Genetic Information In Microbes:

The genetic material of bacteria and plasmids is DNA. Bacterial viruses (bacteriophages or phages) have DNA or RNA as genetic material. The two essential functions of genetic material are replication and expression. Genetic material must replicate accurately so that progeny inherit all of the specific genetic determinants (the genotype**)** of the parental organism. Expression of specific genetic material under a particular set of growth conditions determines the observable characteristics (phenotype) of the organism.

Nucleic Acid Structure:

Nucleic acids are large polymers consisting of repeating nucleotide units (**Fig.1-1**). Each nucleotide contains one phosphate group, one pentose or deoxypentose sugar, and one purine or pyrimidine base. In DNA the sugar is D-2-deoxyribose; in RNA the sugar is D-ribose.

In DNA the purine bases are adenine (A) and guanine (G), and the pyrimidine bases are thymine (T) and cytosine (C). In RNA, uracil (U) replaces thymine. The repeating structure of polynucleotides involves alternating sugar and phosphate residues, with phosphodiester bonds linking the 3'-hydroxyl group of one nucleotide sugar to the 5'-hydroxyl group of the adjacent nucleotide sugar. A purine or pyrimidine base is linked at the 1'-carbon atom of each sugar residue and projects from the repeating sugar-phosphate backbone.

Double-stranded DNA is helical, and the two strands in the helix are antiparallel. The double helix is stabilized by hydrogen bonds between purine and pyrimidine bases on the opposite strands. A on one strand pairs by two hydrogen bonds with T on the opposite strand, or G pairs by three hydrogen bonds with C. The two strands of double-helical DNA are, therefore, complementary. Because of complementarity, double-stranded DNA contains equimolar amounts of purines (A + G) and pyrimidines (T + C), with A equal to T and G equal to C, but the mole fraction of G + C in DNA varies widely among different bacteria. Information in nucleic acids is encoded by the ordered sequence of nucleotides along the polynucleotide chain. The extent of sequence homology between DNAs from different microorganisms is the most stringent criterion for determining how closely they are related.

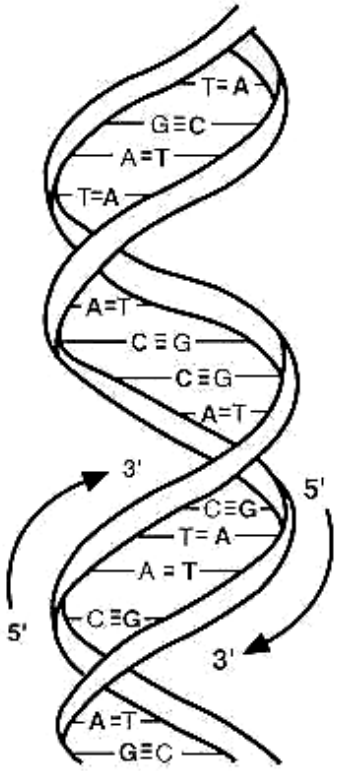


Figure1-1:Double helical structure of DNA represented as a helical ladder. The backbone of each polynucleotide strand consists of alternating phosphate and deoxyribose residues linked by phosphodiester bonds,and the strands have opposite polarities. The purine or pyrimidine base of each nucleotide on one strand projects toward the comple-mentary base of the correspond-ding nucleotide from the other strand and is linked to it by hydrogen bonds. The double helix has a diameter of 2 nm. Each full turn of the double helix contains 10 nucleotide pairs and is 3.4 nm in length.

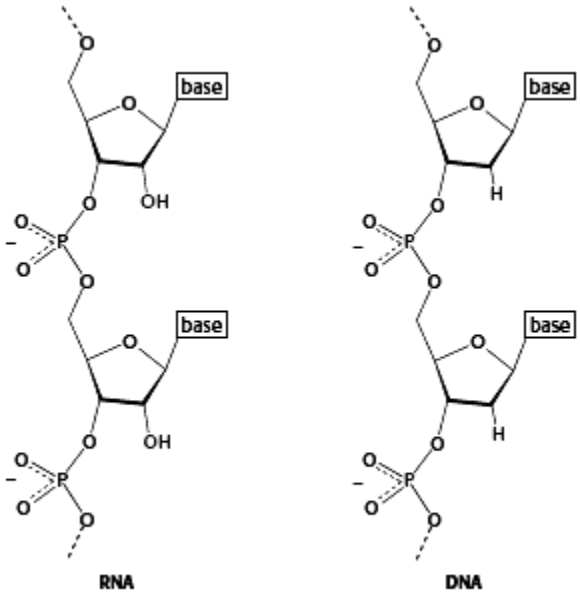


Figure1-2: RNA and DNA compared. Removal of the 2-hydroxyl group from RNA to form DNA results in a backbone that is less suscept-ible to cleavage by hydrolysis and thus enables more-stable storage of genetic information.

DNA Replication:

DNA molecules that replicate as discrete genetic units in bacteria are called **replicons**. In some *Escherichia coli* strains, the chromosome is the only replicon present in the cell. Other bacterial strains have additional replicons, such as plasmids and bacterio-phages. During replication of the bacterial genome, each strand in double-helical DNA serves as a template for synthesis of a new complementary strand. Each daughter double-stranded DNA molecule thus contains one old polynucleotide strand

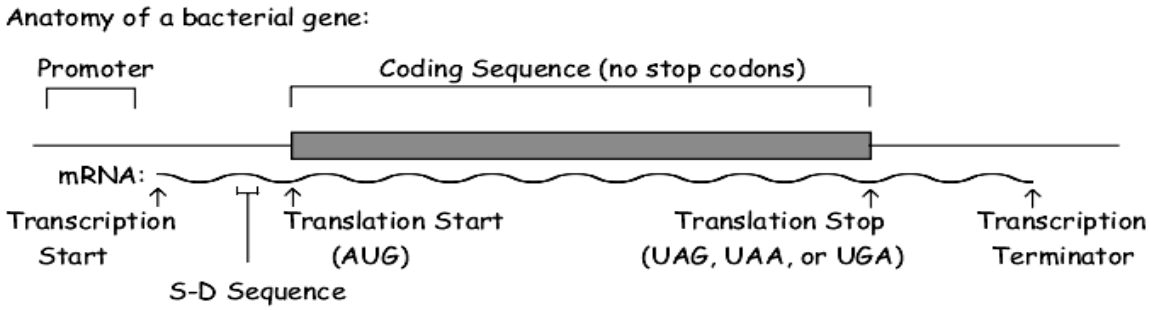
and one newly synthesized strand. This type of DNA replication is called **semi-conservative**. Replication of chromosomal DNA in bacteria starts at a specific chromosomal site called the origin and proceeds bidirectionally until the process is completed. When bacteria divide by binary fission after completing DNA replication, the replicated chromosomes are partitioned into each of the daughter cells.

Gene Expression:

Genetic information encoded in DNA is expressed by synthesis of specific RNAs and proteins, and information flows from DNA to RNA to protein. The DNA-directed synthesis of RNA is called **transcription**. Because the strands of double-helical DNA are antiparallel and complementary, only one of the two DNA strands can serve as template for synthesis of a specific mRNA molecule. Messenger RNAs (mRNAs) transmit information from DNA, and each mRNA in bacteria functions as the template for synthesis of one or more specific proteins. The process by which the nucleotide sequence of an mRNA molecule determines the primary amino acid sequence of a protein is called **translation**.

Genes of higher eukaryotes are interrupted by **introns**, (long sequences that do not code for proteins) which are spliced out of the RNA before translation (post-transcription process).

Ribosomes, complexes of ribosomal RNAs (rRNAs) and several ribosomal proteins, translate each mRNA into the corresponding polypeptide sequence with the aid of transfer RNAs (tRNAs), amino-acyl tRNA synthesases, initiation factors and elongation factors. All of these components of the apparatus for protein synthesis function in the production of many different proteins. A **gene** is a DNA sequence that encodes a protein, rRNA, or tRNA molecule (gene product).



Promoter: A sequence in DNA at which RNA polymerase bind to DNA and start transcription of a mRNA copy of the gene sequence.

S-D: A sequence in mRNA that will load ribosomes to begin translation.

The **genetic code** determines how the nucleotides in mRNA specify the amino acids in a polypeptide. Because there are only 4 different nucleotides in mRNA (containing U, A, C and G), single nucleotides do not contain enough information to specify uniquely all 20 of the amino acids. In dinucleotides 16 (4 x 4) arrangements of the four nucleotides are possible, and in trinucleotides 64 (4 x 4 x 4) arrangements are possible. Thus, a minimum of three nucleotides is required to provide at least one unique sequence corresponding to each of the 20 amino acids. The "universal" genetic code employed by most organisms (**Table 1-1**) is a triplet code in which 61 of the 64 possible trinucleotides (codons) encode specific amino acids, and any of the three remaining codons (UAG, UAA or UGA) results in termination of translation.

TABLE 5-1 The Genetic Code[a]

First Nucleotide of Codon	Second Nucleotide of Codon				Third Nucleotide of Codon
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Termination	Termination	A
	Leu	Ser	Termination	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	His	Arg	A
	Leu	Pro	His	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	C

[a]Abbreviations: Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Cys, cysteine; Gln, glutamine; Glu, glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Try, tyrosine; Trp, tryptophan; Val, valine.

The chain-terminating codons are also called **nonsense codons** because they do not specify any amino acids. The genetic code is described as degenerate, because several codons may be used for a single amino acid, and as nonoverlapping, because adjacent codons do not share any common nucleotides. Translation of mRNA is initiated at an AUG codon for methionine, and adjacent codons are translated sequentially as the mRNA is read in the 5' to 3' direction. The sequence of amino acids in the polypeptide

chain is, therefore, colinear with the sequence of nucleotides in the mRNA and the corresponding gene.

Expression of genetic determinants in bacteria involves the unidirectional flow of information from DNA to RNA to protein. Studies with the retrovirus group of animal viruses reveal that DNA molecules can be synthesized from RNA templates by enzymes designated as RNA-dependent DNA polymerases (**reverse transcriptases**). This reversal of the usual direction for flow of genetic information, from RNA to DNA instead of from DNA to RNA, is an important mechanism for enabling information from retroviruses to be encoded in DNA and to become incorporated into the genomes of animal cells.

Chromosomal DNA:

Bacterial genomes vary in size from about 0.4×10^9 to 8.6×10^9 daltons (Da). Most bacteria have a haploid genome, a single chromosome consisting of a circular, double stranded DNA molecule. However linear chromosomes have been found in Gram-positive *Borrelia* and *Streptomyces* spp. The single chromosome of *E. coli* is 3×10^9 Da (4,500 kilobase pairs [kbp]) in size, accounting for about 2 to 3 percent of the dry weight of the cell.

The chromosome of *E. coli* has a length of approximately 1.35 mm, several hundred times longer than the bacterial cell, but the DNA is supercoiled and tightly packaged in the bacterial nucleoid. The time required for replication of the entire chromosome is about 40 minutes. DNA replication must be initiated as often as the cells divide, so in rapidly growing bacteria a new round of chromosomal replication begins before an earlier round is completed. The chromosome in rapidly growing bacteria is replicating at more than one point and involves many different proteins.