
CHAPTER

20

CHROMOSOMES AND DNA

*Animation 20:: Growth and development
Source & Credit: Wikipedia*

Chromosomes are thread like structures that appear inside the nucleus at the time of cell division. They were first observed by the German embryologist Walther Fleming in 1882, when he was examining the rapidly dividing cells of salamander larvae. Since their discovery, chromosomes have been found in the cells of all eukaryotes. Their number however varies from species to species. **Pencillium**, a fungus, has only one pair of chromosomes, while some ferns have more than 500 pairs. A mosquito has 6, honeybee 32, corn 20, sugarcane 80, frog 26 and a mouse has 40 chromosomes. Human cells have 46 chromosomes, consisting of 23 pairs (Fig 20.1). Each of these 46 chromosomes contains hundreds or thousands of genes that play important roles in determining how a person's body develops and functions. The possession of all these chromosomes is therefore, essential for survival. Missing of a part or whole of chromosome leads to serious consequences, and death occurs.

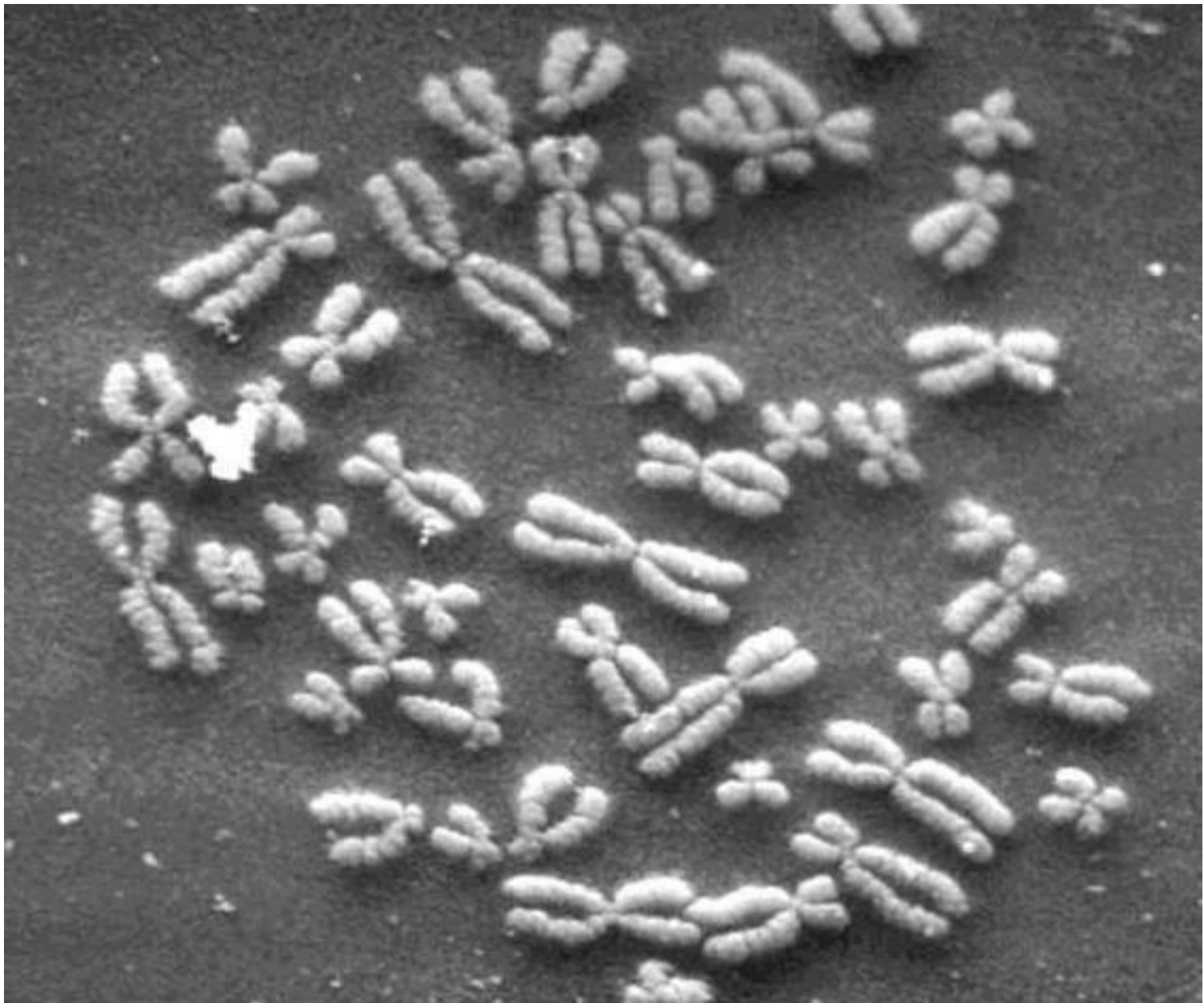


Fig. 20.1 Human chromosomes

TYPES OF CHROMOSOMES

Typically, a chromosome is made of chromatids, centromere, (primary constriction), and a secondary constriction (Fig 20.2).

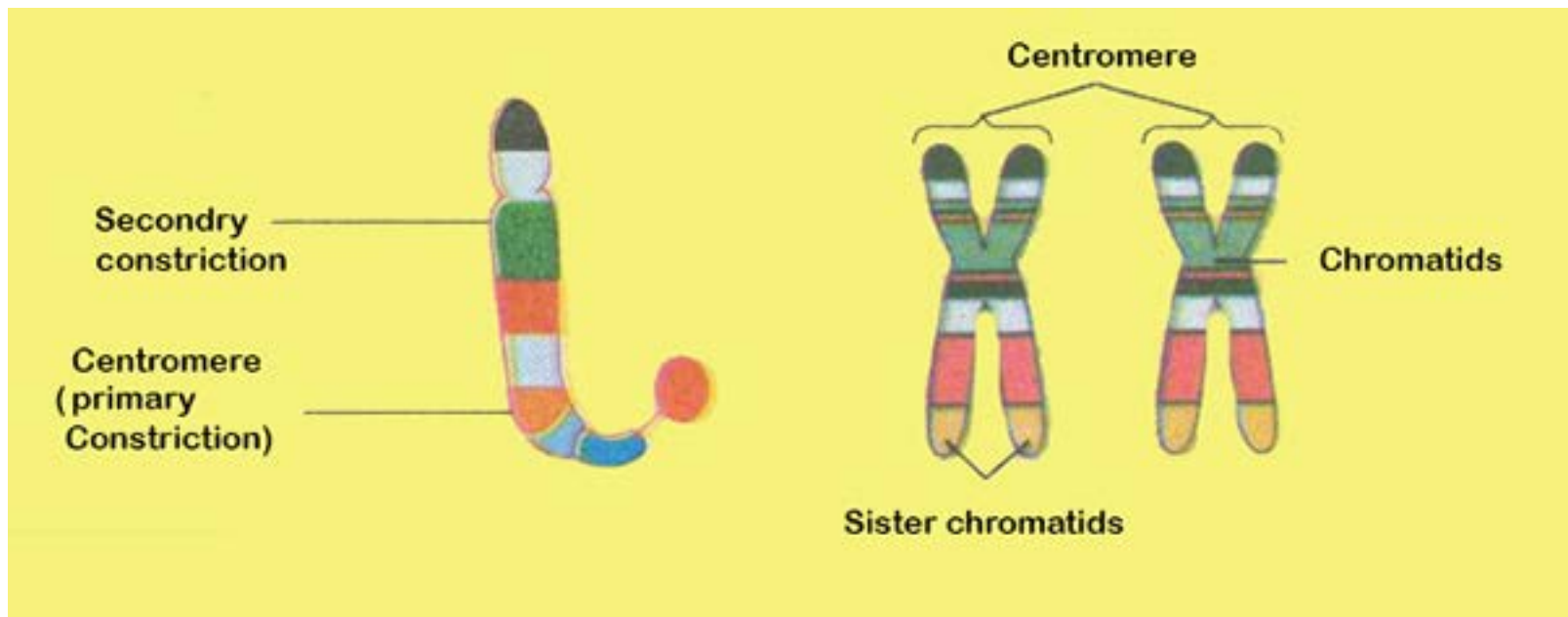


Fig 20.2 Structure of chromosome

Chromosomes may widely differ in appearance. They vary in size, staining properties, the location of centromere, relative length of two arms on either side of centromere, and the position of constricted regions along the arms. The particular array of chromosomes that an individual possesses is called its karyotype (Fig 20.3). Karyotypes show marked differences among species and sometimes even among individuals of the same species.

The chromosomes are called telocentric, acrocentric, sub metacentric and metacentric depending upon the location of centromere between the middle and tip of the chromosomes.

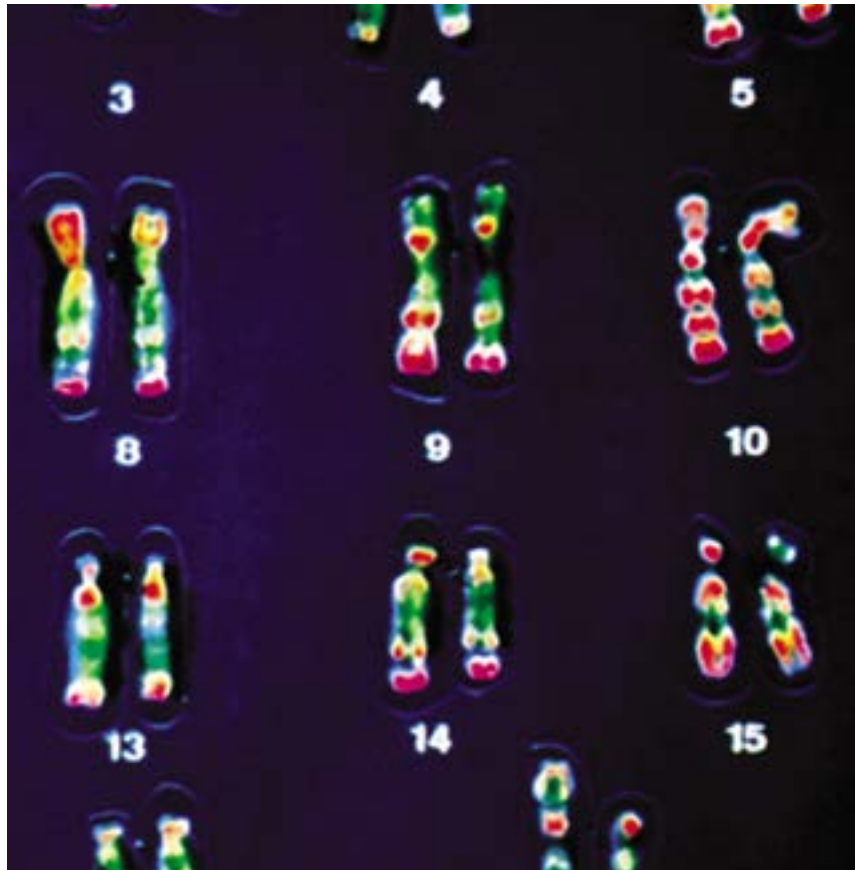


Fig. 20.3 A human karyotype

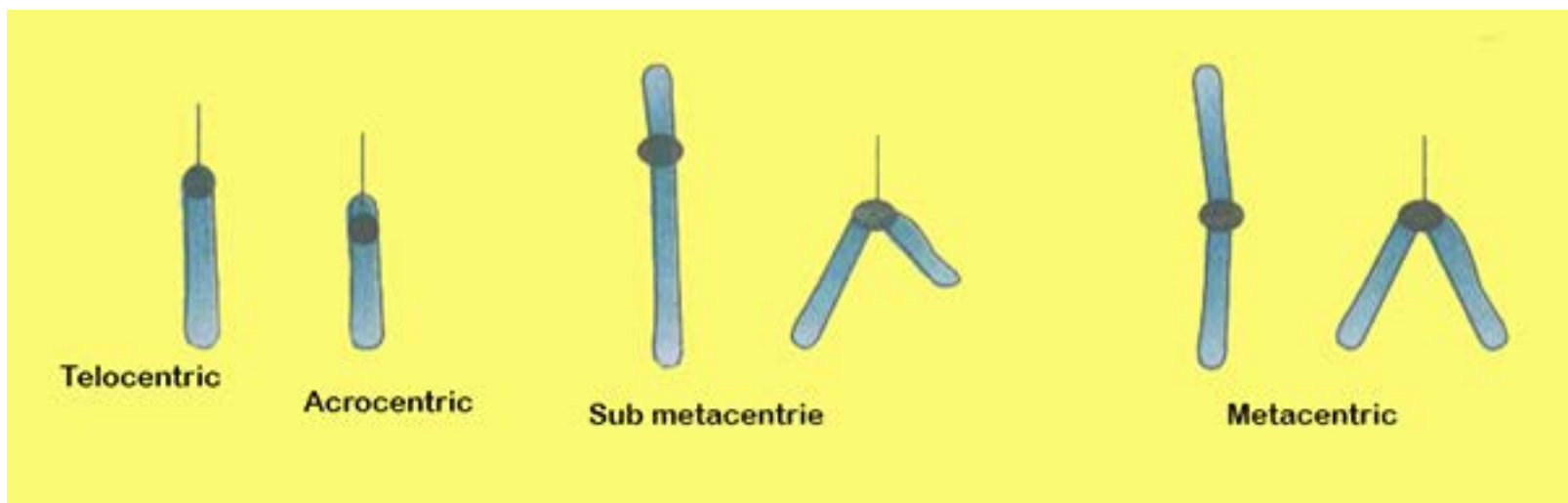


Fig. 20.4 Shapes of chromosomes depends upon the location of centromere

These chromosomes- acquire different shapes at the time of anaphase during cell division. The usual shapes are i, j and v.

COMPOSITION OF CHROMOSOME

Chromosomes are composed of DNA and protein. Most are about 40% DNA and 60% protein. A significant amount of RNA is also associated with chromosomes, because these are the sites of RNA synthesis. The DNA of a chromosome is one very long, double stranded fiber that extends unbroken through the entire length of the chromosome. A typical human chromosome contains about 140 million (1.4×10^8) nucleotides in its DNA. The amount of information, one chromosome contains would fill about 280 printed books of 1000 pages each, if each nucleotide corresponds to a word and each page had about 500 words on it. Further more, if the strand of DNA from a single chromosome were laid out in a straight line, it would be about 5 centimeter long. Fitting such a strand into a small space of nucleus is nature's marvel - and that's only 1 of 46 chromosomes. In the cell, however, the DNA is coiled allowing it to fit into a much smaller space than would otherwise be possible.

How can this long DNA fibre coil so tightly? If we gently disrupt a eukaryotic nucleus and examine the DNA with an electron microscope, we find that it resembles a string of beads (Fig 20.5). Every 200 nucleotides, the DNA duplex is coiled around a core of eight histone proteins forming a complex known as a nucleosome. Unlike most proteins, which have an overall negative charge, histones are positively charged due to an abundance of the basic amino acids arginine and lysine. They are thus strongly attracted to the negatively charged phosphate groups of the DNA. The histone cores thus act as magnetic forms that promote and guide the coiling of the DNA. Further coiling occurs when the string of nucleosomes wraps up into higher order coils called supercoils.

Highly condensed portions of the chromatin are called **heterochromatin**. Some of these portions remain permanently condensed, so that their DNA is never expressed. The remainder of the chromosome called **euchromatin** is condensed only during cell division,

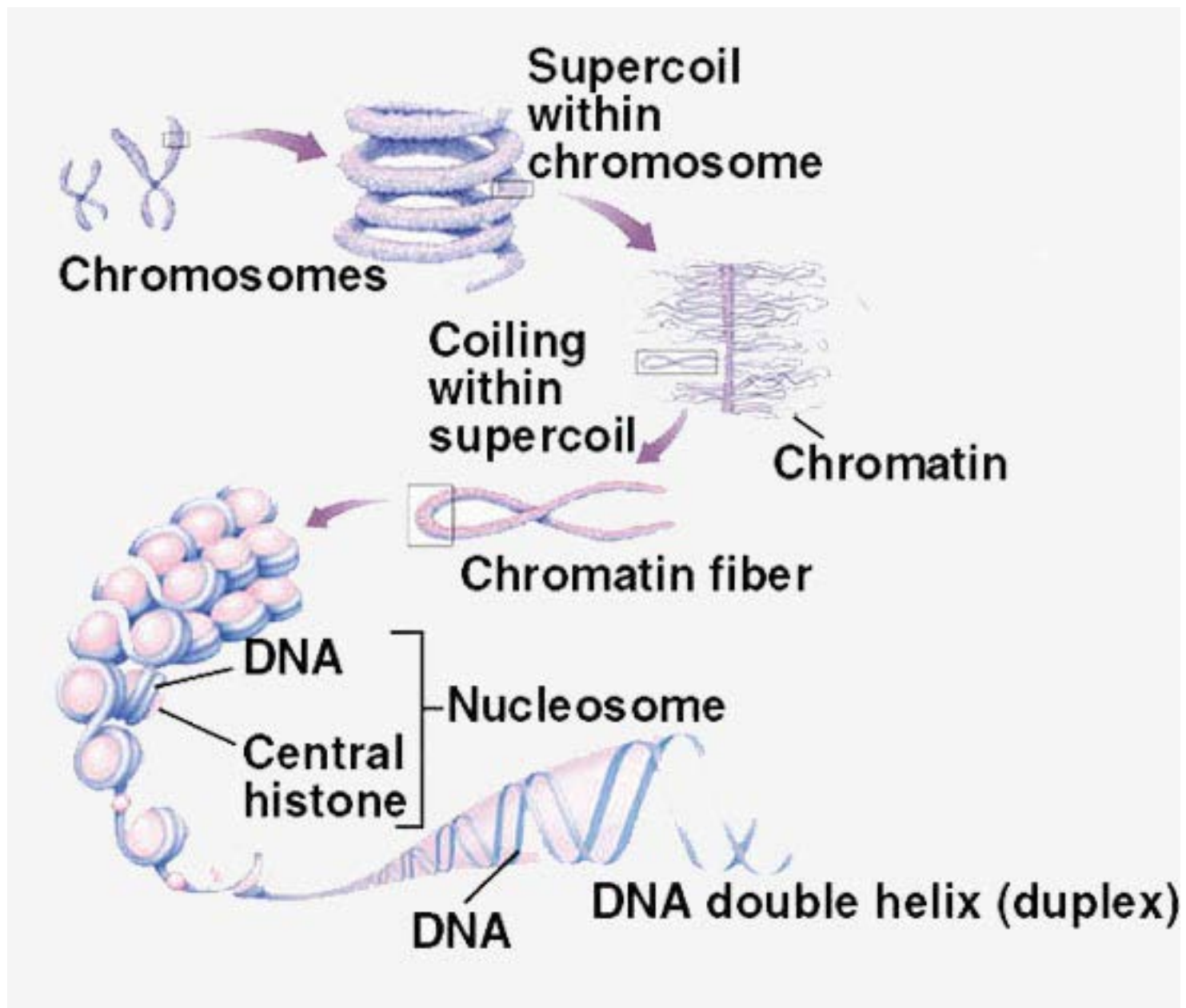


Fig. 20.5 Levels of eukaryotic chromosomal organization

when compact packaging facilitates the movement of the chromosomes. At all other times, euchromatin is present in an open configuration and its genes can be expressed. The way, chromatin is packaged when the cell is not dividing is not well understood beyond the level of nucleosomes and is a topic of intensive research.

THE CHROMOSOMAL THEORY OF INHERITANCE

A central role for chromosomes in heredity was first suggested in 1900 by the German geneticist Karl Correns, in one of the papers announcing the rediscovery of Mendel's work. Soon after, observations that similar chromosomes paired with one another during meiosis led directly to the chromosomal theory of inheritance, first formulated by the American Walter Sutton in 1902.

Several pieces of evidence supported Sutton's theory. One was that reproduction involves the initial union of only two cells, egg and sperm. If Mendel's model was correct, then these two gametes must make equal hereditary contributions. Sperm, however, contain little cytoplasm, suggesting that the hereditary material must reside within the nuclei of the gametes. Furthermore, while diploid individuals have two copies of each pair of homologous chromosomes, gametes have only one. This observation was consistent with Mendel's model, in which diploid individuals have two copies of each heritable gene and gametes have one. Finally, chromosomes segregate during meiosis, and each pair of homologous chromosomes orients on the metaphase plate independently of every other pair.

There is however one problem with this theory. If Mendelian characters are determined by genes located on the chromosomes, and if the independent assortment of Mendelian traits reflects the independent assortment of chromosomes in meiosis, why does the number of characters that assort independently in a given kind of organism often greatly exceed the number of chromosome pairs the organism possesses? This has led many early researchers to have serious reservations about Sutton's theory.

In 1910 Thomas Hunt Morgan, studying the fruit fly, *Drosophila melanogaster*, detected a mutant male fly, one that differed strikingly from normal flies of the same species its eyes were white instead of red.

Morgan crossed mutant male to a normal female. All F_1 progeny had red eyes. He then crossed red eyed flies from F_1 generation with each other. Of the 4252 F_2 progeny Morgan examined, 782 (18%) had white eyes. Although the ratio of red eyes to white eyes in the F_2 progeny was greater than 3:1, the results of the cross nevertheless provided clear evidence that eye colour segregates.

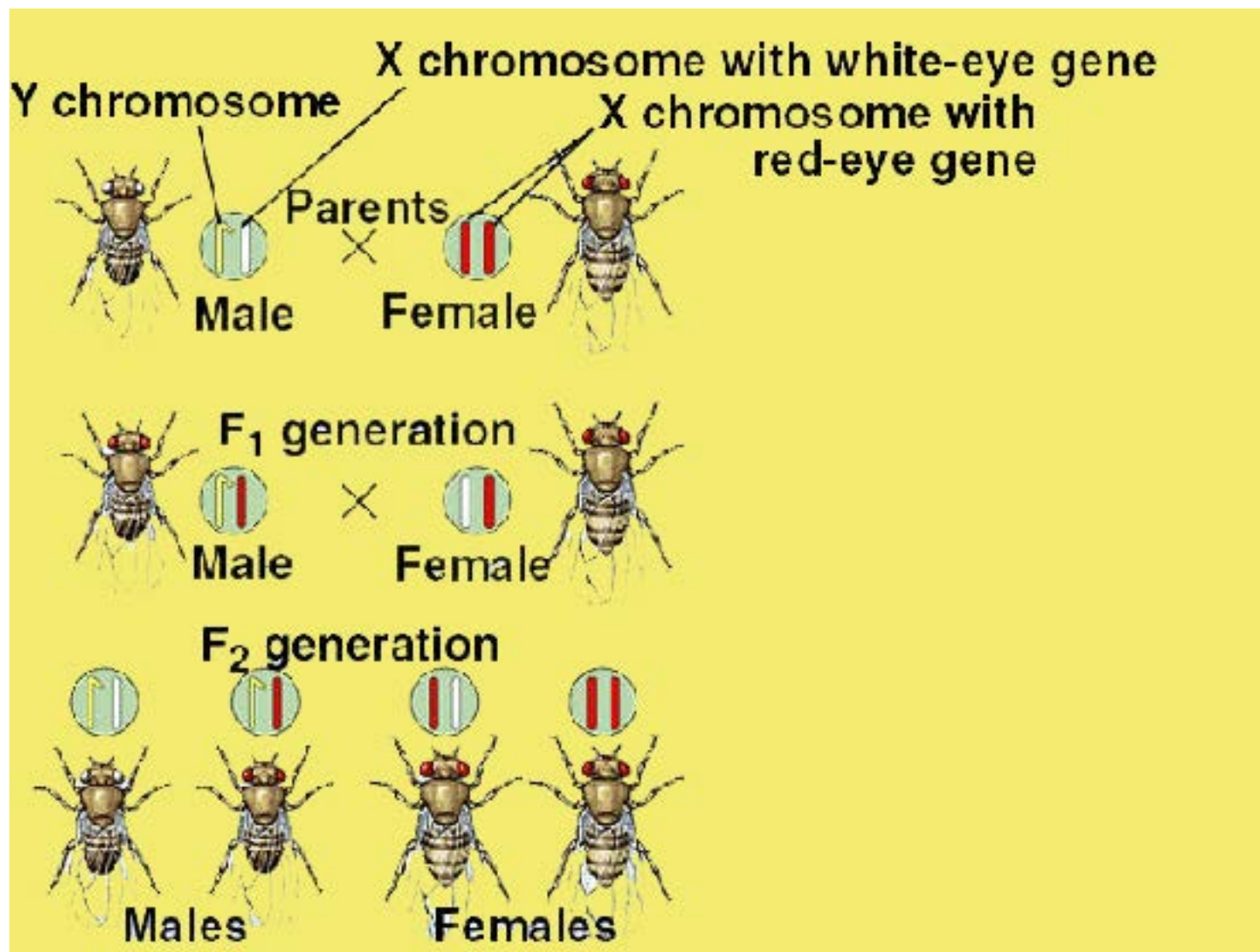


Fig. 2Q.6 Morgan's experiment demonstrating the chromosomal basis of Sex linkage.

However, there was something about the outcome that was strange and totally unpredicted by Mendel's theory - all of the white eyed F₂ flies were male!

How could this result be explained? Perhaps it was impossible for a white eyed female fly to exist; such individuals might not be viable for some unknown reason. To test this idea, Morgan test crossed the female F_1 progeny with the original white eyed male. He obtained both white-eyed and red-eyed males and females in a 1:1:1:1 ratio, just as Mendelian theory predicted. Hence a female could have white eyes. Why, then were there no white eyed females among the progeny of the original cross?

The solution to this puzzle involved sex. The gene causing the white eye trait in *Drosophila* resides only on the X chromosome. It is absent from the Y chromosome. A trait determined by a gene on the X chromosome is said to be sex linked. Knowing that the white eye trait is recessive to the red eye trait, the Morgan's result was a natural consequence of the Mendelian assortment of chromosomes (Fig 20.6).

Morgan's experiment was one of the most important in the history of genetics because it presented the first clear evidence that the genes determining Mendelian traits do indeed reside on the chromosomes, as Sutton had proposed. The chromosome theory of inheritance therefore, propounds that genes are located on chromosomes. The segregation of the white-eye trait has one-to-one correspondence with the segregation of the X chromosome, in other words, Mendelian traits such as eye colour in *Drosophila* assort independently because chromosomes do.

DNA AS HEREDITARY MATERIAL

The first evidence of hereditary nature of DNA was provided by a British microbiologist Frederick Griffith who made some unexpected observations while experimenting with pathogenic bacteria. When he infected mice with a virulent strain of *Streptococcus pneumoniae* bacteria (then known as *Pneumococcus*), the mice died of blood poisoning. However, when he infected similar mice with a mutant strain of *S. pneumoniae* that lacked the virulent strains polysaccharide coat, the mice showed no ill effects. The coat was apparently necessary for virulence.

The normal pathogenic form of this bacterium is referred to as the S form because it forms smooth colonies on a culture dish. The mutant forms, which lacks an enzyme needed to manufacture the polysaccharide coat, is called the R form because it forms rough colonies.

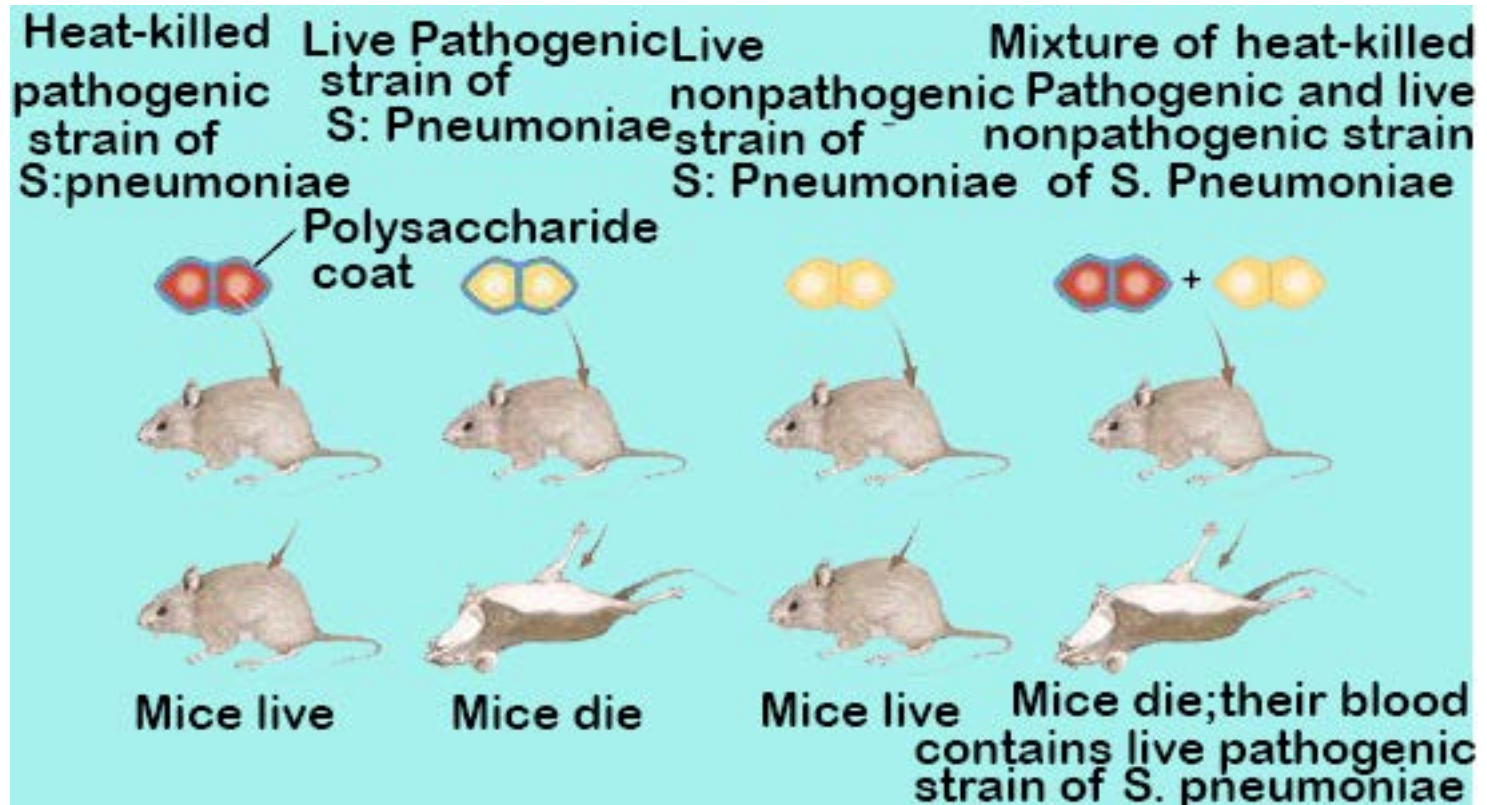


Fig 20.7 Griffith's discovery of transformation

To determine whether the polysaccharide coat itself had a toxic effect. Griffith injected dead bacteria of the virulent S strain into the mice; the mice remained perfectly healthy. As a control, he injected mice with a mixture containing dead S bacteria of the virulent strain and live coatless R bacteria, each of which by itself did not harm the mice (Fig 20.7). Unexpectedly, the mice developed the disease symptoms and many of them died. The blood of the dead mice was found to contain high levels of live, virulent streptococcus type S bacteria, which had surface proteins characteristic of the live (previously R) strain. Somehow, the information specifying the polysaccharide coat had passed from the dead, virulent S bacteria to the live, coatless R bacteria in the mixture, permanently transforming the coatless R bacteria into the virulent S variety. Transformation is the transfer of genetic material from one cell to another and can alter the genetic make up of the recipient cell.

The agent responsible for transforming *Streptococcus* went undiscovered until 1944. In a classic series of experiments, Oswald Avery along with Colin Macleod and Maclyn McCarty characterized what they referred to as the “Transforming principle”. They first prepared mixture of dead S *Streptococcus* and live R *Streptococcus* that Griffith had used. Then they removed as much of the protein as they could from their preparation, eventually achieving 99.98% purity. Despite removal of nearly all the protein, the transforming activity was not reduced. Moreover, the properties of transforming principle resembled those of DNA. The protein digesting enzymes or RNA digesting enzymes did not affect the principle’s activity, but the DNA digesting enzyme DNase destroyed all the transforming activity.

Additional evidence supporting Avery’s conclusion was provided in 1952 by Alfred Hershey and Martha Chase who experimented with bacteriophages T2. In some experiments they labelled viruses with radio isotope ^{32}P , which was incorporated into the newly synthesized DNA of growing phage. In other experiments, the viruses were grown on a medium containing ^{35}S , an isotope of sulphur which is incorporated into the amino acids of newly synthesized protein coats.

After the labelled viruses were permitted to infect bacteria, the bacterial cells were agitated violently to remove the protein coats of the infecting viruses from the surfaces of the bacteria. This procedure removed nearly all of the ^{35}S label from the bacteria. However the ^{32}P label had transferred to the interior of the bacteria (Fig 20.8) and was found in viruses subsequently released from the infected bacteria. Hence, the hereditary information injected into the bacteria that specified the new generation of viruses was DNA and not protein.

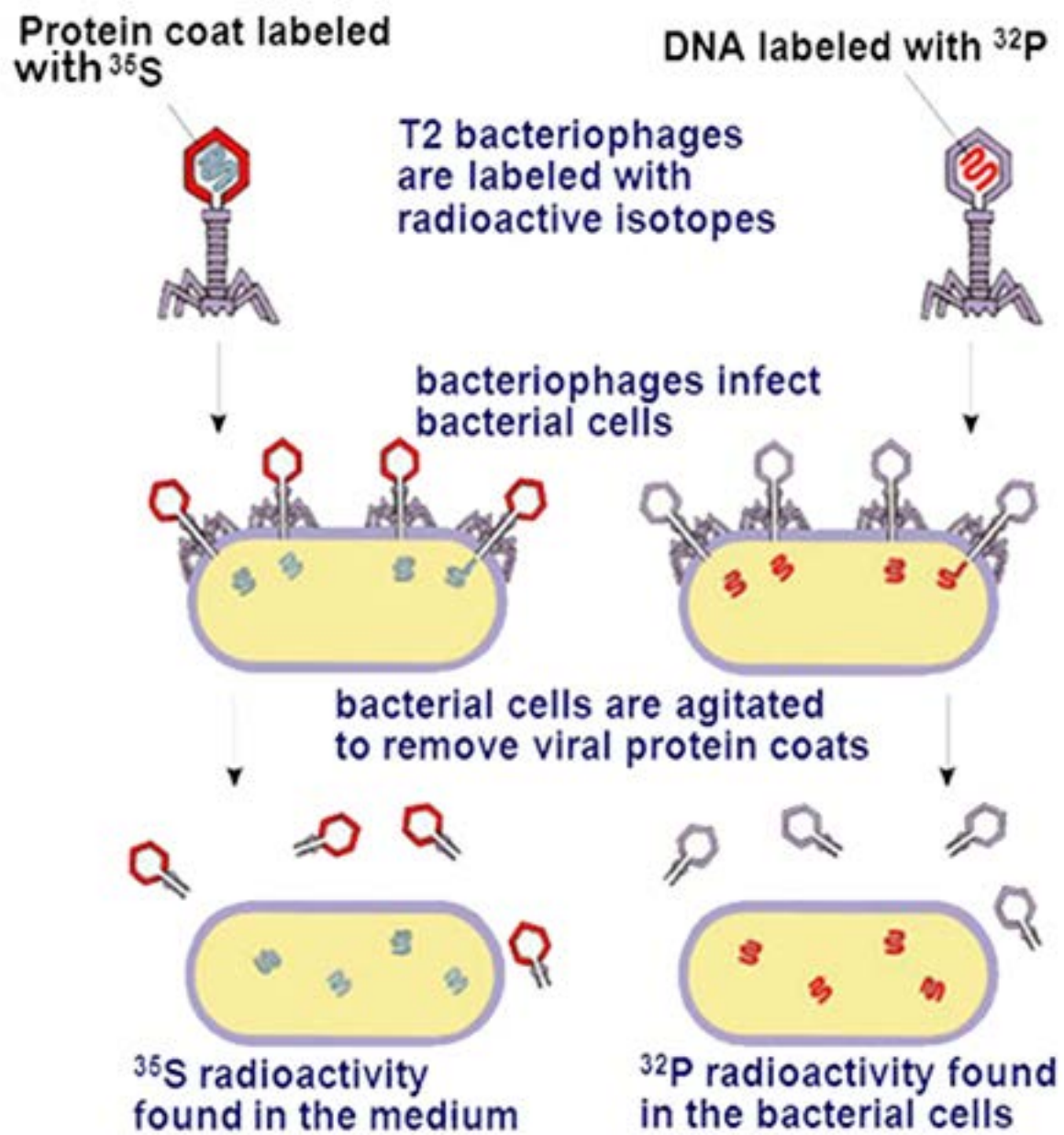


Fig 20.8 the Hershey and chase experiment

Chemical Nature of DNA

A German Chemist, Friedrich Miescher, discovered DNA in 1869, only four years after Mendel's work was published. Miescher extracted a white substance from the nuclei of human cells and fish sperm. He called this substance "nuclein" because it seemed to be specifically associated with the nucleus.

Since nuclein was acidic, it came to be known as nucleic acid. For 50 years biologists did little research on the substance, because nothing was known of its function in cells. In 1920's, the basic structure of nucleic acids was determined by the biochemist P.A. Levene, who found that DNA contains three main components (Fig 20.9): (1) phosphate (P₀₄) groups, (2) five carbon sugars, and (3) nitrogen containing bases called purines (adenine, A, and guanine, G) and pyrimidines (thymine, T and cytosine, C, RNA contains uracil, U instead of T). Levene concluded that DNA and RNA molecules are made of repeating units called nucleotides. In a nucleotide nitrogen base is attached to carbon number 1 of a pentose sugar and phosphate group is attached to carbon number 5 of the sugar. In addition a free hydroxyl (-OH) group is attached to the 3' carbon atom (Fig 20.10). The 5' phosphate and 3' hydroxyl groups allow DNA and RNA to form long chains of nucleotides, because these two groups can react chemically with each other. The reaction between the phosphate group of one nucleotide and the hydroxyl group of another is a dehydration synthesis, eliminating a water molecule and forming a covalent bond that links the two groups (Fig 20.11). The linkage is called a phosphodiester bond because the phosphate group is now linked to the two sugars by means of a pair of ester (P-O-C) bonds. The two unit polymer resulting from this reaction still has a free 5' phosphate group at one end and a free 3' hydroxyl group at the other, so that it can link to other nucleotides. In this way, many thousands of nucleotides can join together in long chains. Linear strands of DNA or RNA no matter how long, will almost always have a free 5' phosphate group at one end and a free 3' hydroxyl group at the other.

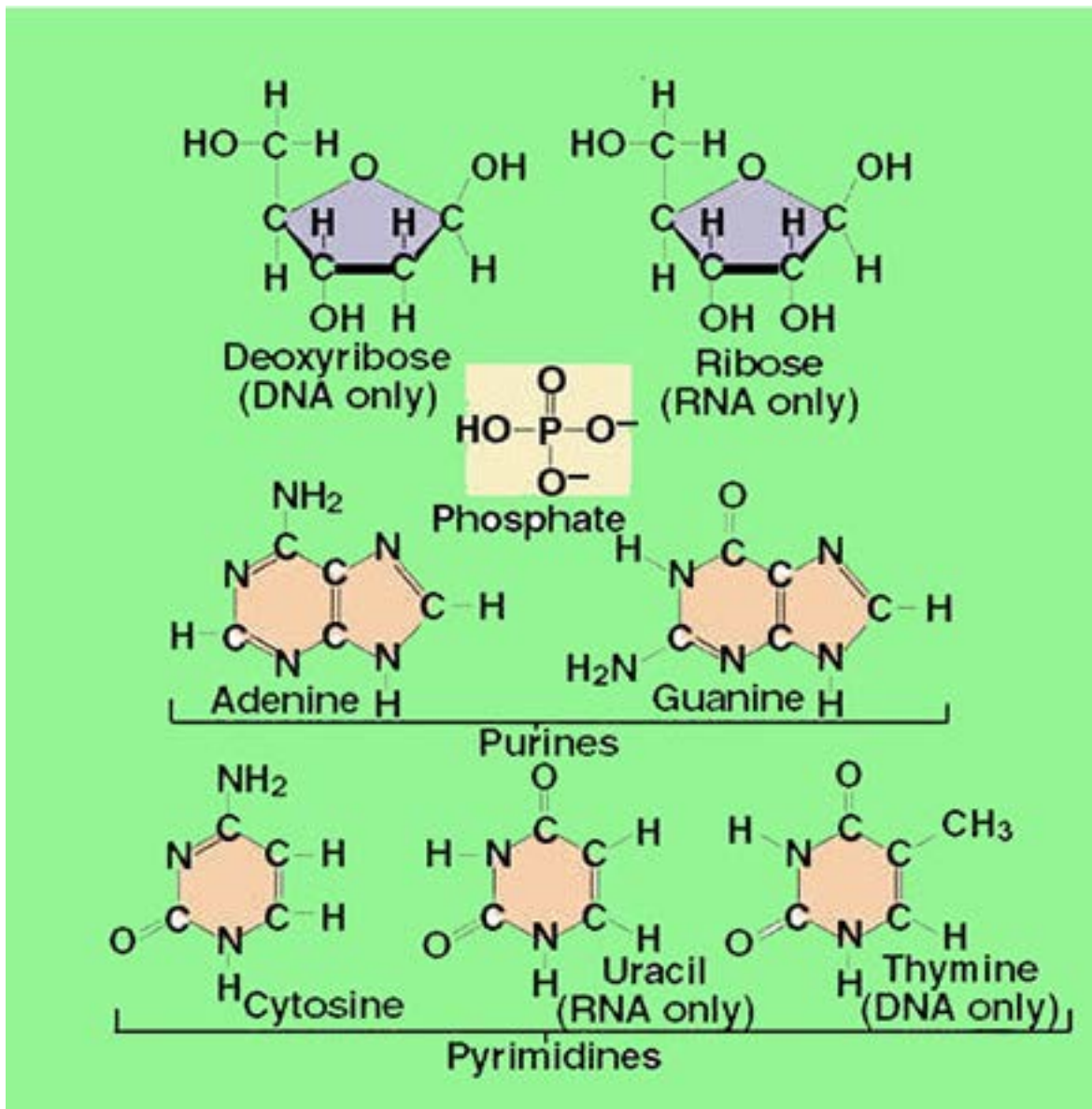


Fig 20.9 Nucleotide subunits of DNA and RNA

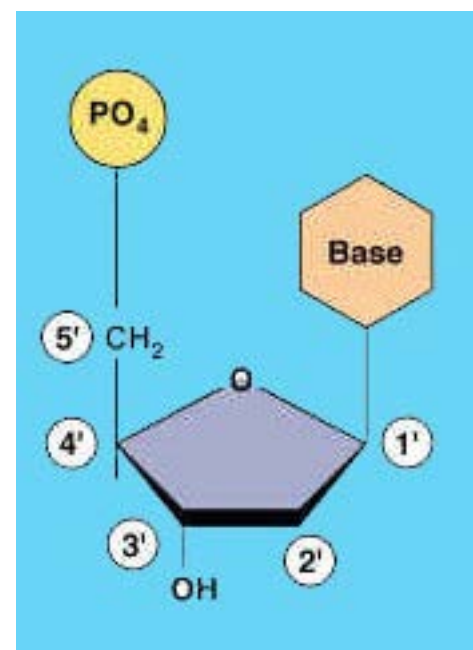


Fig 20.10 Numbering the carbon atoms in a nucleotide

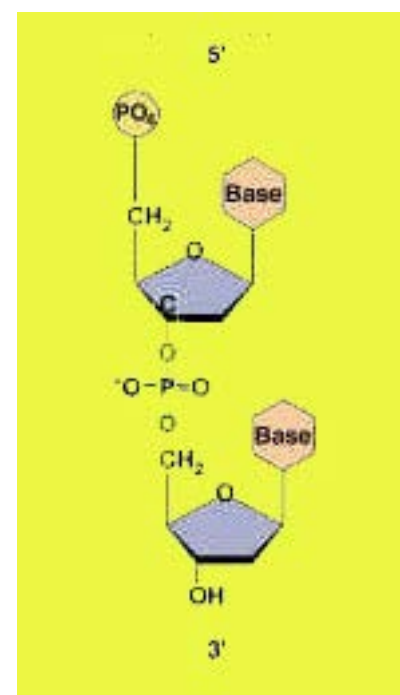


Fig 20.11 A phosphodiester bond.

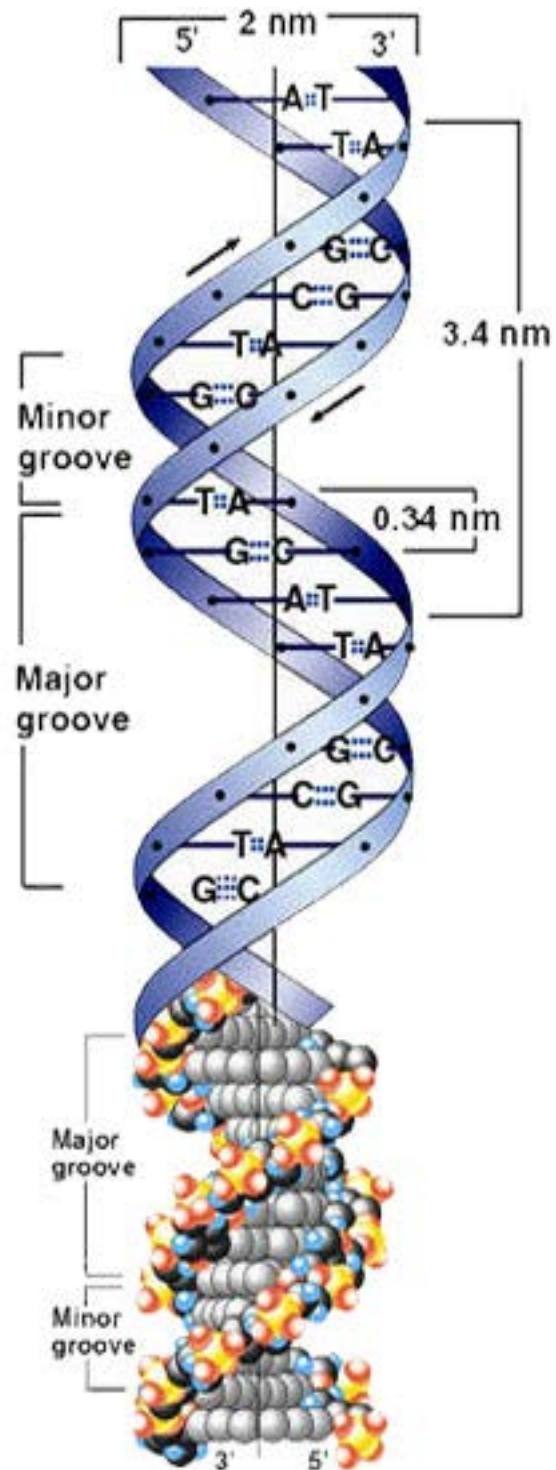
Erwin Chargaff later showed that the amount of adenine in DNA always equals the amount of thymine, and the amount of guanine always equals the amount of cytosine. It also implies that there is always equal proportion of purine (A+G) and pyrimidine (C+T).

The significance of the regularities pointed out by Chargaff became obvious when a British chemist Rosalind Franklin carried on an X-ray diffraction analysis of DNA. In this analysis, a molecule is bombarded with a beam of X-rays. When individual rays encounter atoms their path is bent or diffracted and the diffraction pattern is recorded on the photographic film. When carefully analyzed this pattern gives three dimensional structure of a molecule.

Rosalind Franklin prepared this X-ray diffraction pattern of DNA in the laboratory of British Biochemist Maurice Wilkins, who prepared DNA fibers. The diffraction pattern suggested that the DNA molecule had a shape of a helix with a diameter of 2 nm and a complete helical turn every 3.4 nm (Fig 20.12).

Double Helical Structure of DNA (Watson and Crick's Model)

Learning informally of Franklin's results, before they were published in 1953, James Watson and Francis Crick, two young researchers in University of Cambridge, quickly worked out a likely structure of the DNA molecule (Fig 20.12) which we now know was substantially correct. They proposed that molecule is a simple double helix, with the basis of two strands pointed inward toward each other, forming base-pairs. In their model, base pairs always consist of purines, which are large, pointing toward pyrimidines which are small, keeping the diameter of the molecule a constant 2 nm. Because hydrogen bonds exist between the bases in a base pair, the double helix is stabilized as a duplex DNA molecule composed of two antiparallel strands, one chain running 3' to 5' and the other 5' to 3'. The base pairs are planar (flat) and stack 0.34nm apart as a result of hydrophobic interactions contributing to the overall stability of the molecule (Fig. 20.3). In the double helix, adenine forms two hydrogen bonds with thymine, while guanine forms three hydrogen bonds with cytosine. Adenine will not form proper hydrogen bonds with cytosine, and guanine will not form hydrogen bonds with thymine. Consequently adenine and thymine will always occur in the same proportion in any DNA molecule, as well guanine and cytosine, because of this base pairing.



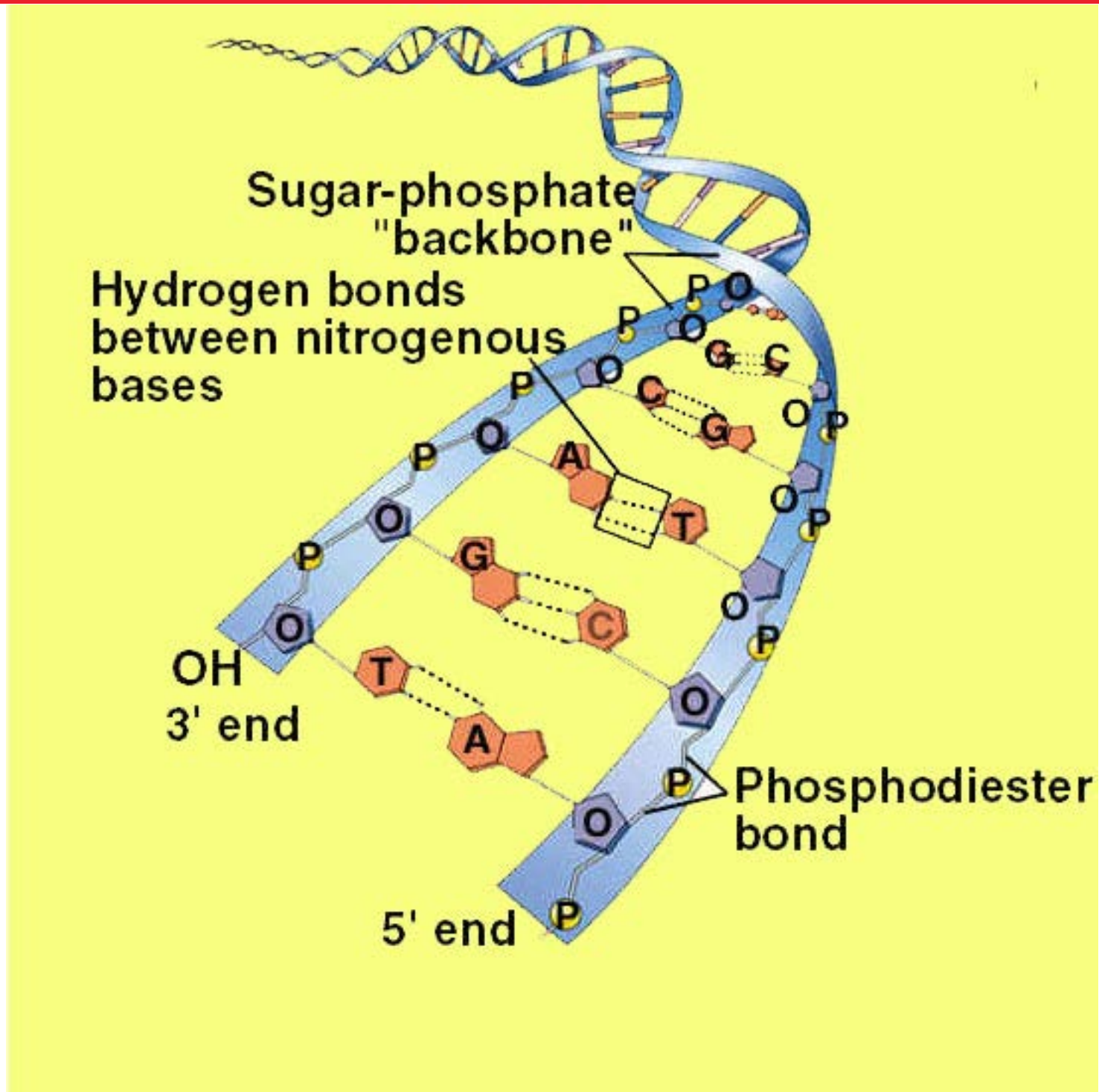


Fig 20.13 DNA is a double helix

DNA Replication

The Watson - Crick model immediately suggested that the basis for copying the genetic information is complementarity. If one were to unzip the molecule, one would need only to assemble the appropriate complementary nucleotides on the exposed single strands to form two daughter complexes with the same sequences. This form of DNA replication is called **semi-conservative**, because while the sequence of the original duplex is conserved after one round of replication, the duplex itself is not. Instead, each strand of the duplex becomes part of another duplex. In semi-conservative replication, the two strands of the duplex separate out each acting as a model or mold, along which new nucleotides are arranged thus giving rise to two new duplexes. In this process by separation of two strands, primary structure has been conserved, whereas the secondary structure has been disrupted.

The other hypotheses of DNA replication were also proposed. The conservative model stated that the parental double helix would remain intact and generate DNA copies consisting of entirely new molecules. The dispersive model predicted that parental DNA would become completely dispersed and that each strand of all the daughter molecules would be a mixture of old and new DNA.

Animation 20.3: DNA Replication
Source & Credit: weloveteaching

The Meselson - Stahl Experiment

The three hypothesis of DNA replication were evaluated by Mathew Meselson and Franklin Stahl of the California Institute of Technology in 1958. They grew bacteria in a medium containing heavy isotope of nitrogen, ^{15}N , which became incorporated into the bases of the bacterial DNA. After several generations, the DNA of these bacteria was denser than that of bacteria grown in a medium containing the lighter isotope of nitrogen, N^{14} . Meselson and Stahl then transferred the bacteria from the N^{15} medium to the N^{14} medium and collected the DNA at various intervals.

They dissolved the DNA in cesium chloride and then spun it at a very high speed in an ultra-centrifuge. DNA strands of different densities got separated. The enormous centrifugal forces generated by the ultracentrifuge caused the cesium ions to migrate toward the bottom of the centrifuge tube, creating a gradient of CsCl , and thus of density. Each DNA floats or sinks in the gradient until it reaches the position where its density exactly matches the density of cesium there. Because N^{15} strands are denser than N^{14} strands, they migrate farther down the tubes to a denser region of the cesium chloride gradient.

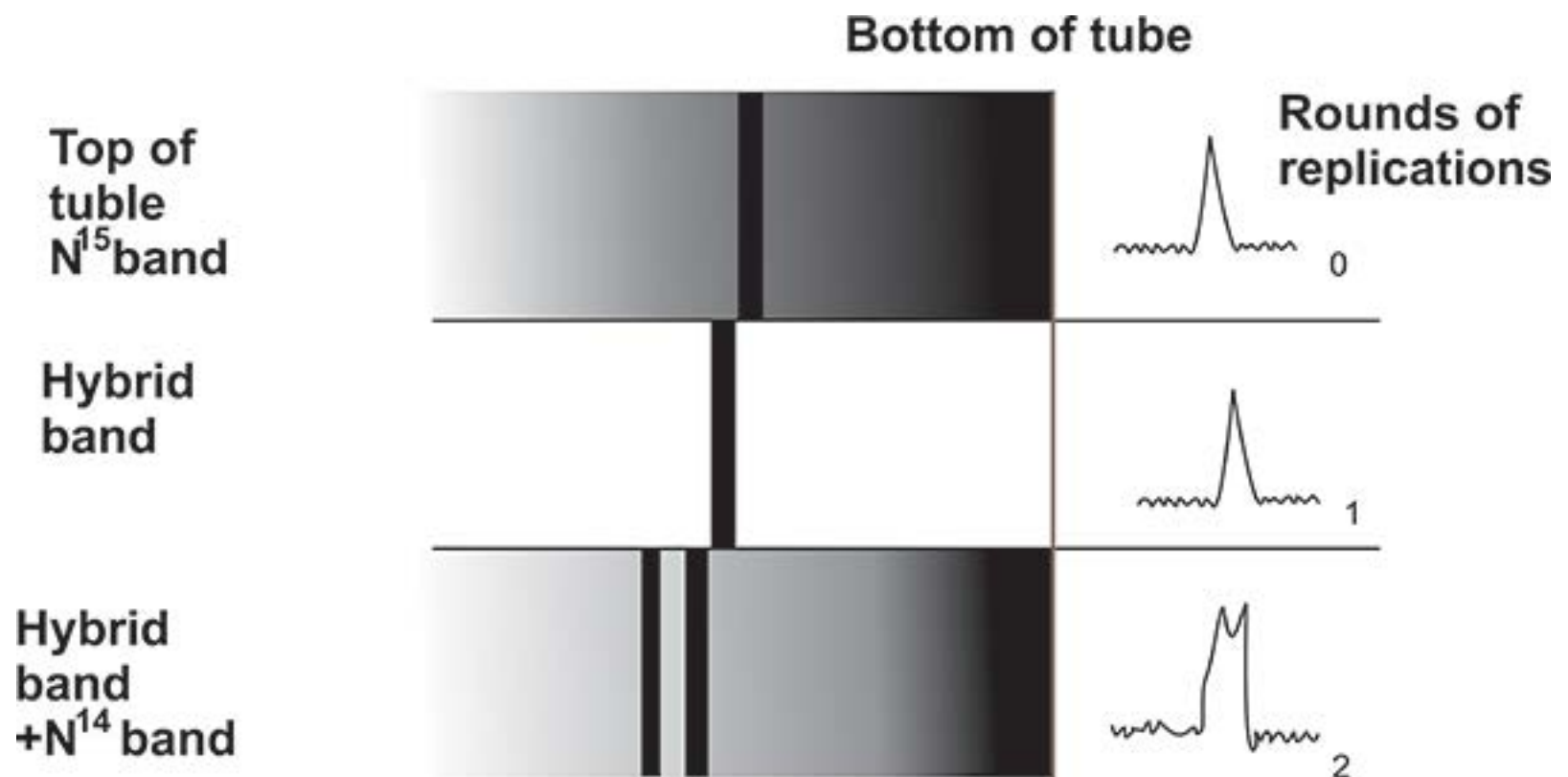


Fig. 20.14 The key result of the Meselson and Stahl experiment. The bands on the left side of the figure show N^{15} DNA which is heavier and is present towards the bottom of the tube. The middle band is a hybrid DNA band of N^{15} and N^{14} and hence lies above the N^{15} band. This is after first round of replication. In the second round of replication, two bands are visible one at the level of hybrid band and the other lighter band which is N^{14} band.

The DNA collected immediately after the transfer, was all dense. However, after the bacteria completed their first round of DNA replication in the N^{14} medium, the density of their DNA had decreased to a value intermediate between N^{15} - DNA and N^{14} - DNA. After the second round of replication, two density classes of DNA were observed one intermediate and one equal to that of N^{14} - DNA (Fig 20.14).

Meselson and Stahl interpreted their results as follows: after the first round of replication, each daughter DNA duplex was a hybrid possessing one of the heavy strands of parent molecule and one light strand. When this hybrid duplex replicated, it contributed one heavy strand to form another hybrid duplex and one light strand to form a light duplex (Fig 20.15). Thus, this experiment clearly confirmed the prediction of the Watson-Crick model that DNA replicates in a semi-conservative manner.

*Animation 20.3: Meselson-Stahl Experiment
Source & Credit: weloveteaching*

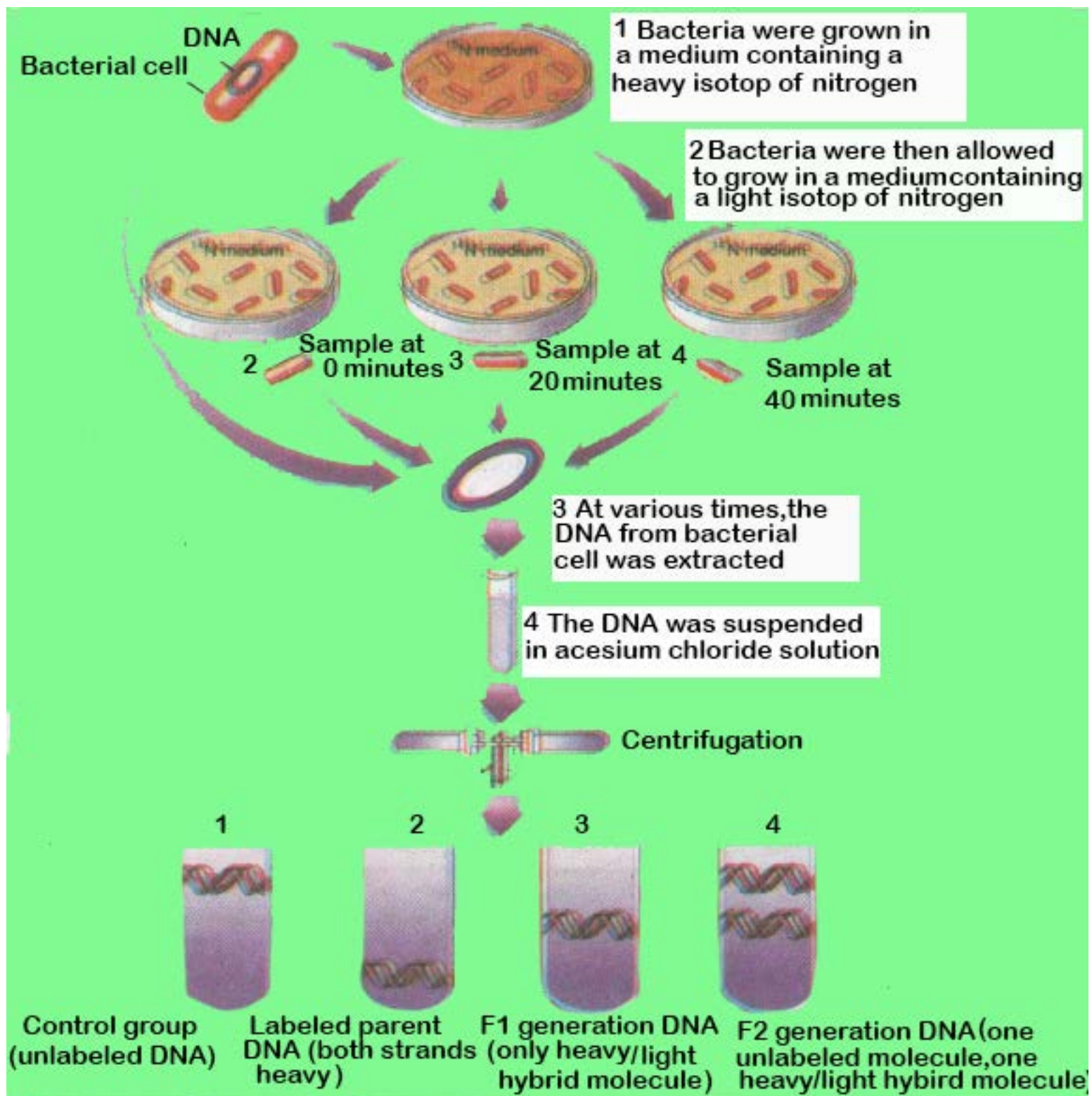


Fig 20.15 The Meselson and Stahl experiment : evidence demonstrating semi-conservative replication

The Replication Process

The DNA replication begins at one or more sites on the DNA molecule, where there is a specific sequence of nucleotides (Fig 20.16), The DNA polymerase III and other enzymes begin a complex process that catalyzes the addition of nucleotides to the growing complementary strands of DNA (Fig 20.17).

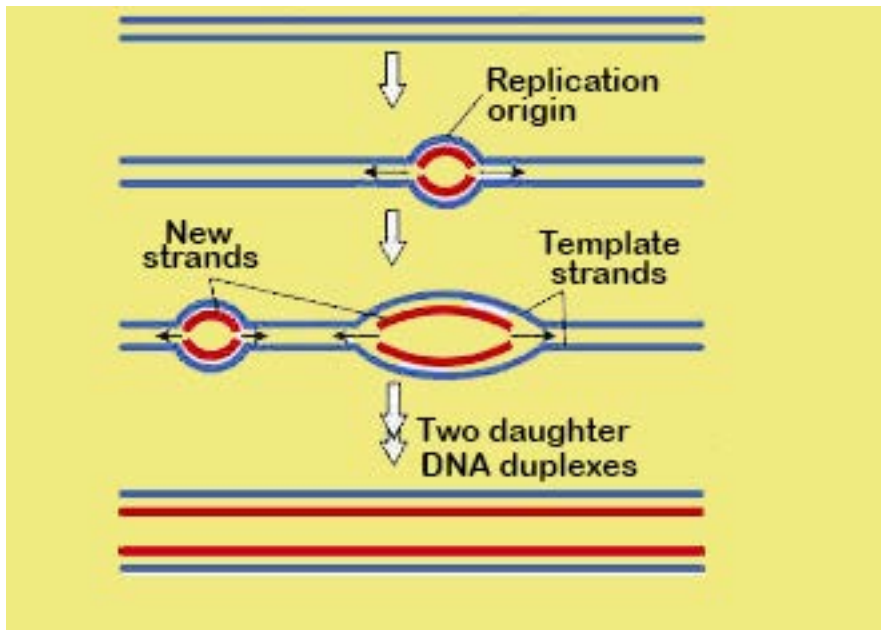


Fig 20.16 Origins of replication

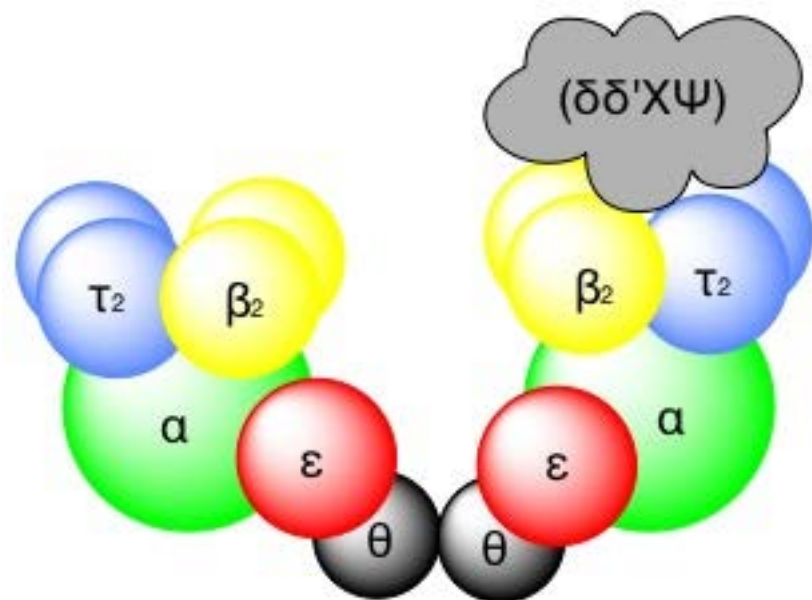


Fig 20.18 Molecular structure of DNA polymerase III complex

There are three DNA polymerases namely I, II and III in bacteria. DNA polymerase I is a relatively small enzyme that plays a supporting role in DNA replication. The true E.coli replicating enzymes is DNA polymerase III which is 10 times larger and far more complex in structure (Fig 20.18). The enzyme is a dimer and catalyzes replication of one DNA strand. Polymerase III progressively threads the DNA through the enzyme complex, moving at a rapid rate, some 1000 nucleotides / second. One of the features of the DNA polymerase III is that it can add nucleotides only to a chain of nucleotides that is already paired with the parent strands. Hence DNA polymerase cannot initiate synthesis on its own. Instead another enzyme, primase, constructs an RNA primer, a sequence of about 10 RNA nucleotides complementary to the parent DNA template. DNA polymerase III recognizes the primer and adds DNA nucleotides to it to construct the DNA strands. The RNA nucleotides in the primers are then replaced by DNA nucleotides.

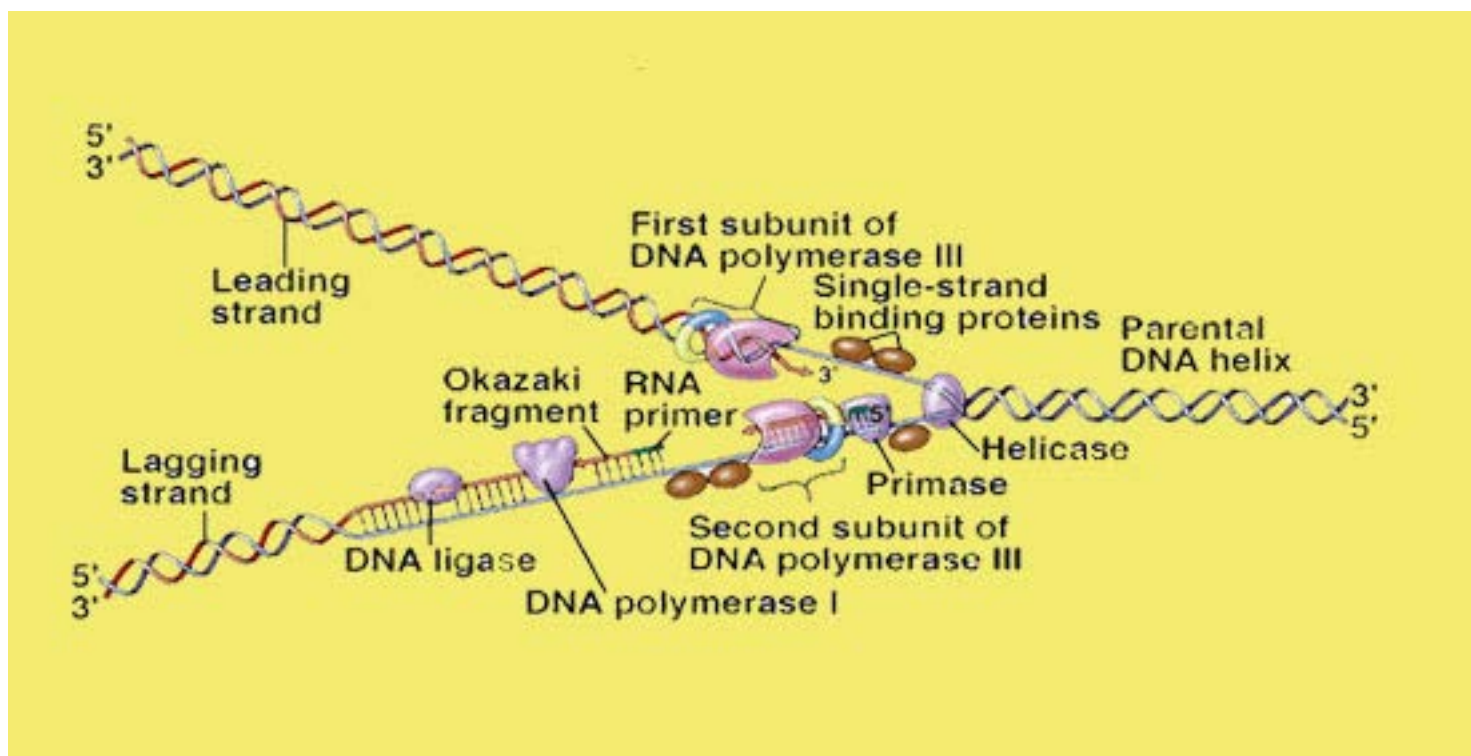


Fig 20.19 A DNA replication fork

Another feature of DNA polymerase III is that it can add nucleotides only to the 3' end of a DNA strand. This means that replication always proceeds 5' → 3' direction on a growing DNA strand. Because the two parent strands of a DNA molecules are antiparallel, the new strands are oriented in opposite directions (Fig 20.19). Therefore, the new strands must be elongated by different mechanisms. Leading strand, which elongates toward the replication fork, is built up simply by adding nucleotides continuously to its growing 3' end. In contrast the lagging strand which elongates away from the replication fork, is synthesized discontinuously as a series of short segments that are later connected. These segments, called Okazaki fragments are about 100 - 200 nucleotides long in eukaryotes and 1000 - 2000 nucleotides long in prokaryotes. Each Okazaki fragment is synthesized by DNA polymerase III in 5' → 3' direction, beginning at the replication fork and moving away from it. When the polymerase reaches the 5' end of the lagging strand, another enzyme, DNA ligase, attaches the fragment to the lagging strand. The DNA is further unwound, new RNA primers are constructed, and DNA polymerase III then jumps ahead 1000 - 2000 nucleotides (toward the replication fork) to begin constructing another Okazaki fragment.

WHAT IS A GENE?

Archibald Garrod and William Bateson concluded in 1902 that certain diseases among their patients were more prevalent in particular families. By examining several generations of these families, Garrod found that some of the diseases behaved as if they were the product of simple recessive alleles. He concluded that these disorders were Mendelian traits and that they had resulted from changes in the hereditary information in an ancestor of the affected families.

Garrod investigated several of these disorders in detail. In alkaptonuria the patients produced urine that contained homogentisic acid. This substance oxidized rapidly when exposed to air, turning the urine black. In normal individuals, homogentisic acid is broken down into simpler substances. With considerable insight Garrod concluded that patients suffering from alkaptonuria lacked the enzyme necessary to catalyze this breakdown. He speculated that many other inherited diseases might also reflect enzyme deficiencies.

From Garrod's finding, it could be inferred that the information encoded within the DNA of chromosomes acts to specify particular enzymes. This point was not actually established, however, until 1941, when a series of experiments by Stanford University geneticists George Beadle and Edward Tatum provided definitive evidence on this point. Beadle and Tatum deliberately set out to create Mendelian mutations in chromosomes and then studied the effect of these mutations on the organisms (Fig 20.20).

Beadle and Tatum exposed *Neurospora* spores to X-rays, expecting that DNA in some of these spores would experience damage in the regions encoding the ability to make compounds needed for normal growth (Fig 20.20). DNA changes of this kind are called mutations and the organisms that have undergone such changes are called mutants. Initially, they allowed the progeny of the irradiated spores to grow on a defined medium containing all of the nutrients necessary for growth, so that any growth deficient mutants resulting from the irradiation would be kept alive.

To determine whether any of the progeny of the irradiated spores had mutations causing metabolic deficiencies, Beadle and Tatum placed subcultures of individual fungal cells on a "minimal" medium that contained only sugar, ammonia, salts, a few vitamins and water. Cells that had lost the ability to make other compounds necessary for growth would not survive on such a medium. Using this approach, Beadle and Tatum succeeded in identifying and isolating many growth deficient mutants.

Next the researchers added various chemicals to the minimal medium in an attempt to find one that would enable a given mutant strain to grow. This procedure allowed them to pinpoint the nature of the biochemical deficiency that strain had. The addition of arginine, for example, permitted several mutant strains, dubbed arg mutants, to grow. When their chromosomal positions were located, the arg mutations were found to cluster in three areas.

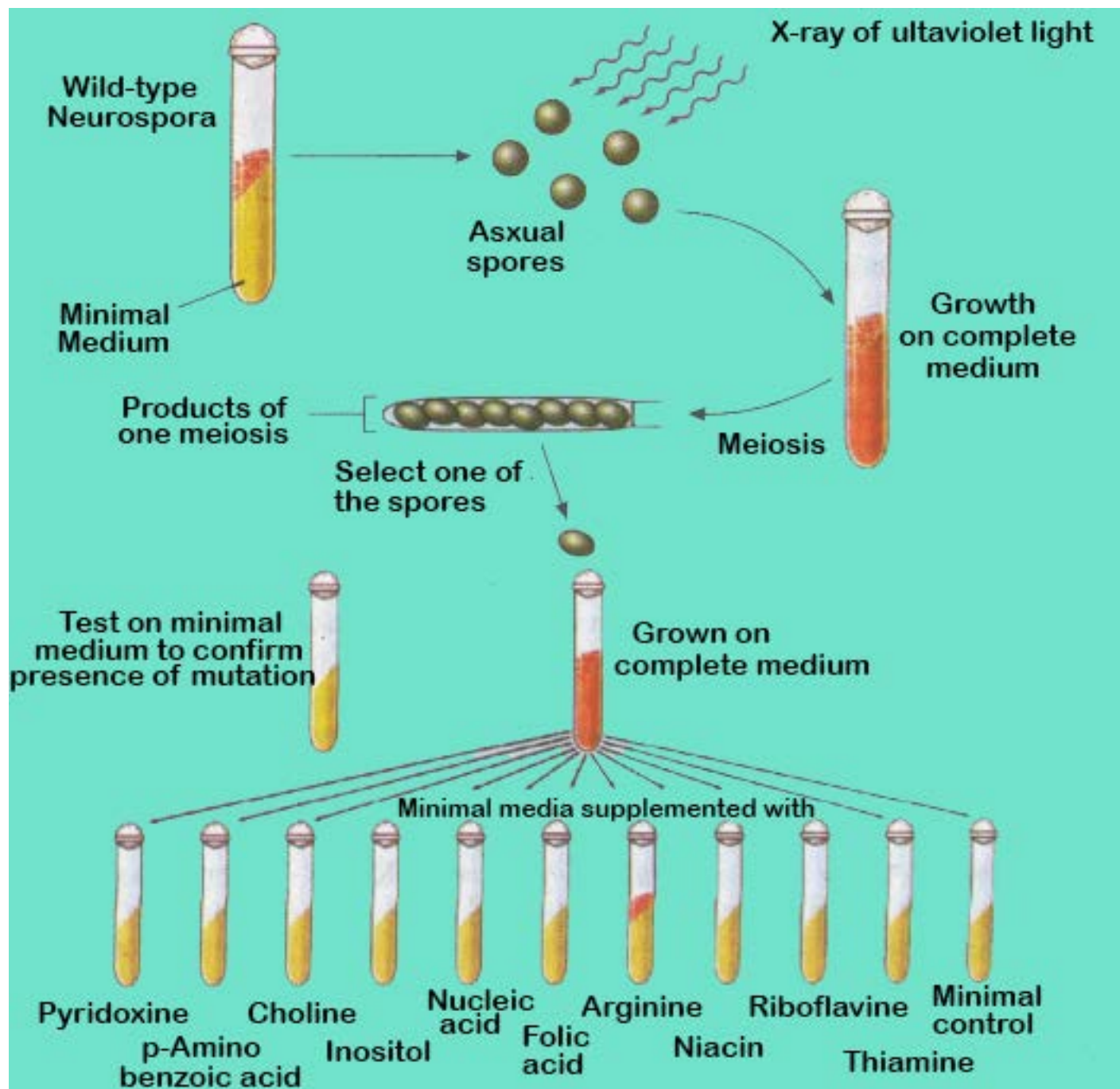


Fig 20.20 Beadle and Tatum's procedure for isolating nutritional mutants in *Neurospora*

One - gene / one - polypeptide Hypothesis

For each enzyme in the arginine biosynthetic pathway, Beadle and Tatum were able to isolate a mutant strain with a defective form of that enzyme, and the mutation was always located at one of a few specific chromosomal sites. Most importantly, they found there was a different site for each enzyme. Thus, each of the mutant they examined had a defect in a single enzyme, caused by a mutation at a single site on one chromosome. Beadle and Tatum concluded that genes produce their effects by specifying the structure of enzymes and that each gene encodes the structure of one enzyme (Fig 20-21). They called this relationship one - gene / one - enzyme hypothesis. Because many enzymes contain multiple protein or polypeptide subunits, each encoded by a separate gene, the hypothesis is today more commonly referred to as “one gene / one- polypeptide”.

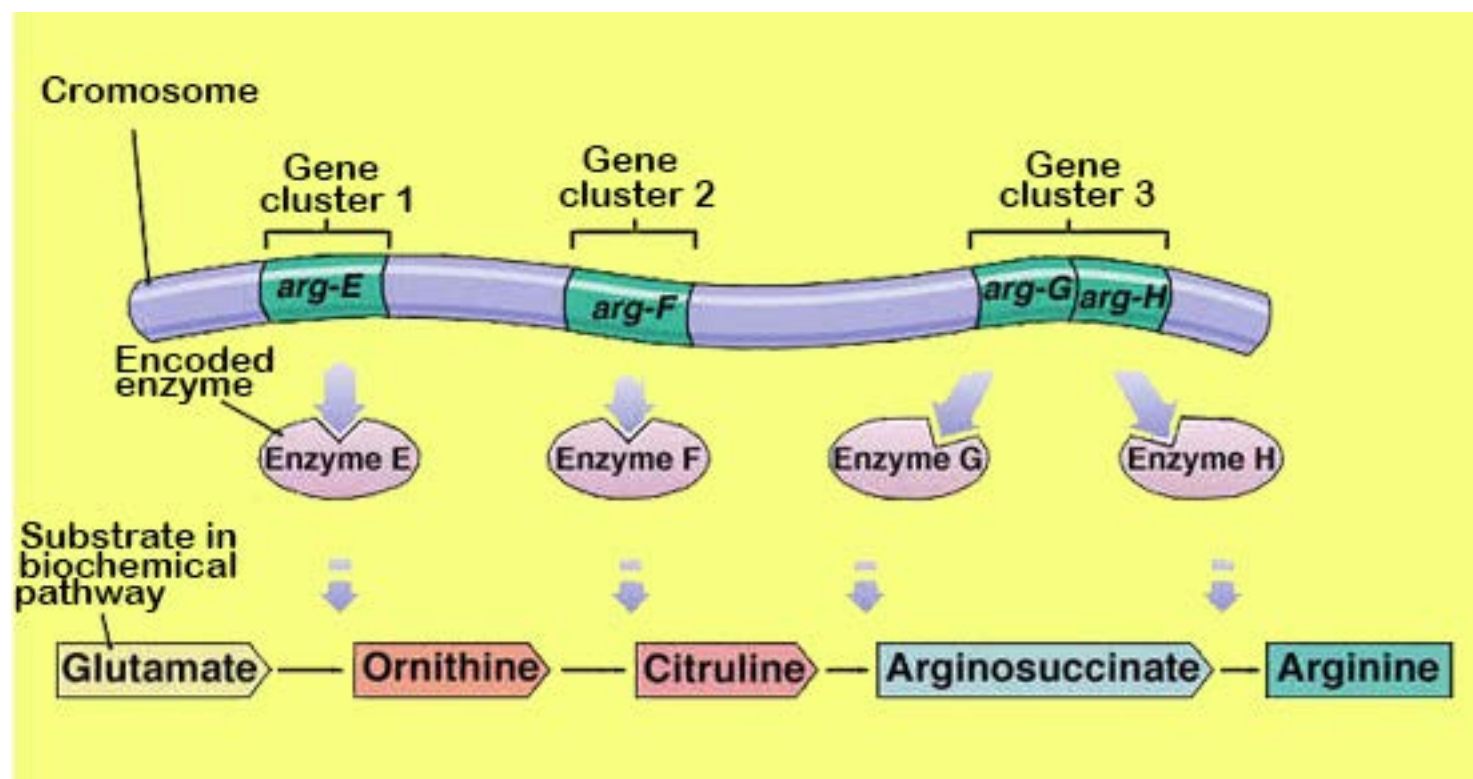


Fig. 20.21 Evidence for the “one-gene/one-polypeptide” hypothesis.

Enzymes are responsible for catalyzing the synthesis of all the parts of an organism. They are also responsible for the assembly of nucleic acids, proteins, carbohydrates and lipids. Therefore, by encoding the structure of enzymes and other proteins, DNA specifies the structure of the organism itself.

How DNA encodes protein structure?

In 1953, an English biochemist Frederick Sanger, described the complete sequence of amino acids of insulin. Sanger's achievement was significant, as it was demonstrated for the first time that proteins consisted of definable sequences of amino acids. Soon it was revealed that all enzymes and other proteins are strings of amino acids arranged in a certain definite order.

Following Sanger's pioneering work, Vernon Ingram in 1956 discovered the molecular basis of sickle cell anemia, a protein defect inherited as a Mendelian disorder. By analyzing the structure of normal and sickle cell haemoglobin, Ingram, working at Cambridge University, showed that sickle cell anemia is caused by a change from glutamic acid to valine at a single position in the protein (Fig 20.22). The alleles of the gene encoding hemoglobin differed only in their specification of this one amino acid in the hemoglobin amino acid chain.

These experiments and other related ones have finally brought us to a clear understanding of the unit of heredity. The characteristics of sickle cell anemia and most other hereditary traits are defined by changes in protein structure brought about by an alteration in the sequence of amino acids that make up the protein. This sequence in turn is dictated by the order of nucleotides in a particular region of chromosome. For example, the critical change leading to sickle cell disease is a mutation that replaces a single thymine with an adenine at the position that codes for glutamic acid converting the position to valine. The sequence of nucleotides that determines the amino acid sequence of a protein is called a gene.

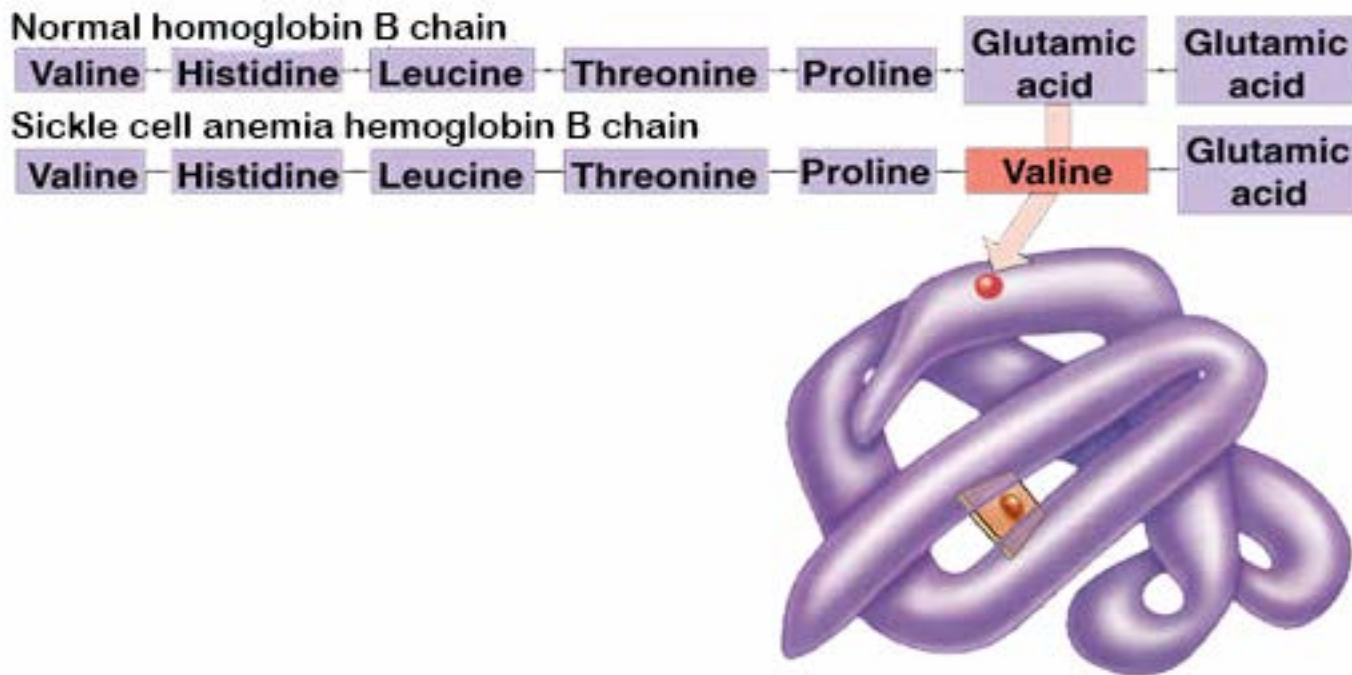


Fig 20.22 The molecular basis of a hereditary disease, sickle cell anemia

CELLS USE RNA TO MAKE PROTEIN

All organisms use the same basic mechanism of reading and expressing genes, which is often referred to as central dogma. The genetic information resides in DNA, which is also the main fountain head. The genetic information flows down into RNA, which is then converted into protein (Fig 20.23).

The first step of central dogma is the transfer of information from DNA to RNA, which occurs when an mRNA copy of the gene is produced. The process is called transcription. Transcription is initiated when the enzyme RNA polymerase binds to a particular binding site called a promoter located upstream of the gene. The enzyme then moves along the strand into the gene and mRNA is synthesized. At stop signal on the other end of gene, the enzyme disengages itself from the DNA and releases the newly assembled RNA chains. This chain is a complementary transcript of the gene from which it was copied.

The second step of the central dogma is the transfer of information from RNA to proteins, which occurs when the information contained in the mRNA is used to direct the synthesis of polypeptides by ribosomes. This process is called translation, because the nucleotide sequence of the mRNA is translated into an amino acid sequence in the polypeptide.

The two steps of central dogma taken together are also means of gene expression.

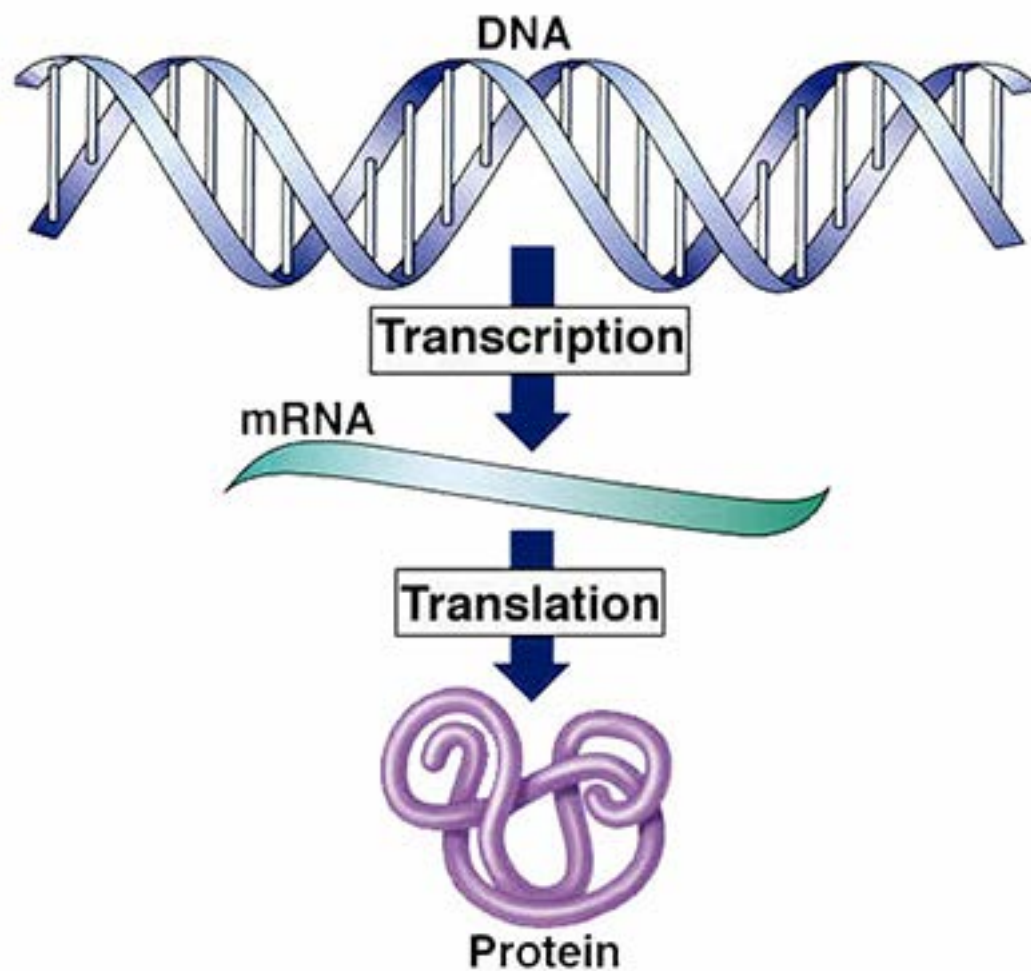


Fig. 20.23 The central dogma of gene expression.

Three types of RNA

The class of RNA found in ribosome is called ribosomal RNA (rRNA). During translation, rRNA provides the site where polypeptides are assembled. In addition to rRNA, there are two other major classes of RNA in the cells : transfer RNA (tRNA) and messenger RNA (mRNA). Transfer RNA molecules transport the amino acids to the ribosomes for use in building the polypeptides and also position each amino acid at the correct place on the elongating polypeptide chain (Fig 20.24). Human cells contain about 45 different kinds of tRNA molecules. Messenger RNA are long strands of RNA that are transcribed from DNA and that travel to the ribosomes to direct precisely which amino acids are assembled into polypeptides.

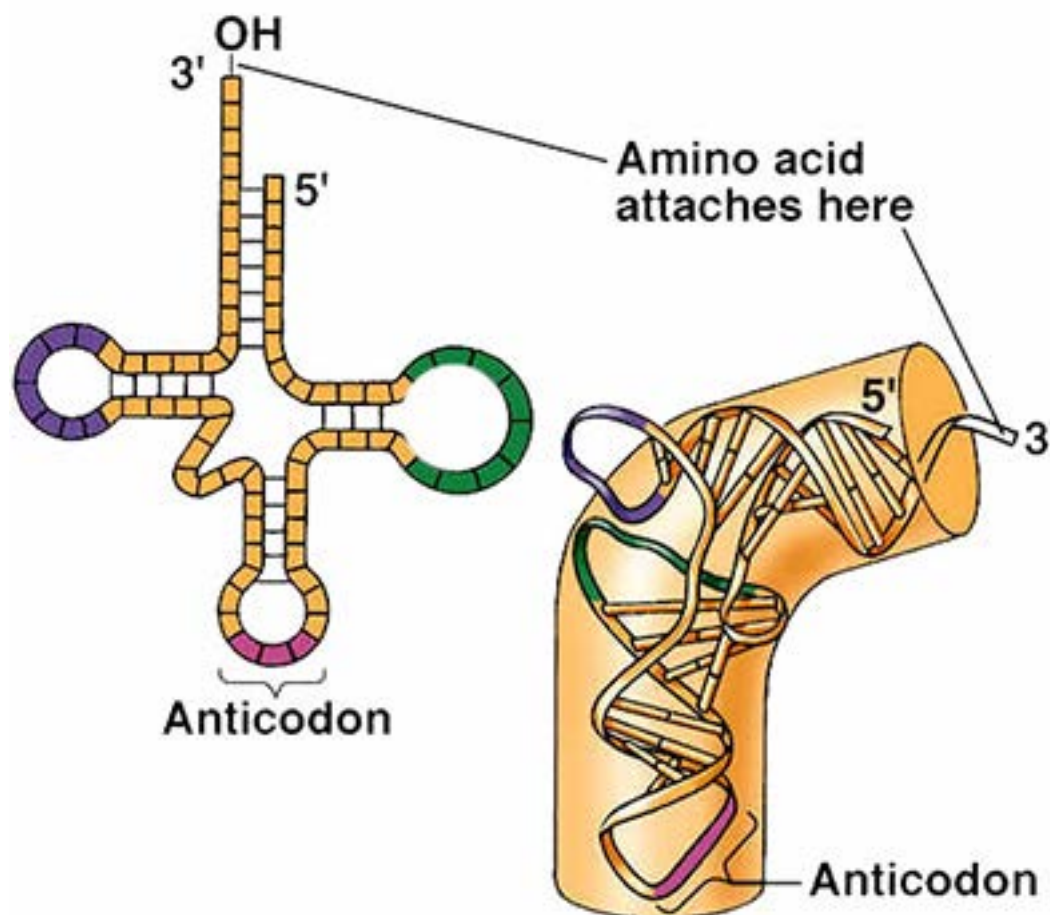


Fig. 20.24 The structure of tRNA
(a) two-dimensional schematic
(b) three dimensional structure.

Transcription

This is the process in which an RNA copy of the DNA sequence encoding the gene is produced with the help of an enzyme, RNA polymerase. Only one of the two strands of DNA are transcribed. This strand is called template strand or the antisense strand. The opposite strand is called coding strand or the sense strand. The RNA polymerase enzymes synthesize RNA from 5' → 3' direction. There is only one type of RNA polymerase in prokaryote which is responsible for the synthesis of all the three types of RNAs viz. rRNA, mRNA and tRNA. On the other hand there are three types of RNA polymerases in eukaryotes namely RNA polymerase I, which synthesizes rRNA, RNA polymerase II, which synthesizes mRNA and RNA polymerase III which synthesizes tRNA.

Transcription starts at the RNA polymerase binding site called promoter on the DNA template strand. In prokaryotes within promoter there are two binding sites TTAGACA also called -35 sequence and TATAAT sequence also called -10 sequence, which have affinity for the RNA polymerase. In eukaryotes these sites are at -75 and -25 sites, respectively

The binding of RNA polymerase to the promoter is the first step in gene transcription. One of the subunits of RNA polymerase sigma factor, is responsible for correct initiation of transcription process. Once the transcription has started the sigma factor is released and the remaining part of the enzyme (core enzymes) moves over template; strand and completes the transcription of the gene. The DNA strands open up at the place where enzyme is attached to the template strand forming transcription bubble. The transcription bubble moves down the DNA, leaving the growing strand protruding from the bubble (Fig 20.25). The stop sequences at the end of the gene terminate the synthesis of mRNA. The simplest stop signal is a series of GC base pairs followed by a series of AT base pairs. The RNA formed in this region forms a GC hairpin (Fig 20.27) followed by four or more U ribonucleotides. The hairpin causes RNA polymerase to stop synthesis.

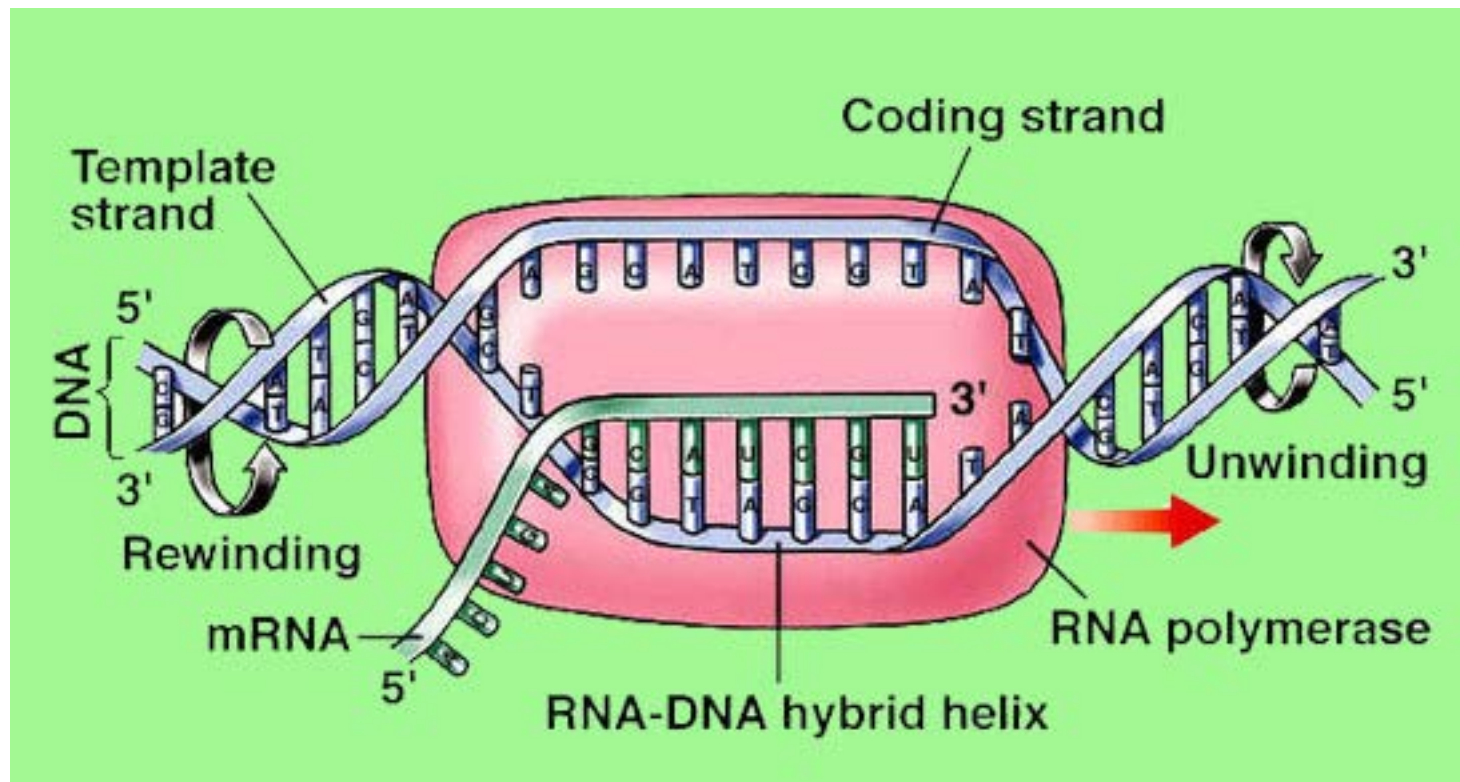


Fig. 20.25 Model of a transcription bubble.

In bacteria the newly synthesized mRNA is directly released into the cytoplasm, when it is converted into polypeptide chain. In eukaryotes however, it has to travel long distance from inside the nucleus to ribosomes outside in the cytoplasm. The eukaryotic mRNA is therefore modified in several ways to aid this journey. A cap and a tail is added so that the molecule may remain stable during long journey to ribosome. The cap is in the form of 7 methyl GTP, which is linked 5' to 5' with the first nucleotide, whereas tail is in the form of poly A tail linked to 3' end of the RNA. These caps and tails save the mRNA from variety of nucleases and phosphatases.

GENETIC CODE

Genetic code is a combination of 3 nucleotides, which specify a particular amino acid. There are three nucleotides in a codon, because a two nucleotide codon would not yield enough combinations to code for the 20 different amino acids that commonly occur in proteins. With four DNA nucleotides (G,C, T and A) only 42 or 16, different pairs of nucleotides could be formed. However, these same nucleotides can be arranged in 43 or 64 different combinations of three, more than enough to code for the 20 amino acids. The genetic code is a triplet code and the reading occurs continuously without punctuation between the three nucleotide units.

After Crick's initial experiments, Marshall Nirenberg, Philip Leader and Har Gobind Khorana tested all the 64 codons by making artificial mRNAs and triplet codons and using them to synthesize a protein or aminoacyl-tRNA complexes in cell free systems. '

The full genetic code was determined during mid 60s (Table 20.1). Out of 64 codons, three codons UAA, UAG and UGA do not code for any amino acid and hence are known as nonsense codons. These codons are usually present at the end of the gene and hence are also called stop codons. Every gene starts with initiation codon AUG, which encodes the amino acid methionine.

Table 20.1 The Genetic Code

First letter	Second Letter								
	U	C	A	G	Third letter				
U	UUU UUC	phenylalanine	UCU UCC	Serine	UAU UAC	Tryosine	UGU UGC	Cysteine	U C
	UUA UUG	Leucine	UCA UCG		UAA UAG	Stop Stop	UGA UGG	Stop Tryptophan	A G
C	CUU CUC	Leucine	CCU CCC	Proline	CAU CAC	Histidine	CGU CGC	Arginine	U C
	CUA CUC		CCA CCG		CAA CAG	Glutamine	CGA CGG		A G
A	AUU AUC	Isoleucine	ACU ACC	Treonine	AAU AAC	Asparagine	AGU AGC	Serine	U C
	AUA AUG	Methionine; Start	ACA ACG		AAA AAG	Lysine	AGA AGG	Arginie	A G
G	GUU GUC	Valine	GCU GCC	Alanine	GAU GAC	Aspartate	GGU GGC	Glycine	U C
	GUA GUG		GCA GCG		GAA GAG	Glutamate	GGA GGG		A G

The genetic code is universal. It is the same in almost all the organisms. For example AGA specifies arginine in bacteria, in humans and all other organisms whose genetic code has been studied. Because of the universality of codon, the genes can be transferred from one organism to another and be successfully transcribed and translated in their new host.

The study of genetic code of mitochondrial DNA however, showed that genetic code is not that universal. For example UGA codon is normally a stop codon but, in mitochondria it reads as tryptophan. Likewise AUA was read as methionine instead of isoleucine and AGA and AGG for termination of protein synthesis is instead of arginine. Thus it appears that genetic code is not quite universal.

TRANSLATION

In prokaryotes, translation begins when the initial portion of an mRNA molecule binds to rRNA molecule in a ribosome. The mRNA lies on the ribosome in such a way that only one of its codons is exposed at the polypeptide site at any time.

A tRNA molecule possessing the complementary three nucleotide sequence or anticodon, binds to the exposed codon on the mRNA. As the ribosome moves along the messenger RNA, successive codons on the mRNA are exposed and the series of tRNA molecules bind one after another to the exposed codons. Each of these tRNA molecules carries an attached amino acid, which is added to the end of the growing polypeptide chain.

Particular tRNA molecules become attached to specific amino acids through the action of activating enzymes called aminoacyl-tRNA synthetase, one of which exists for each of the 20 common amino acids (Fig 20.26).

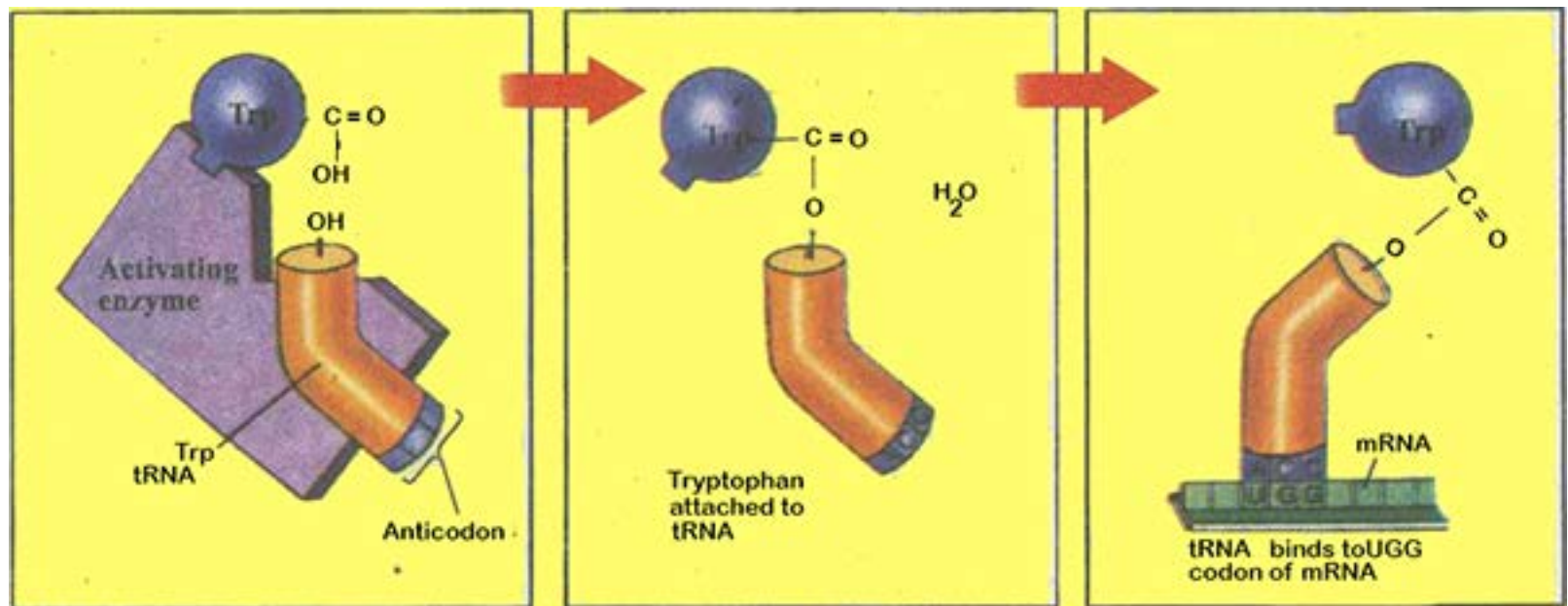


Fig. 20.26 Activating enzymes "read" the genetic code.

In prokaryotes, polypeptide synthesis begins with the formation of initiation complex (Fig. 20.27). First a tRNA molecule carrying a chemically modified methionine (called N-formyl methionine) binds to the small ribosomal subunit. Proteins called initiation factor position the tRNA on the ribosomal surface at the P site (peptidyl site) where peptide bonds will form. Nearby two other sites will form. A site (for aminoacyl site), where successive amino acid bearing tRNAs will bind and the E site (for exit site) where empty tRNAs will exit the ribosome (Fig 20.27). This initiation complex, guided by another initiation factor, binds to AUG on the mRNA.

After the initiation complex has formed, the large ribosome subunit binds tRNA molecule with the appropriate anticodon appears, proteins called elongation factors assist in binding it to the exposed mRNA codon at the A site. The two amino acids which now he adjacent to each other undergo a chemical reaction, catalyzed by the large ribosomal subunit, which releases the initial methionine from its tRNA and attaches it instead by a peptide bond to the second amino acid (Fig. 20.28).

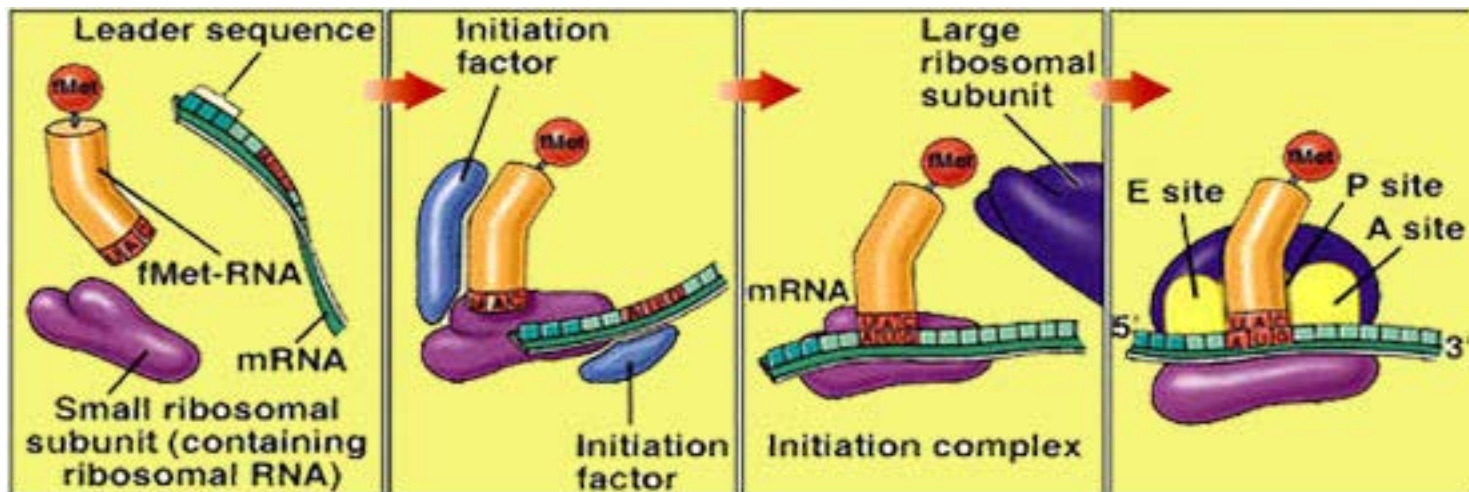


Fig. 20.27 Formation of the initiation complex.

The ribosome now moves (translocates) three more nucleotides along the mRNA molecule in the 5' → 3' direction, guided by other elongation factors. This movement translocates the initial tRNA to the E site and ejects it from the ribosome, repositions the growing polypeptide chain (at this point containing two amino acids) to the P site, and exposes the next codon on the mRNA at the A site (Fig 20.28). When a tRNA molecule recognizing that codon appears, it binds to the codon at the A site, placing its amino acid adjacent to the growing chain. The chain then transfers to the new amino acid, and the entire process is repeated.

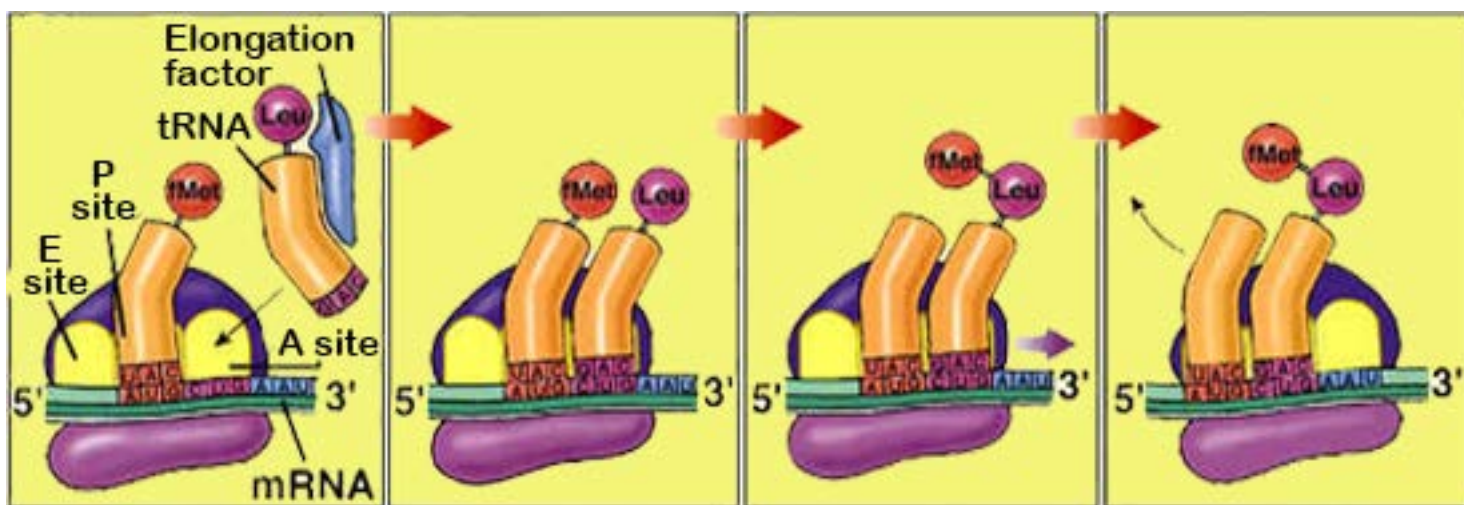


Fig. 20.28 The translocation process.

Elongation continues in this fashion until a chain-terminating non sense codon is exposed (for example UAA in Fig 20.29). Nonsense codons do not bind to tRNA, but they are recognized by release factors, proteins that release the newly made polypeptide from the ribosomes.

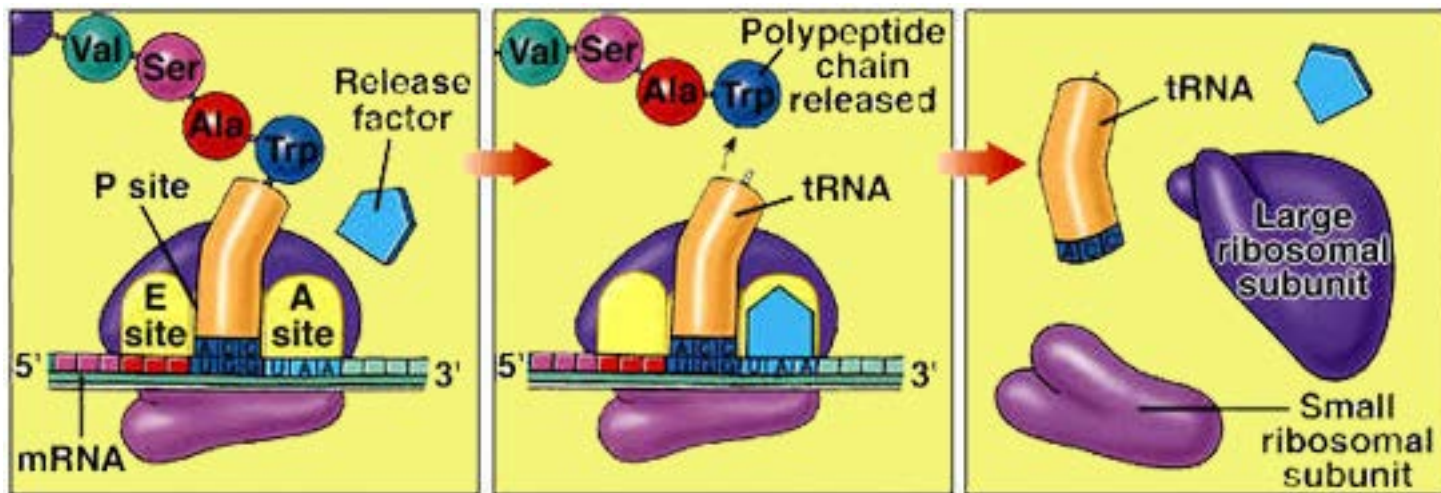


Fig. 20.29 Termination of protein synthesis

MUTATIONS

The cells of eukaryotes contain an enormous amount of DNA. If the DNA in all of the cells of an adult human were lined up end to end, it would stretch nearly 100 billion kilometers - 60 times the distance from Earth to Jupiter.

Changes in the DNA occur either due to mistake in replication or damage to the genetic message causing mutations. The mutations in somatic cells do not pass on to offspring and so have little evolutionary consequence than germ line changes. The mutation in germ line cell is passed to subsequent generations thus providing the raw material from which natural selection produces evolutionary change.

Mutations can broadly be classified as (i) chromosomal aberration and (ii) point mutation. Chromosomal aberrations are mega changes which involve presence of an extra chromosome or loss of a chromosome from the diploid number of chromosomes, or changes like deletions, insertions, inversions etc in the parts of the chromosome, Such chromosomal aberrations lead to syndromes like Down's syndrome, Klinefelter's syndrome etc.

Point mutations are mutational changes which affect the message itself, producing alterations in the sequence of DNA nucleotide (Table 20.2). If alterations involve only one or a few base pairs in the coding sequence they are called point mutations. While some point mutations occur due to spontaneous pairing errors that occur during DNA replication, others result from damage to the DNA caused by mutagens, usually radiations or chemicals. The latter class of mutations is of particular practical importance because modern industrial societies often release many chemical mutagens into the environment. Sickle cell anemia and phenylketonuria are well known examples of point mutation, both of which have been discussed in previous page. In sickle cell anemia a point mutation leads to the change of amino acid glutamic acid into valine at position 6 from N terminal end in hemoglobin P chain. This consequently alters the tertiary structure of the hemoglobin molecule, reducing its ability to carry oxygen.

In phenylketonuria, phenylalanine is not degraded because of defective enzyme phenylalanine hydroxylase. Phenylalanine consequently accumulates in the cells leading to mental retardation, as the brain fails to develop in infancy. This disorder is because of the point mutation.

EXERCISE

Q.1 Fill in the blanks.

1. Particular tRNA molecules become attached to specific amino acids through the action of activating enzymes called_____.
2. _____ is the transfer of genetic material from one cell to another and can alter the genetic make up of the recipient cell.
3. In a bacteria, a subunit of RNA polymerase called_____ recognizes-10 sequence in the promoter and binds RNA polymerase there.
4. A typical human chromosome contain about nucleotides in its DNA.
5. Miescher extracted a white substance from the nuclei of human cells and fish sperm and called this substance_____ .

Q.2 Write whether the statement is true or false and write the correct statement if it is false.

1. The strand of DNA that is not transcribed is called the coding strand.
2. TA TAAT sequence called - 35 sequence is part of promoter, where transcription actually starts.
3. Rosalind Franklin carried out an x-ray diffraction analysis of DNA.
4. The base pairs in DNA helix are planar and stack 34 nm apart as a result of hydrophobic interactions.

Q.4 Short Questions

1. What are the three major classes of RNA?
2. What is the function of RNA polymerase in transcription?
3. How did Crick and his colleagues determine how many nucleotides are used to specify each amino acid?
4. What is anticodon?

Q.5 Extensive Questions

1. How did Hershey and Chase determine which components of bacterial viruses contain the hereditary information?
2. 'What is the three dimensional shape of DNA? How does three dimensional shape of DNA fit with Chargaff's observations on the proportions of purines and pyrimidines in DNA?
3. How did Meselson and Stahl show that DNA replication is semi conservative?
4. What is the basis for the requirement that the leading and lagging strands be replicated by different mechanisms?
5. What hypothesis did Beadle and Tatum test in their experiments on Neurospora ?