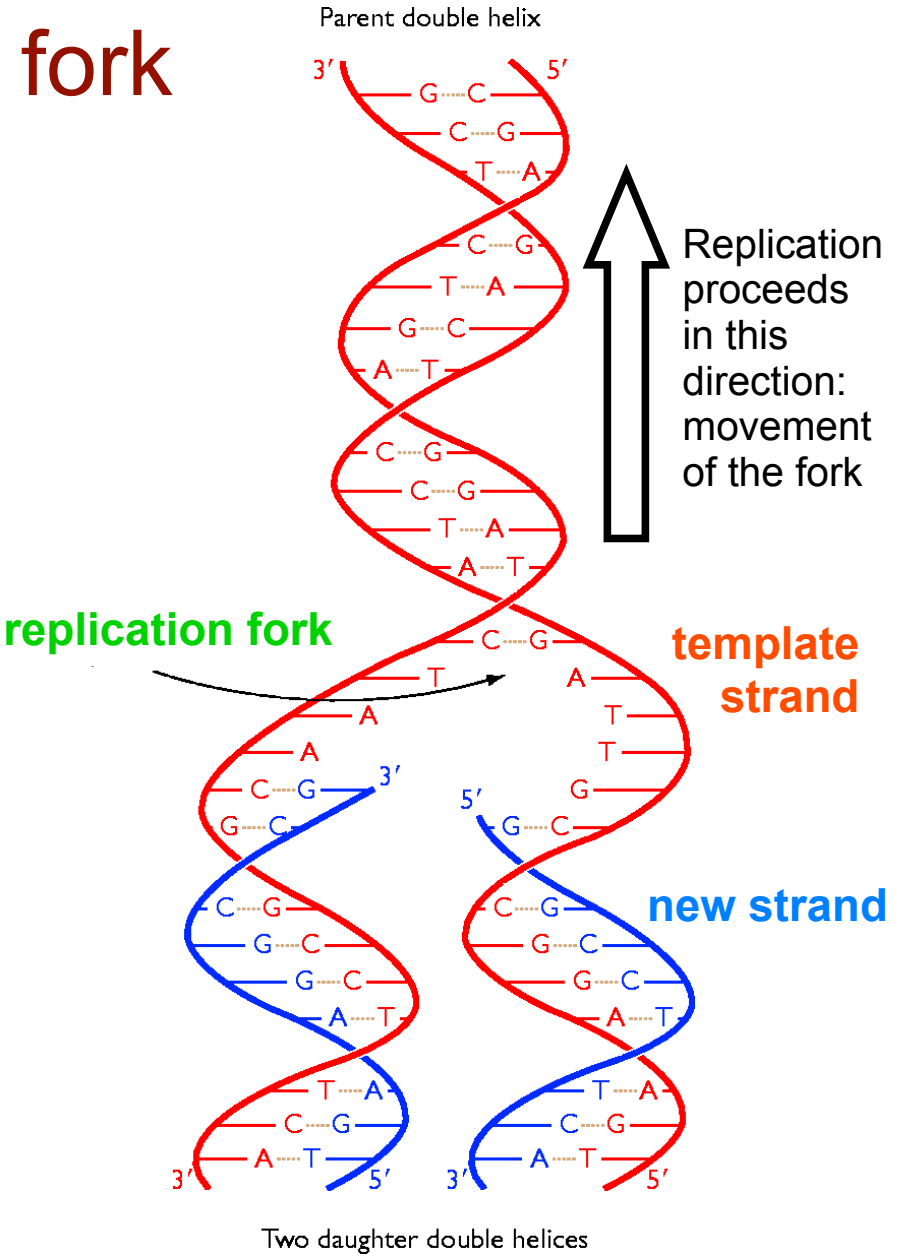
Used in Chemotherapy

The replication fork

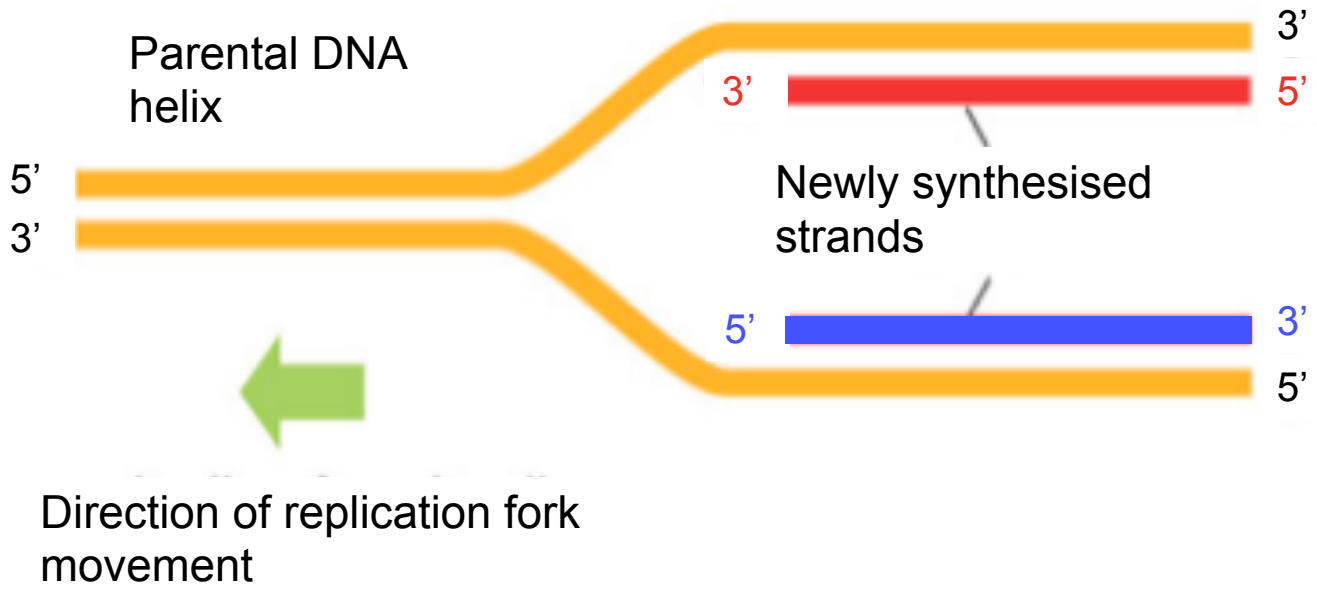
Replication begins at discrete points on the DNA molecule called origin of replication.

The site of DNA synthesis is called a replication fork: the fork moves along during the process.

The two daughter molecules are identical, each containing an old and a new strand.

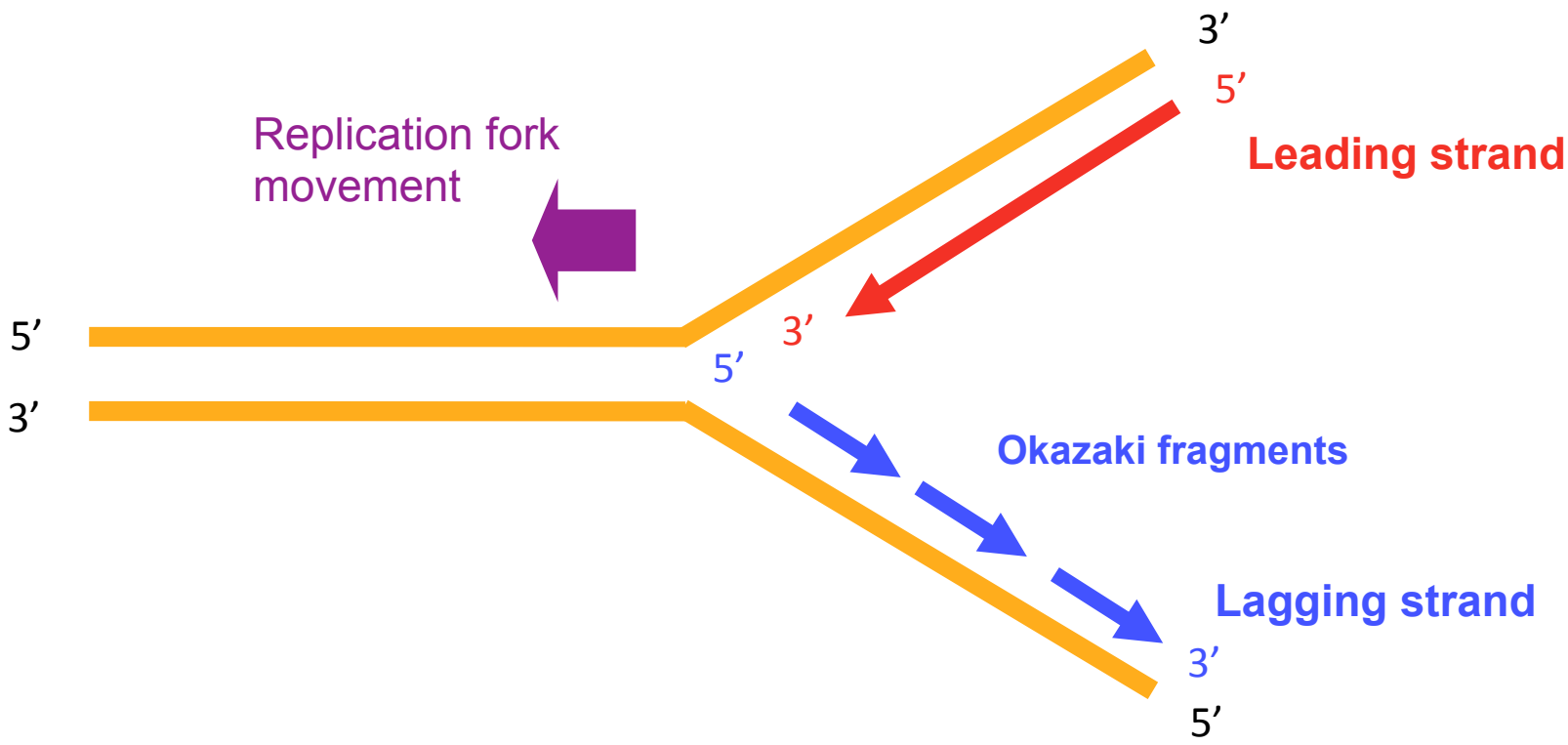


Asymmetry of replication fork



The templates for the two new daughter strands have opposite orientations: 3’ to 5’ and 5’ to 3’

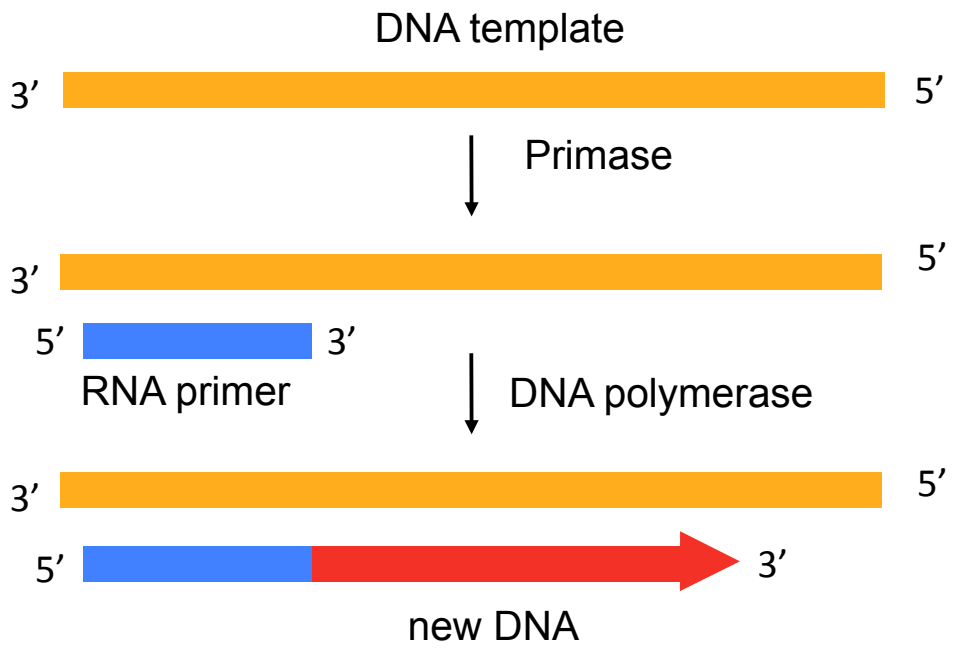
Leading and lagging strands



The replication fork is asymmetric. Both strands are synthesised in a 5’-3’ direction. The leading strand is synthesised continuously, whereas the lagging strand is synthesised in short pieces termed Okazaki fragments.

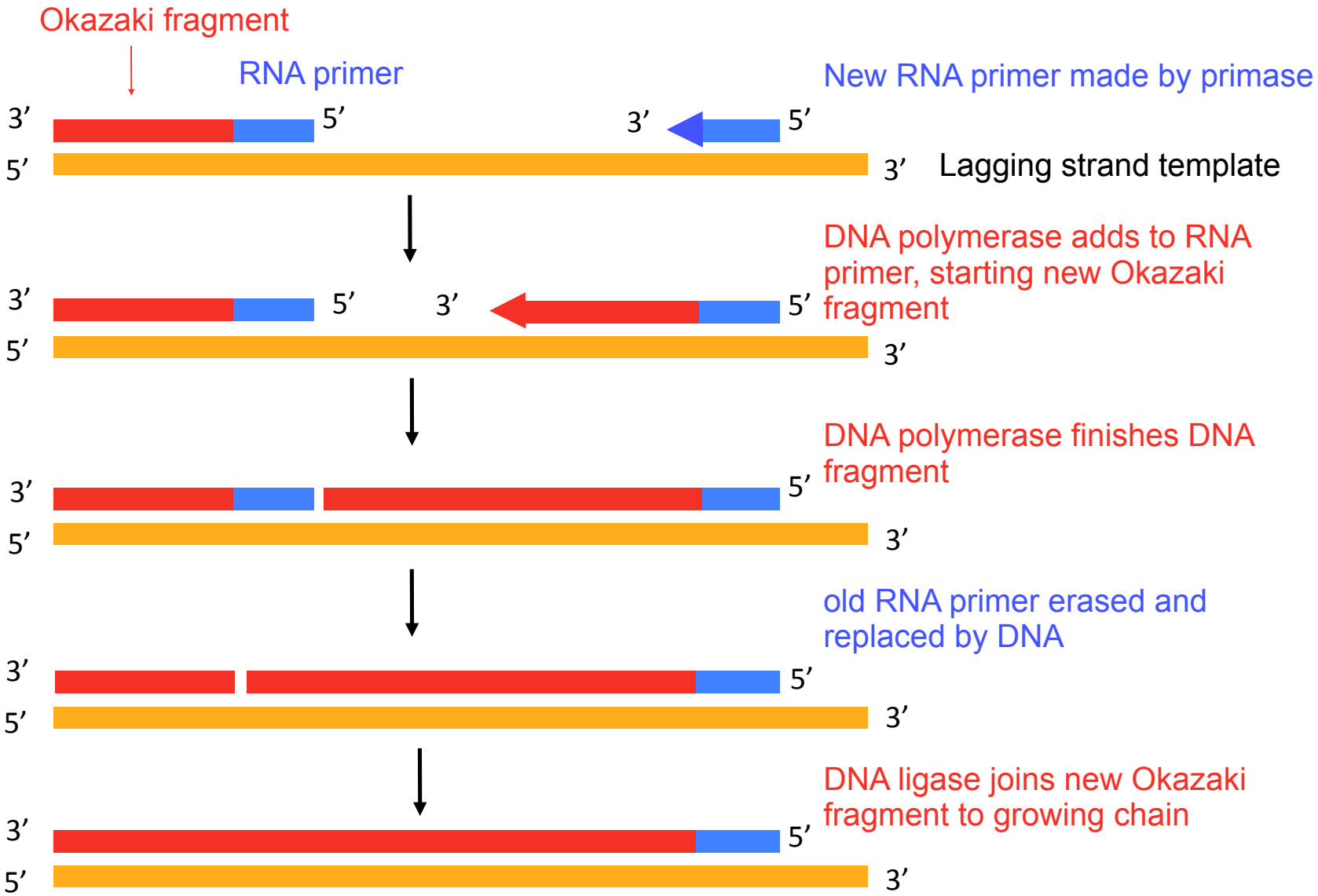
RNA primer starts new chain

RNA primes the synthesis of new DNA. A specialised RNA polymerase called primase synthesises a short RNA fragment (~ 5 nucleotides). The RNA primer is only transient and removed at a later stage of replication.

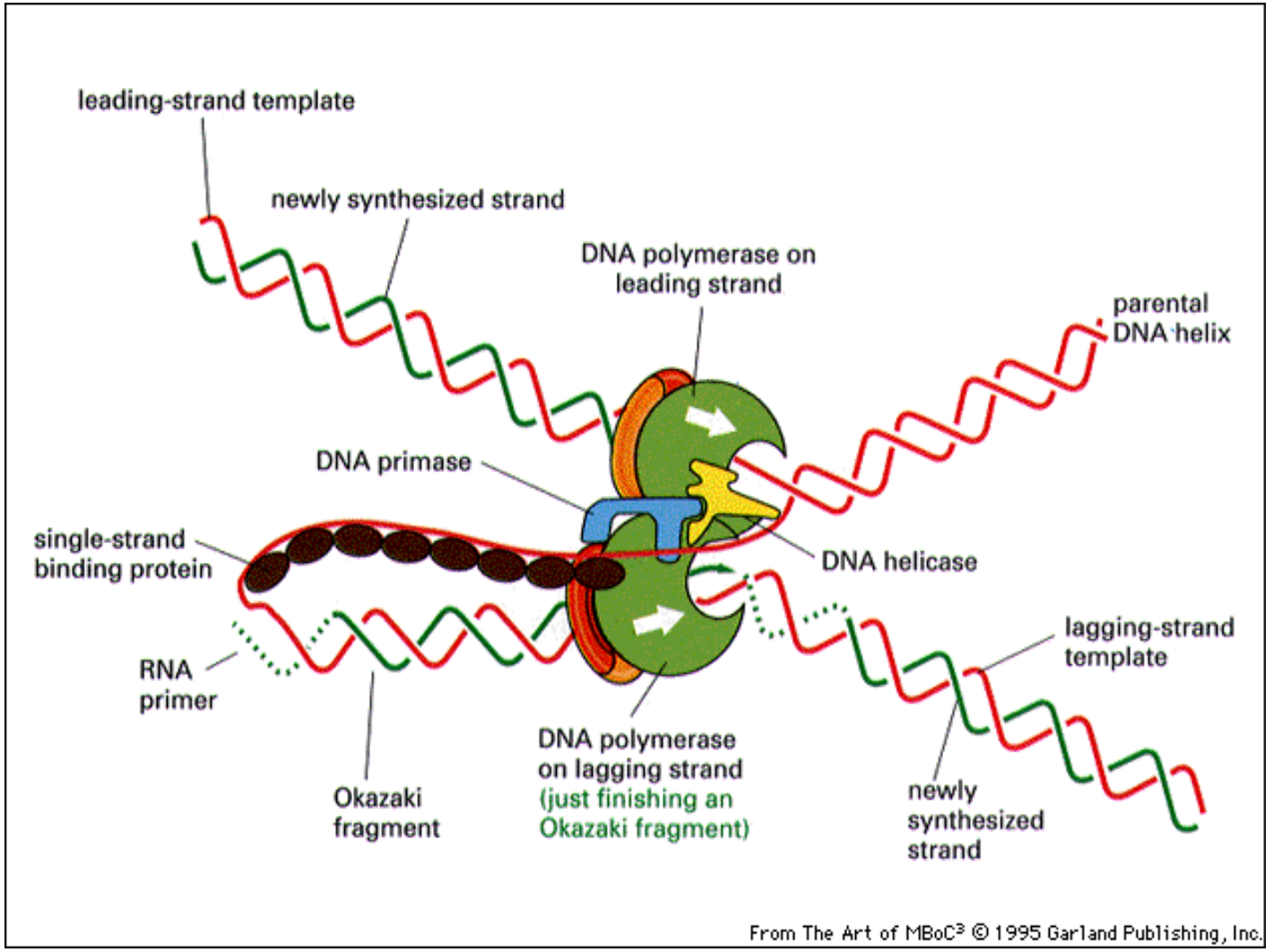


For the synthesis pf the leading strand, an RNA primer is needed only to start replication at a replication origin.

Synthesis of lagging strand



The Replication Complex

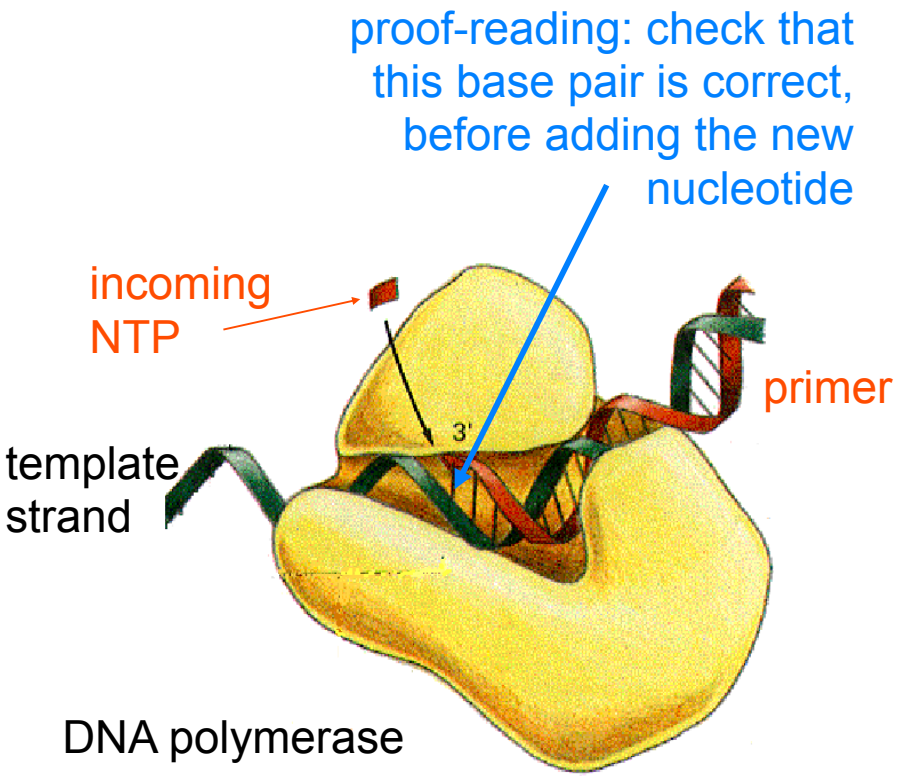


Proofreading mechanism

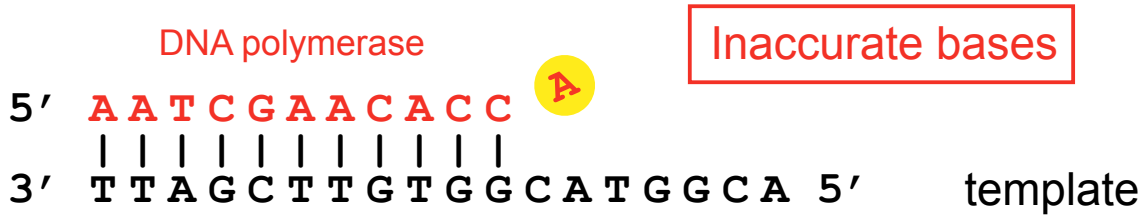
The high fidelity of DNA replication requires a proof-reading mechanism to ensure no mistakes are made. Mutations (changes in DNA sequence) are very dangerous to the organism. Any errors in replication cannot be repaired.

DNA replication has an error frequency of about 1 change per 10^9 base pairs.

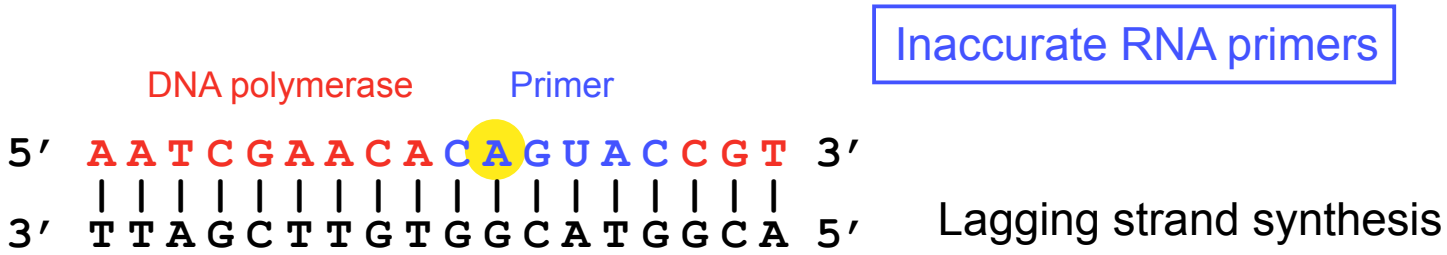
Before a new nucleotide is added, the previous nucleotide is checked for correct base-pairing.



Proofreading



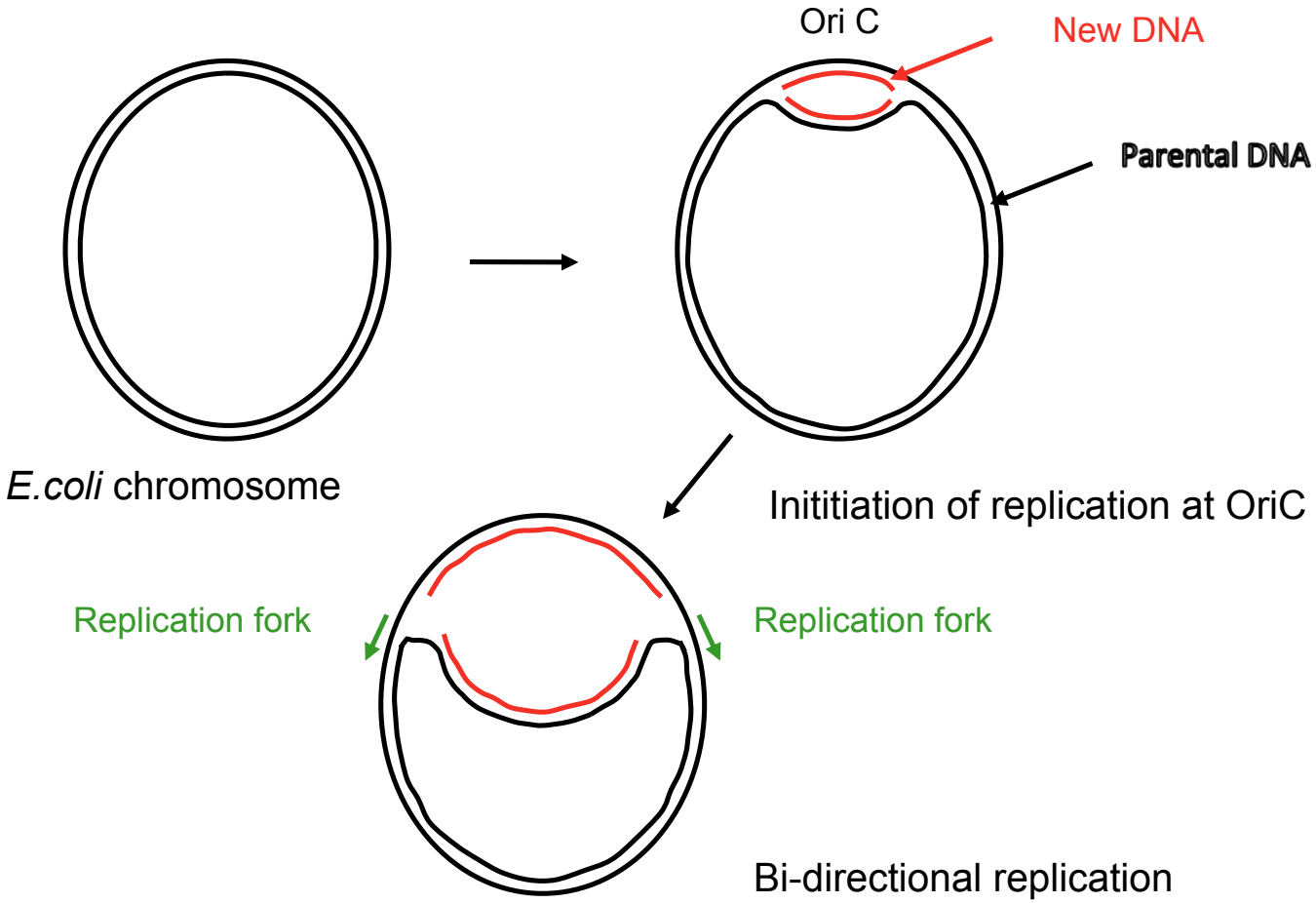
Incorrect base does not fit correctly. Phosphodiester bond is hydrolysed by 3' to 5' exonuclease activity of DNA polymerase. A new, correct nucleotide is then added.



The first few bases of RNA synthesis are inaccurate as there is no pre-existing double helix to be extended. This inaccurate RNA is replaced by accurate DNA sequence.

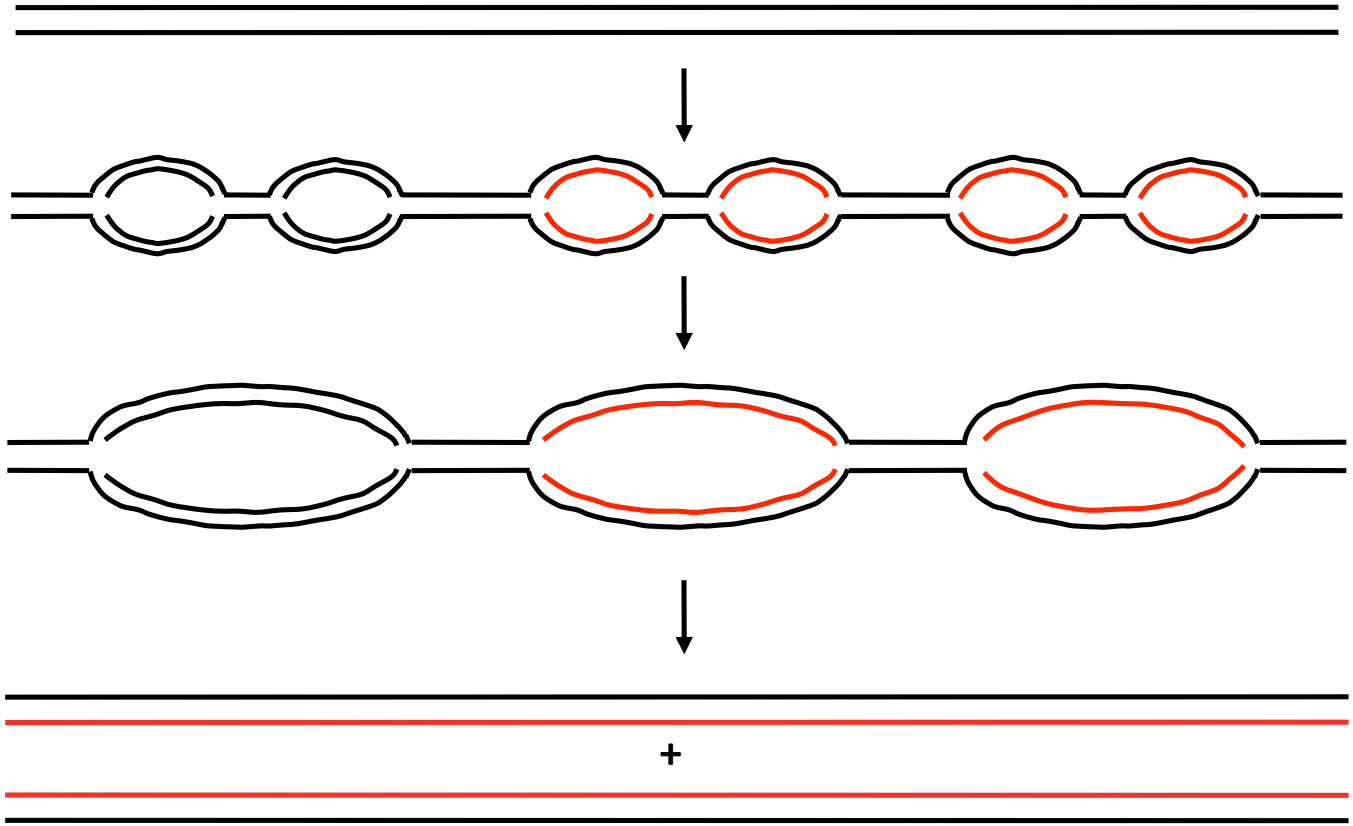
Replication of the *E.coli* chromosome

In *E.coli*, replication starts at a unique origin, OriC. Two replication forks proceed simultaneously in opposite directions. The two forks meet at the other side of the circular chromosome.



Replication of eukaryotic genome

Eukaryotic chromosomes are linear and very long. Multiple replication origins are distributed at intervals of about 100 kilobase pairs. Each replication origin gives bi-directional replication forks. Replication is finished when all the forks have met.

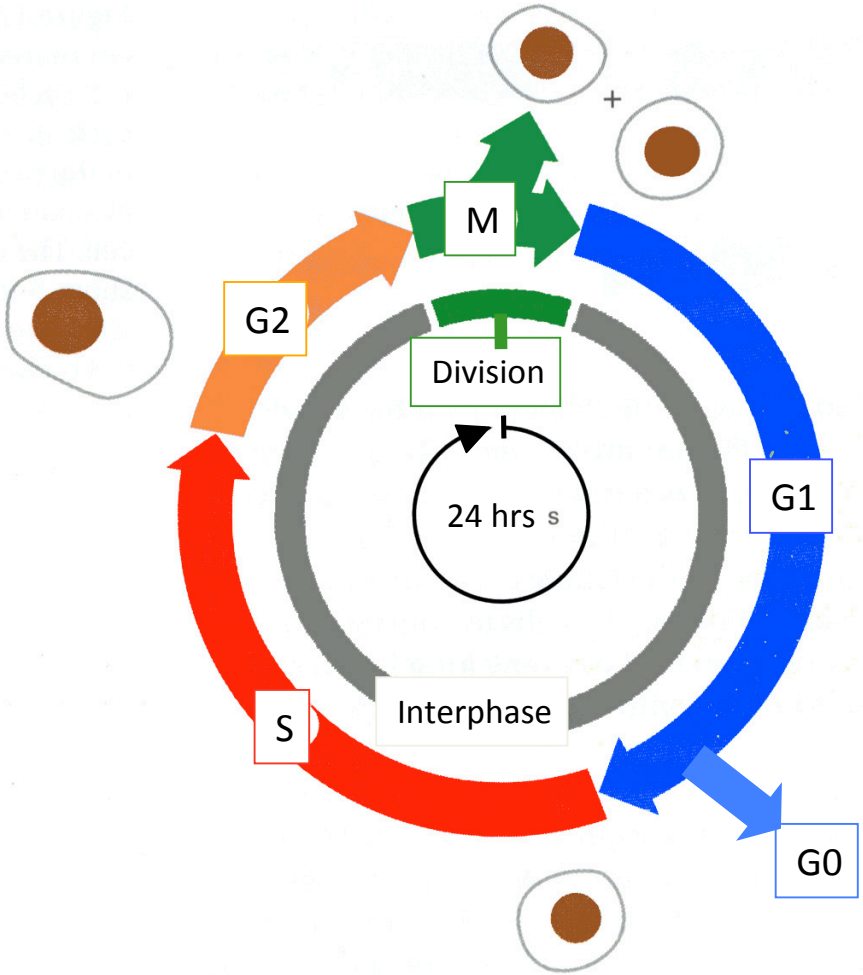


Replication Enzymes

DNA polymerase	5’ to 3’ polymerase activity 5’ to 3’ exonuclease activity (removes RNA primers and repair) 3’ to 5’ exonuclease ‘proofreading’ activity
DNA ligase	seals gaps in ds DNA
DNA helicase	unwinds double helix for replication
RNA primase	synthesises RNA primer to initiate replication
Telomerases	ensure integrity of linear DNA molecules

The Cell Cycle

G1, G0, S, and G2 are Interphase



M phase: Mitosis; cell division; 1hr

G1 phase: Gap phase 1 (prior to DNA syntesis); 10 hrs

S phase: period of DNA **s**ynthesis (replication); 9hrs

G2 phase: Gap phase 2 (between DNA synthesis and mitosis); 4 hrs

G0: cells which have stopped dividing

SBA - Single Best Answer

1. The name given to the DNA strand of a gene that is identical in sequence to mRNA is the:

a) Protein strand

b) Nonsense strand

c) Antisense strand

d) Sense strand

e) Lagging strand

2. Which enzyme catalyses the unwinding of DNA strands during replication?

a) DNA ligase

b) DNA polymerase

c) DNA helicase

d) Telomerase

e) RNA primase

SBA - Single Best Answer

3. Which one of the following best describes gene structure?

a) All exons in a gene code for protein

b) All introns start with the sequence 5'-GU- and end with -AG-3'

c) Mammalian mRNA can encode multiple sequential protein open reading frames

d) Alternatively spliced mRNAs are produced by splicing of the genomic DNA

e) Introns are redundant sequences, which never contribute to gene regulation

4. In the genetic code:

a) CCG is the initiator codon

b) Each amino acid is encoded by 4 bases in the mRNA which is called a codon

c) Each amino acid is encoded by only one possible codon

d) UAG is a termination codon

e) The anticodon is a triplet sequence present on the mRNA

SAQ 1 *Nucleic acids*

What are the building blocks of nucleic acids called? What is the chemical composition of these building blocks? *(2 marks)*

Give similarities and/or differences (4 in total) between DNA and RNA, in terms of chemical composition and structure. *(2 marks)*

```
atccgactccttggatactaGACTCCTTACTATGCGTGCTGCATACTGgtactgac--

--tagca (sequence left out) --ctgagagtgacatgagCATATATACAGTA
```

Above is the sequence of the 5' region of a gene. Only the sense strand of the DNA is shown. The first and second exons are shown in capital letters. The regions on either side are in small letters.

i) Mark and label the start of transcription on the above diagram. (1 mark)

ii) Mark and label the splice acceptor signal on the above sequence. (1 mark)

iii) Mark and label the most likely position where the ribosome will start translation of the above sequence (once it has been used to make mRNA). (1 mark)

SAQ 2 *Tissues*

What are basement membranes? (1 mark). Name 3 locations where they are found. (1½ marks)

Describe the structure of the collagen triple helix. Include and name the 3-amino acid motif and explain its significance. (2½ marks)

Aggrecan is the major proteoglycan of cartilage. Define a proteoglycan. (2 marks)