Evidence that DNA is the genetic material: Deoxyribonucleic acid (DNA) was described as a biochemical substance in 1871, but its role as the carrier of genetic information was not published until 1944. DNA was shown to be associated with chromosomes in the late 19th century, but nuclear proteins were generally considered to the the more likely carrier of genetic information. The following are some of the early lines of evidence that DNA carries genetic information.

- In 1928, Frederick Griffith found that non-living material derived from a "smooth" strain of *Streptococcus pneumoniae* (which has a polysaccharide on its outer surface that protects it from the immune system of mice and makes it virulent) was able to transform a living non-virulent "rough" strain that lacked the protective coat into a virulent "smooth" strain.
- In 1944, Avery, McLeod and McCarty reported that the active substance in the non-living material was DNA (see boxed example 2.1 in textbook).
- In 1952, Hershey and Chase used radioactive labels to show that when bacteriophage infect bacteria, bacteriophage DNA labeled with ^{32}P enters the bacterial cell whereas bacteriophage protein labeled with ^{35}S does not (textbook fig. 2.4)
- More recent studies utilizing a wide variety of techniques, including recombinant DNA and transgenic organisms, have fully verified the conclusions of these early experiments.

"Central dogma" of molecular biology (textbook pages 57-58)

- Genetic information is stored in DNA as a sequence of nucleotide bases (adenine, cytosine, guanine, thymine, abbreviated A,C,G,T) read sequentially in a 5' to 3' direction (or in RNA, with uracil, abbreviated U, replacing thymine).
- The most common form of DNA (present in all cellular genomes, as well as many viral genomes) is double stranded. The 5' to 3' polarity of the two strands is opposite, and they are held together by hydrogen bonding between nucleotide base pairs, A to T and G to C. The ***sense*** strand carries the coded genetic information. The ***antisense*** strand consists of a complementary sequence of bases oriented in the opposite 5' to 3' direction.
- During DNA ***replication***, the two strands separate and each is used as a template for synthesis of a new complementary strand. This allows genetic information to be replicated with a high level of precision. Because replication is bidirectional, but new DNA can only be synthesized in a 5' to 3' direction, the overall pattern of replication is rather complex as will be discussed in the next lecture.
- Genetic information is ***transcribed*** from DNA to RNA, with the antisense strand of the DNA serving as a template for synthesis of an RNA with the same base sequence (5' to 3') as the sense strand of the double helical DNA, except that uracil (U) replaces thymine (T).
- Genetic information contained in messenger RNA (mRNA) is ***translated*** into a sequence of amino acids in a polypeptide chain during protein synthesis (translation). A redundant nucleotide triplet code, read 5' to 3' on the mRNA (and

on the sense strand of the DNA), specifies the amino acid sequence of the protein, read from N-terminal to C-terminal.

Prokaryotic and eukaryotic cells: Be sure that you fully understand appreciate the distinctions between these two basic types of cells. The textbook appears to assume that its readers already have a full understanding of the differences. The first mention by name of prokaryotic and eukaryotic cells that I have found is in question 8 for chapter 1. We will encounter substantial differences between prokaryotic and eukaryotic cells is lectures 3, 4, and 5, and in many other parts of this course.

- There are two basic types of cells, called ***eukaryotic*** and ***prokaryotic*** respectively (sometimes spelled ***eucaryotic*** and ***procaryotic***).
- Eukaryotic cells have their genetic material separated from the cytoplasm in a membrane-bound nucleus. The nuclear envelope restricts and regulates the passage of materials between the two compartments. DNA synthesis and transcription (RNA synthesis) occur within the nucleus. Messenger RNA (mRNA) is then processed and transported to the cytoplasm for use as a template for translation (protein synthesis). This causes translation to be separated both in space and in time from transcription.
- Prokaryotic cells, which include bacteria and some primitive algae, are much smaller and have a simpler structure, with no true separation of DNA from the cytoplasm. Because of this, transcription and translation occur in the same space and overlap in time. In fact, translation of an mRNA will normally begin soon after its transcription has been initiated, and long before synthesis of a full-length message has been completed (see figure 3.3, page 60).

Current understanding of the roles of DNA and RNA as carriers of genetic information

- Double stranded DNA is the genetic material in all eukaryotic and prokaryotic cells, as well as mitochondria, chloroplasts, plasmids and many types of viruses..
- Some types of viruses have single stranded DNA genomes.
- Some types of viruses have single stranded or double stranded RNA genomes

Note: Because of the large amount of text material covered and the review nature of the material, the following sections are presented only as brief outlines.

Structural components of nucleic acids

- Pentose sugars
  - Ribonucleic acid (RNA) contains ribose.
  - Deoxyribonucleic acid (DNA) contains 2-deoxyribose, which is often referred to simply as deoxyribose.
  - The only difference between the two sugars is that ribose has a hydroxyl group on carbon 2, whereas deoxyribose has only a hydrogen in that position. Both have hydroxyl groups on carbons 3 and 5.

    - In the configuration found in nucleic acids, carbons 1 through 4 are part of a ring structure, whereas carbon 5 is on a side chain (see figure 2.6).
- phosphate
    - In both RNA and DNA, the phosphate groups form phosphodiester linkages between the 3' hydroxyl group on one pentose sugar and the 5' hydroxyl group on the next pentose sugar. (In nucleic acids the "prime" designation indicates a position on a ribose or deoxyribose, as opposed to a position on a purine or pyhrimidine base).
    - The polymeric structure consisting of alternating sugar and phosphate residues provides nucleic acids with an extended backbone structure.
    - The pentose sugars in the nucleic acid backbones all have the same 5' to 3' orientation, giving each chain a 5' to 3' directionality.
- Purine bases
    - Purines are flat planar molecules consisting of a six member ring fused to a five member ring, with two nitrogens in each of the rings.
    - Adenine (A) has an amino nitrogen at position 6 of its larger ring.
    - Guanine (G) has an amino nitrogen at position 2 and an oxygen at position 6 of the larger ring.
    - A and G occur in both DNA and RNA.
    - The the bond between the purines and the nucleic acid backbone is from position 1' on the pentose sugar to position 9 (a nitrogen in the smaller ring) of the purine.
- Pyrimidine bases
    - Pyrimidines are flat planar molecules consisting of a single six member heterocyclic ring containing two nitrogens. Please note that the numbering of positions in pyrimidines is different than in the six-membered ring of purines (fig. 2.6 in textbook).
    - Cytosine (C) has an amino nitrogen it position 4 and an oxygen at position 2. It is found both in DNA and in RNA.
    - Uracil (U), which is only in RNA, has oxygens at positions 2 and 4 of its ring.
    - Thymine (T), which is only in DNA, is identical to uracil, except that it has an added methyl group at position 5 of its ring.
    - The bond between pyrimidines and the nucleic acid backbone is from position 1' on the pentose sugar to position 1 (a nitrogen) in the pyrimidine ring.

Nucleosides and Nucleotides: The terminology and abbreviations used to describe the various precursors and hydrolysis products of nucleic acids are not always explained clearly in our textbook. Be sure that you are familiar with all of the following.

- The term **nucleoside** refers to one purine or pyrimidine base attached to one ribose or deoxyribose: adenosine, guanosine, cytidine, uridine, deoxyadenosine, deosyguanosine, deoxycytidine, thymidine.
    - Be careful with the nomenclature of cytosine, which is the free pyrimidine, and cytidine, which is the nucleoside.

    - Thymidine is generally assumed to be derived from DNA, and thus usually not called "deoxythrmidine". When thymine is found attached to ribose, as is the case in some of the transfer RNAs (see page 99 of textbook), it is referred to as "ribothymidine".
- The term **nucleotide** refers to a nucleoside with one attached phosphate group (nucleoside monophosphate).
    - Nucleic acids are often described as polymers of nucleotides or polynucleotides
    - The ribonucleotides are commonly abbreviated AMP, CMP, GMP, UMP.
    - The deoxynucleotides are commonly abbreviated dAMP, dCMP, dGMP, TMP.
    - Depending on the method of hydrolysis of the nucleic acid (or other means of preparation of nucleotides), the phosphate can be either in the 5'- or 3'-position. When not specified otherwise, the 5'- position is usually assumed.
    - In some cases the phosphate may be esterified to both the 3' and 5' hydroxyl groups of ribose, generating a cyclic nucleotide, such as cyclic AMP (cAMP).
- **Nucleoside diphosphates** are nucleosides with two attached phosphates.
    - When not specified otherwise, the two phosphates are tandemly linked to the 5'-position of the ribose or deoxyribose. Two phosphates linked to each other are often referred to as a pyrophosphate group.
    - Nucleoside 5'-diphosphates are abbreviated ADP, GDP, CDP, UDP, dADP, dCDP, dGDP, TDP.
    - It is also possible to have nucleoside 3'-,5'-diphosphates.
- **Nucleoside triphosphates** are nucleosides with three tandemly linked phosphates attached at the 5'- position.
    - The nucleoside triphosphates are commonly abbreviated ATP, CTP, GTP, UTP, dATP, dCTP, dGTP, TTP.
    - The third phosphate in nucleoside triphosphates is attached by a high-energy bond, whose hydrolysis can drive a variety of biosynthetic processes.
    - Nucleoside triphosphates are the immediate precursors for synthesis of DNA and RNA

Structural features of double helical DNA

- Nucleotide sequences in DNA are always read 5' to 3' when not specified otherwise.
- The 5' to 3' orientations of the two strands in a DNA double helix are in the opposite directions. The two strands are described as being **antiparallel** in their orientation.
- Purine and pyrimidine bases form hydrogen bonded base pairs in double stranded DNA (textbook figure 2.10)
    - The normal pattern of base pairing in DNA is A to T and G to C.

    - Abnormal pairing during DNA synthesis can result in mutation (chapter 5 of textbook)
    - The base pairs in the DNA double helix are flat planar structures
    - In the DNA double helix, the flat base pairs are tightly stacked on top of one another, with the deoxyribose phosphate backbones wound around the outside of the stack.
    - The base pairs are 3.4 Å (0.34 nm) thick, with ten base pairs per complete turn of the double helix.
- The two deoxyribose phosphate backbones are unevenly spaced around the helix, such that there is a wide groove and a narrow groove between them.
- Separation of DNA strands (melting) occurs at high salt levels and/or high temperature.
    - GC base pairs have 3 hydrogen bonds and are more stable than AT base pairs, which only have 2.
    - Stability of double-helical DNA increases with higher GC content.
    - Separated strands can reform a double helix (anneal) if the conditions that caused denaturation are slowly reversed.
    - The amount of time required to renature increases with the compexity of the mixture of sequences present.
    - Complexity of DNA (total number of nucleotides of unique sequence) is often measured in terms of the product of total nucleotide concentration and time required for 50% renaturation (C_ot values) (Described in more detail in Chapter 10, Pages 298 - 300)
    - Complex genomes, such as those of mammals, contain a mixture of repetitive and unique sequence DNA.
- In addition to the common B configuration, double stranded DNA can also assume other configurations, some of which may be biologically significant.

Structural features of RNA

- RNA is much like a single strand of DNA, and can pair with a single complementary strand of DNA to form an antiparallel double helical structure.
- RNA molecules may fold back on themselves and form limited double-stranded regions, which appear to be important for achieving active configurations.
- Three major classes of RNA are involved in protien synthesis, messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA).
- RNA sequences are read 5' to 3' in the same manner as DNA.
- The coding sequence in mRNA is the same as that of the sense strand of DNA, except that U replaces T.
- Small nuclear RNA (snRNA) molecules have a variety of biological functions.

Electrophoresis

(Described on pages 278-283 of textbook, but presented in outline form at this time so references can be made to it in some of the lectures that follow).

- Nucleic acids have strong negative charges at neutral pH, and thus move toward the anode (positive pole) in an electric field.
- Because negative charges are uniformly distributed along nucleic acids, the charge to mass ratio is essentially the same for nucleic acids of different sizes.
- The rate of movement through a porous gel (polyacrylamide or agarose) is faster for smaller molecules and slower for larger molecules.
- Gel electrophoresis is a highly effective technique for separating mixtures of nucleic acids by size.
- Electrophoretic separations play major roles in many of the molecular techniques we will be studying