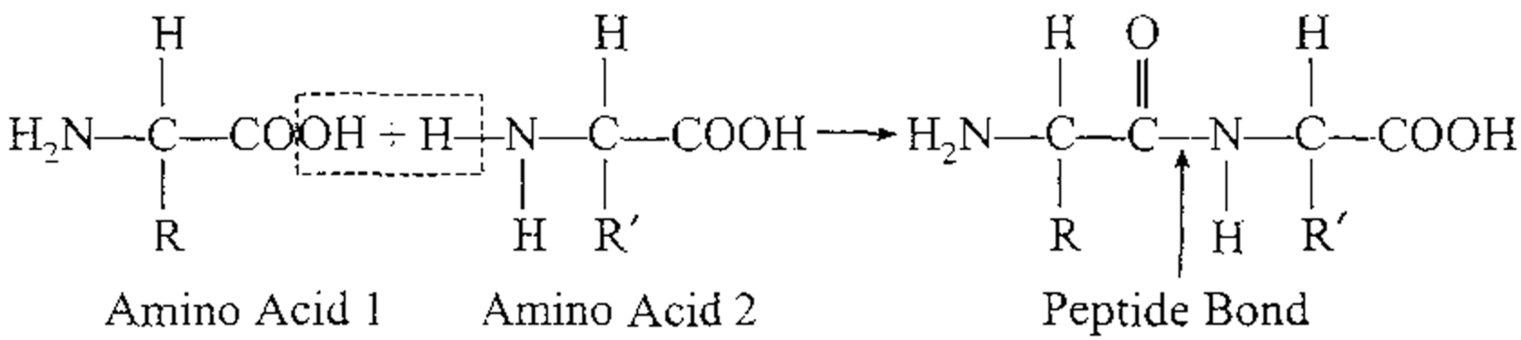
Lecture 10

Cells:
Its Organelles and Molecules:

Proteins and DNA

Proteins

- Proteins are the most important molecules in organism,
- 45-50% by dry weight
- In humans there are more than a hundred thousand proteins
- Proteins are composed of amino acids linked together into a long chain by peptide bonds
- Peptide bond is a bond between COOH group of one amino acid and NH2 group of another amino acid



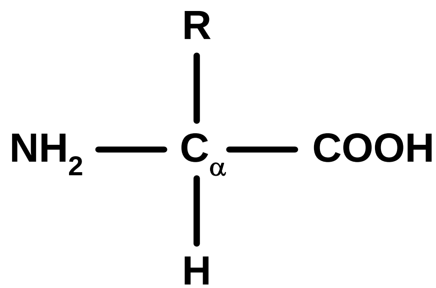
- Protein is a chain of at least 40 or more amino acids
- Proteins are high molecular weight molecules with molecular mass from 4000 to several million

Amino Acids

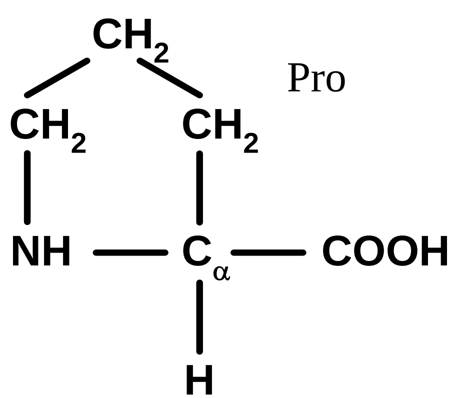
- **Amino acids** (~200 aminoacids identified, 60 are found in humans)

- Fundamental units of proteins – alpha **amino acids**
- There are 20 major amino acids, that make proteins
-

- Nineteen
- have general
- structure of amino acid:

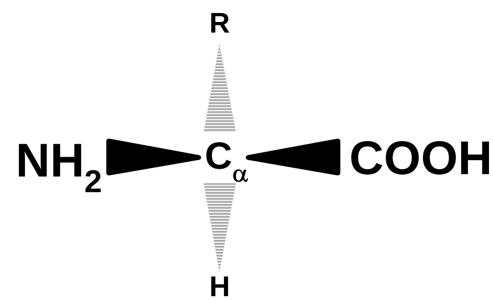


- Only **Proline**
- is imino acid:



Pro

Amino acids are stereoisomers:
Can exist in both L-and D-form
Natural amino acids all are **L-stereoisomers**



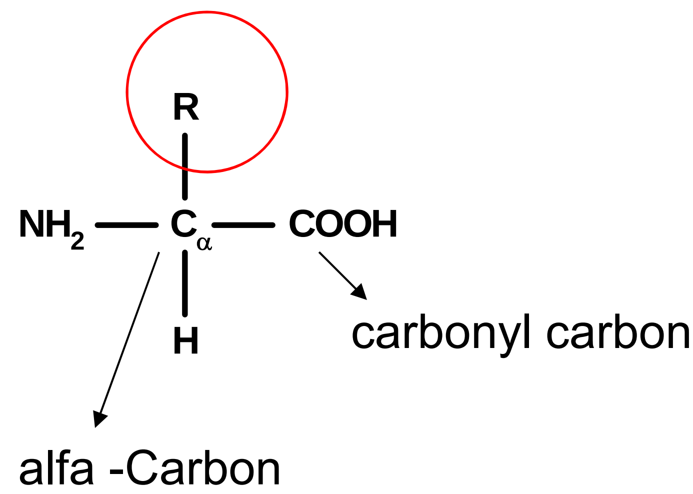
Classification of amino acids

Depends on R (side group)

Acidic – contain additional COOH in R

Basic – contain basic group NH2 in R

Polar or non-polar groups in R



Essential – cannot be synthesized in the organism– brought with food

Half essential – produced in the organism but not enough

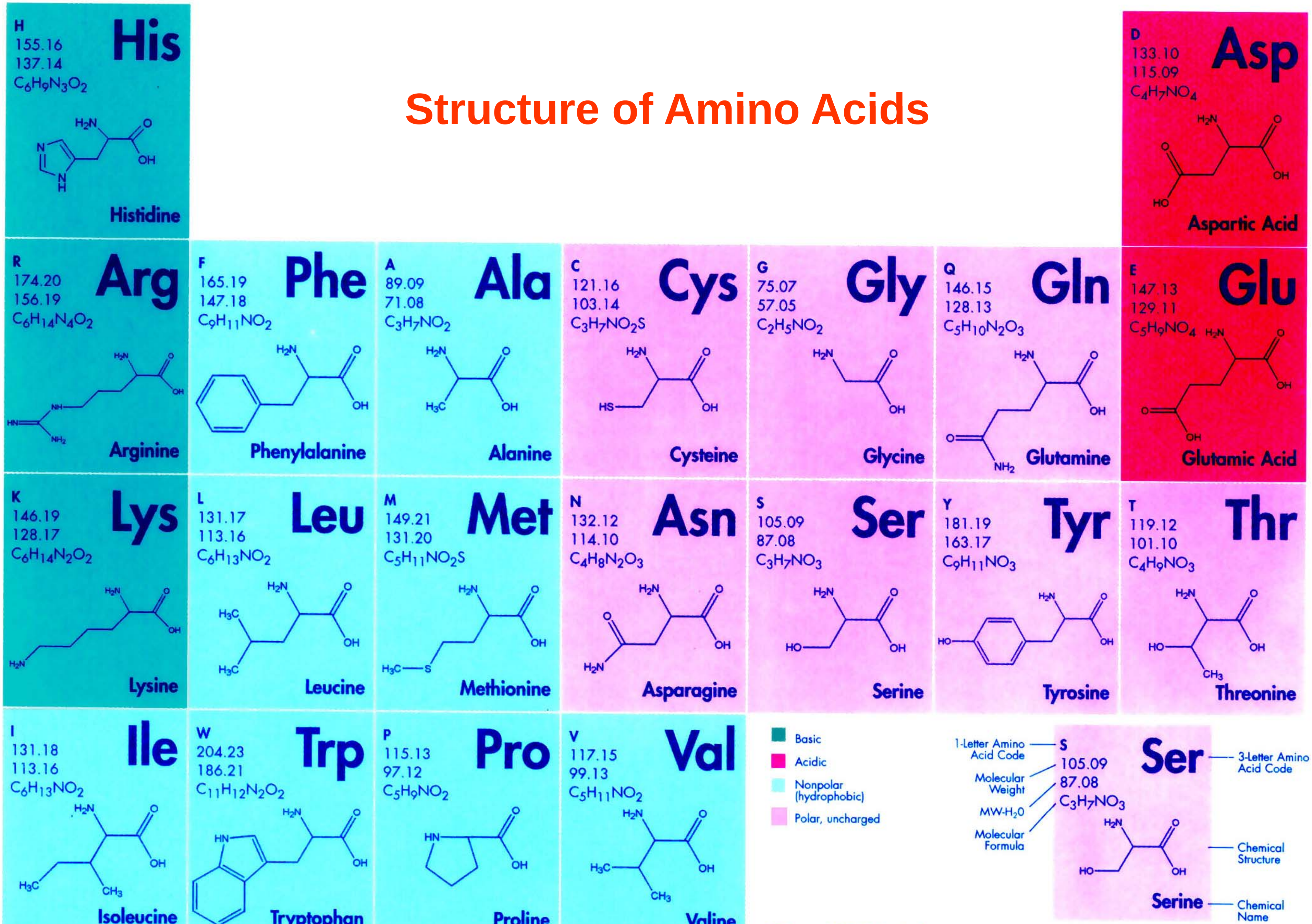
Non-essential are synthesized in organism in sufficient amount

Twenty important amino acids

Non-polar Groups (MW)	Uncharged polar groups (MW)	Charged polar groups (MW)
Alanine (Ala, 89)	Glycine (Gly,75)	Acidic
Valine (Val,117)	Serine (Ser,105)	Aspartic Acid (Asp,133)
Leucine (Leu,131	Threonine (Thr,119)	Glutamic Acid (Glu,147)
Isoleucine (Ile,131)	Cysteine (Cys,121)	Basic
Proline (Pro,115)	Tyrosine (Try,181)	Lysine (Lys,146)
Phenylalanine (Phe,165)	Aspargin (Asn,132)	Arginine (Arg,174)
Triptophan (Trp, 204)	Glutamine (Gln,146)	Histidine (His,155)
Methionine (Met, 149)		

Number in bracket is molecular weight

Structure of Amino Acids



Structure of molecules

Configuration of molecules – can not be changed without braking the bond

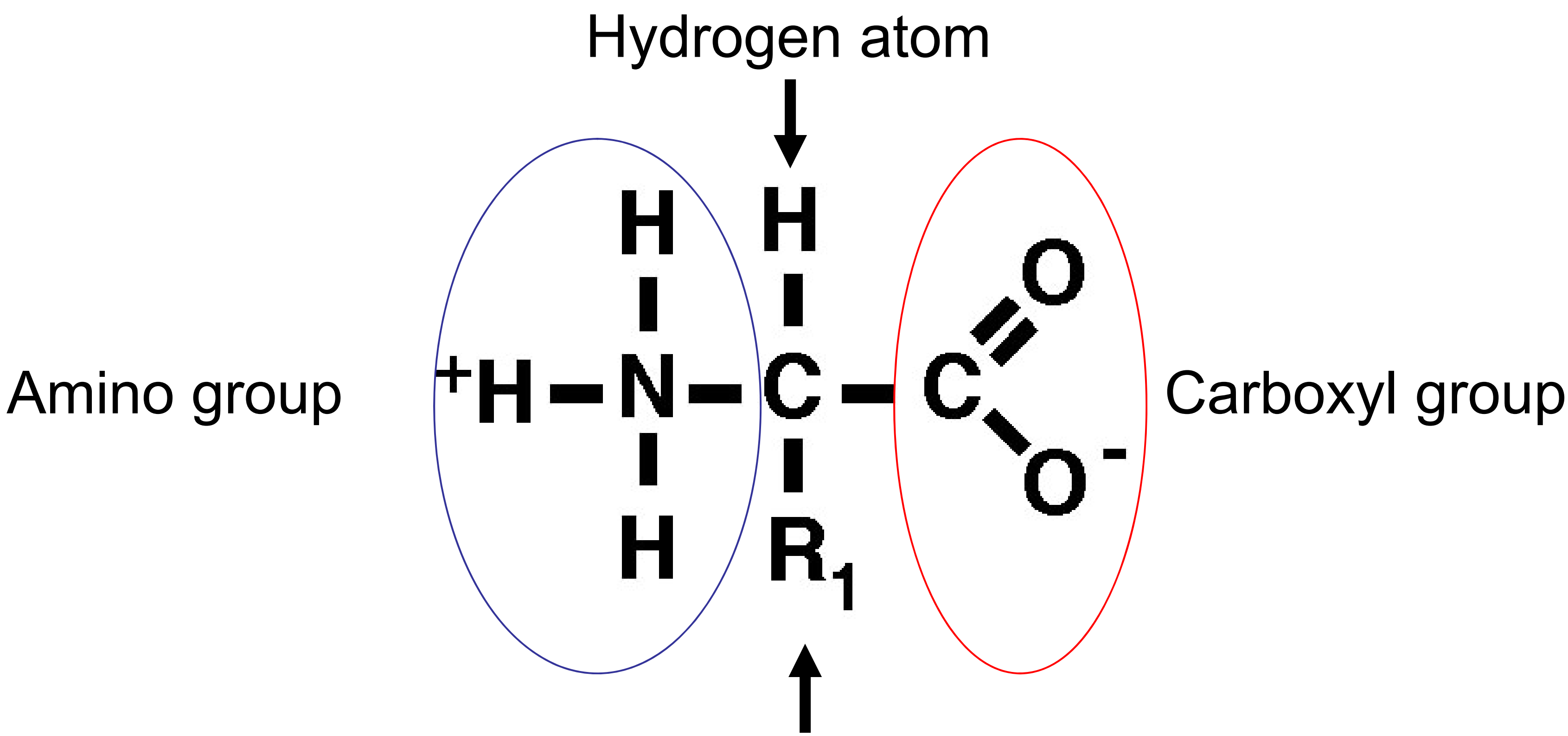
Conformation of molecules – can be changed by rotation around the bond

Levels of structure

1^{o} primary – sequence
2^{o} secondary – local structure
3^{o} tertiary – global 3D structure
4^{o} quaternary – arrangement of subunits

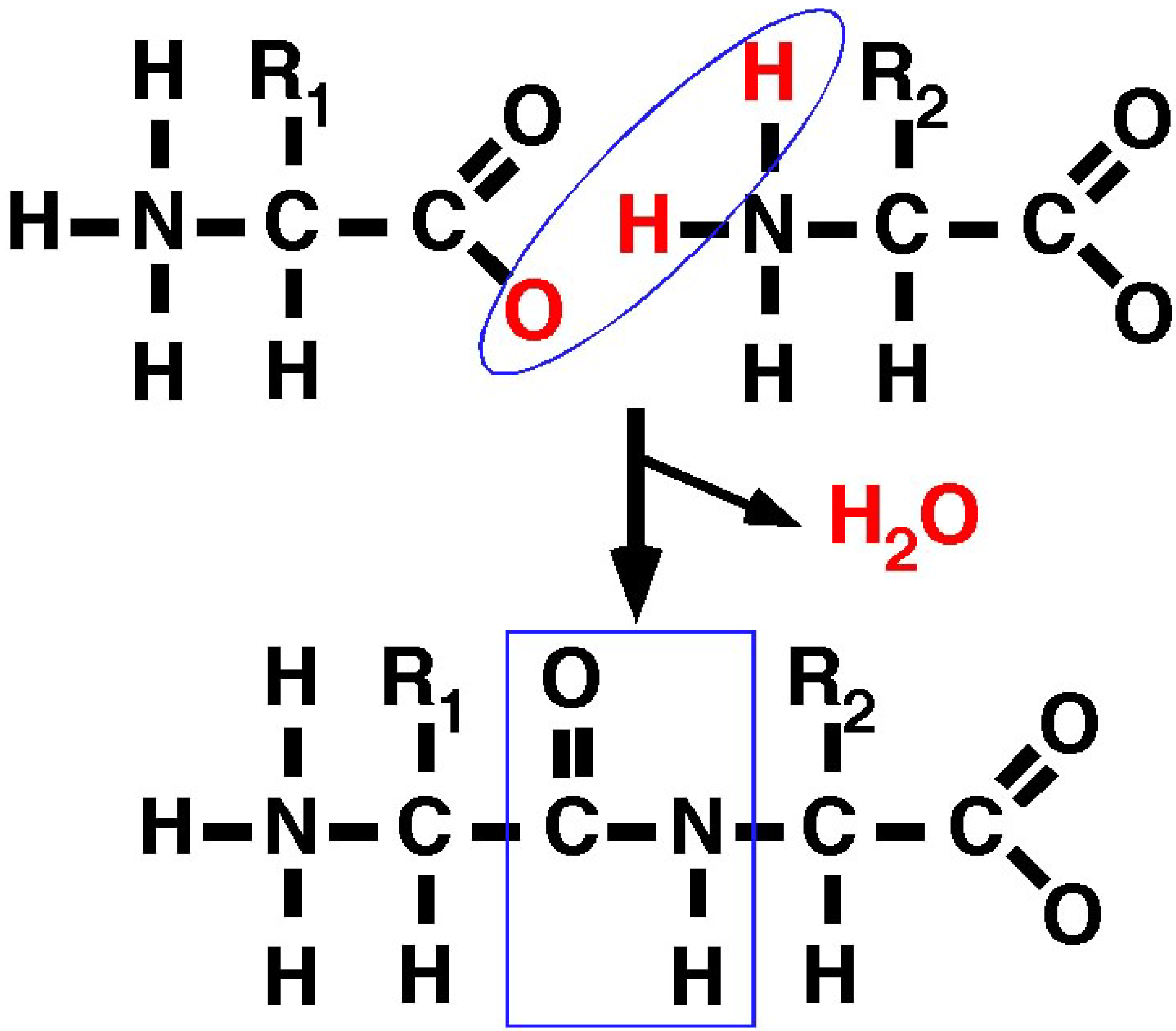
Amino acids are the building blocks of proteins:

Common to all:



Distinguishing feature: R group; side chain conveys specific chemical properties

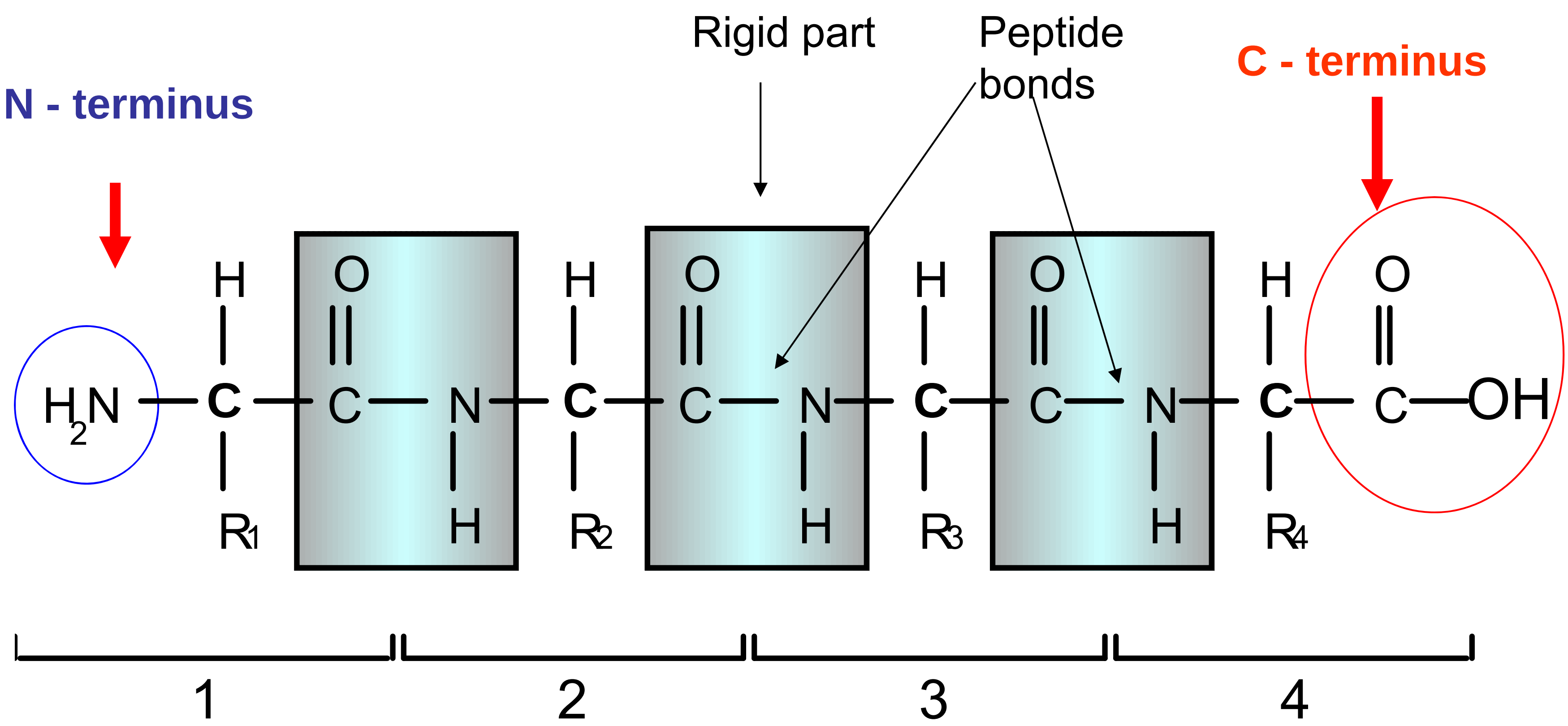
Protein synthesis--chain elongation mechanism



Proteins are formed when two amino acids react to form a peptide bond

Protein structure

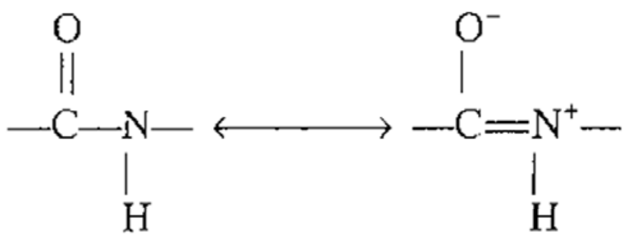
Primary structure – linear sequence of amino acids



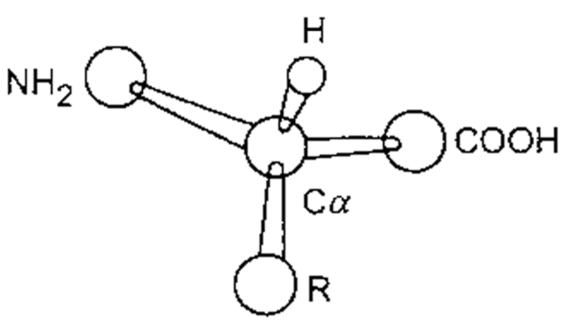
In the free form at neutral pH amino acids are zwitterions: **$NH3^{+}$—CHR—COO^{-}**
Four amino acids have charged side chains at pH 7: **Asp** & **Glu** are negative
Lys & **Arg** are positive

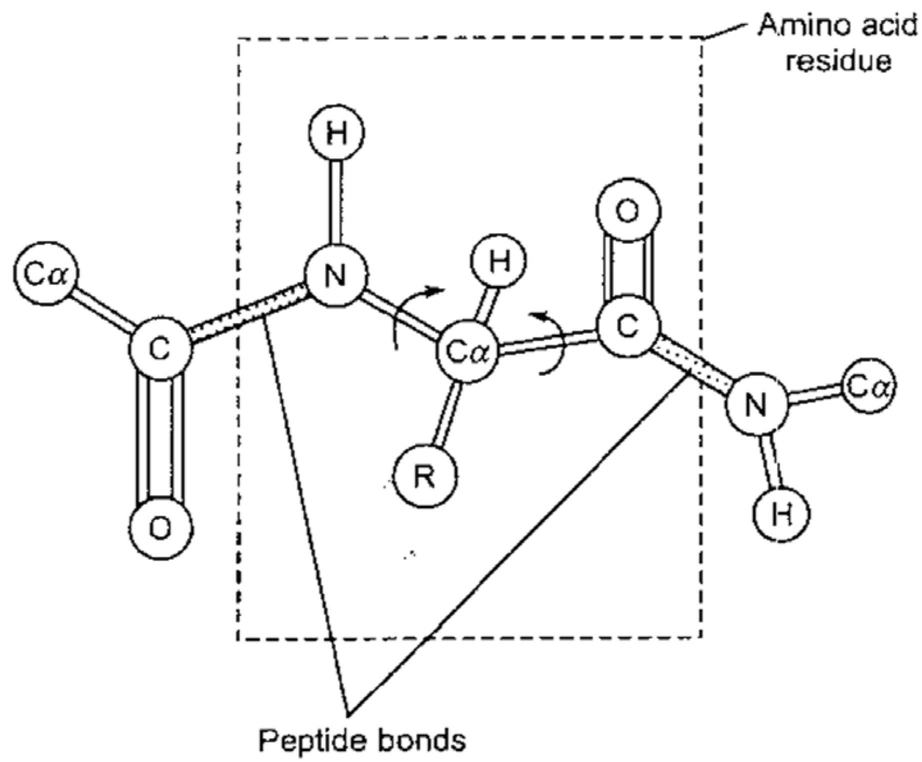
• Secondary structure

- Peptide unit is rigid and planar –no rotation around peptide bond (C-N bond, which is next to double bond) - it is partially double bond



Link between α-C and carbonyl C and link between α-Carbon and N are single bonds – rotation is possible





This results in spatial structures which are stabilized by intra peptide H-Bonds between CO and NH groups

- Primary structure is linear – long chain of amino acids connected by peptide bonds, the number and sequence of amino acids varies.

- Biological activity and functions are defined by the higher spatial structures.
- In fact, stretched or randomly arranged polypeptide chain is biologically inactive.

- **Linear structure** ⟶ **3D structure**

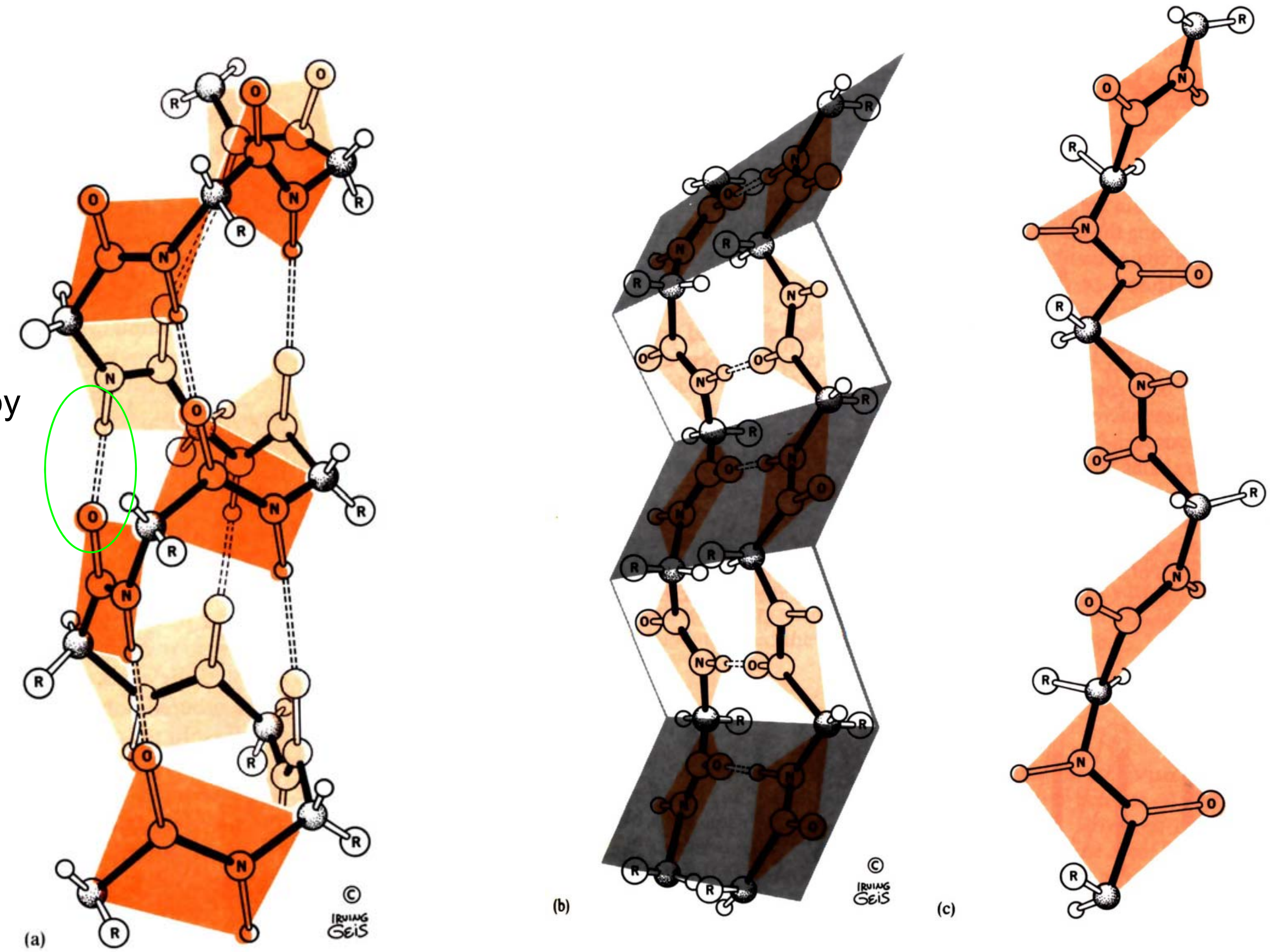
  inactive biologically active

- **Secondary structure** arises due to rotation around single C-N and C-C bonds and is stabilized by ionic and H-bond interactions inside the molecule or between two peptide chains

Alfa - helix or beta sheet - typical secondary structures

Stabilized by
H-bonds

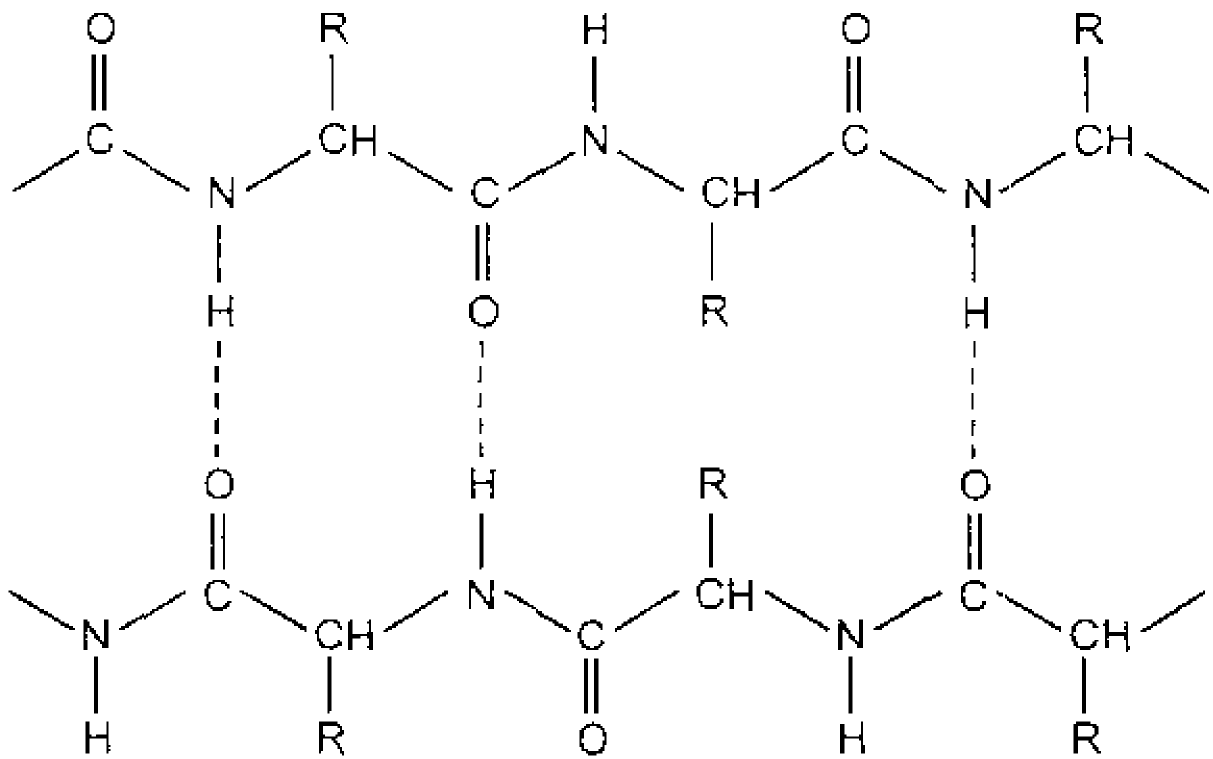
NH ….CO



right-handed α-helix β-sheet strands left-handed polyproline II helix

Cantor & Schimmel, 1980

Hydrogen bonds in Beta sheets



Beta - sheet structure is characterized by linking two or more linear polypeptide chains by hydrogen bonds

Misfolding from alpha helix to beta-sheets results to disease:

Alzheimer’s, Parkinson’s

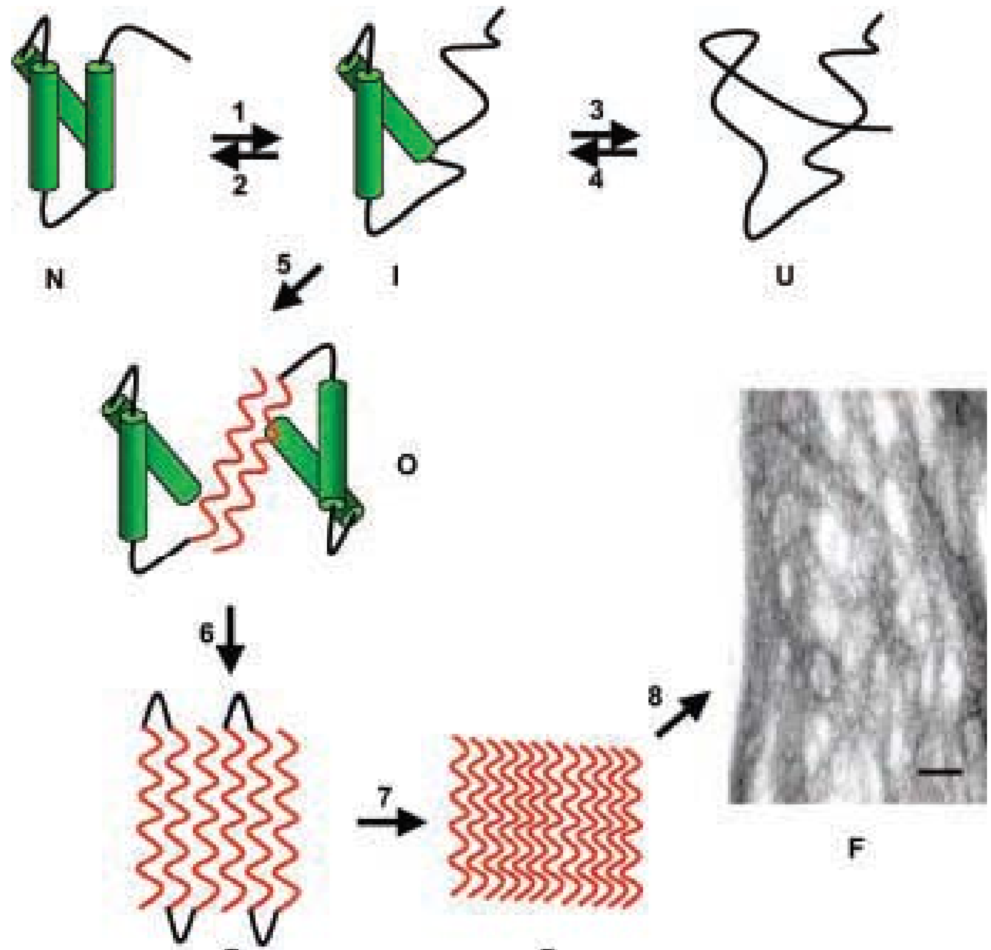
Type II Diabetes

Mad Cow disease

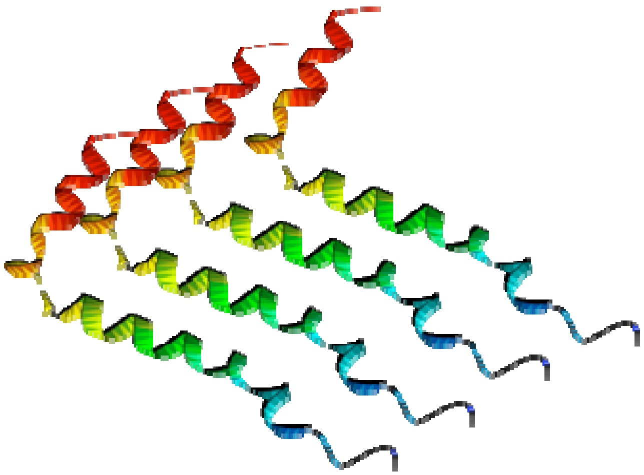
and many others

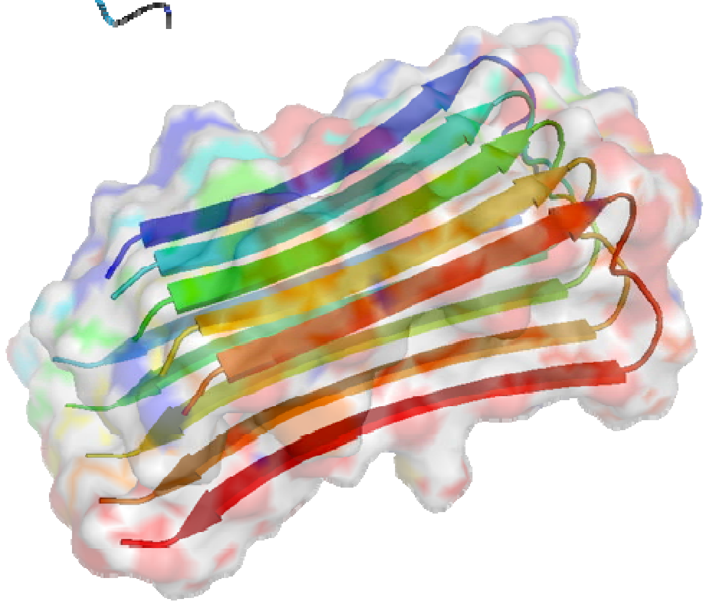
Proposed amyloid fibril formation:

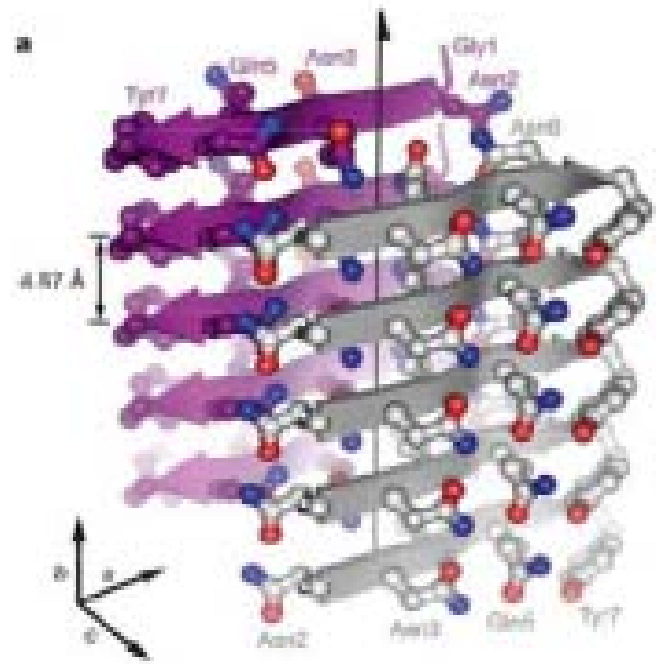
involves uncoiling of alfa-helix and formation of beta-sheets



Amyloid beta (1-42)

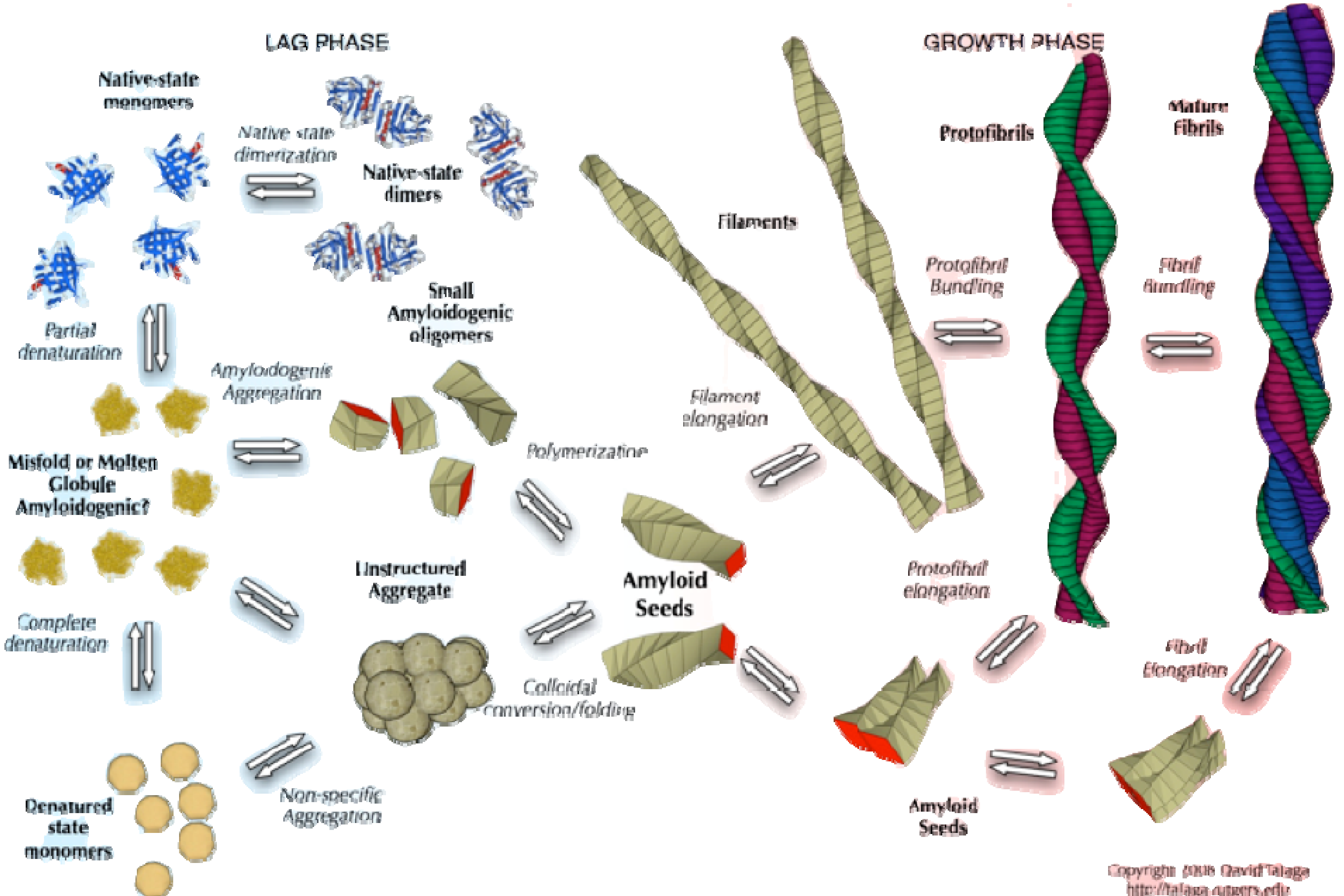






Crescenzi et al, 2002 (RCSB PDB)

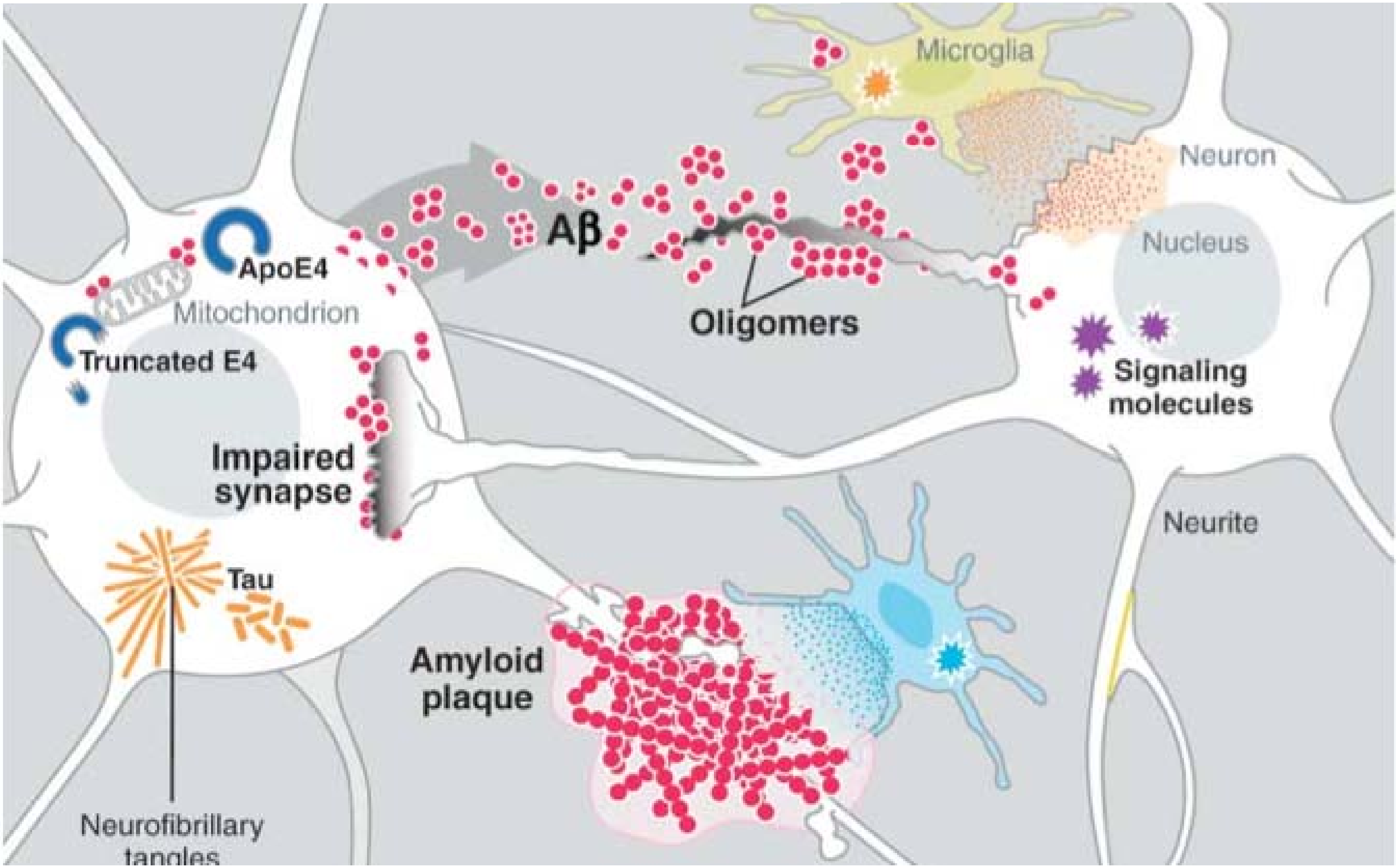
Kage (2008), Nelson et al, *Nature,* 2005


LAG PHASE
GROWTH PHASE
Native-state monomers
Native state dimerization
Native-state dimers
Partial denaturation
Small Amyloidogenic oligomers
Amyloidogenic Aggregation
Misfold or Molten Globule Amyloidogenic?
Polymerization
Filaments
Filament elongation
Protofibril Bundling
Protofibrils
Fibril Bundling
Mature Fibrils
Complete denaturation
Unstructured Aggregate
Amyloid Seeds
Colloidal conversion/folding
Protofibril elongation
Fibril Elongation
Denatured state monomers
Non-specific Aggregation
Amyloid Seeds

(Talaga, 2008)

Science AAAS

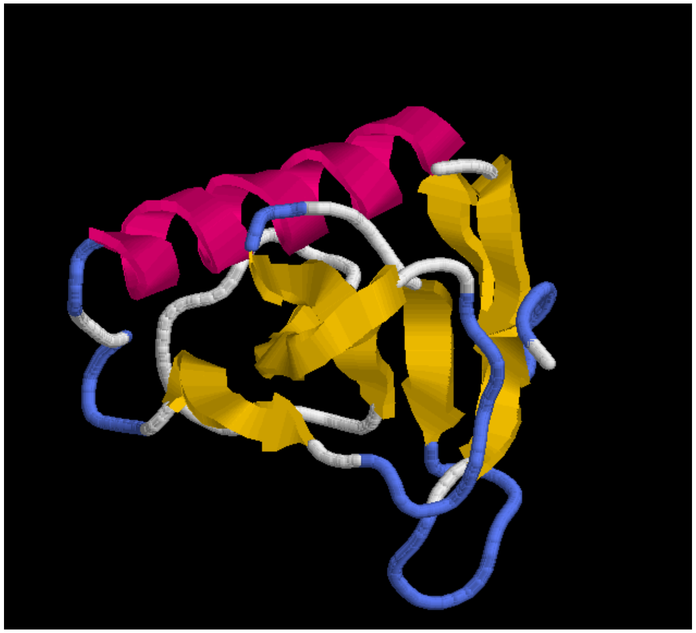
100 Years and Counting: Prospects for Defeating Alzheimer's Disease
Erik D. Roberson, *et al.*
Science **314**, 781 (2006);
DOI: 10.1126/science.1132813



Roberson & Mucke, 2006

Conformations

- **Tertiary structure** : three dimensional structure resulting out of folding of protein into globules due to interactions between different segments which are far apart
- This defines biological activity – only in particular conformation given protein binds to a particular structure

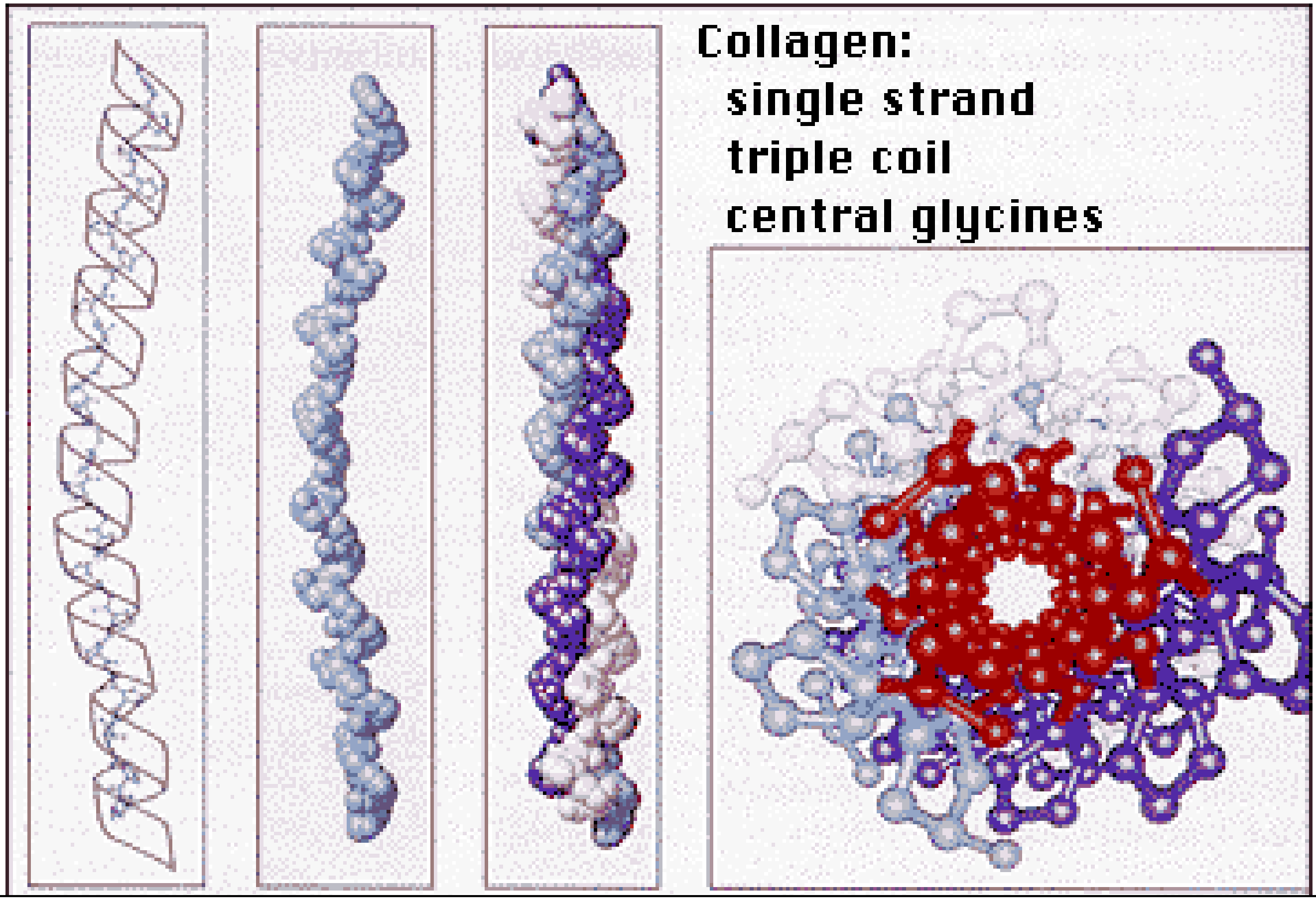


RNAse F1 **glutathione reductase**

Conformations

- **Quaternary Structure** : Proteins often have more than one polypeptide chain (subunit):

- Collagen:
- 3 subunits
- Hemoglobin:
- 4 subunits



High tensile strength – load of 10 kg to break a collagen fiber 1nm in diameter

- **Summarize:**
- **Primary structure** – amino acid sequence
- **Secondary structure** - due to H-bonds between inter-peptide CO and NH, and between two peptide chains
- **Tertiary structure** – H-bonds between distant segments
- **Quaternary structure** – two or more polypeptide chains in tertiary forms make one protein

- **Structure of proteins. Available Internet resources**
- Molecular Biology Server @ **http://us.expasy.org/**
- Protein Structure Laboratory, Kinemage @ **http://kinemage.biochem.duke.edu/**
- Molecular Visualization Freeware @ **http://www.umass.edu/microbio/rasmol/**
- Protein Data Bank @ **http://www.pdb.org/**

Physical properties of proteins

- Defined by primary, secondary structures as well higher 3 D organization:
- Proteins can be acidic or basic or hydrophobic
- Due to bulk globular structures and inability to diffuse through membrane create osmotic pressure
- Due to globular structure and various conformations proteins increase the viscosity of solutions, can even force gelation of the solution, characteristic for fibril proteins. Collagen protein, found in bones and cartilage, is a fibril protein and exists in a gel phase – very strong elastic fibers
- Proteins can be hydrophilic – bind water molecules around them and form hydration shell
- Temperature, mechanical force, ionization, and some chemical agents can break H-bond and protein unfolds into primary chain and becomes inactive this is called **denaturation**

Biological function of proteins

- Enzymes – act as catalysts to accelerate certain chemical processes in cells
- Carriers – transport other molecules in blood and in cells (for example oxygen, lipids)
- Sense receptor molecules are proteins, they transform:
- light into chemical energy - vision
- chemical energy into mechanical energy in muscles
- Immune system – antibody binds foreign substances and protect the cell –example immunoglobulines
- Structural proteins – constituents of hair, skin (keratin), bones (collagen)
- Participate in Genetic information transfer – regulate replication of DNA and assist in transfer information from DNA to RNA

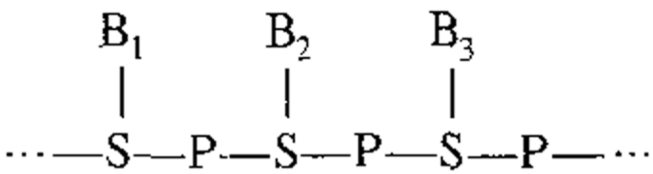
Nucleic Acids

- Largest molecules
- Fundamental constituents of genetic material

- **Roles of nucleic acids**

- Storage of genetic information (DNA, RNA)
- Protein synthesis (messenger RNA, ribosomal RNA, transfer RNA)
- Enzymes
  RNA with enzymatic properties: ribozymes, ribonuclease P;

Configuration (primary sequence)

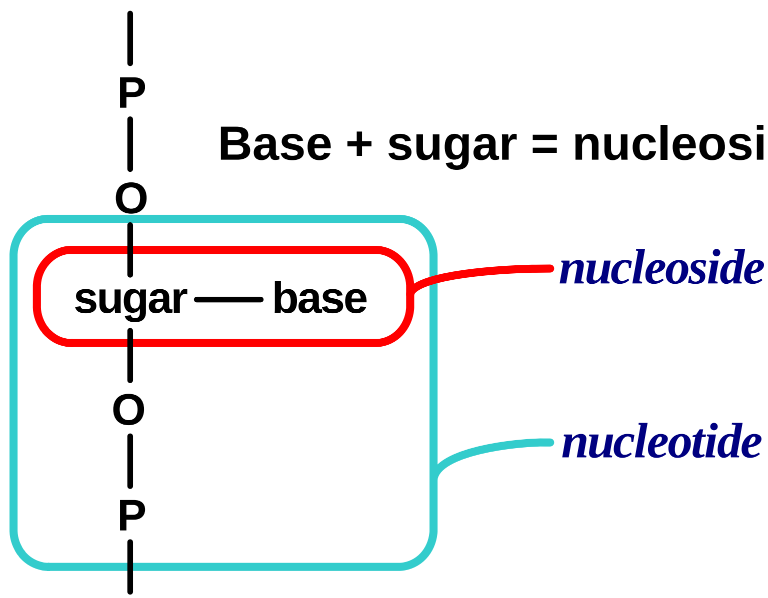


Sugar determines type of polymer: DNA or RNA

Four bases determine the sequence

Bases: pyrimidines and purines:

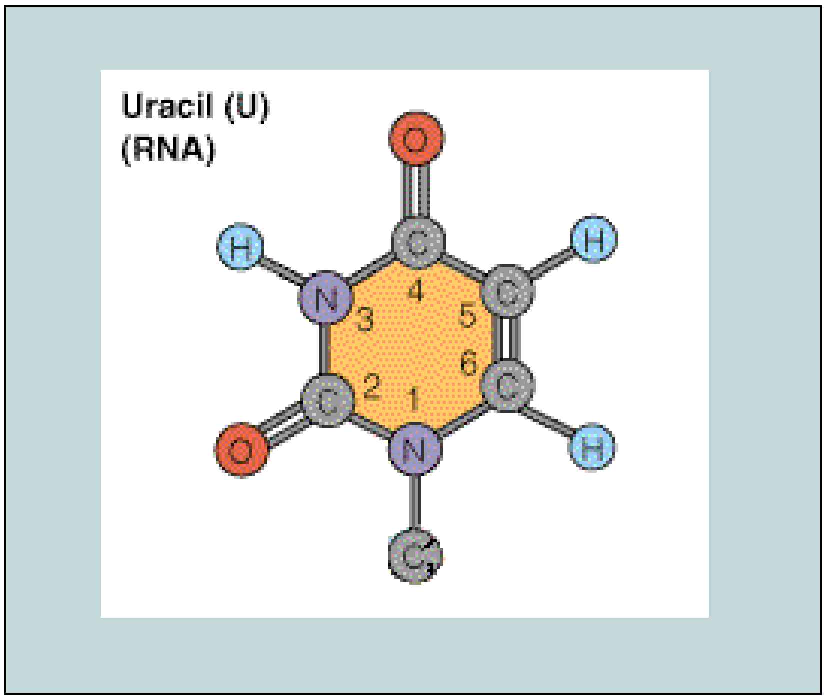
Base + sugar = nucleoside

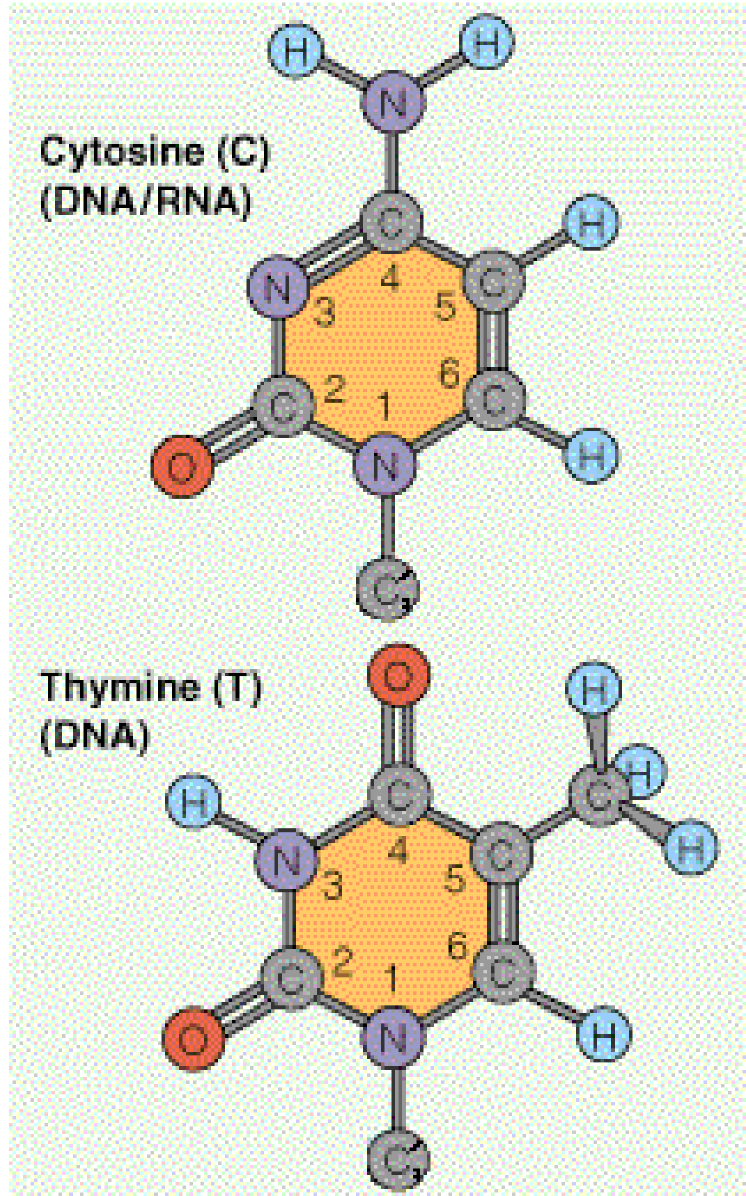


Nucleoside + phosphate group = <u>nucleotide</u>

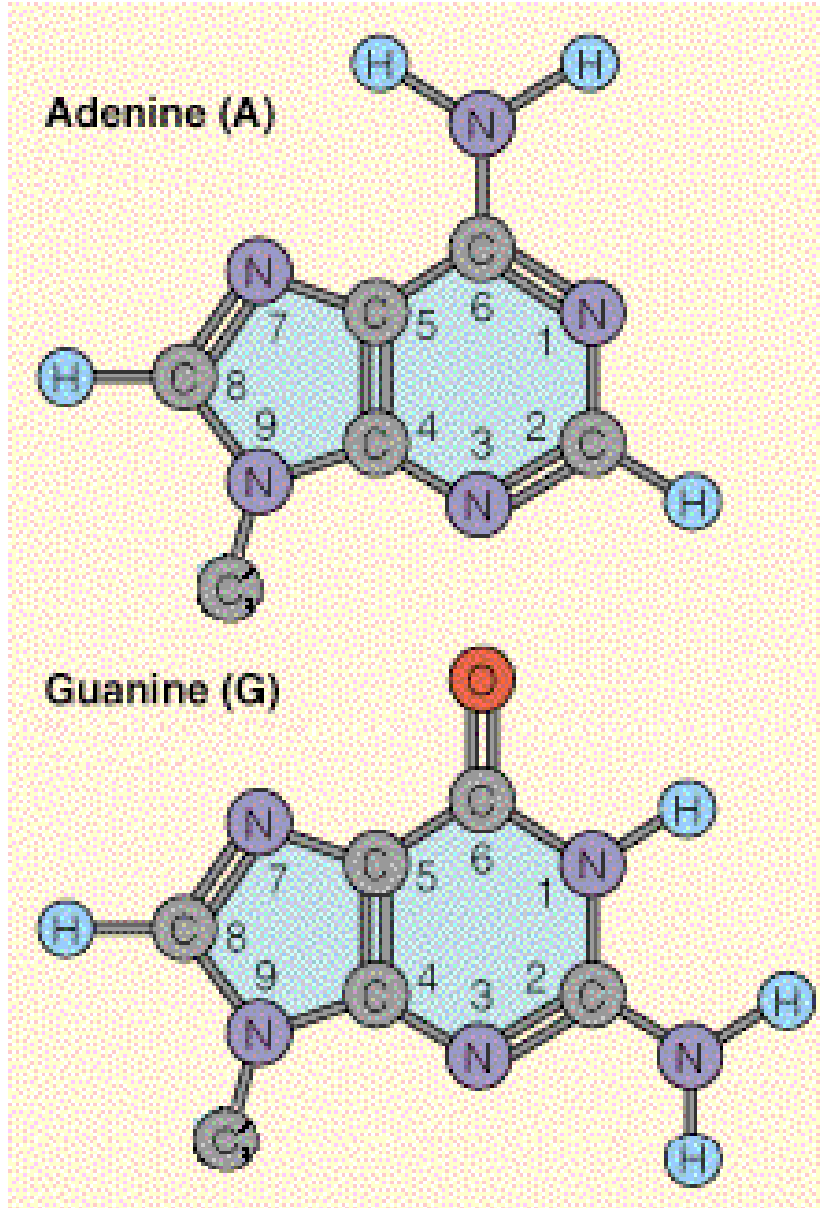
Bases: pyrimidines and purines:

- Pyrimidines



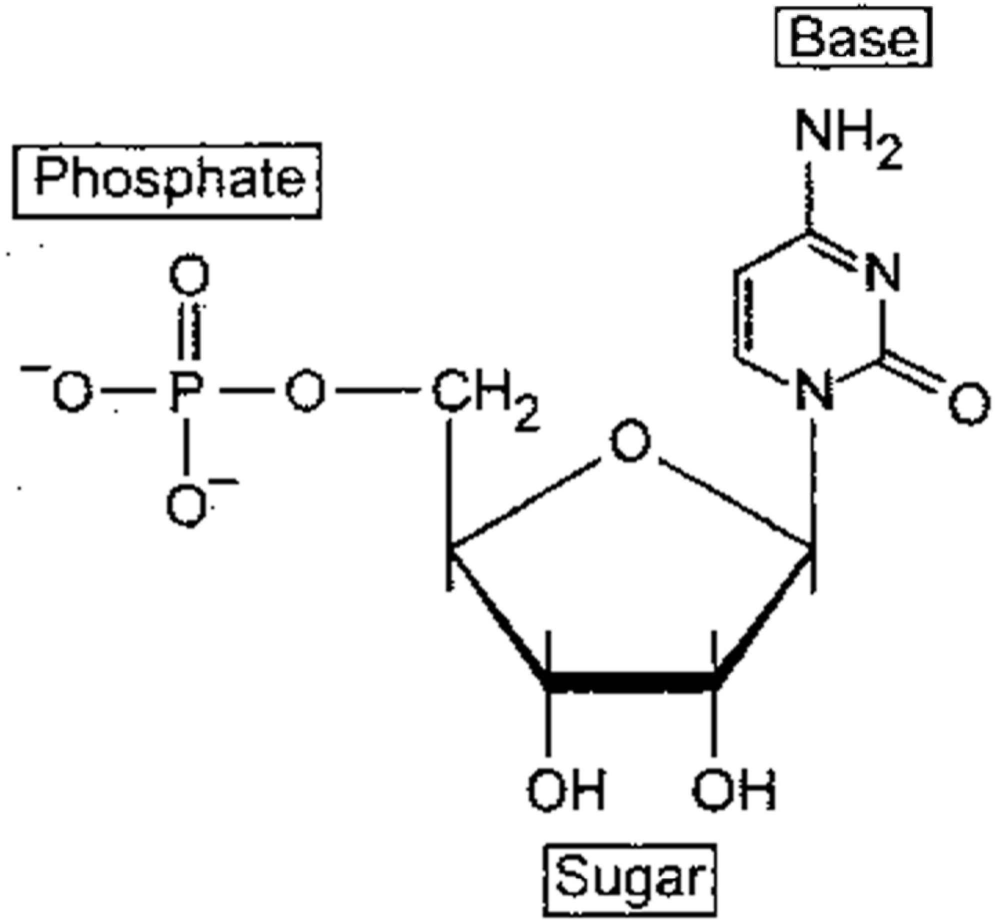


Purines



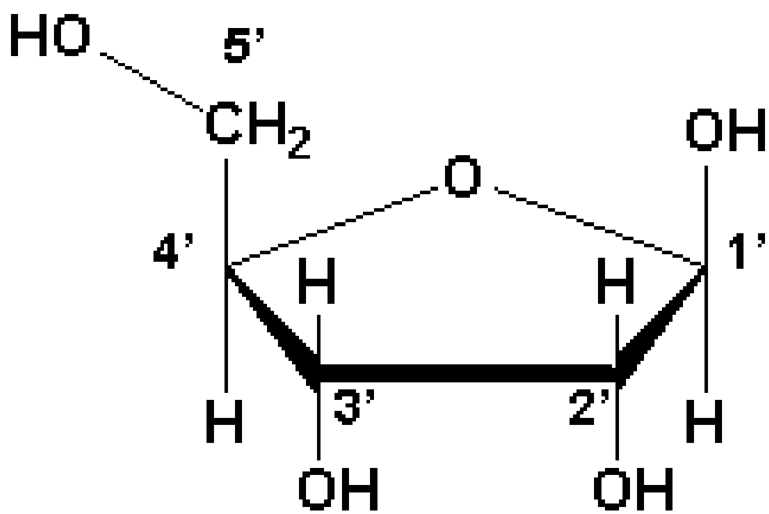
Nucleotide = Phosphate + Sugar + Base

nucleotide - repeat unit of DNA

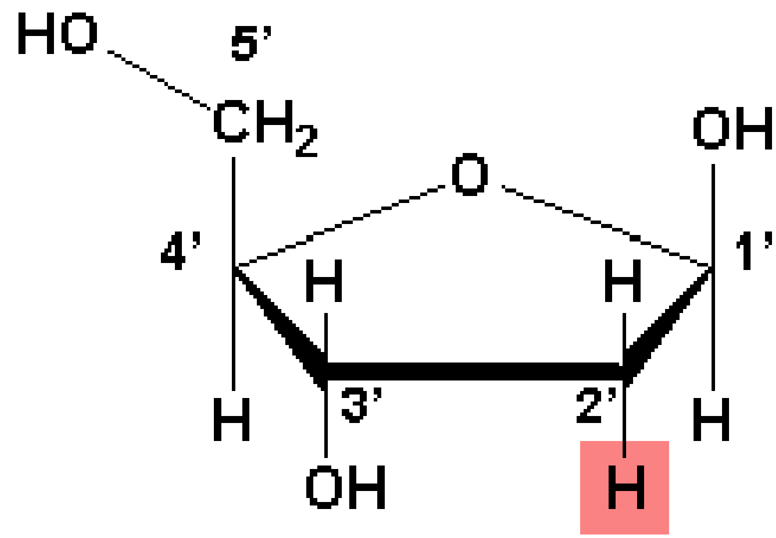


Sugars in nucleotide defines DNA or RNA:

Ribose (RNA)



Deoxyribose (DNA)



DNA *vs.* RNA: DNA is less susceptible to hydrolysis, has more conformational flexibility, lacks extra H-bonding site

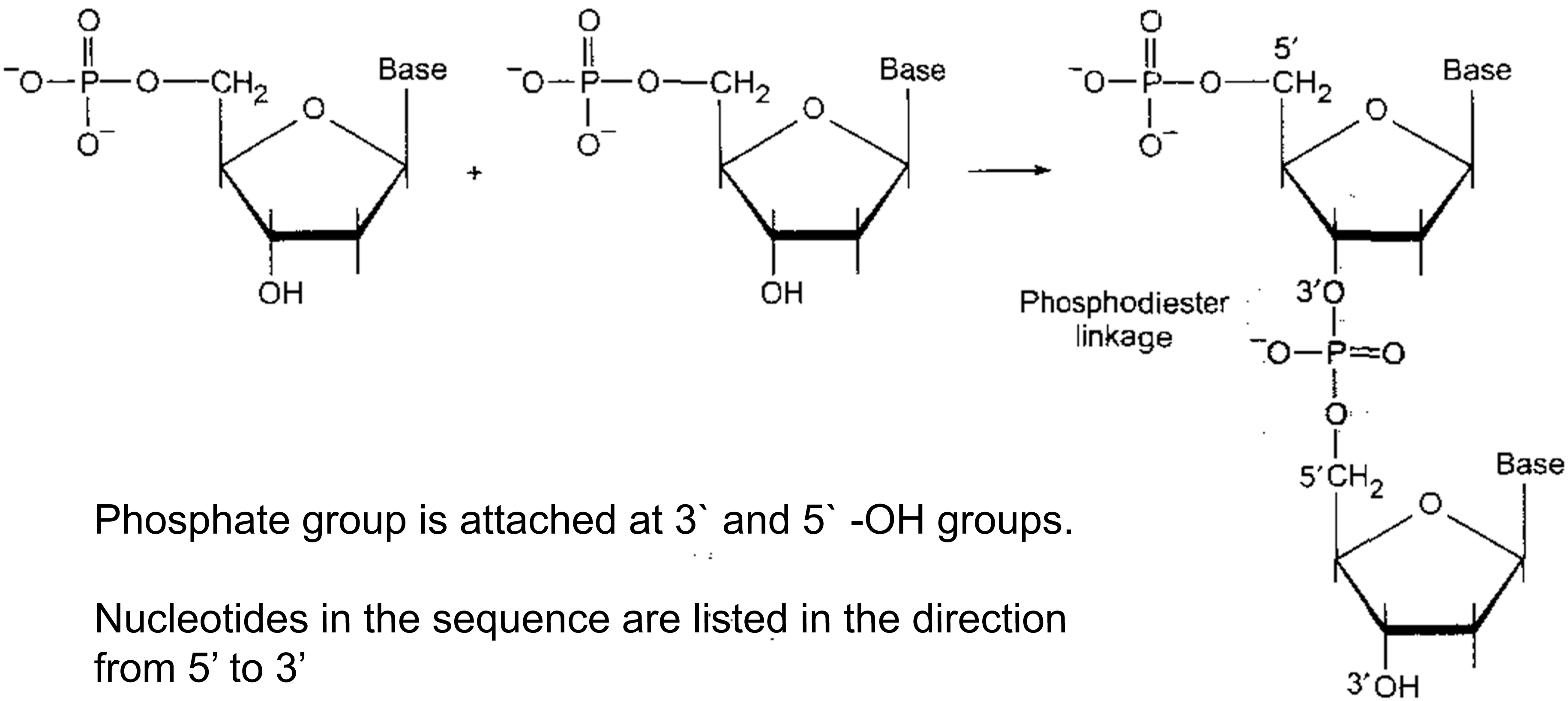
Base + sugar = nucleoside
Bases are attached at C1` atom of sugar.

Nucleosides: (Base + sugar)

ribonucleosides	2` deoxyribonucleosides
Adenosine Cytidine Guanosine Uridine	Deoxyadenosine Deoxycytidine Deoxyguanosine Thymidine

Nucleotides (DNA or RNA) (Nucleoside + phosphate group):

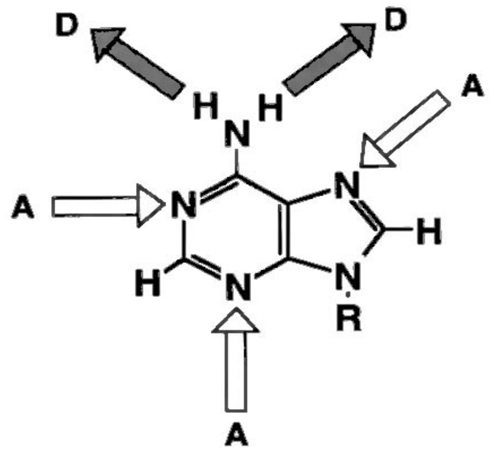
ribonucleotides	2` deoxyribonucleotides
Adenylic acid Cytidylic acid Guanylic acid Uridylic acid	Deoxyadenylic acid Deoxycytidylic acid Deoxyguanylic acid Thymidylic acid



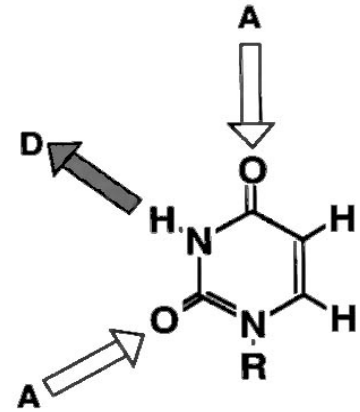
Phosphate group is attached at 3` and 5` -OH groups.

Nucleotides in the sequence are listed in the direction from 5’ to 3’

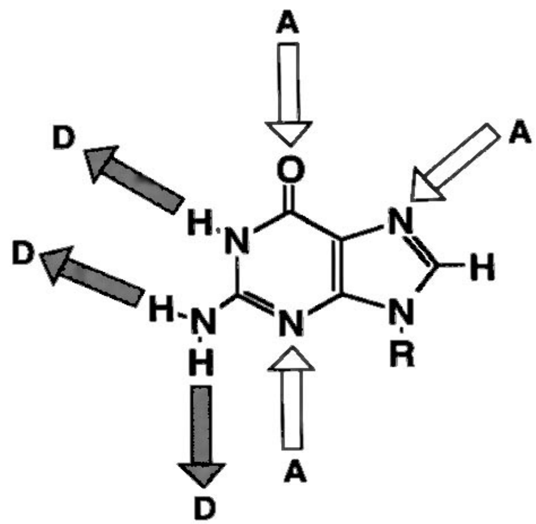
Hydrogen bonding sites in the bases of nucleosides



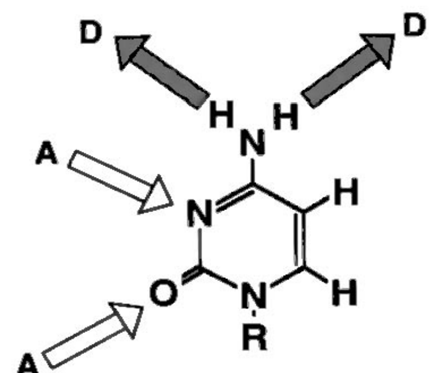
Adenosine



Uridine



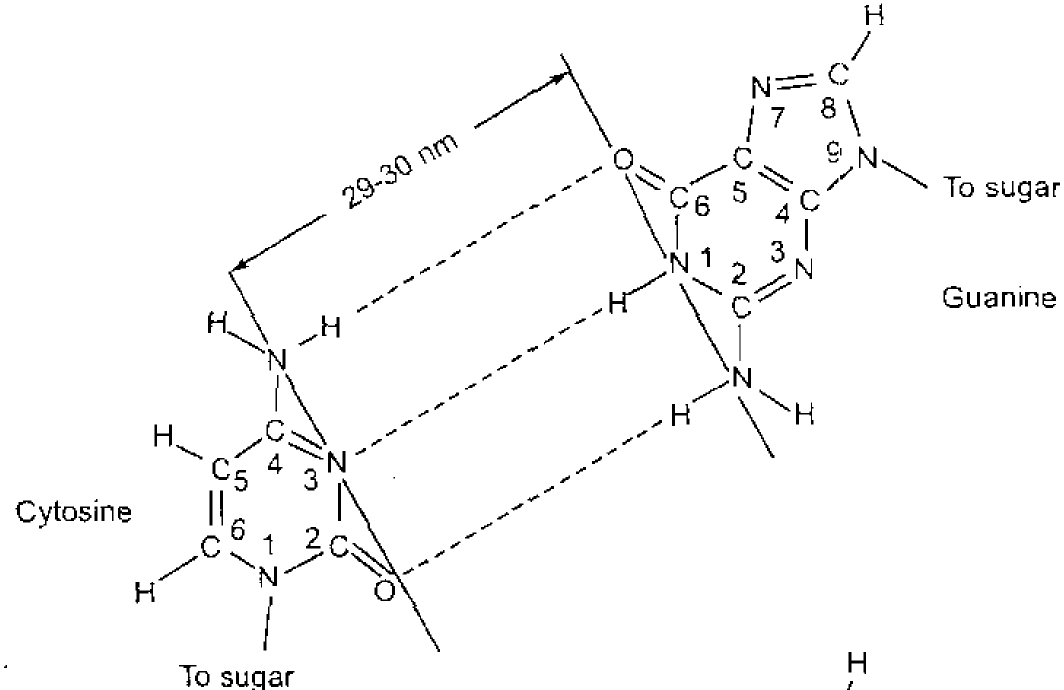
Guanosine



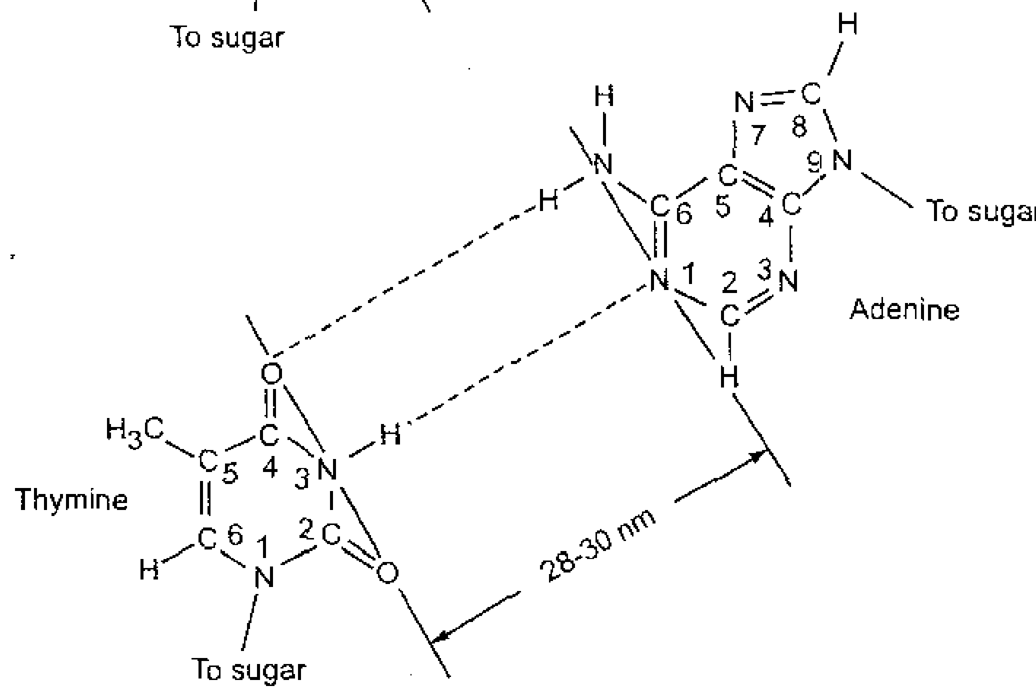
Cytidine

DNA is double-stranded--
two polynucleotide chains

Hydrogen bonds between
bases hold them together

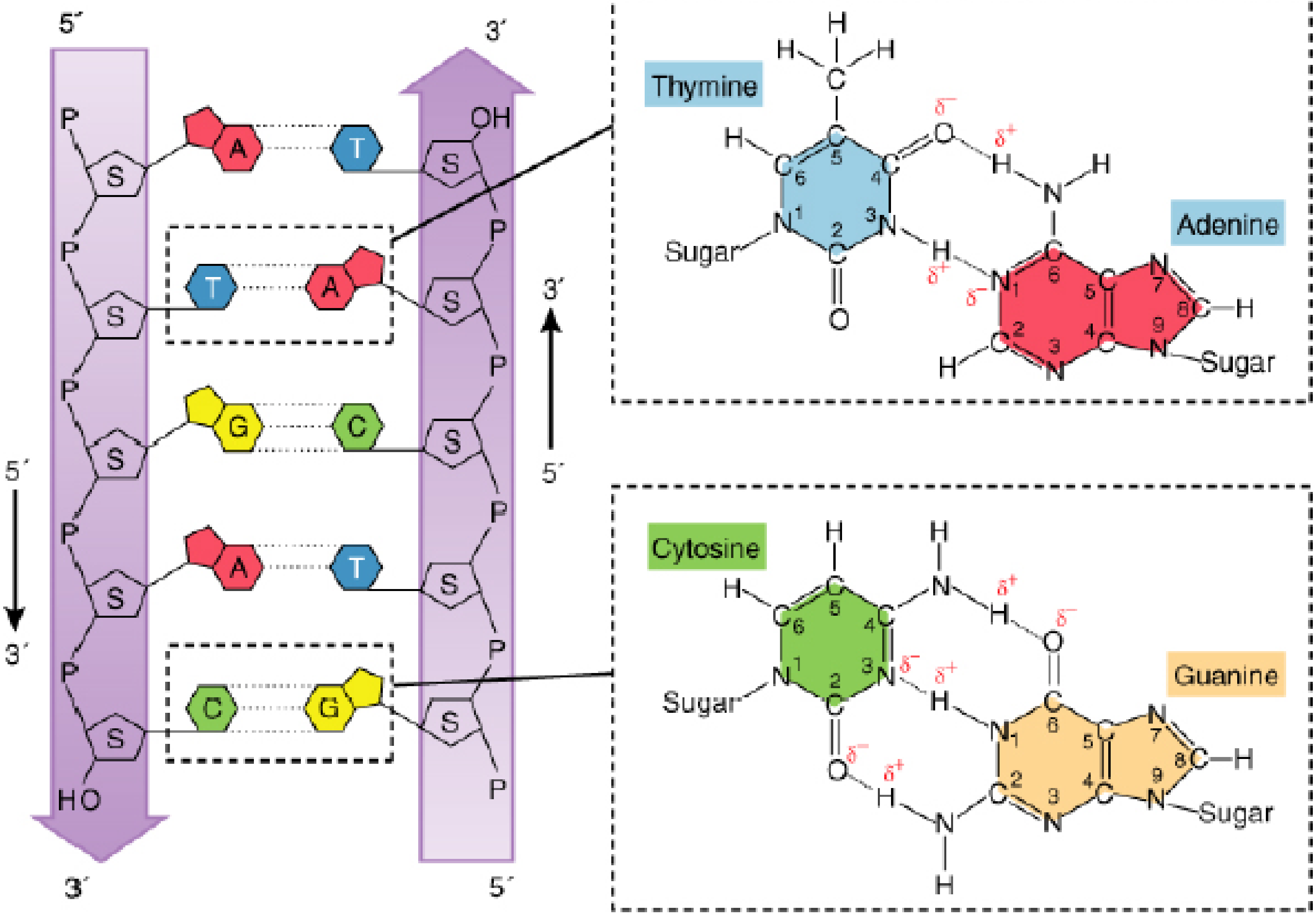


G and C have 3 H-bonds



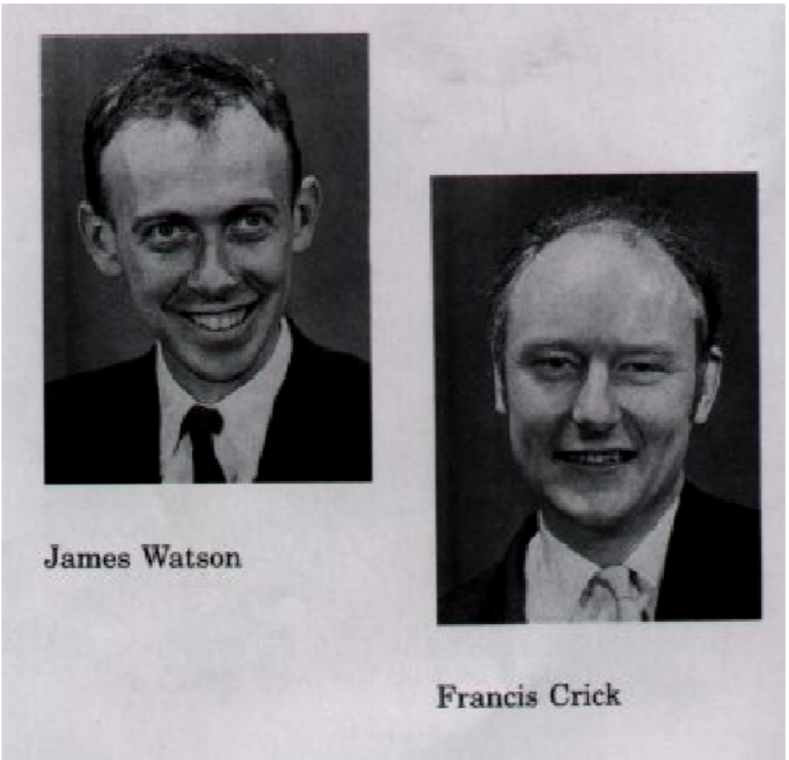
A and T make 2 H-bonds

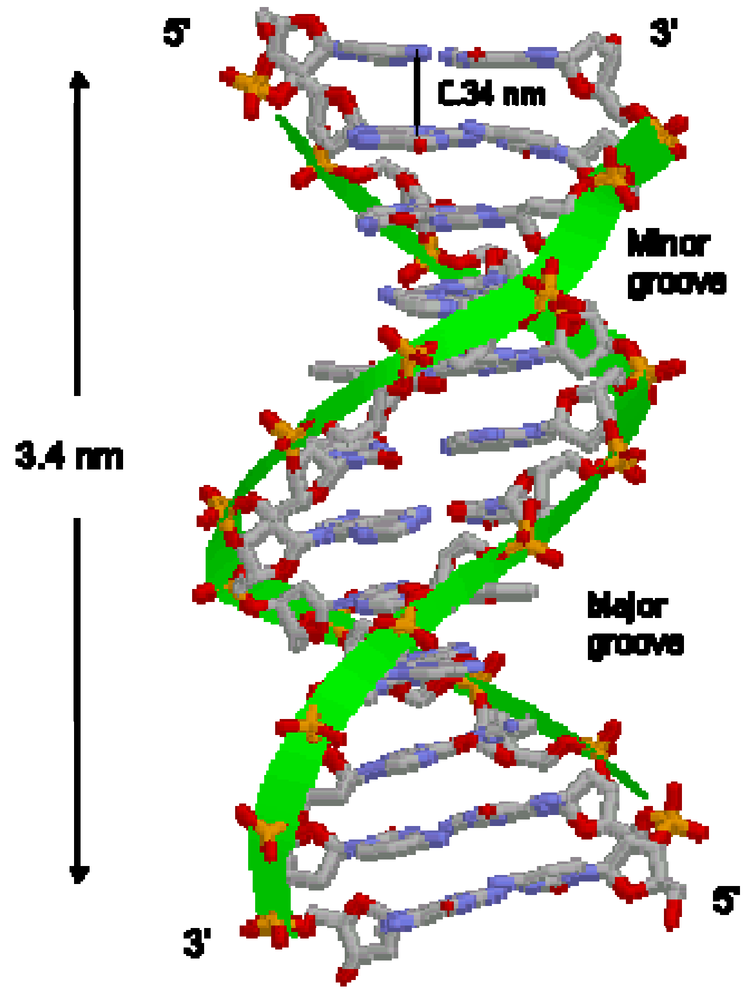
Hydrogen bonding in A-T and G-C base pairs



DNA is a double helix

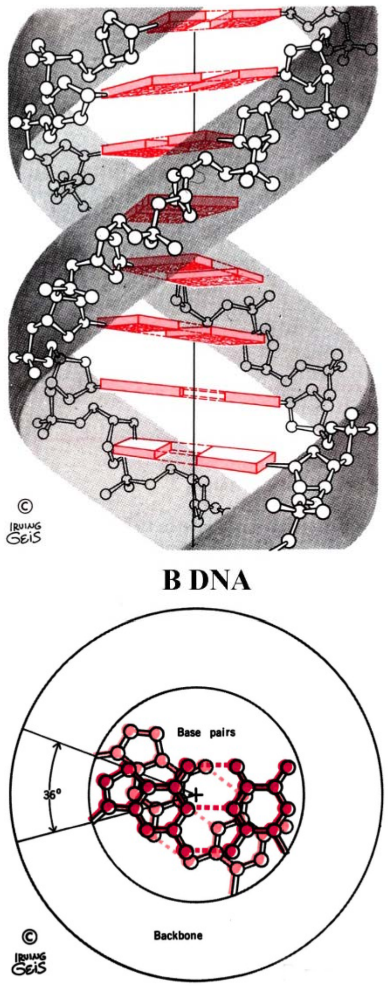
•Watson and Crick , 1953

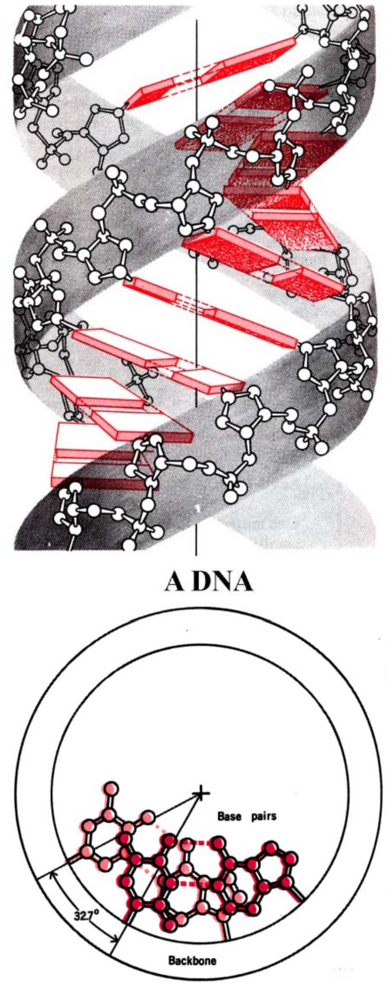


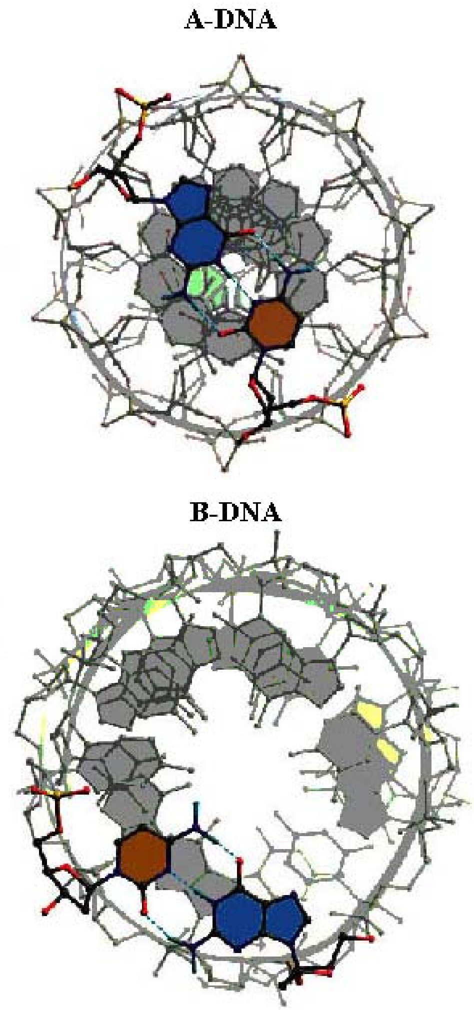


In this structure, discovered by Watson and Crick and Wilkins, also known as the **B form**, the helix makes a turn every 3.4 nm, and the distance between two neighboring base pairs is 0.34 nm. Hence, there are about 10 pairs per turn. The intertwined strands make two grooves of different widths, referred to as the **major groove** and the **minor groove**, which may facilitate binding with specific proteins.

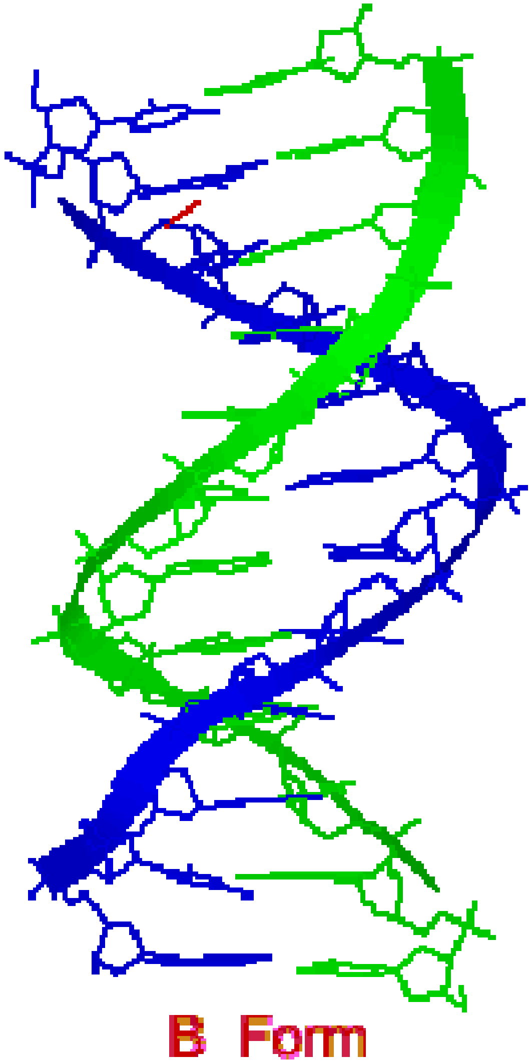
Comparison of A- and B- forms

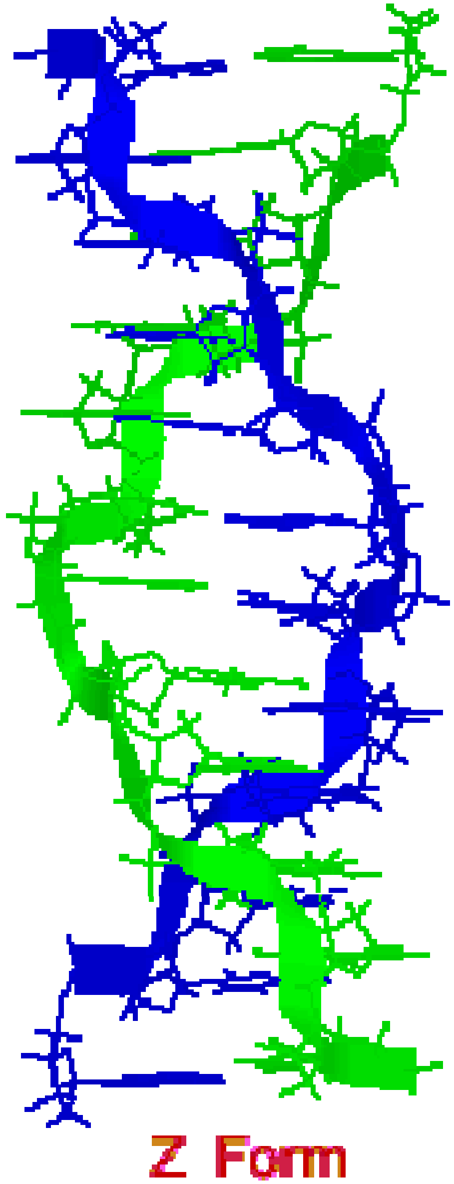






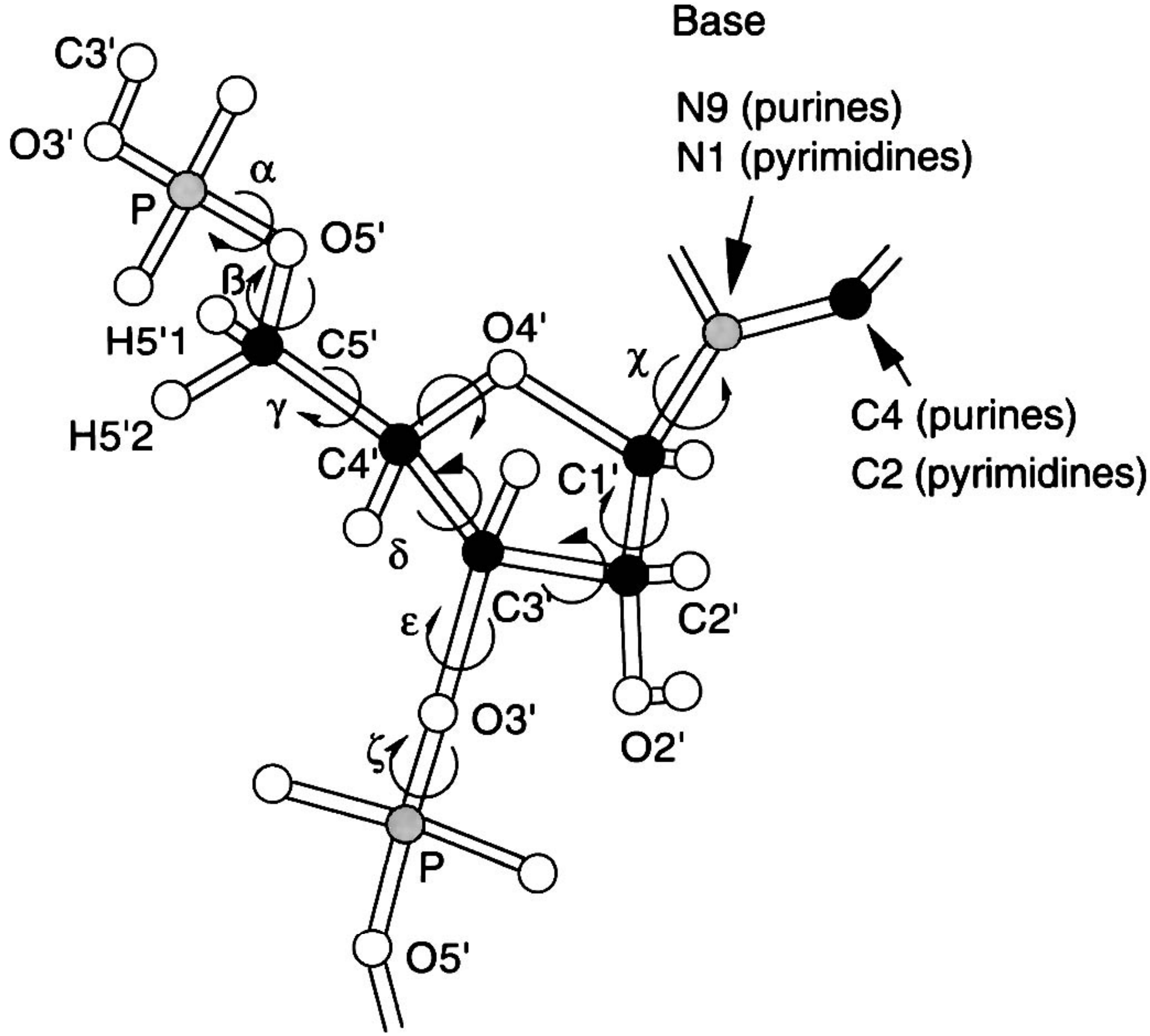
In a solution with higher salt concentrations or with alcohol added, the DNA structure may change to an **A form**, which is still right-handed, but every 2.3 nm makes a turn and there are 11 base pairs per turn.



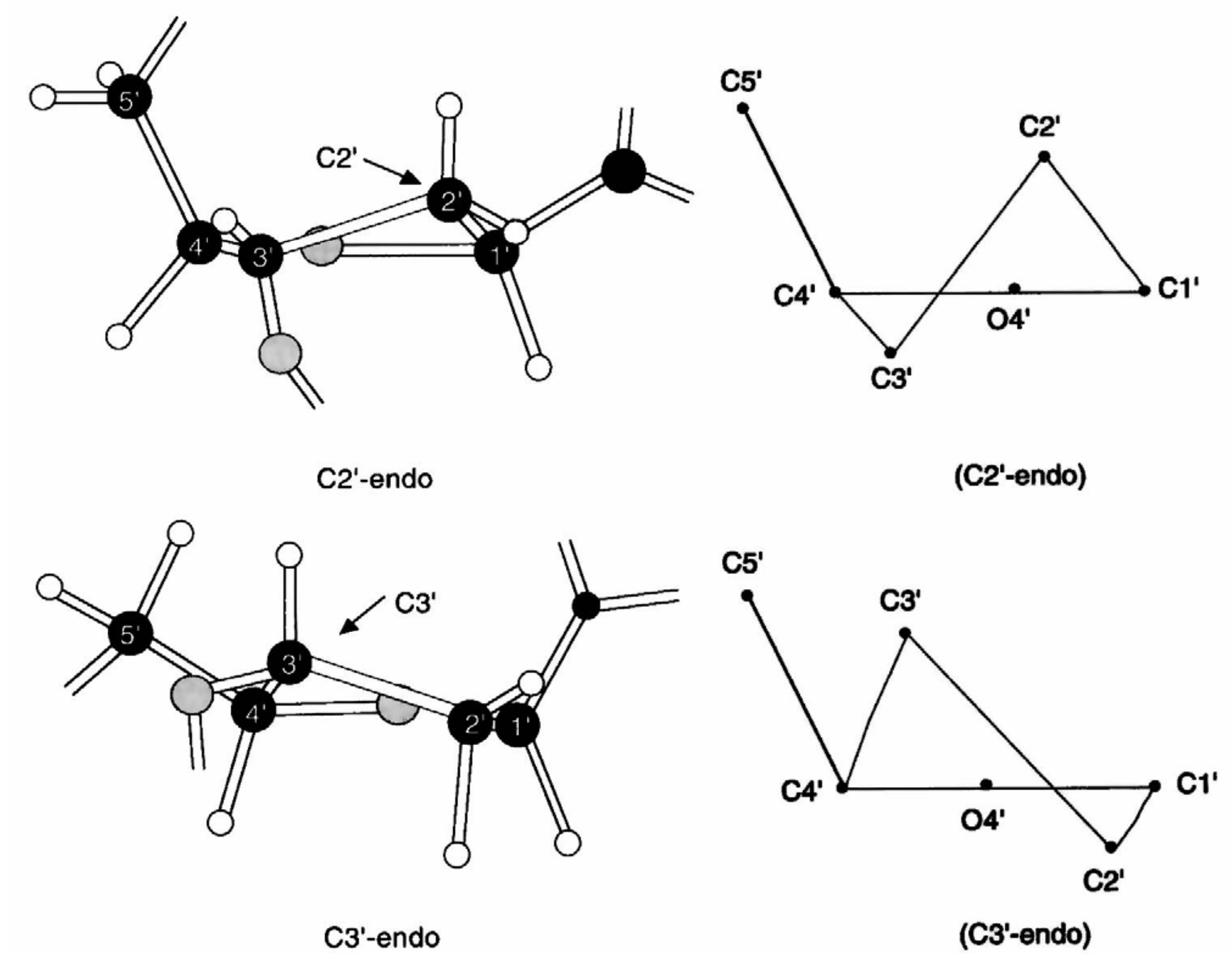


Another DNA structure is called the **Z form**, because its bases seem to zigzag. Z DNA is left-handed. One turn spans 4.6 nm, comprising 12 base pairs. The DNA molecule with alternating G-C sequences in alcohol or high salt solution tends to have such structure.

Torsion angles of the backbone



conformations of sugar ring

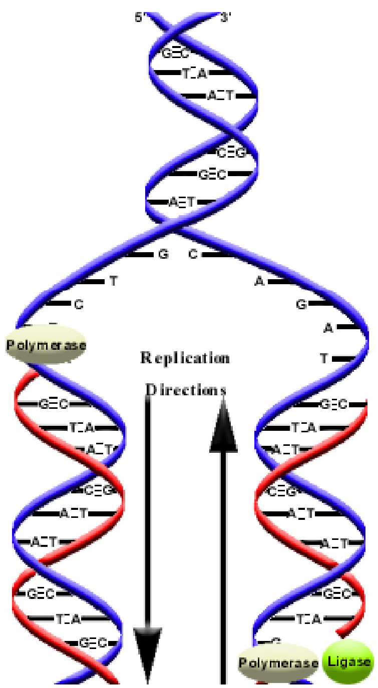


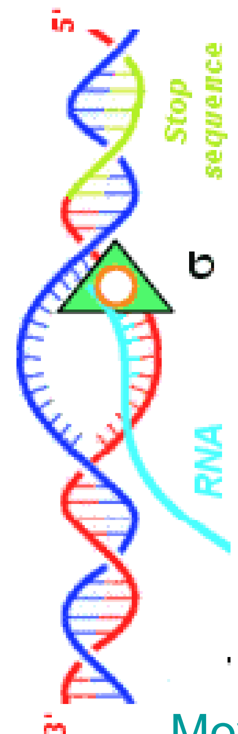
Transition from C2'-endo to C3'-endo corresponds to B-form → A-form transition

DNA is a flexible structure
Dynamics of DNA is essential

Replication

Transcription

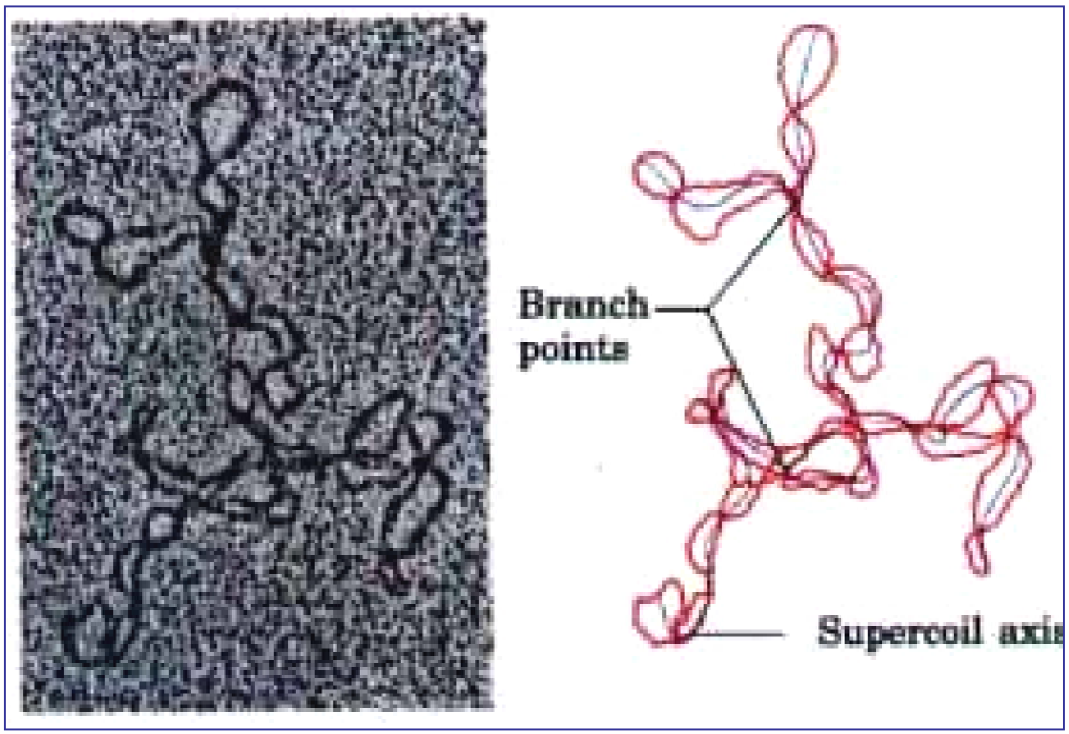


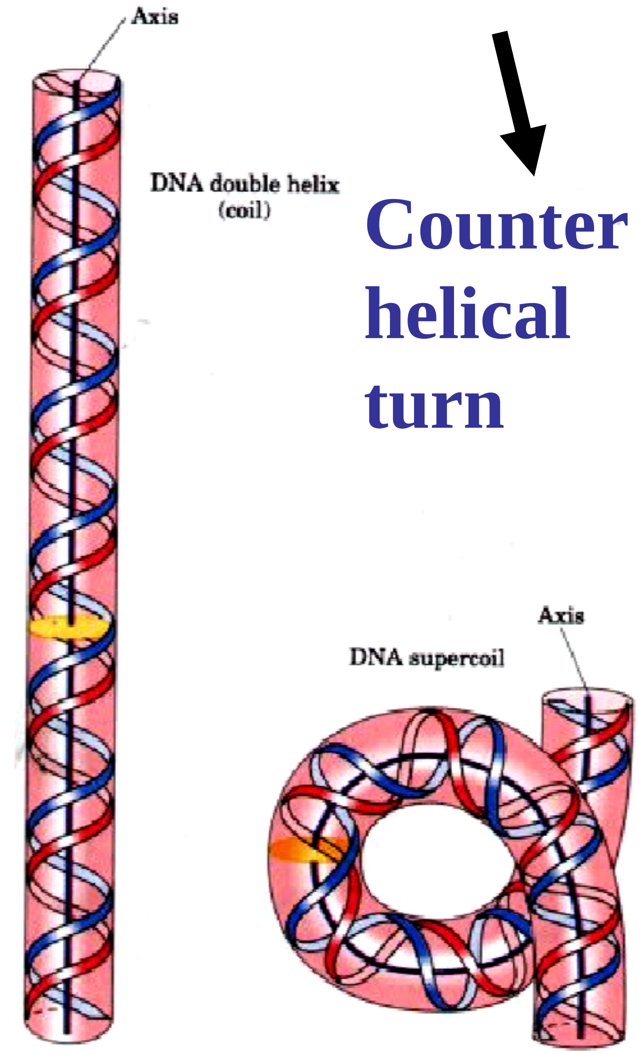


[Movie]

K.Rasmusen, MD

DNA molecules in cells are mostly circular, and are on average negatively supercoiled.





Physical properties of DNA

- Similar to proteins DNA show high viscosity,
- Colloidal and osmotic pressure
- Optical activity – dextrorotary behavior
- Denaturation under T, mechanical stress, irradiation, chemical agents

Heteromacromolecules

- Heteromolecules are present in cells
- They are composed of two compounds of different type.
- Examples:
- **Nucleoproteins** – nucleic acid + protein
- **Glycoproteins** - carbohydrate-protein complexes found in blood , more resistive to heat
- **Lipoproteins** – lipid-protein complexes – blood, plasma, milk
- **Glycolipids** - carbohydrate-lipid complexes

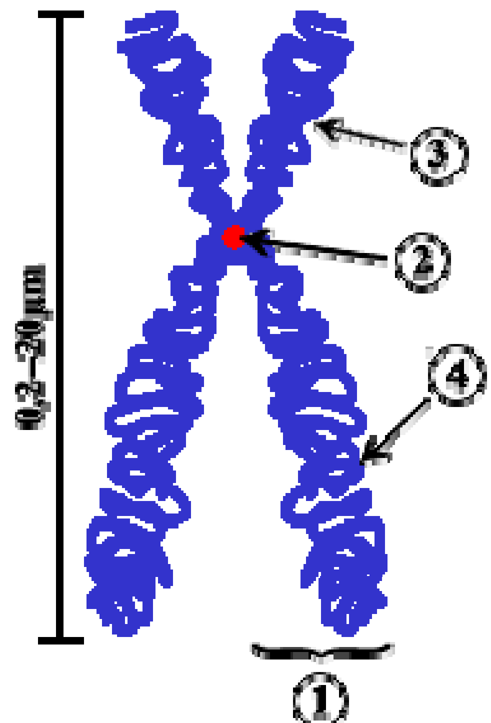
- There are many other complexes, such as cofactor proteins (enzymes, chlorophyll proteins), hemoproteins (hemoglobine, myoglobin), etc.

Additional material: Chromosome

Chromosomes are organized structures of DNA and Proteins that are found in cells. Chromosomes contain a single continuous piece of DNA, which contains many genes, regulatory elements and other nucleotide sequences. Chromosomes also contain DNA-bound proteins, which serve to package the DNA and control its functions. The word *chromosome* comes from the Greek χρῶμα (*chroma*, color) and σῶμα (*soma*, body) due to their property of being stained very strongly by some dyes.

Chromosomes vary extensively between different organisms. The DNA molecule may be circular or linear, and can contain anything from tens of kilobase pairs to hundreds of megabase pairs. Typically eukaryotic cells (cells with nuclei) have large linear chromosomes and prokaryotic cells (cells without nuclei) smaller circular chromosomes,

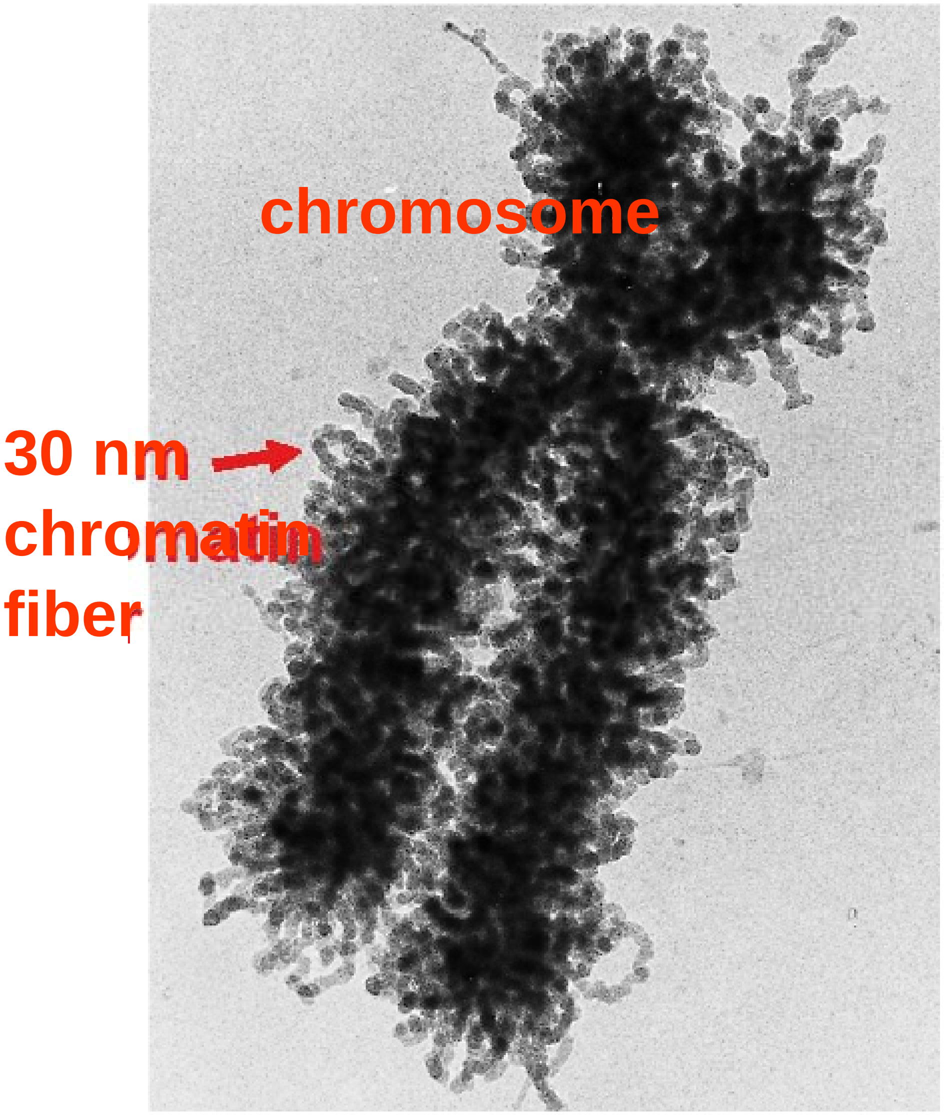
In eukaryotes, nuclear chromosomes are packaged by proteins into a condensed structure called chromatin. This allows the massively-long DNA molecules to fit into the cell nucleous.



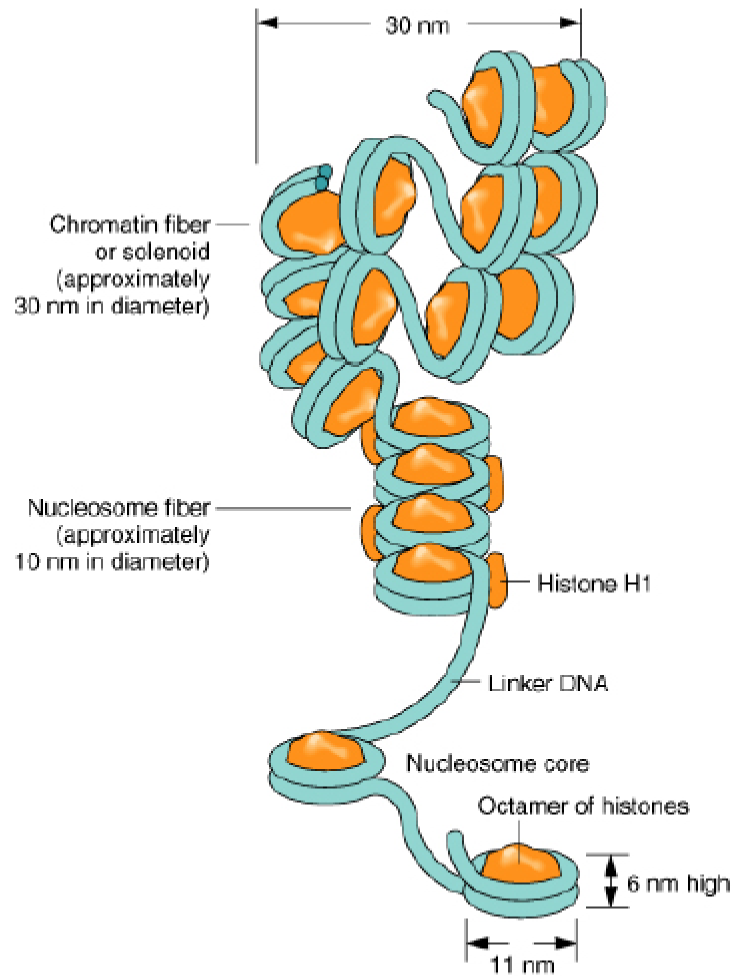
A scheme of a condensed chromosome. (1) Chromatid - one of the two identical parts of the chromosome

(2) Centromere - the point where the two chromatids touch, and where the microtubules attach. (3) Short arm. (4) Long arm.

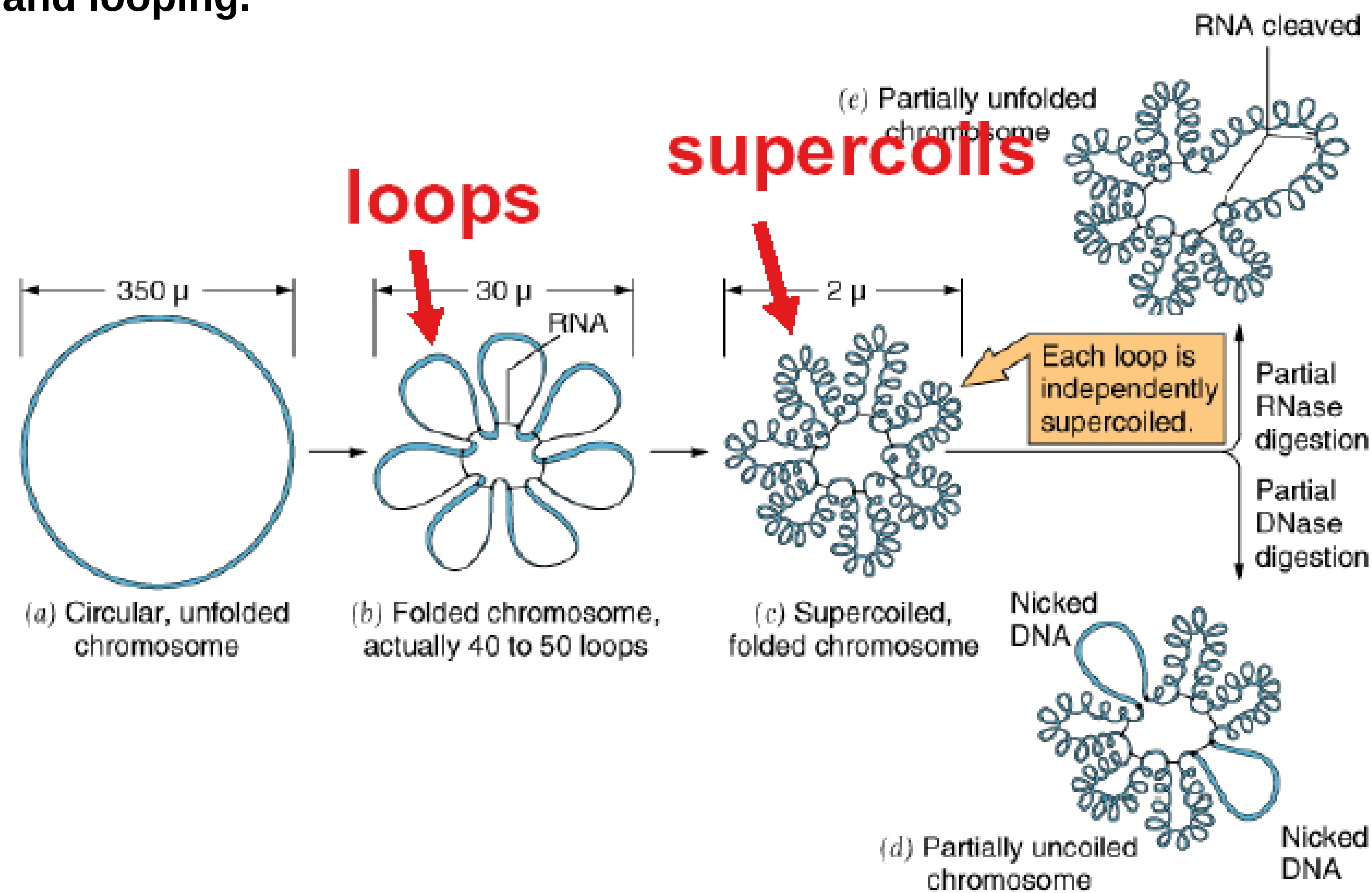
Human genome = 2000 mm long (diploid) Two meters!!!
Eukaryotes need great genome compaction! - chromatin



From E. DuPraw, *DNA & Chromosomes*, Holt, Rinehart, & Winston, New York, 1970. Original photo courtesy E. DuPraw



Typical prokaryotic circular chromosome compacted by supercoiling and looping.



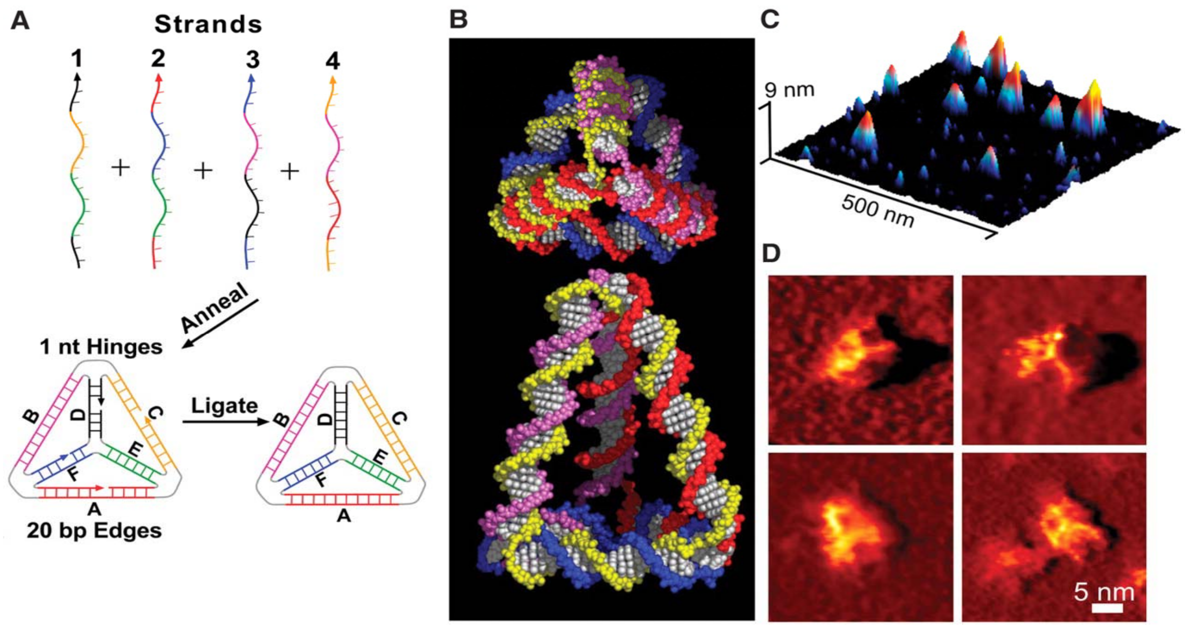
(a) Circular, unfolded chromosome

(b) Folded chromosome, actually 40 to 50 loops

(c) Supercoiled, folded chromosome

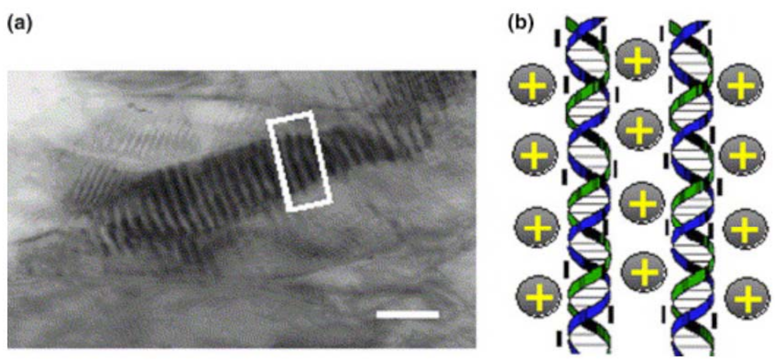
(d) Partially uncoiled chromosome

DNA is A NANOTECH SYSTEM



DNA templating gold particles

Curr. Appl. Phys. 5, 102 (2005)



DNA Building Blocks for Molecular Nanofabrication

Science 310, 1661 (2005)