Astronomy 330

<u>This class (Lecture 15):</u>

Life on Earth

<u>Next Class:</u>

Origin of Life

HW 6 due next Wednesday

Music: *Bring Me to Life*– Evanescence

Outline

- Monomers and polymers
- Proteins and nucleic acid?
- Where did the monomers of life come from?

Drake Equation

That's 2.68 x ? Life-liking systems/year

Frank Drake

$$N = R_* \times f_p \times n_e \times f_l \times f_i \times f_c \times L$$

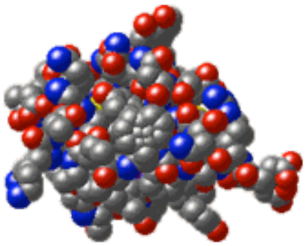
N	R_*	f_p	n_e	f_l	f_i	f_c	L
# of advanced civilizations we can contact in our Galaxy today	Star formation rate	Fraction of stars with planets	# of Earthlike planets per system	Fraction on which life arises	Fraction that evolve intelligence	Fraction that commun-icate	Lifetime of advanced civilizations
	9 stars/ yr	0.29 systems/ star	1.03 x ? = ? planets/ system	life/ planet	intel./ life	comm./ intel.	yrs/ comm.

Why Carbon Based Life?

- Carbon's electronic structure allows it to form long chains
  - Chains of atoms and chains of molecules– complexity
  - Life needs bonds to be stable but breakable
- Good for us, at temperatures at which water is liquid, carbon bonds are stable but breakable
- Organic chemistry is the special branch devoted to carbon chemistry.

Insulin

$C_{254}H_{377}N_{65}O_{76}S_6$



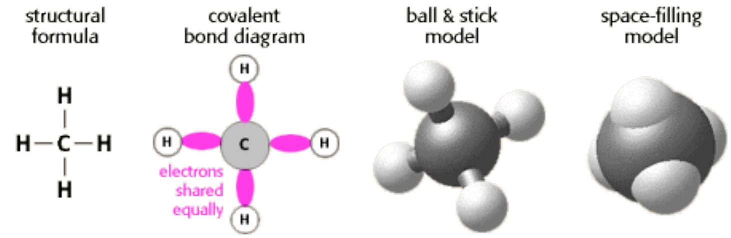
http://www.biology.arizona.edu/biochemistry/tutorials/chemistry/page2.html

Bond, Carbon Bond

Carbon has 6 protons, 6 neutrons, and 6 electrons

- Electrons distribute themselves in "shells"
  - Pauli exclusion principle
  - 1st (inner-most) shell wants to be filled by 2 electrons
  - 2nd shell wants to be filled with 8 electrons
  - BUT, Carbon only has 6 electrons!
    - So, Carbon has 2 electrons in inner shell and 4 in 2nd shell
    - It likes to bond: to "fill" second shell by sharing with four other electrons

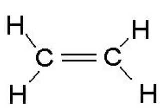
The Simplest C Bond–Methane

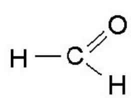


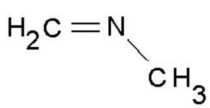
Not many other elements can share 4 bonds.
Silicon, which is much more abundant, can.
Silicon based life?

http://www.biology.arizona.edu/biochemistry/tutorials/chemistry/page2.html

More Bonds

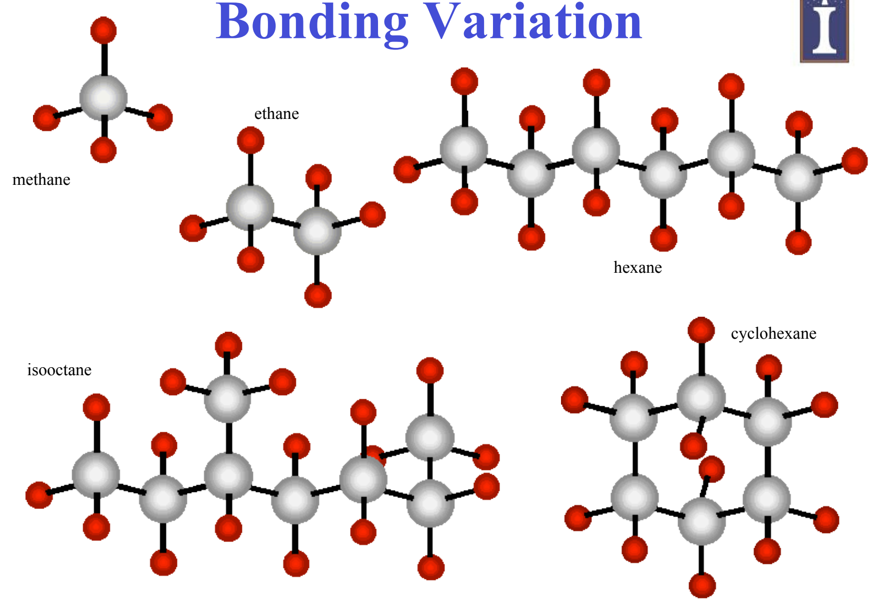






http://www.colossusblog.com/mt/archives/images/drno5.jpg

Bonding Variation



Unique?

As far as we know, the complexity of terrestrial biochemistry can only be achieved with carbon-based molecules.

- Especially considering the need for liquid water
  - Which puts restrictions on the temperature in which the chemical reactions occur

Question

Life uses carbon for making long molecular chains because

a) it is much more abundant than silicon.
b) it likes to share 4 electrons.
c) it is abundant in the ocean.
d) it makes chains that are not easily broken.
e) it is the most abundant element.

Molecular Basis of All Life

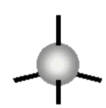
- Great diversity of Life on Earth, but still it is 70% water and 24% four large molecules:
  - Proteins
  - Nucleic Acids
  - Lipids
  - Carbohydrates

In this class, we will focus on the 2 most important molecules

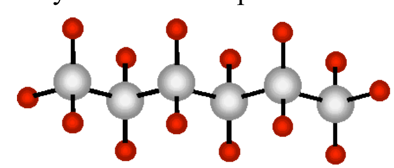
Not completely true. The simplest life, viruses, can have a single molecule of nucleic acid surrounded by a protein coating.

Monomers and Polymers

- All of the fundamental chemicals of life are organic polymers
  - A monomer is a small molecule (like carbon bonds we have seen).
  - A polymer is a number of monomers joined together to form larger, more complex molecules.
  - Polymers are nice for life, as they can form complex and repetitive sequences

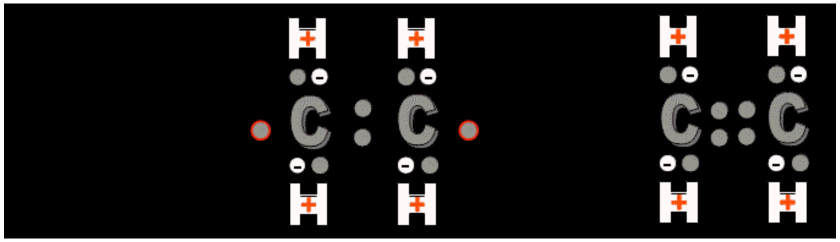


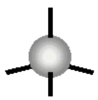
Monomer of C



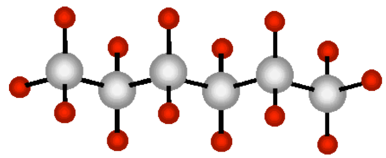
Polymer of hexane

Making A Polymer





Monomer of C



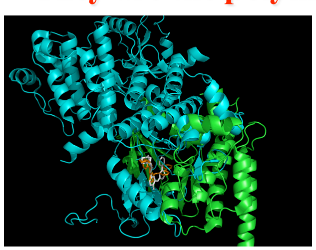
Polymer of hexane

Polymers

- Can form complex, repetitive sequences.
- The order of the monomers determines the function of the polymers.
- Monomers are the letters and words in the molecular basis of life, and polymers are the messages.

Proteins & Nucleic Acids

- Proteins are either structural elements or provide catalytic reactions (enzymes).
- Nucleic acids carry the genetic information– Replication of nucleic acid is crucial to reproduction of organism.
- **They are the polymers of life!**



How is Life Put Together?

- Living things are not just bags of large molecules and polymers mixed in a big soup
  - Living things have structure
  - Plants, animals have different parts
    - Skin, Hair, Leaves, Hearts, etc.

How do these structures relate to the complex organic polymers and nucleic acids?

DNA Based Life

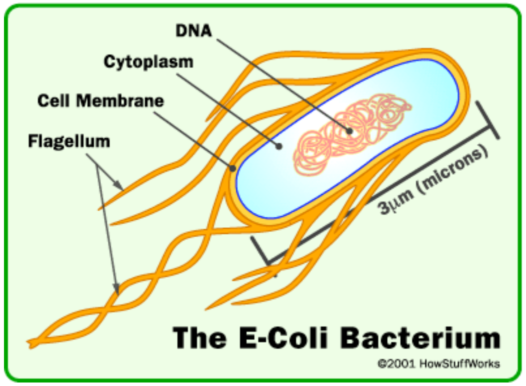
- All life is based on DNA/RNA. What does this mean?
  - The basic reproducible unit of all living organisms is centered around the complex DNA molecule.
  - DNA lives in cells
    - Except in viruses, which are basically pure DNA
  - Cells of different types form different parts of each organism
    - Heart cells different from blood cells.
    - Leaf cells different from root cells.

Cell Bits

- The cell function directly relates to a different organic polymer:
  - <u>Proteins</u>: They form the structural components of the cell or form enzymes that do all the real chemical work inside the cell. Polymers of amino acid monomers.
  - <u>DNA</u>: The genetic coding molecules that controls enzyme and cell reproduction. Polymers of a sugar, phosphate, and nucleotides monomers.

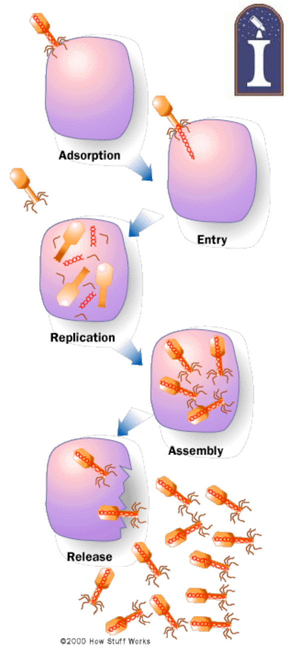
Bacteria Cells

- Simplest cell that exists today.
- Completely self-contained organism.
- Human cells are much more complicated.
- 1 trillion cells in a typical human and they're usually 10 microns in diameter.



Viruses

- Straddles between the living and non-living
- The protein protects the virus until it enters a living cell, where the nucleic acid is released.
- Using the cell's machinery, the nucleic acid reproduces itself.
- They are all parasites, so thought to be from free-living organisms and not descendents of early life.



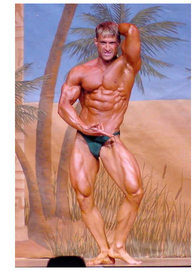
Question

What are the two important polymers for life?

a) Amino acid and sugar
b) Ribonucleic acid and deoxyribonucleic acid
c) Proteins and nucleic acids
d) Deoxyribonucleic Acid and proteins
e) Enzymes and proteins

General Protein Types

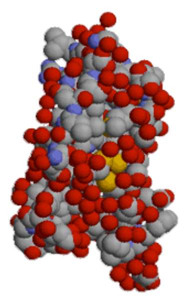
Type	Examples
• Structural	tendons, cartilage, hair, nails
• Contractile	muscles
• Transport	hemoglobin
• Storage	milk
• Hormonal	insulin, growth hormone
• Enzyme	catalyzes reactions in cells
• Protection	immune response



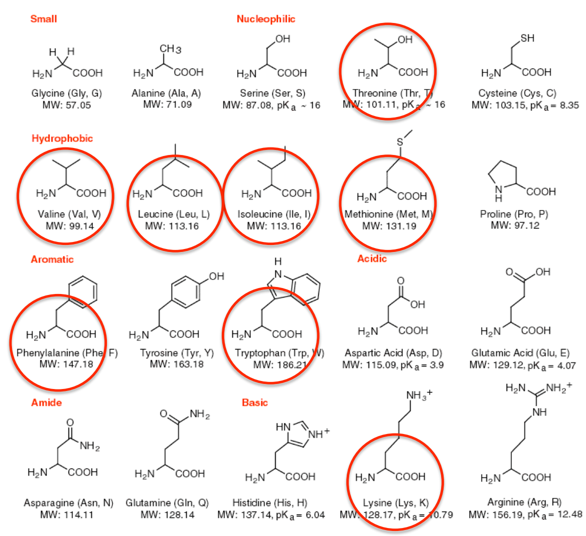
http://66.41.139.241:8000/fitam/muscle.JPG

Focus on Proteins

- Proteins are large, very complex, and very numerous.
- All proteins in living organisms are made from combinations of **20 types** of amino acids (about 100 available though).



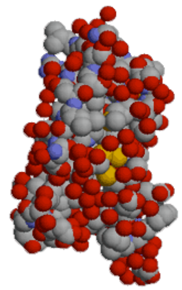
Amino Acids Are for Lovers



http://www.neb.com/neb/tech/tech_resource/miscellaneous/amino_acid.html

Focus on Proteins

- Proteins are made up of 100s to 1000s of ONLY those 20 amino acids, with a particular sequence and shape.
  - This gives 20^{100+} possible combinations
    - How many 100 character sequence can you form from the alphabet?
- BUT, only about 10,000 proteins are used.
- Note, the human body is about 20% protein.



Protein Desert

- The fact that only 10,000 of the billions+ of proteins are used, suggests that life is a little picky.
- Only certain combinations seem to work?
- Does this mean that ET life would find the same useful permutations as Earth life found.
  - Many options were available
  - But, only a small fraction actually worked?

A Type of Protein: Enzymes

- All of the day-to-day work of life is being done by enzymes. Enzymes are little chemical-reaction machines.
- The purpose of an enzyme is to allow the cell to carry out chemical reactions very quickly.
- These reactions allow the cell to build things or take things apart as needed– grow and reproduce.

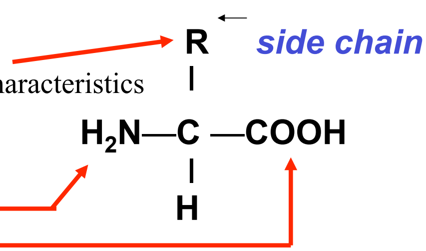
A Type of Protein: Enzymes

- E. coli has about 1,000 different types of enzymes floating around in it at any given time.
- To understand enzymes is to understand cells. To understand cells is to understand life on Earth.
- Maybe similar to life in space?
- Enzymes are made from 3-D structures of amino acids orchestrated by the DNA.

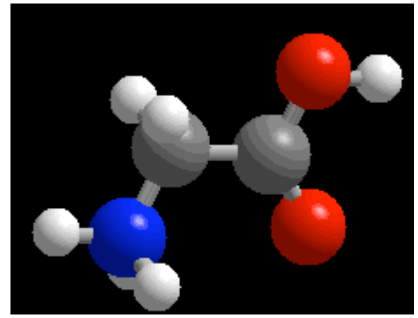
http://res2.agr.ca/lethbridge/emia/images/SEMproj/Ecoli.jpg

Amino Acids

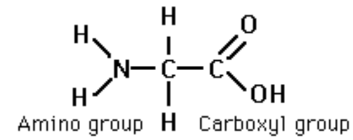
- Are the monomers from which proteins (polymers) are made– building blocks.
- Combinations of the amino acids make the proteins needed– only 20 amino acids used by life.
- Carboxylic acid group
- Amino group
- Side group R gives unique characteristics



Glycine

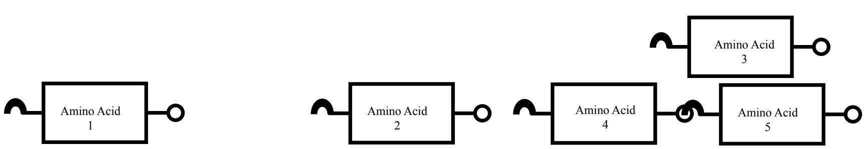


- Simplest amino acid. Just an H in the R position.
- Main ingredients are HONC– other amino acids contain Sulfur (S) as well.



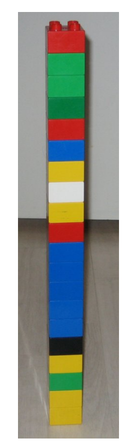
Getting Hooked Up

- Proteins are polymers, made of the monomer, amino acids.
- A number of specific amino acids "hook up" to form a specific protein.
- As a chain grows, there is always a hook (the amino group) on one end and an eye (the carboxyl group) on the other.



Amino Acids

- Can think of the 20 amino acids as different color Legos.
- Each color is a different piece, but they can all be put together into a tower.
- This tower is a specific protein.
- The function depends on the colors used.



http://www.frepa.org/wp/wp-images/lego-torn.jpg

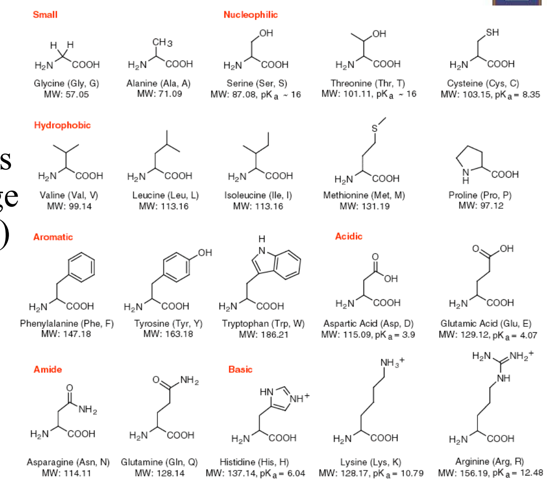
Question

If we think of a protein as a stack of Legos, then what are the blocks?

a) Amino acids
b) Ribonucleic Acids
c) Proteins
d) Deoxyribonucleic Acids
e) Enzymes

Build Me Up

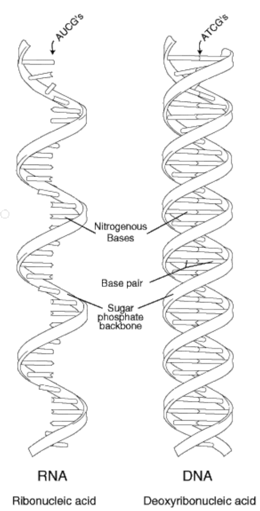
- Amino acids are essential for life– building blocks.
- But who orchestrates or writes the message (the special proteins) that the amino acids make up?
- Need something to teach them how to spell.



http://www.neb.com/neb/tech/tech_resource/miscellaneous/amino_acid.html

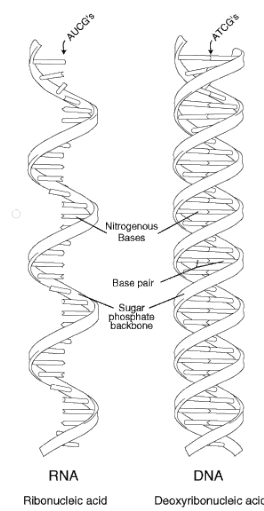
Nucleic Acid: DNA and RNA

- Two types of nucleic acid.
- A polymer built up from monomers we'll come back to which ones.
- RNA (RiboNucleic Acid) is usually a long strand
- DNA (DeoxyriboNucleic Acid) is the double helix– visualize as a spiral ladder.



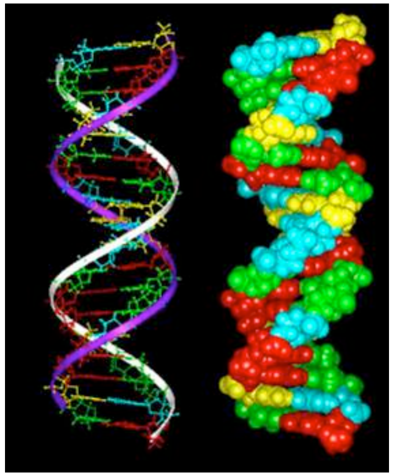
Nucleic Acid: DNA and RNA

- These molecules carry the genetic information of the organism– the message that gets coded into the amino acid chain.
- It is very much like computer code in many ways– and teaches how to spell useful word (proteins) out of the letters of the available amino acids.



DNA / RNA

- The origins of DNA and RNA are mysterious and amazing
- DNA/RNA are complex: Built from three basic types of monomers
  1. Sugar (deoxyribose or ribose)
  2. A phosphate PO_4
  3. One of four "nitrogenous bases"
     - Adenine (A)
     - Guanine (G)
     - Cytosine (C)
     - Thymine (T) in DNA / Uracil (U) in RNA



Question

What is the DNA molecule's role?

a) Telling proteins what to do.

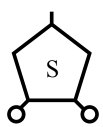
b) A dictionary for protein spellings.

c) Sitting around all day and lording over the other parts of the cell.

d) Deoxyribonucleic Acid

e) Invading other animals.

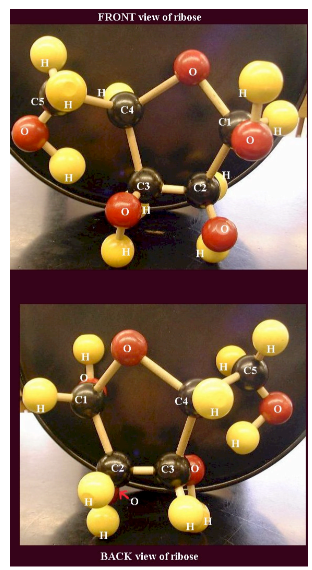
Sugars: Ribose or Deoxyribose

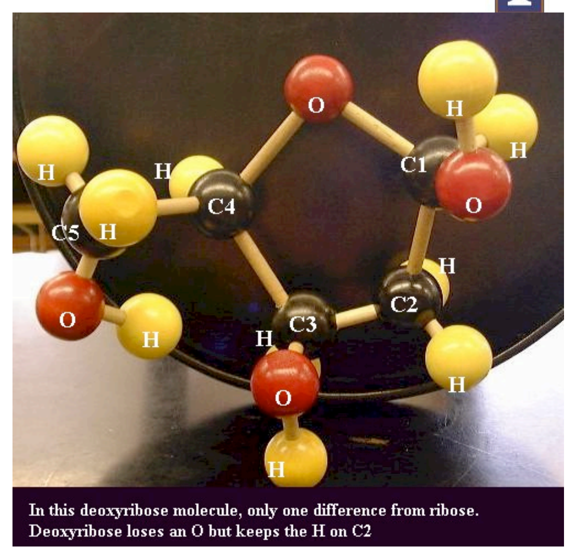


We will represent the sugar molecule (either ribose or deoxyribose) as a pentagon with two eyes.

http://www.dscc.edu/bwilliams/Biology/biology1molemodels.htm

Sugars: Ribose or Deoxyribose

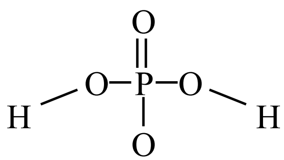


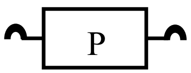


http://www.dscc.edu/bwilliams/Biology/biology1molemodels.htm

Phosphates

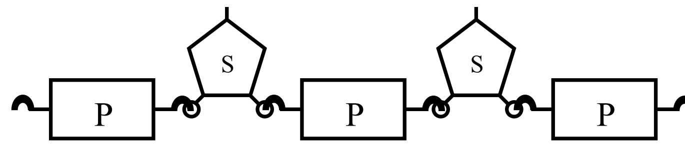
- Is often referred to as phosphoric acid.
- Makes five bonds with oxygen.





Phosphates and Sugars

- Make the sides of the twisted DNA ladder structure.
- Sugars and phosphates connect up in alternating bonds.   P-S-P-S-P-S-P
- These are phosphodiester bonds.



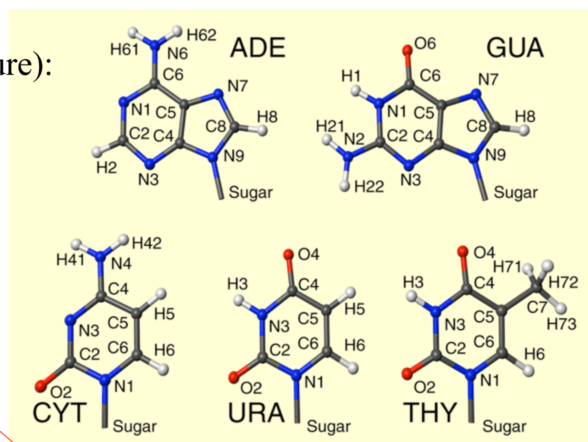
And the Bases

5 types in 2 groups (based on structure):

- Purines:
  - Adenine
  - Guanine
- Pyrimidines:
  - Cytosine
  - Uracil
  - Thymine

For DNA

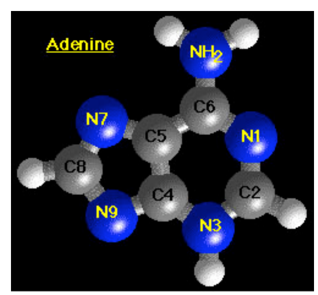
For RNA

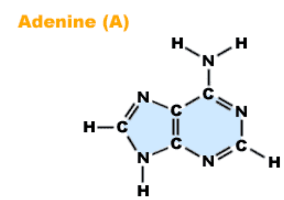


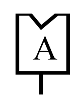
http://www.bmrb.wisc.edu/referenc/nomenclature/figures/bases.gif

Purines: Adenine

- 5-sided ring built on the side of a 6-sided ring.







Adenine

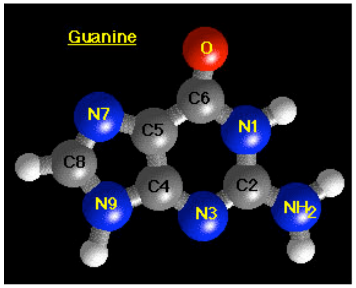
$H_5C_5N_5$

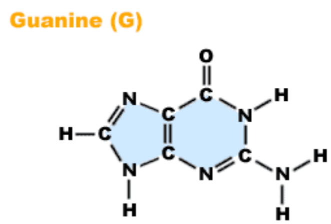
http://resources.emb.gov.hk/biology/english/inherit/genetics.html

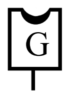
http://dlm.tmu.edu.tw/phase2/glossary/image/adenine.gif

Purines: Guanine

- 5-sided ring built on the side of a 6-sided ring.







Guanine

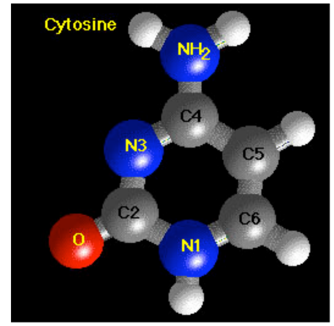
$H_5C_5N_5O$

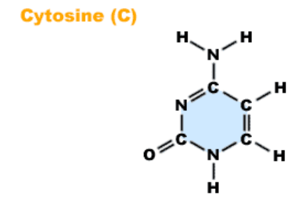
http://resources.emb.gov.hk/biology/english/inherit/genetics.html

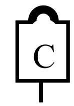
http://dlm.tmu.edu.tw/phase2/glossary/image/adenine.gif

Pyrimidines: Cytosine

- 6 sided rings (without a 5 sided ring)







Cytosine

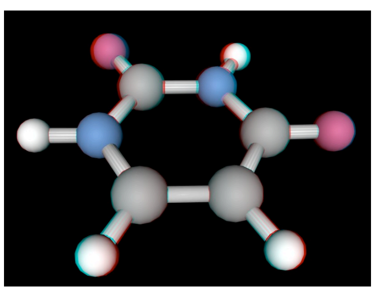
$H_5C_4N_3O$

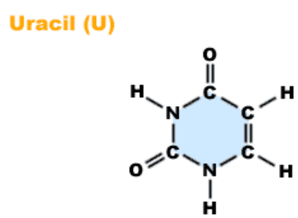
http://resources.emb.gov.hk/biology/english/inherit/genetics.html

http://dlm.tmu.edu.tw/phase2/glossary/image/adenine.gif

Pyrimidines: Uracil

- 6 sided rings (without a 5 sided ring)





Uracil

$H_4C_4N_2O_2$

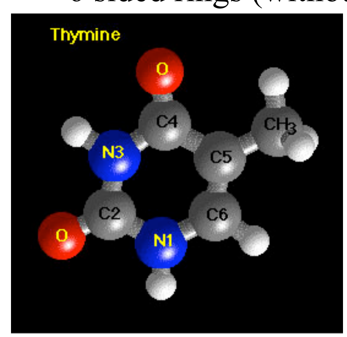
For RNA

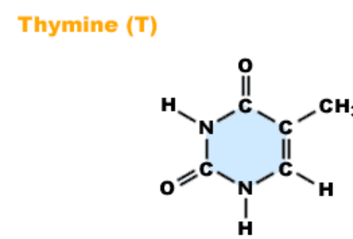
http://nautilus.fis.uc.pt/molecularium/stereo/

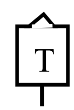
http://dlm.tmu.edu.tw/phase2/glossary/image/adenine.gif

Pyrimidines: Thymine

- 6 sided rings (without a 5 sided ring)







Thymine

$H_6C_5N_3O_2$

For DNA

http://resources.emb.gov.hk/biology/english/inherit/genetics.html

http://dlm.tmu.edu.tw/phase2/glossary/image/adenine.gif

Monomers and Polymers

Monomer:

1. Amino acids →
2. Sugar →
   phosphate
   nitrogenous bases

Polymer:

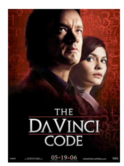
1. Proteins
2. Nucleic acids

Question

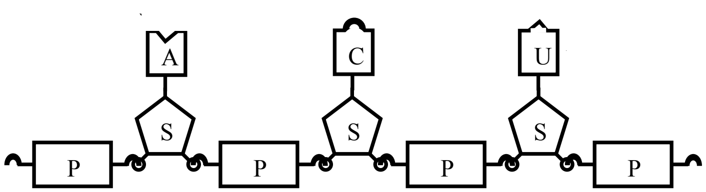
Which of the following is not a polymer?

a) Amino acid
b) Ribonucleic Acid
c) Protein
d) Deoxyribonucleic Acid
e) Enzyme.

Making RNA Mean Something



- Schematic of an RNA molecule.
- This segment can be read from left to right as ACU– called a <u>codon</u> (a three letter word, so to speak)
- Can be translated to a specific amino acid (**the code!**) – this corresponds to the amino acid Threonine.
  GGU is gylcine.
- By building up these amino acid codons, we can spell out (and thus construct) a protein.



Meaning in Mystery

FIRST LETTER	SECOND LETTER: U	C	A	G	THIRD LETTER
U	Phenylalanine	Serine	Tyrosine	Cysteine	U
	Phenylalanine	Serine	Tyrosine	Cysteine	C
	Leucine	Serine	Stop	Stop	A
	Leucine	Serine	Stop	Tryptophan	G
C	Leucine	Proline	Histidine	Arginine	U
	Leucine	Proline	Histidine	Arginine	C
	Leucine	Proline	Glutamine	Arginine	A
	Leucine	Proline	Glutamine	Arginine	G
A	Isoleucine	Threonine	Asparagine	Serine	U
	Isoleucine	Threonine	Asparagine	Serine	C
	Isoleucine	Threonine	Lysine	Arginine	A
	(Start) Methionine	Threonine	Lysine	Arginine	G
G	Valine	Alanine	Aspartate	Glycine	U
	Valine	Alanine	Aspartate	Glycine	C
	Valine	Alanine	Glutamate	Glycine	A
	Valine	Alanine	Glutamate	Glycine	G

For DNA replace U with T

http://library.thinkquest.org/C004535/PF_amino_acids.html

Question

The Codon code is

a) a bad movie starring Tom Hanks.
b) how DNA encodes the 20 amino acids by using bases.
c) how RNA encodes the 10,000 proteins by using bases.
d) a three letter word using the 20 letters of the Wasibi alphabet.
e) a bad book starring Tom Hanks.

Overconstrained

- 4 options for each letter in the Codon
- 4 x 4 x 4 = 64 options (can think if it as bits)
- But only 20 amino acids ⇒ over constrained
- 4 x 4 = 16 wouldn't work.
- Life picked the next highest number and copes with redundancy.

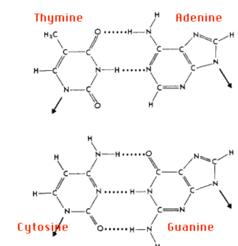
FIRST LETTER	SECOND LETTER: U	C	A	G	THIRD LETTER
U	Phenylalanine	Serine	Tyrosine	Cysteine	U
	Phenylalanine	Serine	Tyrosine	Cysteine	C
	Leucine	Serine	Stop	Stop	A
	Leucine	Serine	Stop	Tryptophan	G
C	Leucine	Proline	Histidine	Arginine	U
	Leucine	Proline	Histidine	Arginine	C
	Leucine	Proline	Glutamine	Arginine	A
	Leucine	Proline	Glutamine	Arginine	G
A	Isoleucine	Threonine	Asparagine	Serine	U
	Isoleucine	Threonine	Asparagine	Serine	C
	Isoleucine	Threonine	Lysine	Arginine	A
	(Start) Methionine	Threonine	Lysine	Arginine	G
G	Valine	Alanine	Aspartate	Glycine	U
	Valine	Alanine	Aspartate	Glycine	C
	Valine	Alanine	Glutamate	Glycine	A
	Valine	Alanine	Glutamate	Glycine	G

DNA

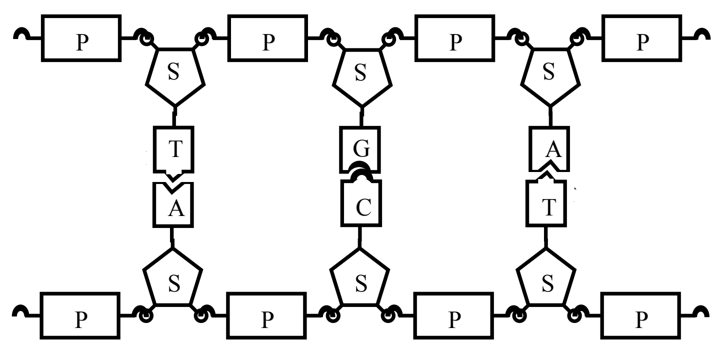
- For life more complicated than viruses, the genetic code is stored in DNA.
- Differs from RNA in a few ways: uses deoxyribose sugar rather than ribose sugar and it uses thymine instead of uracil.
- Forms the double strand where two complementary bonds are held together with weaker hydrogen bonding– allowing easier separation.
- In that case, bases form unique pairs:
  – AT, TA, GC, CG

http://mbsu.sus.mcgill.ca/POST_MIDTERM PICS/DNA is my life.jpg

DNA

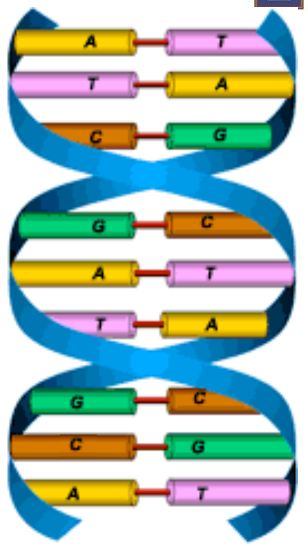


- A codon of DNA: AT, CG, TA
  – purine to pyrimidine connections



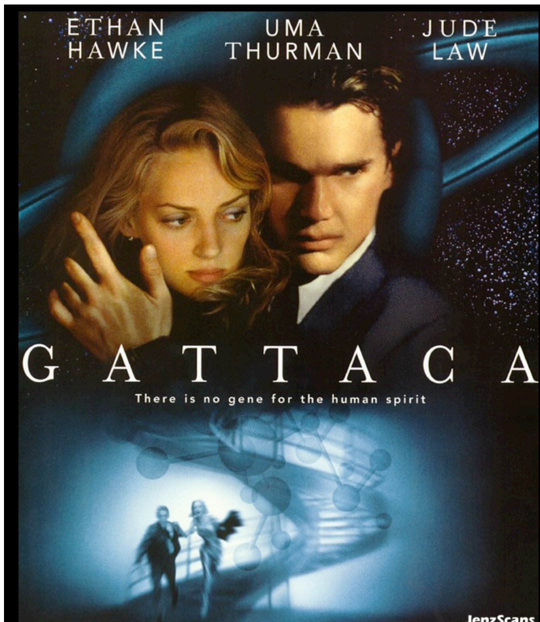
The Double Helix

- Resembles a twisted ladder
- The sides of the DNA ladder are made of the sugar and phosphate.
- The steps or rungs of the ladder are composed of one of the 4 nitrogenous base pairs.
  – AT, TA, GC, CG
- In other words, if you know the sequence on one side, you can deduce the sequence on the other side.



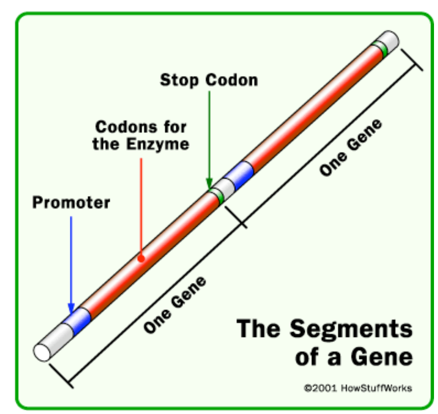
The Double Helix

- The ladder is twisted into the helix shape since the hydrogen bonds are at an angle.
- 3 pairs make up a codon, like RNA (4x4x4 = 64)
- Each codon is info on the amino acid, but only 20 of those– again over constrained.



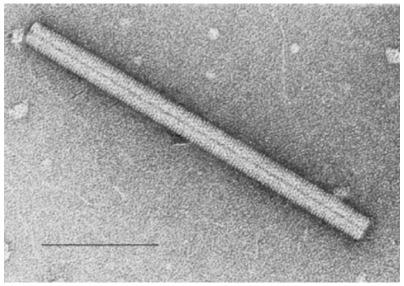
Genes

- Each codon specifies an amino acid, and a sequence of condons specifies a protein or enzyme.
- E. coli bacterium has about 4,000 genes, and at any time those genes specify about 1,000 enzymes. Many genes are duplicates.

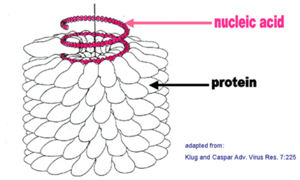


Ta-Backy

- Different organisms have different number of genes.
- Tobacco mosaic virus has 4 genes.
- A small bacterium has about 1000 genes– average sized bacterium has 4000 genes.



TOBACCO MOSAIC VIRUS



http://pathmicro.med.sc.edu/mhunt/intro-vir.htm