AST 205. Lecture 17. November 17, 2003
Biochemical Basis of Life: Nucleic Acids & Proteins

- Context
- Importance of carbon
- Four major classes of biological compounds
  - Proteins
  - Nucleic acids
  - Carbohydrates/polysaccharides
  - Lipids
- General features of terrestrial biochemistry
- Implications for extraterrestrial life

$$N = f(p)n(e)f(l)f(i)f(c)R_{*}L$$

- Radial velocity techniques have provided first direct clues about f(p), ≥ 5 -10%
- No direct information on n(e) yet
- R_{*} measured by astronomical observations
- f(l) depends on complex biochemistry
- Detailed knowledge of *one & only one* case
- But/thus no first principles general theory
- "Life as we know it" is the practical option

Convergence or Divergence of Cosmic and Biological Evolution? (How similar to here?)

- Large/coarse scales -> convergence
- But on some small/fine scales -> divergence
- Divergence *might* begin on the scale of planetary systems since known extrasolar systems are unlike the Solar System
- However it *might* not occur until far finer levels of detail <- assumption!

Contents of the "Chemistry Set(s)"

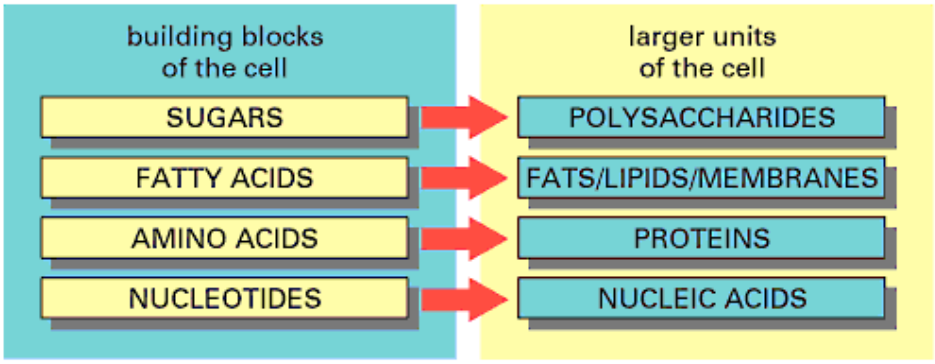
Solar	(%)	Terrestrial Mantle	(%)
Hydrogen	70.7	Oxygen	48.9
Helium	27.4	Silicon*	26.3
Oxygen	0.96	Aluminum*	7.7
Carbon	0.31	Iron	4.7
Neon	0.17	Calcium	3.4
Iron	0.14	Sodium	2.7
Nitrogen	0.11	Potassium	2.4
Silicon	0.07	Magnesium	2.0
Magnesium	0.06	(Hydrogen)	0.74
Sulfur	0.04	(Carbon)	0.02

Carbon (C) is a unique element, "da man" of organic chemistry & molecular biology

- Strong C-C bonds provide the structural support for very large 3D molecules
- C can simultaneously form strong bonds with H and O thus allowing large and complex molecules
- CO_2 is a gas allowing easy C transport and interactions
- Organic (C structured) compounds are 50x more numerous than inorganic ones
- *Not* particularly abundant in the Earth's mantle

Four major classes of bio-organic compounds

- *Proteins*: chains of amino acids that are the functional "machines" of biology
- *Nucleic Acids*: lengthy sequences of nucleotides which store, copy & implement protein structures
- Carbohydrates: energy storage & structure
- Lipids: energy storage & cell membranes



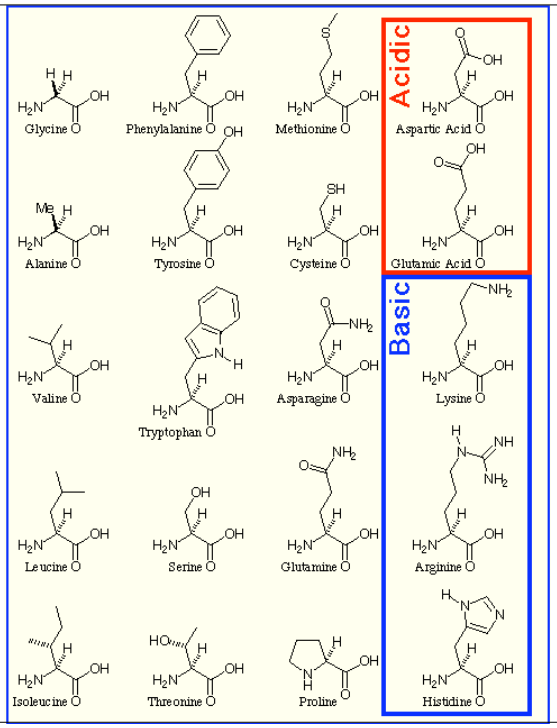
Three key self-propagation mechanisms

- DNA (info archive) replication
- DNA to RNA (blueprint) transcription
- RNA to protein (hardware) construction

The functioning of these mechanisms requires a genetic code/language, an energy supply or carrier (ATP), a "building tool" (ribosome) and water as a medium.
All above are *universal* features of terrestrial life!

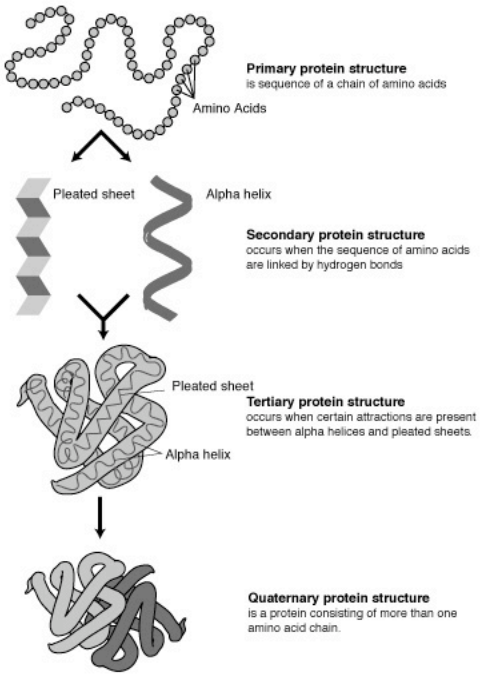
Amino acids are the "lego" building block components of proteins

- 13 to 27 atoms of C, O, N, H & S
- A COOH (carboxy) end that loses a H^+ ion
- A NH_2 (amino) end that takes a H^+ ion
- More than 170 known, but only 20 are coded by nucleic acids & "used" to make proteins
- 19 are l-chiral (left-handed) & one is symmetric
- Carboxy & amino ends "plug" together to form a peptide bond and thus make long chains:
- $H_3N^+ + COO^- \rightarrow OC\text{-}NH + H_2O$



Protein structure

- Peptide bond chains of 100s of amino acids
- Chain winds to form an α-helix or folds to form a β-sheet stabilized by H bonds
- Fold into specific 3D shapes set by disulfide bonds and hydrophobic interactions
- Also such proteins may combine as subunits to form a larger and more complex protein



Protein functions = many and diverse

- Structure
- Enzymes
- Hormones
- Transportation
- Protection
- Sensors
- Toxins
- Gates
- Movement

Proteins comprise >50% of the mass of many cells (the rest being largely water).

More than 10^4 human proteins are known.

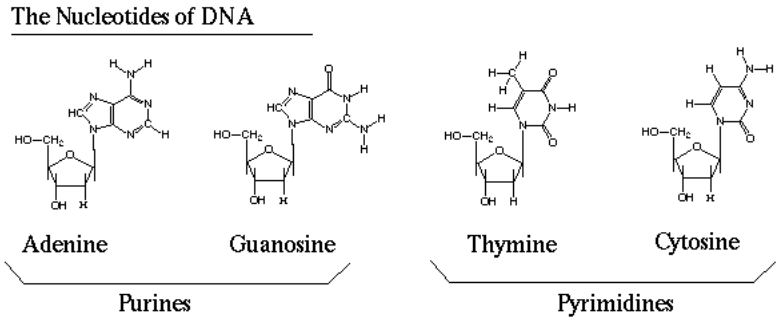
Genetic information specifies proteins and nothing else.

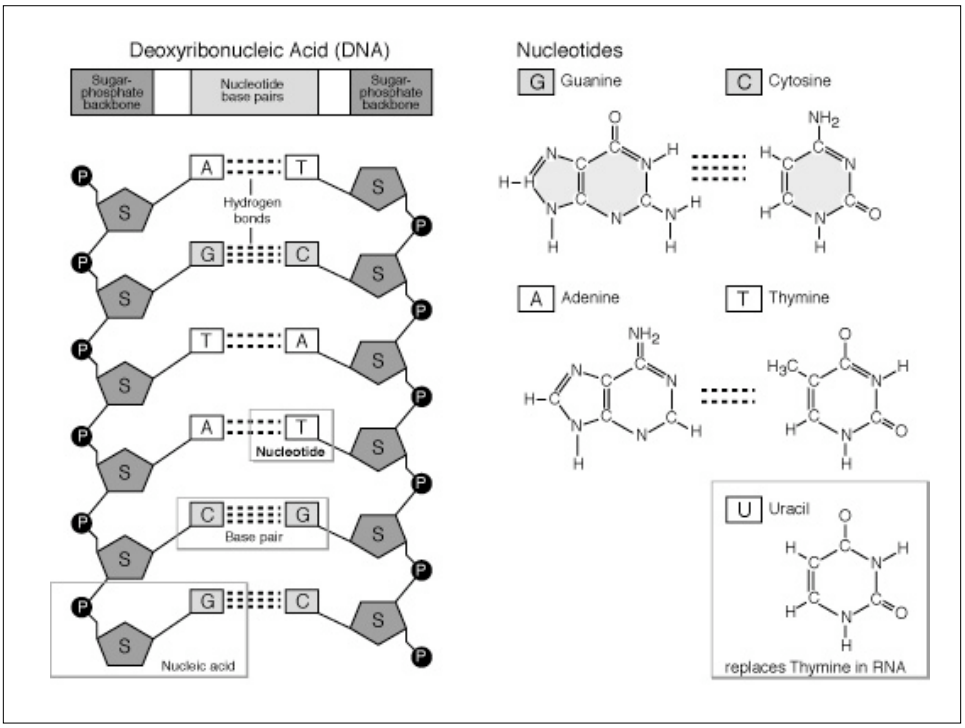
Nucleic acids (DNA and RNA)

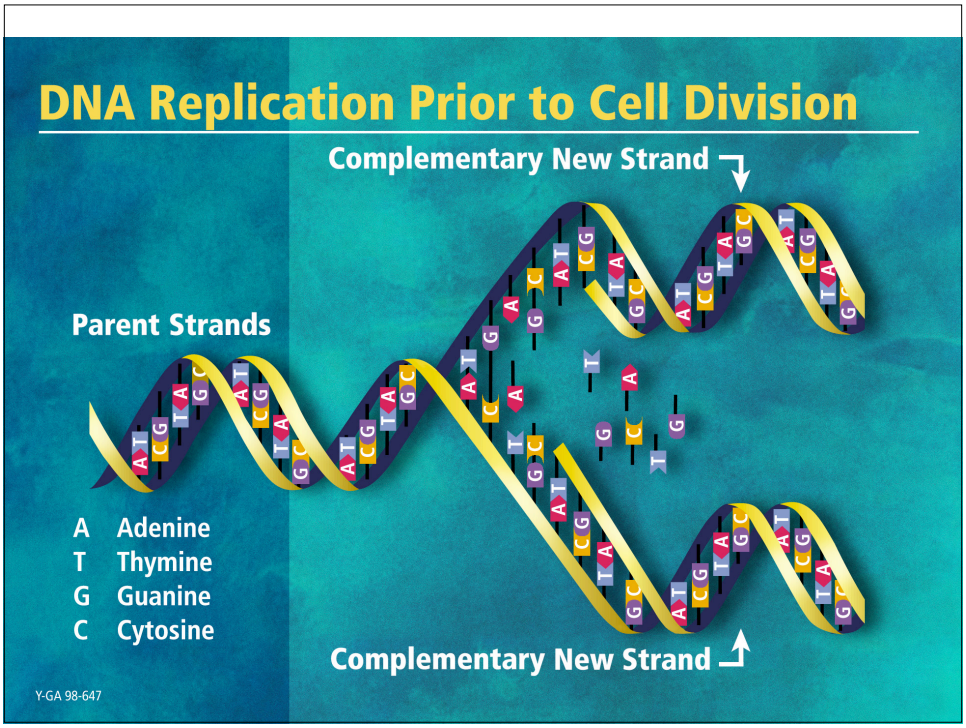
- Very long chains (*again*) of nucleotides
- Each nucleotide is made of a phosphoric acid, a sugar and a base
- Sugar is d-ribose in RNA & deoxy-d-ribose in DNA
- RNA bases are **C**ytosine, **U**racil, **A**denine & **G**uanine; DNA bases = C, A, G & **T**hymine

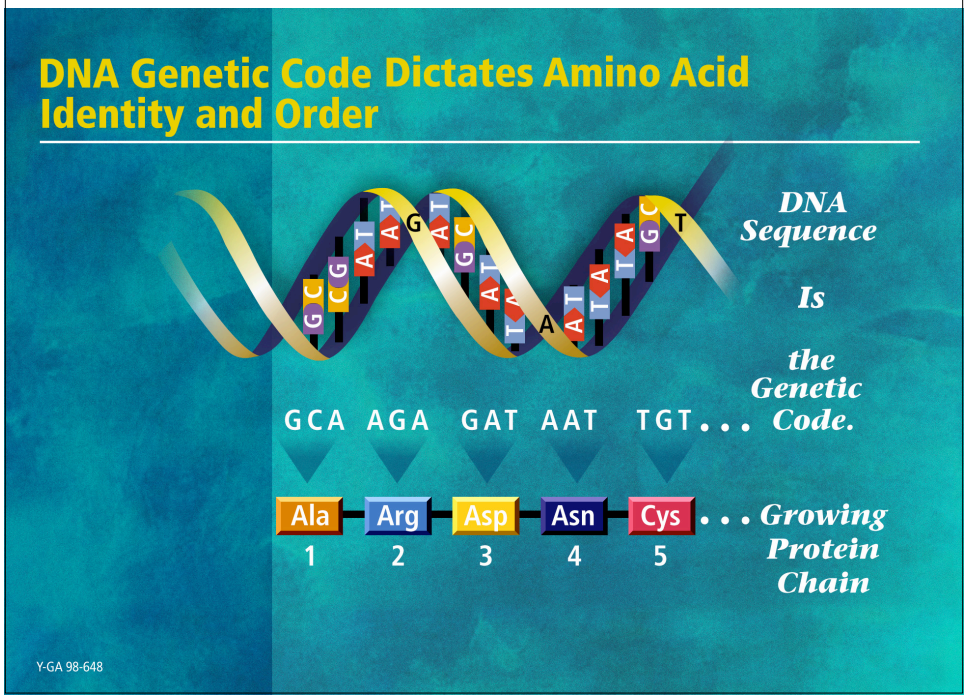
DNA structure & replication

- Consists of two nucleotide chains/strands wrapped around each other in a spiral helix
- **A** on one strand matches **T** on the other
- Similarly **G** and **C** pair between strands
- When the strands are separated, they can each regenerate their partner & thus copy the information they encode
- A *codon* consists of 3 sequential bases and specifies one amino acid (or start/stop)




Deoxyribonucleic Acid (DNA)
Sugar-phosphate backbone
Nucleotide base pairs
Sugar-phosphate backbone
Hydrogen bonds
Nucleotide
Base pair
Nucleic acid
Nucleotides
G Guanine
C Cytosine
A Adenine
T Thymine
U Uracil
replaces Thymine in RNA


DNA Replication Prior to Cell Division
Complementary New Strand
Parent Strands
A Adenine
T Thymine
G Guanine
C Cytosine
Complementary New Strand
Y-GA 98-647

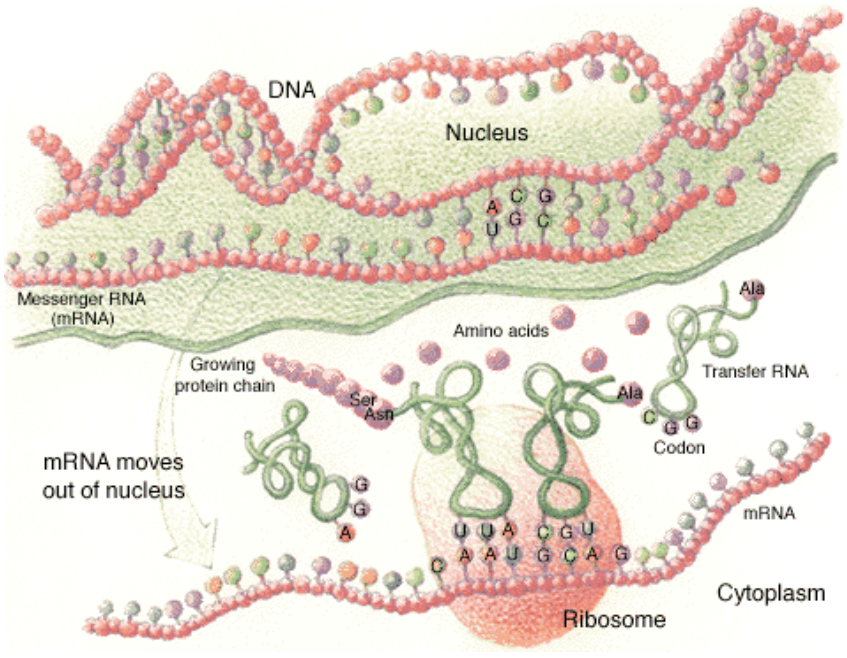


RNA structure & transcription

- Consists of a single chain/strand of nucleotides
- Organized in the same 3 base codons as DNA except that **U** replaces **T**
- DNA generates/transcribes messenger-RNA (mRNA) which provides the "working blueprint" for protein synthesis

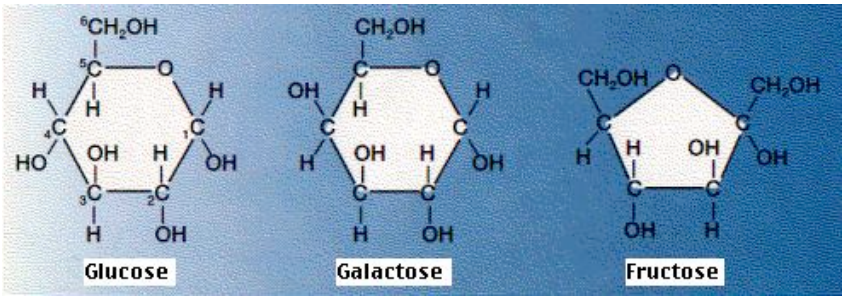
Protein synthesis/construction

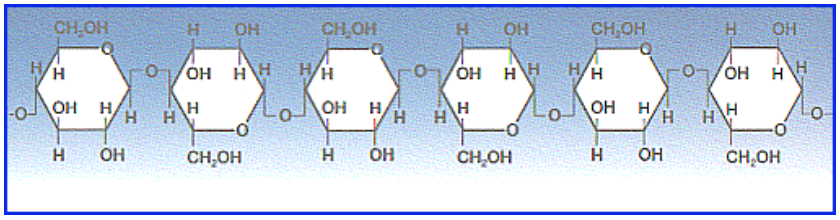
- mRNA carries the amino acid sequence information for the protein
- transfer-RNA (tRNA) provides the raw material amino acids by binding them to anti-codon (complementary base triplets)
- Ribosome macromolecule reads mRNA to select specified amino acids from tRNA and extrudes them in a chain as it moves along the mRNA
- ATP energizes the individual bondings by transfer of a phosphate group to a X-OH component
- mRNA, tRNA & ribosome are reusable for additional syntheses, but ATP degrades to ADP & must be re-energized



Carbohydrates/polysaccharides

- Macromolecules, molecular weights > 10^5
- Polymers/chains made of repeating monomers (identical components)
- Multiples of CH_2O components
- Starches, glycogen & cellulose
- Structural components
- Energy storage

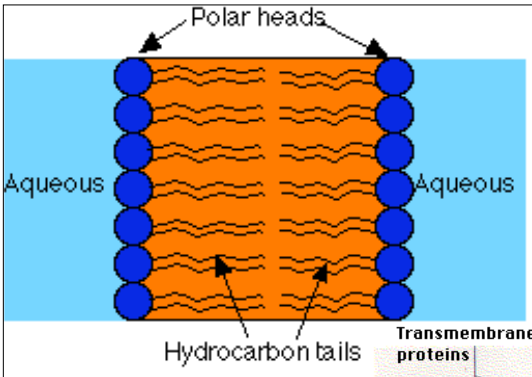




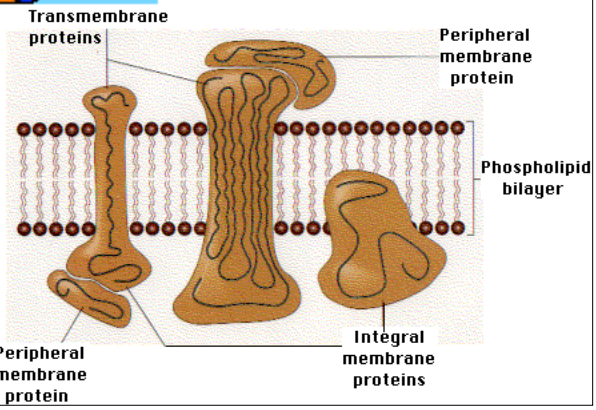
Cellulose - glucose polymer with alternating orientations
(single most abundant organic molecule on Earth)

Lipids

- Composed of fatty acid subcomponents
- Relatively simple/repetitive structures
- Insoluble in water
- Soluble in organic solvents
- Many have hydrophobic properties
- Energy storage
- Cell membranes (essential feature of life!)



Lipid orientation relative to aqueous surroundings forms membranes in which protein structures are imbedded



General Characteristics of the Molecular Biology of Terrestrial Life

- Extraordinarily complex & inter-connected chemical processes, vastly richer than any other known chemical systems
- Basic biochemistry shared by all known terrestrial organisms as well as many of its details
- Carbon based and water dependent
- Hierarchically structured (using much simpler subcomponents), polymerized macromolecules
- Few (4) general classes of compounds but many individual ones with highly specialized and specific biological functions

Implications for (Origin of) Extraterrestrial Life

- Requires no exotic conditions or constituents
- Appears to have happened only once on Earth
- Intricate complexity -> origin problem
  - No obvious route of gradual development
  - Jump in complexity wrt other natural chemistry
  - Absence of theoretical or empirical alternative biochemistries -> one type of life only??
- Physical mechanism for evolutionary adaptation and development once started