The biomolecules of terrestrial life

Planets and Astrobiology (2018-2019)
G. Vladilo

Chemical groups and monomers of large molecules

Phosphate group
Sugars
Nucleotides
Energy exchange molecules
Aminoacids

Functional groups in biomolecules

Groups of atoms that are responsible for the chemical properties of biomolecules

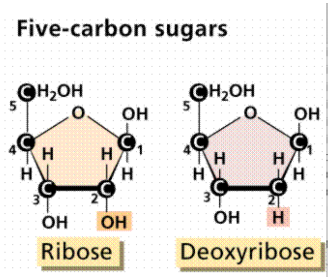
TABLE 2.1 Common Functional Groups Found in Biomolecules

Functional group	Chemical formula	Structure	Chemical property
Hydroxyl	—OH	—O—H	Polar
Sulfhydryl	—SH	—S—H	Polar
Phosphate	$—HPO_4^-$	—O—P(=O)(—O⁻)—OH	Polar
Carboxyl	—COOH	—C(=O)—OH	Acid
Amino	$—NH_2$	—N(—H)—H	Base

Sugars

General formula: $C_x(H_2O)_y$

Sugars in nucleic acids
Ribose $C_5H_{10}O_5$
Deoxyribose $C_5H_{10}O_4$

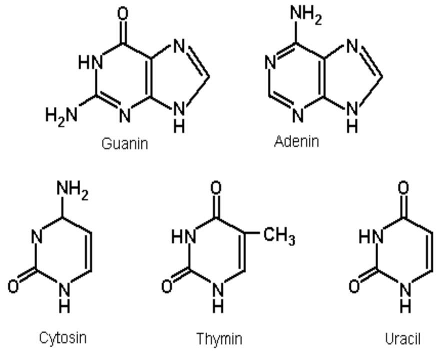


Nitrogenous bases

aromatic rings with N substitutions

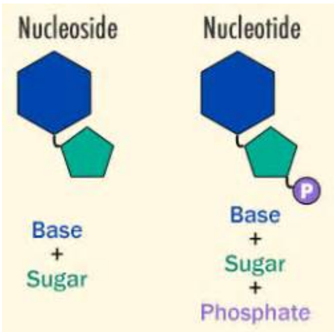
Purines

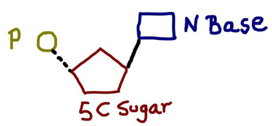
Pyrimidines



Nucleosides and nucleotides

Constituents of the DNA and RNA

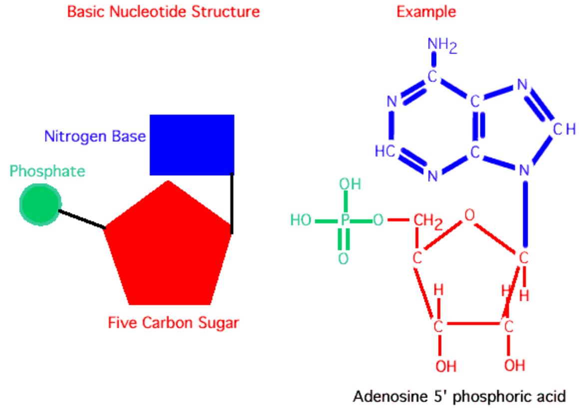




NUCLEO—
SIDE = SUGAR + BASE
TIDE = SIDE + PHOSPHATE

Nucleotides

Constituents of the DNA and RNA

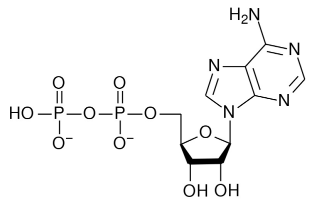


ADP & ATP

Energy exchange molecules

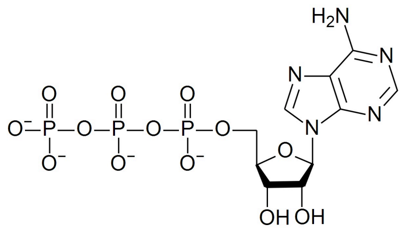
ADP

Adenosine diphosphate



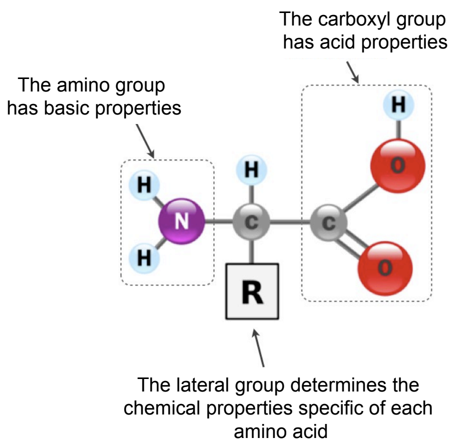
ATP

Adenosine triphosphate



Amino acids

Constituents of proteins



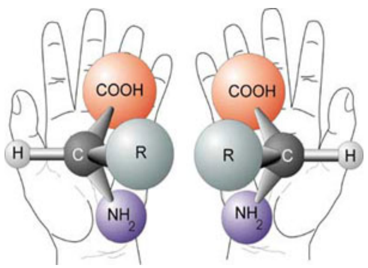
Chirality

Molecules with same chemical formula but different structure are called <u>isomers</u>

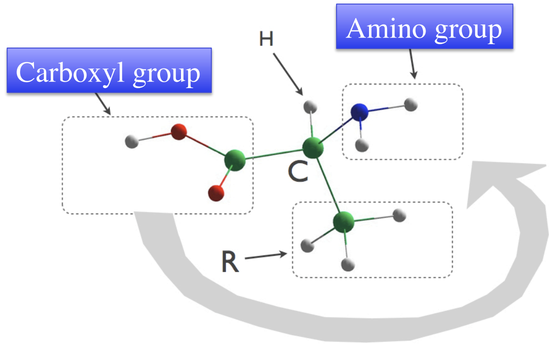
<u>Chiral</u> molecules are isomers with a center of symmetry ("stereocenter")
They cannot be superimposed to their mirror image
The two mirror images of a chiral molecule are called <u>enantiomers</u>

Amino acids are chiral

The carbon atom at the center of the amino acid is the <u>stereocenter</u>



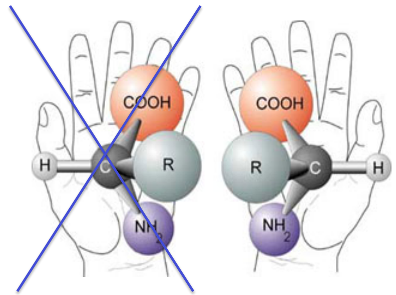
The two enantiomers of amino acids are called L and D according to the "CORN" convention



Example of L-type amino acid: L-alanine

Homochirality of terrestrial biomolecules

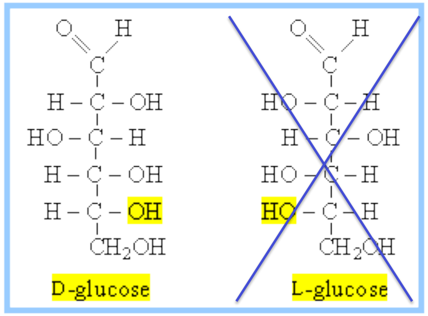
Protein amino acids are <u>homochiral</u>:
they only show the L-type enantiomer



Homochirality of terrestrial biomolecules

Also most biological sugars are homochiral

They are D type (according to a different convention)



Racemic and non-racemic mixtures

Racemic mixture

has equal amounts of left- and right-handed enantiomers of a chiral molecule

The non biological world is racemic

The biological world shows an extreme enantiomeric eccess

(only one enantiomer)

This difference may provide a way to discriminate

biological from non-biological compounds

if we have the possibility to analyse a sample

Biological macromolecules

Created by polymerization of a large number of subunits (monomers)

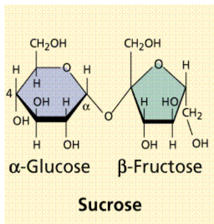
Terrestrial life features 4 types of macromolecules:

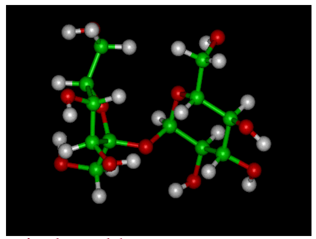
carbohydrates, lipids, proteins and nucleic acids

energy storage | compartmentalization | metabolic | genetic

Biological macromolecules

Carbohydrates (saccharides)





The most abundant molecules in the biological world

Primary source of chemical energy for most organims

General formula: $C_x(H_2O)_y$

Monosaccharides (simple sugars)

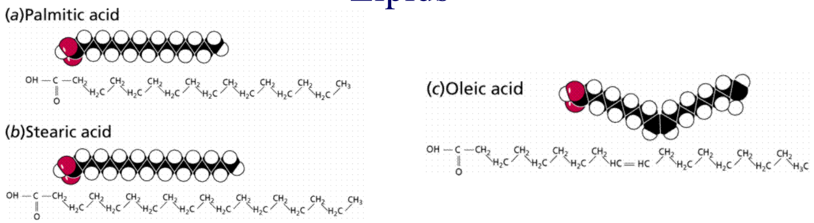
Oligosaccharides

From 2 to 10 units of monosaccharides

Polysaccharides

More than 10 monosaccharides

Lipids



Heterogeneous class of organic molecules with common solubility properties

- Insoluble in water
- Soluble in certain types of non-polar solvents

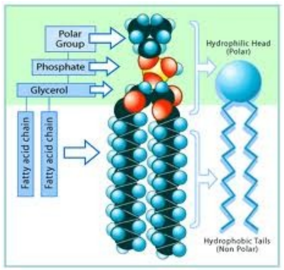
Larger number of C−H bonds with respect to carbohydrates

- Used for long-term storage of energy

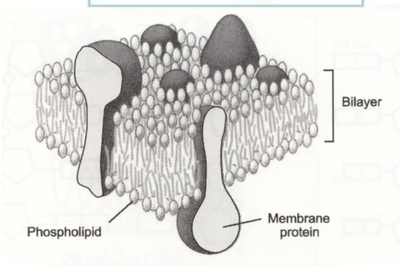
Phospholipids and cell membranes

Phospholipids

- Examples of amphiphilic molecules with a hydrophilic end and a hydrophobic end



In liquid water phospholipids spontaneously form a double layer of molecules (bilayer), with the hydrophobic ends facing each other in the inner part, and the hydrophilic ends facing the water



Bilayers of phospholipids are the main structural components of cell membranes

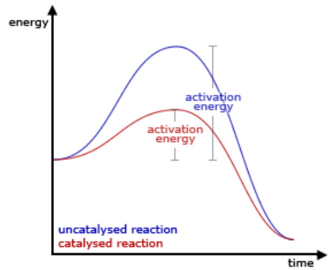
Proteins

Proteins are polymers of amino acids

- Short chains of amino acids are called peptydes
- Long, unbranched peptyde chains are called polypeptides
- Proteins are formed by one or more chains of polypeptides
  - Molecular masses of proteins vary between $\sim 10^3$ e $\sim 10^6$ atomic mass units
  - They contribute to about half the mass of the cell

Proteins play fundamental functions in living organisms

- Mostly structural and enzymatic (i.e., catalytic) functions



The amino acids of terrestrial life

Proteins use only 20 types of amino acids, all of L-type

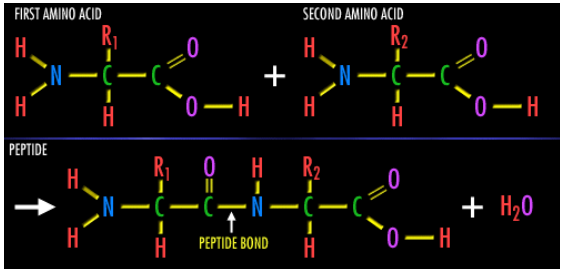
- Organic chemistry allows for the existence of thousands of amino acids

Apparently, terrestrial life has "chosen" a short list of amino acids, sufficiently representative of the different types of chemical properties required to build up the variety of proteins necessary to living organisms

Table 7.2 *The Twenty Amino Acids Found in Living Organisms*

*Amino Acid**	*Chemical Formula*	*Number of Atoms*
L-Alanine	$C_3H_7O_2N$	13
L-Arginine	$C_6H_{15}O_2N_4$	27
L-Asparagine	$C_4H_8O_3N_2$	17
L-Aspartic Acid	$C_4H_6O_4N$	15
L-Cysteine	$C_3H_7O_2NS$	14
L-Glutamic Acid	$C_5H_8O_4N$	18
L-Glutamine	$C_5H_{10}O_3N_2$	20
Glycine	$C_2H_5O_2N$	10
L-Histidine	$C_6H_9O_2N_3$	20
L-Isoleucine	$C_6H_{13}O_2N$	22
L-Leucine	$C_6H_{13}O_2N$	22
L-Lysine	$C_6H_{15}O_2N_2$	25
L-Methionine	$C_5H_{11}O_2NS$	20
L-Phenylalanine	$C_9H_{11}O_2N$	23
L-Proline	$C_5H_9O_2N$	17
L-Serine	$C_3H_7O_3N$	14
L-Threonine	$C_4H_9O_3N$	17
L-Tryptophan	$C_{11}H_{12}O_2N_2$	27
L-Tyrosine	$C_9H_{11}O_3N$	24
L-Valine	$C_5H_{11}O_2N$	19

From amino acids to polypeptides



Amino acids are bound to each other with peptide bonds

The carboxyl end ties to the amino end of the next molecule

A sequence OC-NH is formed (peptide bond)

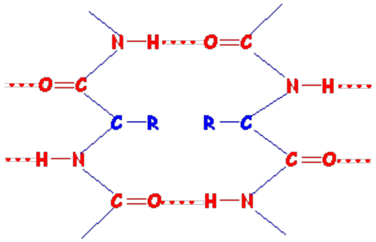
A water molecule is released each time a peptide bond is created

A water molecule is required to break a peptide bond

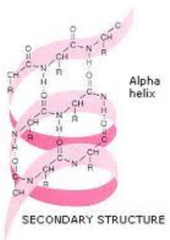
The sequence of lateral groups determine the properties of the protein

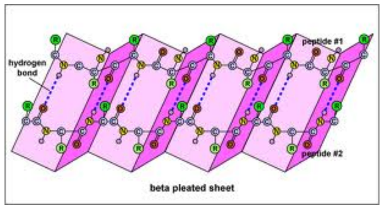
From polypeptides to 3D proteins

Importance of hydrogen bonds as intramolecular forces



Spontaneous formation of geometrical configurations with lowest potential energy



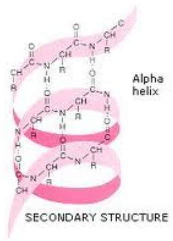


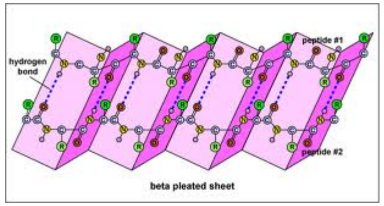
Proteins

The importance of 3D structures

The shape of proteins determines, in large part, their function

Chemistry → Spontaneous folding → 3D Shape → Function





Biological macromolecules

Nucleic acids

Nucleic acids are polymers of nucleotides

Depending on the type of organism,
they may contain ~ $10^6 – 10^8$ nucleotides

Nucleic acids store and use the genetic information

There are two types: DNA and RNA

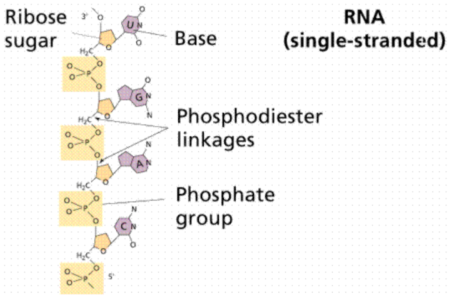
DNA preserves the information

RNA, which comes in different forms, uses the information for driving metabolic/replication processes

Nucleic acids: RNA

RNA has a single strand of nucleotides

The backbone of the strand is made up of a sequence of phosphate groups and <u>ribose</u> sugars



Has 4 types of nucleobases

- Purines
  - Adenine, Guanine
- Pyrimidines
  - Cytosine, Uracyl

RNA drives the synthesis of proteins

The order of the nitrogen bases on the backbone determines the sequence in which amino acids are assembled

Nucleic acids: DNA

DNA has two strands that form a double helix structure

The backbone of each strand is made up of a sequence of phosphate groups and <u>deoxyribose</u> sugars

DNA has 4 types of nucleobases

- 2 purins
  - Adenine e Guanine
- 2 pyrimidins
  - Cytosine e Thymine
    - Thymine replaces Uracyl, which is instead used in the RNA

The complementarity of purines and pyrimidines plays a fundamental role in the pairing between the two strands

FIGURE 14-16
Base pairing. In a DNA molecule, only two base p possible: adenine (A) w mine (T), and guanine (G cytosine (C). A G-C base three hydrogen bonds; an base pair, only two.

Hydrogen bonds in the DNA

The two strands of the DNA are tied up by hydrogen bonds

- G-C pairs have 3 bonds
- T-A pairs have 2 bonds

Hydrogen bonds as intramolecular forces

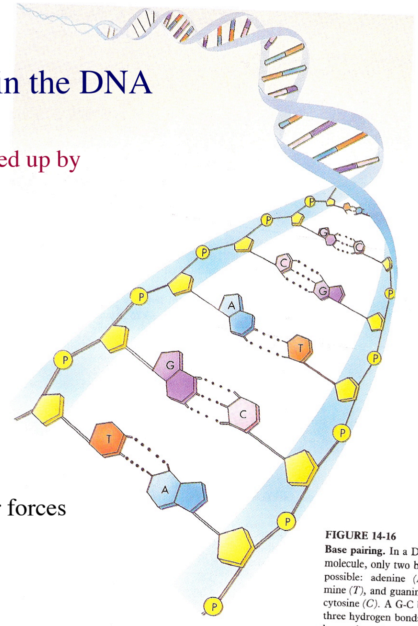


FIGURE 14-16
Base pairing. In a DNA molecule, only two base p possible: adenine (A) w mine (T), and guanine (G cytosine (C). A G-C base three hydrogen bonds; an base pair, only two.