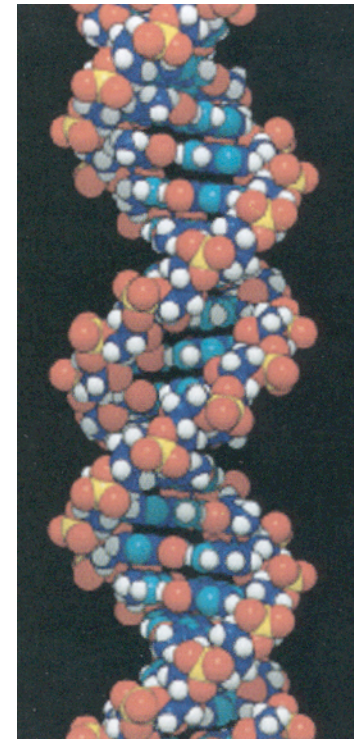


A Crash Course in Genetics

General Overview:

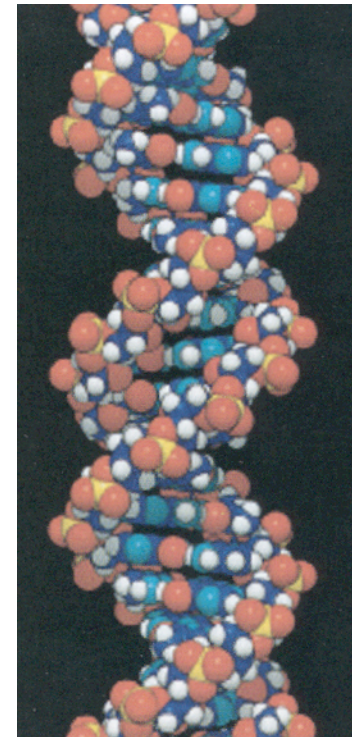
- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR



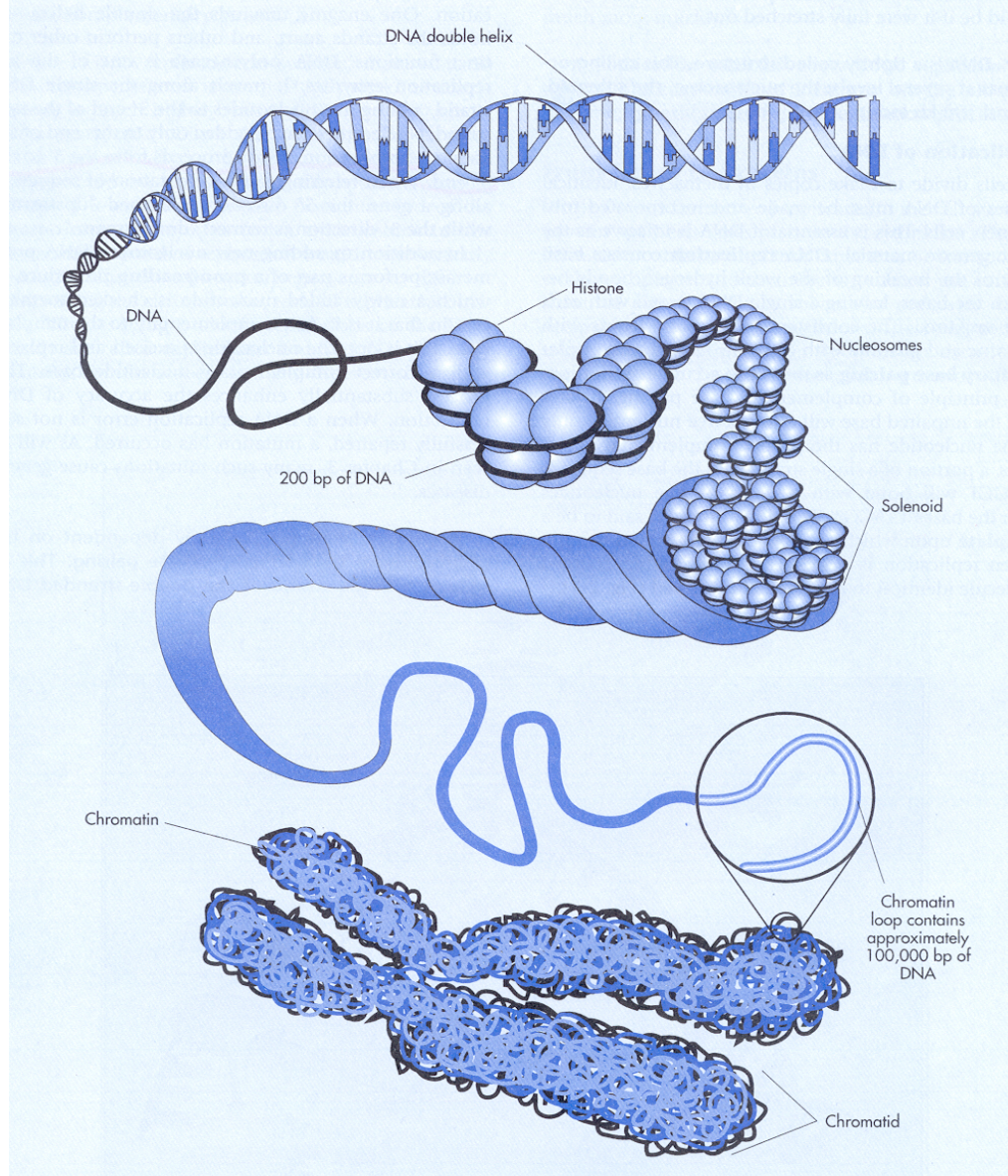
A Crash Course in Genetics

General Overview:

- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR



DNA Is Structured Hierarchically



Levels of Structure

- Double Helix
- Histones / Nucleosomes
- Solenoid Supercoil
- Chromatin
- Chromosomes

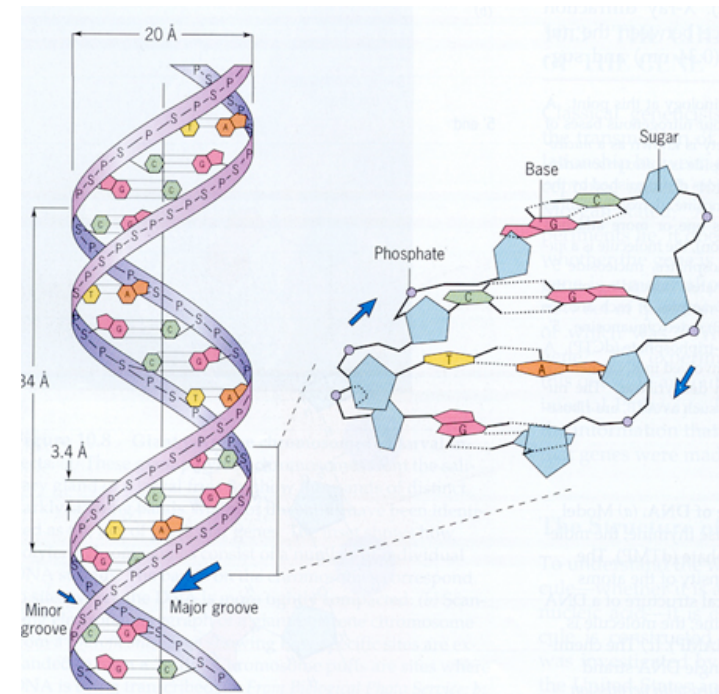
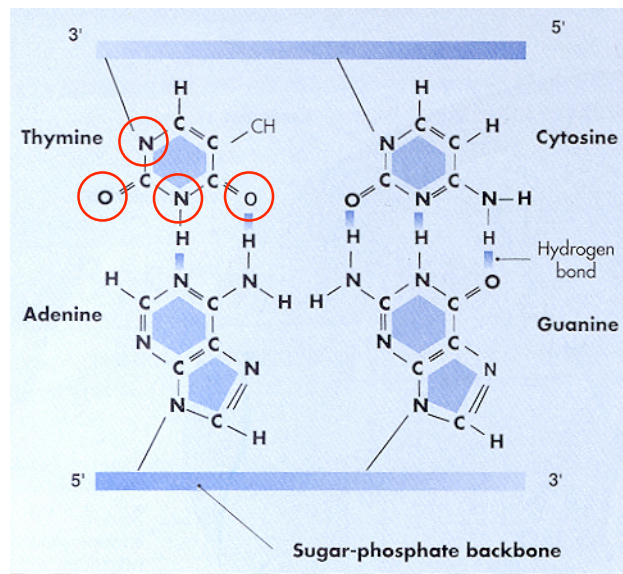
DNA Compacted to Conserve Space

There are several levels at which DNA is compacted:

1. The double helix — the DNA in a single cell contains 2.9×10^9 base pairs and would be a meter long.
2. Nucleosome — DNA is wound around a histone protein core to form a nucleosome. This gives a 5- to 9-fold reduction in length.
3. Solenoids — Nucleosomes (beads on a string) supercoil and form solenoid structures. 4-6-fold reduction in length.
4. Minibands — Solenoid turns loop around a protein-RNA scaffold to form Minibands. 18-fold reduction in length.
5. Chromosomes — Minibands further condense to form Chromosomes, the form of DNA as seen during cell division and genetics studies.

Putting the Puzzle Pieces Together

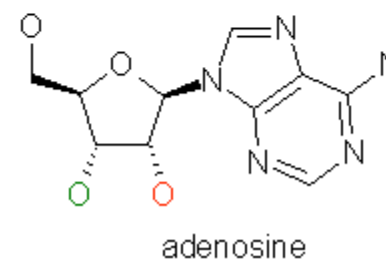
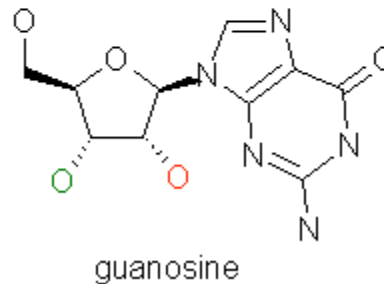
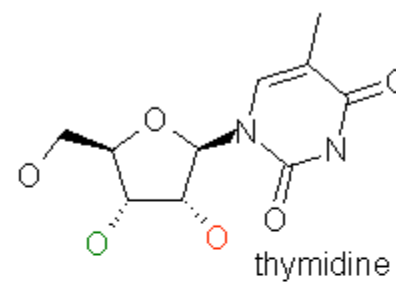
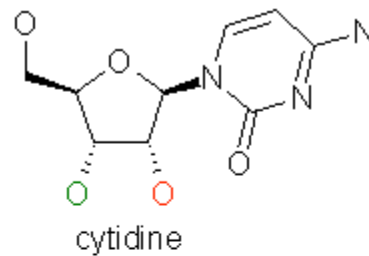
In 1953, James Watson and Francis Crick discovered the structure of the DNA double helix



What DNA is Made Of

DNA = *deoxyribonucleic acid*

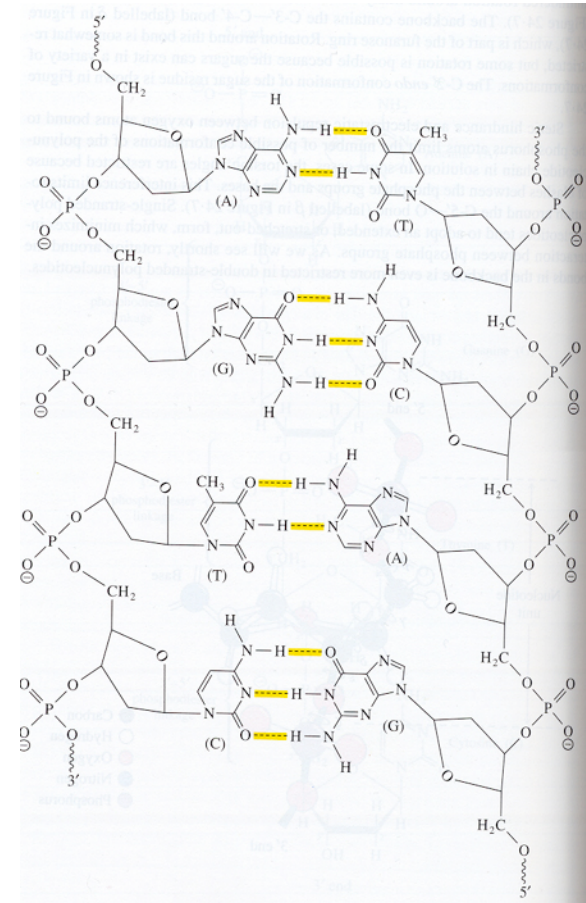
- deoxyribose sugar with the 2'OH (hydroxyl) group missing
- Phosphate group(s) (not shown here, attach to 3'OH)
- Nitrogenous base — Adenine, Guanine, Thymine, Cytosine
- Together these components make up a nucleotide



More About the Bonding

The 5'-phosphate group of one nucleotide joins to the 3'OH group of the next nucleotide (*phosphodiester bond* - very strong)

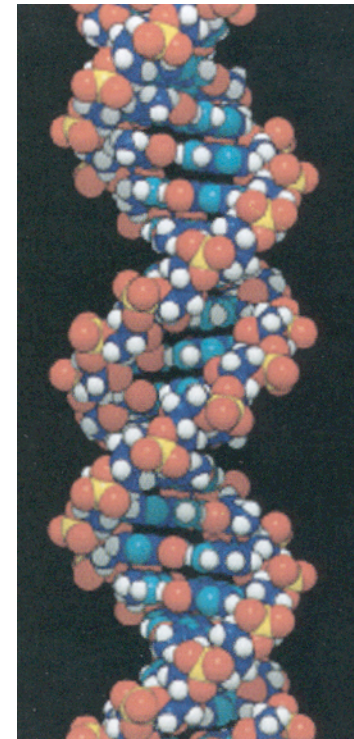
This gives the DNA molecule directionality, which plays a crucial role in DNA replication and transcription.



A Crash Course in Genetics

General Overview:

- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR

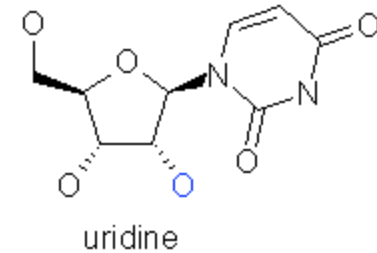


A Side Track: RNA

RNA (ribonucleic acid)

Similar structure to DNA, except for:

1. The 2'OH of all nucleotides are intact
2. All thymidines are replaced by a Uracil
3. Generally single stranded, as the extra hydroxyl group is too bulky to allow base pairing for significant distances.
4. Several forms, all with specific function:
 1. mRNA: messenger RNA
 2. tRNA: transfer RNA
 3. rRNA: ribosomal RNA

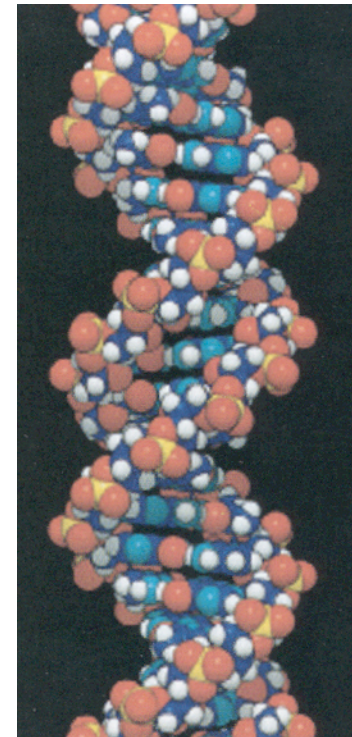


We will see the connection between DNA and RNA shortly...

A Crash Course in Genetics

General Overview:

- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR



DNA Replication — Making Copies

As cells divide, identical copies of the DNA must be made.

The following sequence of events occurs:

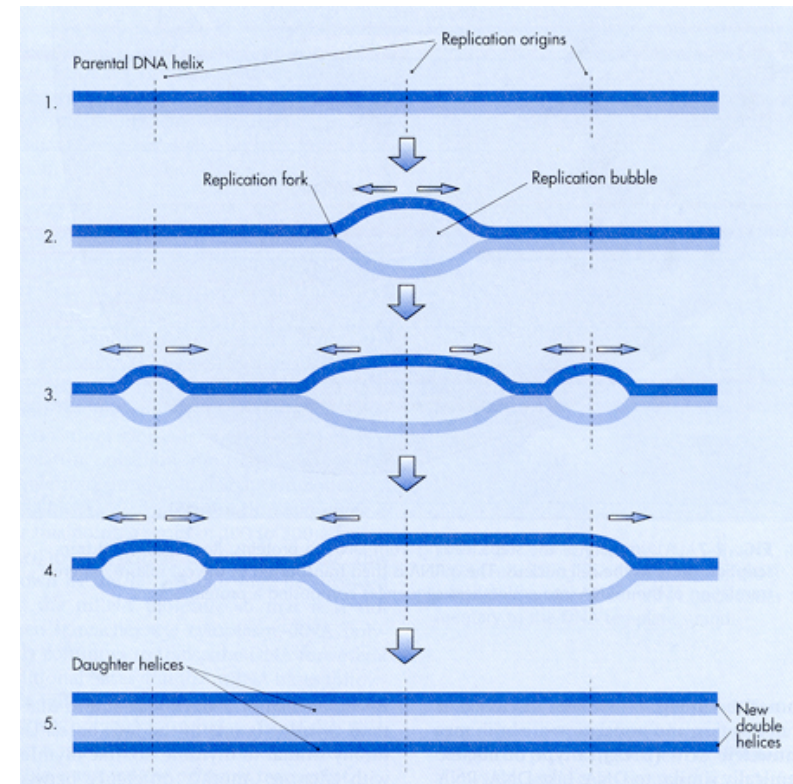
1. The weak hydrogen bonds between the strands break, leaving exposed single nucleotides.
2. The unpaired base will attract a free nucleotide that has the appropriate complementary base.
3. Several different enzymes are involved (unwinding helix, holding strands apart, gluing pieces back together, etc)
4. DNA Polymerase, a key replication enzyme, travels along the single DNA strand adding free nucleotides to the 3' end of the new strand (directionality of 5' to 3').
DNA Polymerase also proofreads the newly built strand in progress, checking that the next added nucleotide is in fact complementary (avoidance of mutations).
5. This continues until a complementary strand is built (semi-conservative model).

More About DNA Replication

The rate of DNA replication is relatively slow, about 40-50 nucleotides per second.

Recalling the length of DNA, it would take 2 months to replicate from one end to the other.

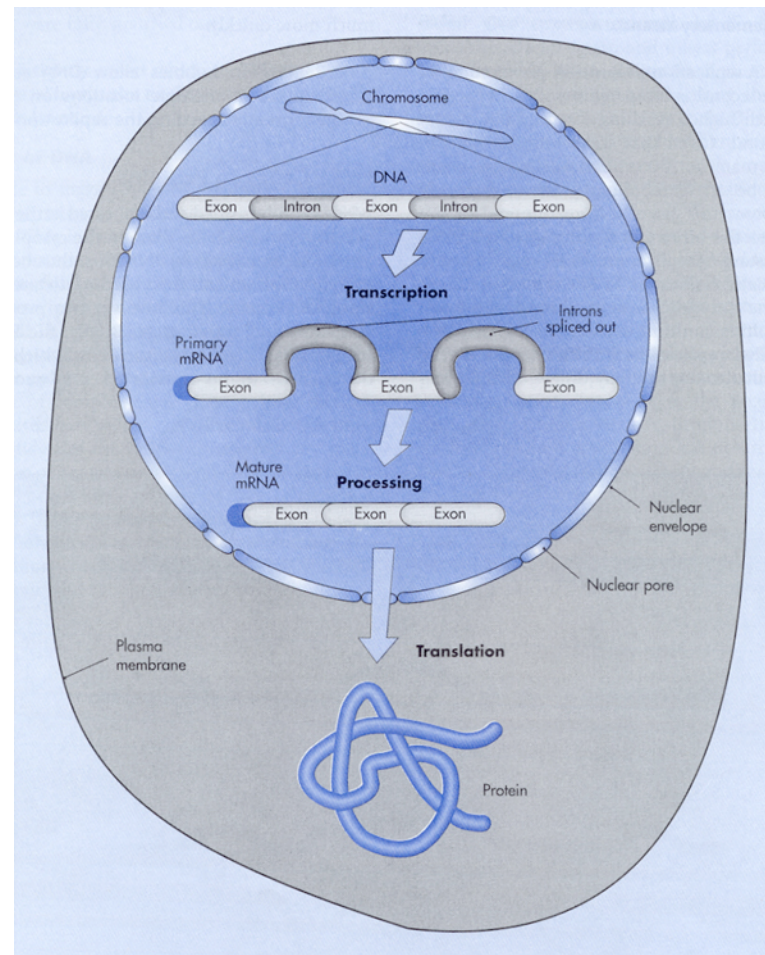
Nature overcomes this by having many replication start points: **replication origins**.



DNA's Purpose in Nature: Encoding Proteins

Before proteins can be assembled, DNA must undergo two processes:

- 1) Transcription
- 2) Translation



Step 1: DNA Transcription

- Process involves formation of messenger RNA sequence from DNA template.
- Although DNA is the same in all tissues, there are different promoters which are activated in different tissues, resulting in different protein products being formed.
- Gene splicing (removing introns) further modifies the sequences that are left to code, ultimately producing different protein products from the same gene.
- RNA polymerase enzymes bind to promoter site on DNA, pull local DNA strands apart.
- Promoter sequence orientates RNA polymerase in specific direction, as RNA has to be synthesized in the 5' to 3' direction (same linking pattern as DNA).
- One DNA strand is used preferentially as template strand, although either could be used.
- Post-transcriptional modifications (**5' methyl cap** and **poly-A-tail** protect mRNA from degradation).

Transcription Example

DNA double-strand sequence:

5'CAG AAG AAA ATT AAC ATG TAA 3'

3'GTC TTC TTT TAA TTG TAC ATT5'

mRNA sequence:

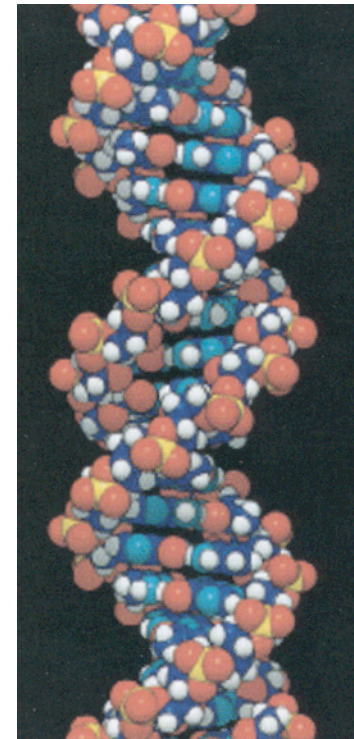
5' CAG AAG AAA AUU AAC AUG UAA3'

NOTE: same as template strand of DNA

A Crash Course in Genetics

General Overview:

- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR



Step 2: Translation & The Genetic Code

Proteins are made of polypeptides, which are in turn composed of amino acid sequences. The body contains 20 different amino acids, but DNA is made up of 4 different bases. Thus we need combinations of bases to denote different amino acids.

Amino Acids are specified by triplets of bases (**codons**):

	T	C	A	G
T	TTT Phe (F) TTC " TTA Leu (L) TTG "	TCT Ser (S) TCC " TCA " TCG "	TAT Tyr (Y) TAC " TAA Ter TAG Ter	TGT Cys (C) TGC " TGA Ter TGG Trp (W)
C	CTT Leu (L) CTC " CTA " CTG "	CCT Pro (P) CCC " CCA " CCG "	CAT His (H) CAC " CAA Gln (Q) CAG "	CGT Arg (R) CGC " CGA " CGG "
A	ATT Ile (I) ATC " ATA " ATG Met (M)	ACT Thr (T) ACC " ACA " ACG "	AAT Asn (N) AAC " AAA Lys (K) AAG "	AGT Ser (S) AGC " AGA Arg (R) AGG "
G	GTT Val (V) GTC " GTA " GTG "	GCT Ala (A) GCC " GCA " GCG "	GAT Asp (D) GAC " GAA Glu (E) GAG "	GGT Gly (G) GGC " GGA " GGG "

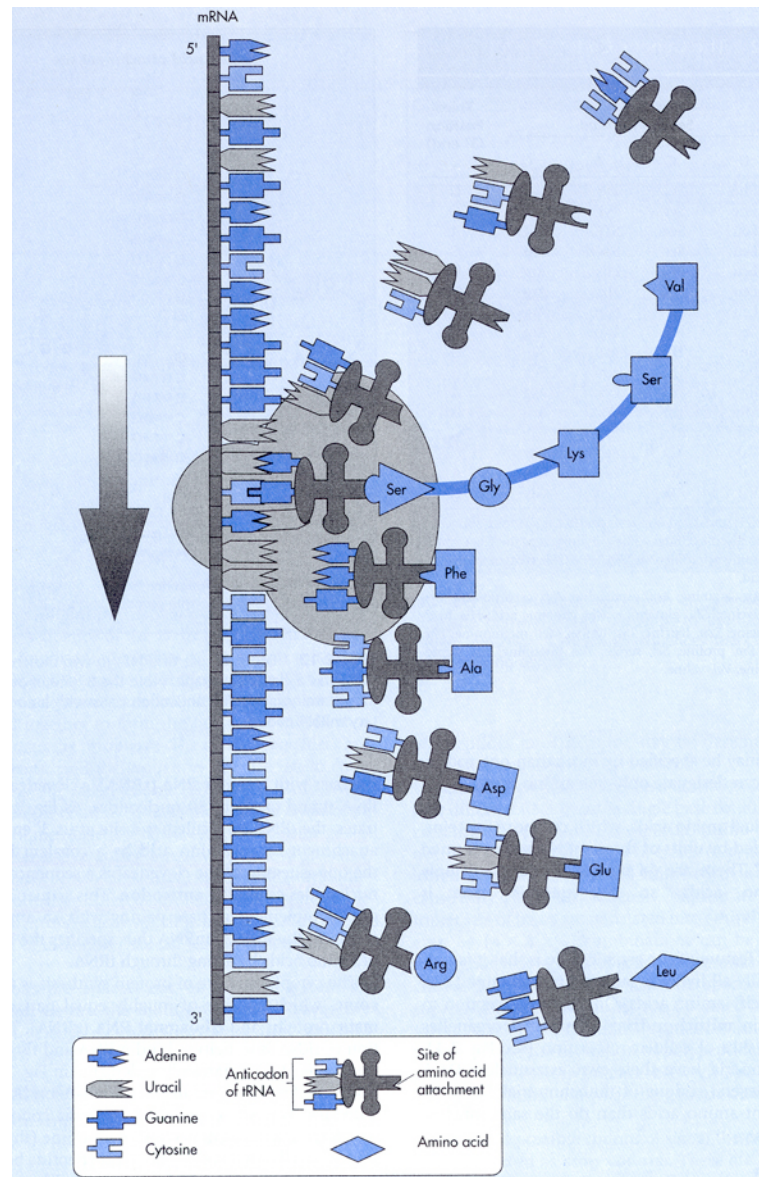
Translation (continued)

- Essentially, mRNA provides a template for the synthesis of a polypeptide (sequence of amino acids).
- mRNA cannot directly bind to amino acids, but instead interacts with tRNA (transfer-RNA), which has a binding site for an amino acid, and a sequence of three nucleotides on another side (**anticodon**).
- mRNA thus specifies amino acid sequence by acting through tRNA
- The site of translation in the cytoplasm is on a **ribosome**, which contains enzymatic proteins (linking amino acids together) and **ribosomal RNA (rRNA)**.
- rRNA helps to bind mRNA and tRNA to the ribosome.

Sequence of Events:

1. Ribosome first binds to initiation site on mRNA sequence (AUG =start), specifying amino acid methionine
2. Ribosome then draws corresponding tRNA (with attached methionine) to its surface, allowing base pairing between tRNA and mRNA
3. Ribosome moves along mRNA sequence codon by codon in 5' to 3' direction until it reaches a STOP codon. The ribosome releases, and we have a polypeptide!

How Translation Works ...



Example continued:

mRNA sequence:

5' CAG AAG AAA AUU AAC AUG UAA 3'

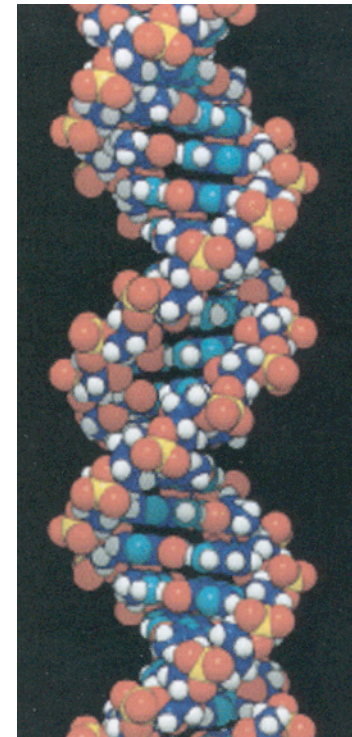
amino acid sequence (using Genetic Code):

Gln-Lys-Lys-Ile-Asn-Met-STOP

A Crash Course in Genetics

General Overview:

- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR



Predictability of Protein Folding

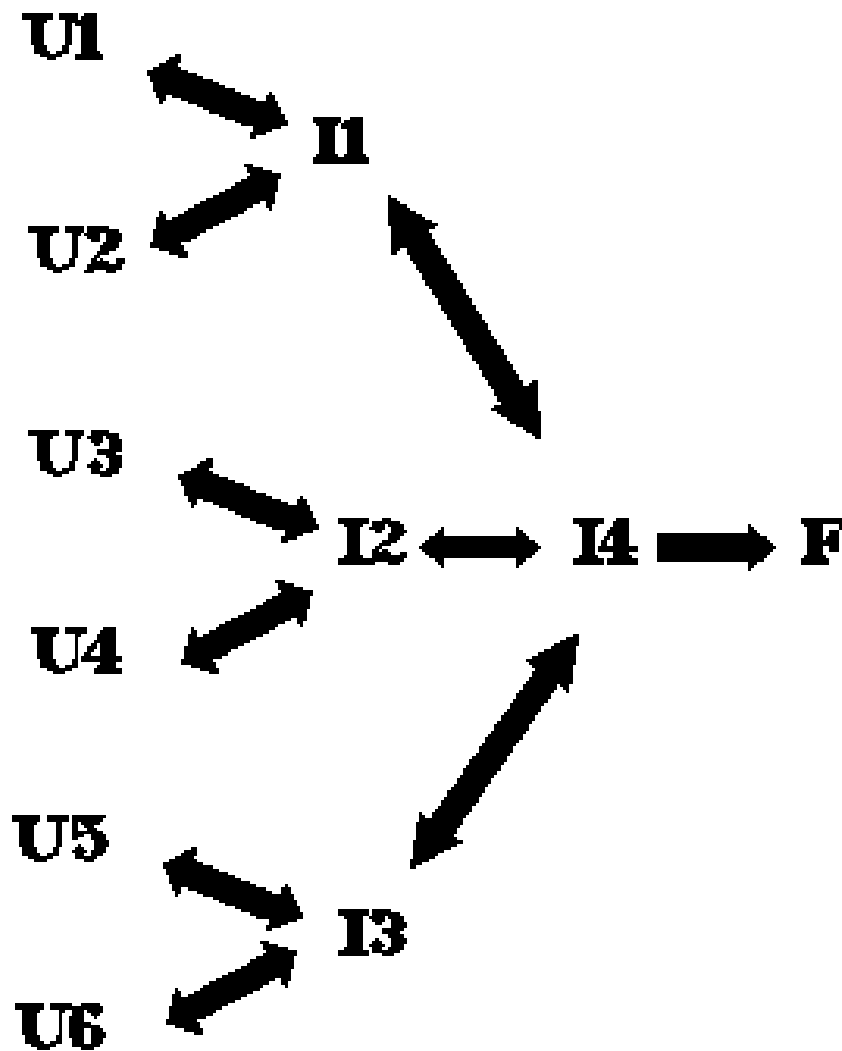
Although protein structure can be determined relatively easily using various crystallography and spectroscopy methods, as of last year, it is impossible to predict protein folding based on the primary amino acid sequence.

Proteins do follow rules in folding, but which rules they apply are unpredictable.

Rules Include:

- Interior is densely packed.
- Minimal exposure of nonpolar groups
- Backbone of polar groups are buried
- Folding with minimal conformational strains preferred
- Elements of secondary structure that are adjacent in sequence tend to be adjacent in tertiary structure

One Model of Protein Folding

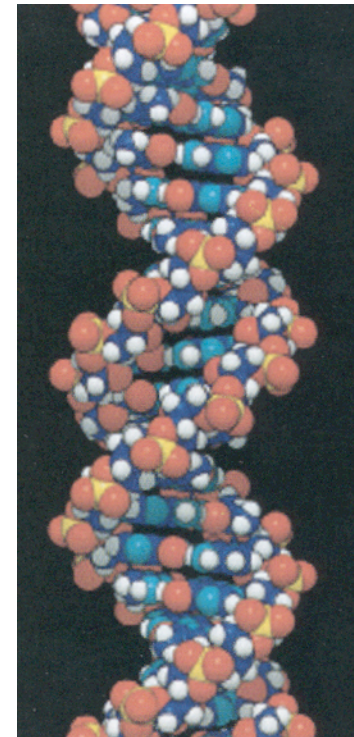


- Many (10^{16}) different unfolded states (U) quickly equilibrate to a small number of partially folded, marginally stable intermediates (I).
- Kinetic restraints under refolding conditions cause (U) to converge to a common folding pathway.
- Intermediates have a preference for partially folded conformations.
- Last transition from I4 to F is a slow equilibrium with a nearly folded transition state.

A Crash Course in Genetics

General Overview:

- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR



Types of DNA

Fewer than 10% of the three billion nucleotide pairs in the human genome actually encodes proteins.

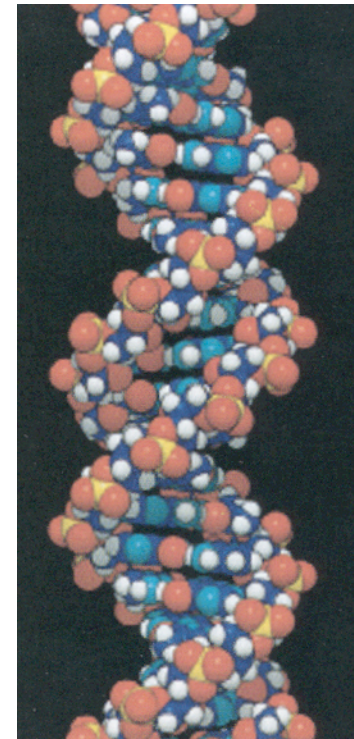
There are several categories of DNA:

1. Single copy DNA - seen only once in a cell, makes up about 75% of the genome, includes protein-coding genes. Most of this DNA is found in introns or in sequences that lie between genes.
2. Dispersed Repetitive DNA - as name suggests, this repetitive DNA is scattered singly throughout the genome.
3. Satellite DNA - repetitive DNA found in clusters around certain chromosome locations. Called so because they can be easily separated by centrifugation.
 - Makes up about 10% of genome.
 - Highly variable, source of differentiation between people.

A Crash Course in Genetics

General Overview:

- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR



Manipulating DNA: Laboratory Uses

Measuring length of DNA molecules:

Knowing the atomic weight of a nucleotide, and markers, *gel electrophoresis* separates pieces of DNA by weight, with the heavier (longer) segments moving slower and the lighter (shorter) segments moving faster through the gel.

These bands are compared with “markers” (pieces of DNA with known molecular weights and lengths), which are run simultaneously with the segments to be measured.

Nature's Secret: Denaturation & Renaturation

Recall that the two strands of DNA are held together by weak hydrogen bonds.

Thus, heating the double-strand DNA increases the kinetic energy, breaking these bonds/

A=T-rich regions separate first (recall, two H-bonds between A and T as opposed to three bonds between G and C).

This property allows researchers to estimate the relative AT vs GC content in a segment of DNA, according to how quickly the DNA denatures

If the temperature is lowered again *slowly*, the DNA can renature.

Process must be done slowly so that correct base pairing can occur.

Consequently, DNA denaturation is REVERSIBLE and is useful in the laboratory!

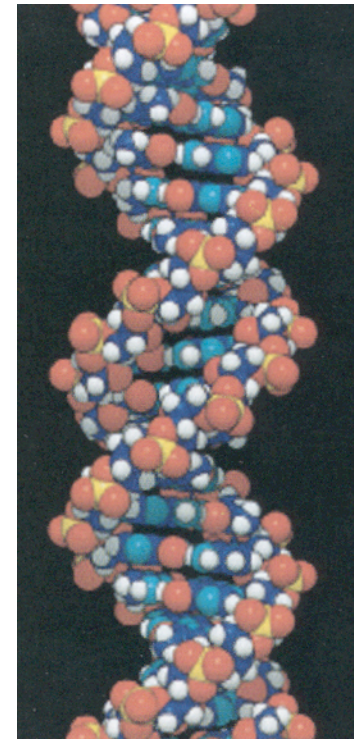
Other Operations on DNA

- **DNA can be lengthened**, through polymerases, providing there is a primer (existing sequence partially bonded to a template) and a free 3' end to which bases can be added.
- **DNA can be shortened**, via DNA nucleases:
 - *Exonucleases* (like Pacman, eating from one side to the other) cut from the ends, removing one nucleotide at a time.
 - *Endonucleases* (like a pair of scissors taken to the middle of a strip of paper) cut from the inside, leaving either “sticky ends” or blunt ends.
 - *Restriction endonucleases* are most useful in genetics research. They cut at specific sites, and only cut ds DNA.

A Crash Course in Genetics

General Overview:

- DNA Structure
- RNA
- DNA Replication
- Encoding Proteins
- Protein Folding
- Types of DNA
- Manipulating DNA
- PCR



Multiplying DNA — What is PCR?

PCR = Polymerase Chain Reaction

Problem:

To be able to use DNA segments in the laboratory, one often needs multiple copies of the segment.

Nature's solution (DNA replication) is too slow, and requires *in vivo* conditions.

Purpose:

Some potential uses for many copies of DNA include:

- *Forensics* - identifying the guilty party through genetic analysis.
- *Genetically Inherited Diseases* - some diseases are inherited through a mutation of a single gene. The presence of that gene could be detected using PCR to exaggerate its presence, allowing detection.

The Scientist's Solution: PCR

PCR (polymerase chain reaction) is a laboratory-based method of immitating nature's DNA replication.

We need:

1. **Two primers**, each 15-20 bases long (oligonucleotides), corresponding to the DNA sequences on either side of the sequence of interest
2. **DNA polymerase**, a thermally stable form (thermophilic bacterium origin) to mimic DNA replication
3. A large collection of free DNA **nucleotides**
4. A **template strand** (Genomic DNA from an individual)

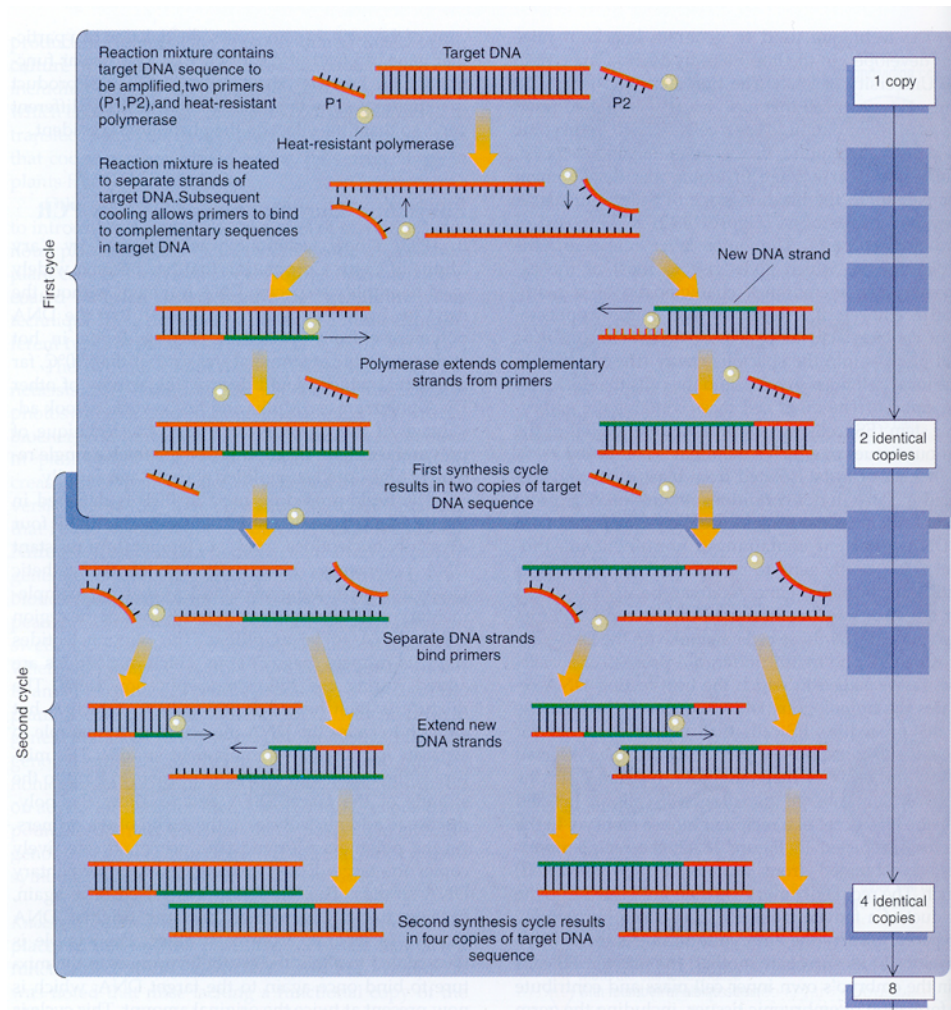
PCR — How It Works

1. Heat the genomic DNA to denature, resulting in a single stranded template.
2. Expose the DNA to the primers, allowing them to hybridize (under cooler conditions) to the appropriate locations on either end of the sequence of choice.
3. Reheat the DNA to an intermediate temperature and expose the mixture to free DNA bases, allowing a new DNA strand to be synthesized by DNA polymerase.

This results in a double stranded sequence of DNA.

4. Heat the double-stranded DNA to a high temperature, causing it to denature.
5. Repeat steps 2-4 to multiply sequence of choice.

Schematic Representation of PCR



The Pros and Cons of PCR

Benefits and Advantages of PCR

- The heating-cooling cycle takes minutes, allowing amplification of sequence to occur quickly.
- PCR can be used with extremely small quantities of DNA (blood stain, single hair, saliva on postage stamp)
- DNA produced is very pure, thus do not need radioactive probes to detect certain sequences or mutations.

Downfalls of PCR

- Primer synthesis requires knowledge of the DNA sequence around the DNA segment of interest.
- PCR is extremely sensitive, therefore prone to contamination in laboratory.
- Limited to short sequences (1000 bases) only

References

- Julie Stromer, *A Crash Course in Genetics*, CPSC 601.73 (W2002) presentation
- Alberts, B., A. Johnson, et al. (2002). Molecular Biology of the Cell. New York, Garland Science.