Life on Earth

- Basic biochemical structure
- Origin and Evolution of Life
- How strange can life be?

Basic biochemical structure built on C-chain molecules:

Simplicity:	Complexity:
Limited number of monomer building blocks	Large number of polymer molecules

20 Amino acids → proteins

4 Nucleotides → genetic code (DNA cross-links)

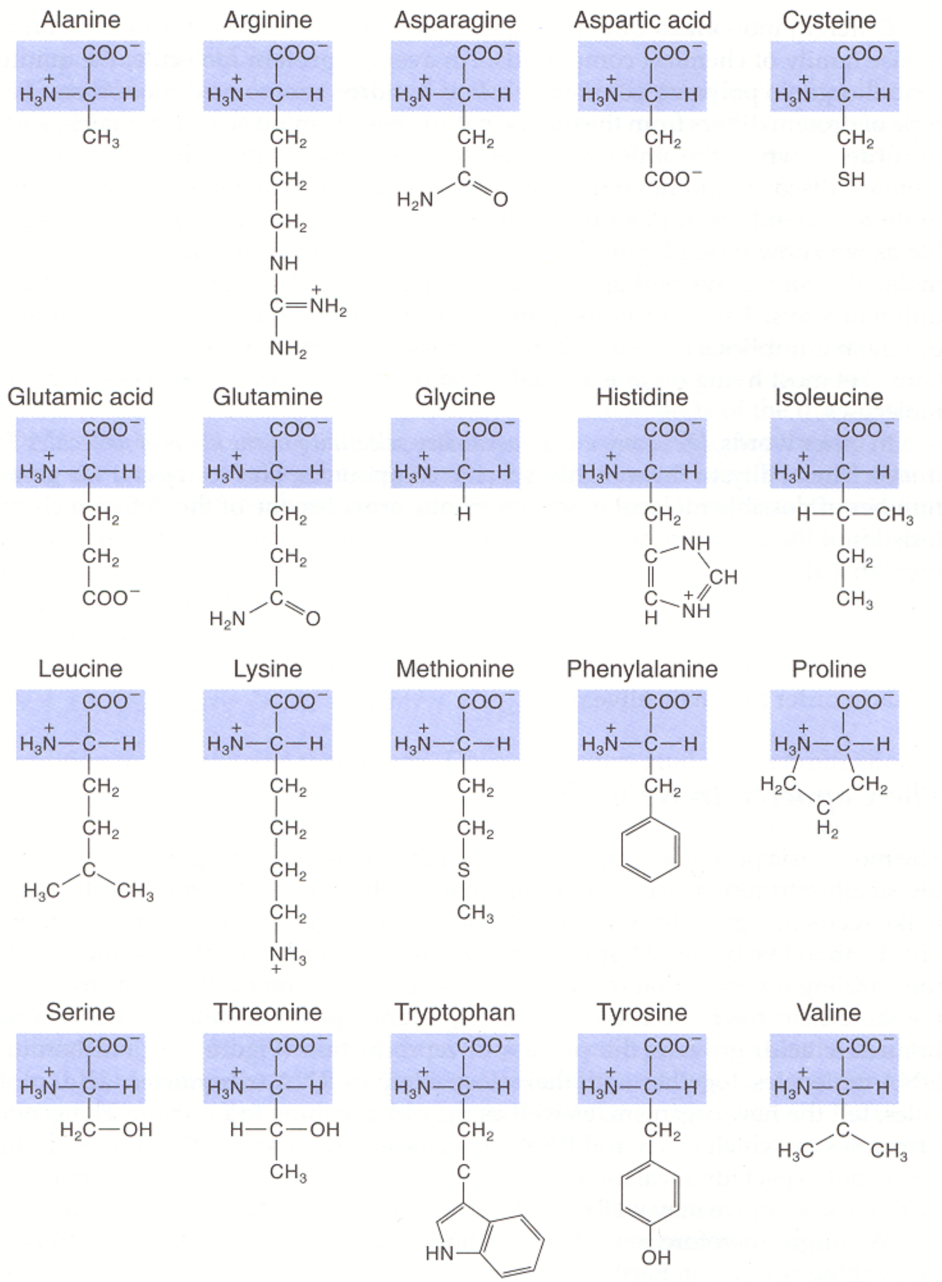
Sugars → structural and energy storing molecules

Distinguish between

- Polymers repeating without ordering (e.g. cellulose, gycogen)
- Polymers with specific ordering of monomers → complex shapes and highly specific functionality (e.g. proteins)

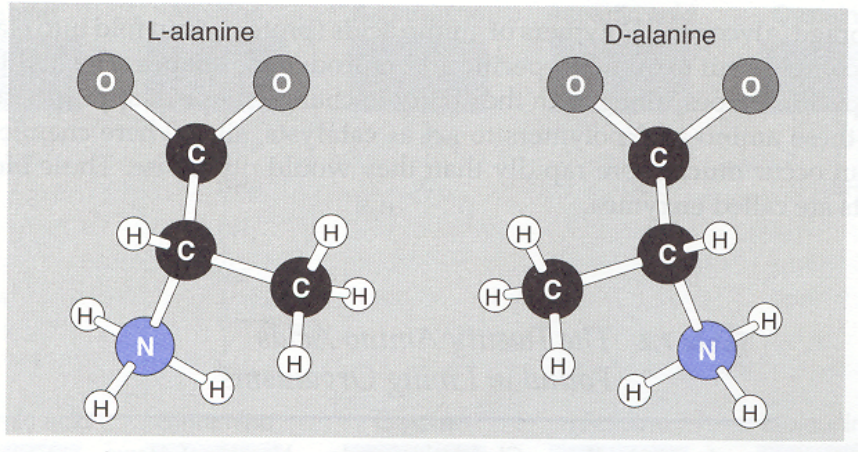
20 amino acids → 10,000 proteins << “theoretical number” (see Lecture 1)

20 amino acids in terrestrial life



<u>All</u> amino-acid monomers found in terrestrial life are “left handed” (in sense of arrangement of bonds to CO_2, NH_3, H and 4th branches)

Note: glycine is actually ambidextrous because of having two H atoms

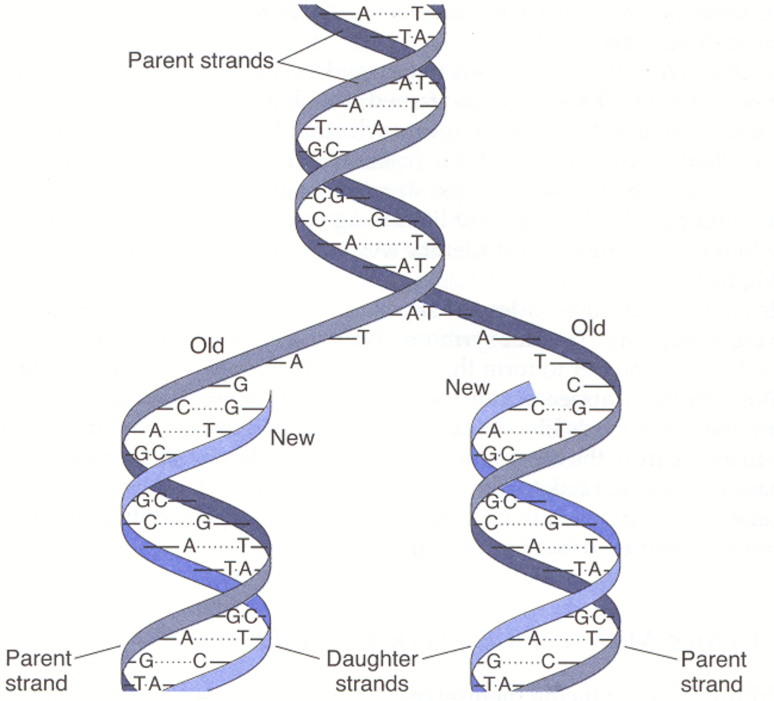


- System of right-handed monomers would “work” equally well
- Having only one handedness allows greater structural definition of molecules
- Evidence that amino-acids in Murchison meteorite are 2-9% enhanced in Left vs. Right (<u>not</u> terrestrial contamination) – see Science Feb 14, 1997

Role of DNA in replication and protein synthesis

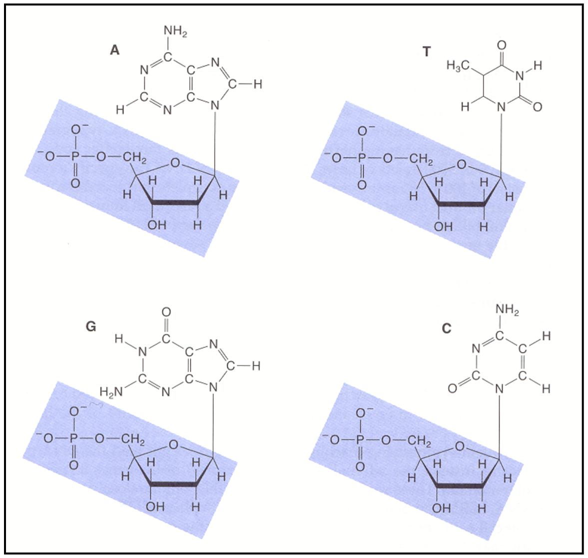
Double helix structure of DNA with ladder of sugar+phosphate and "rungs" composed of nucleotide pairs:

Adenine (A) + Thymine (T)
Cytosine (C) + Guanine (G)

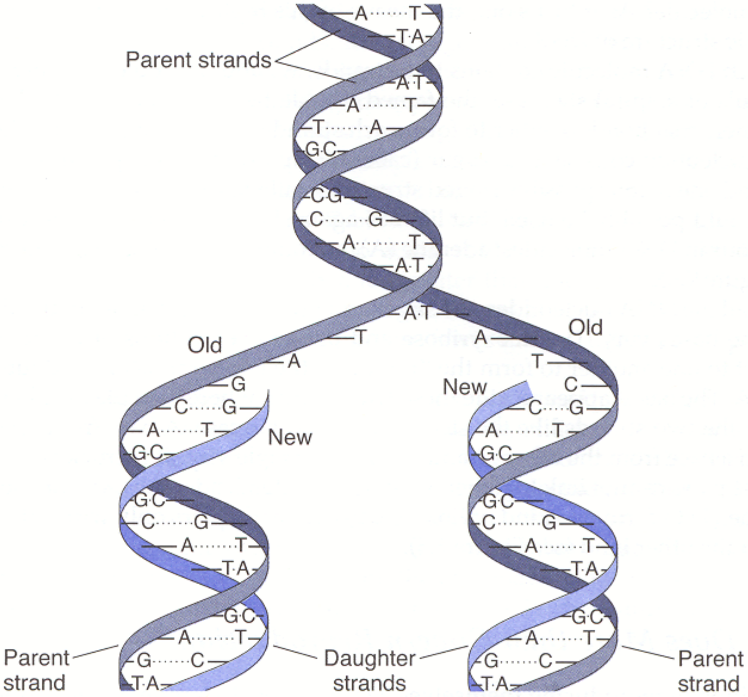


When split in half:

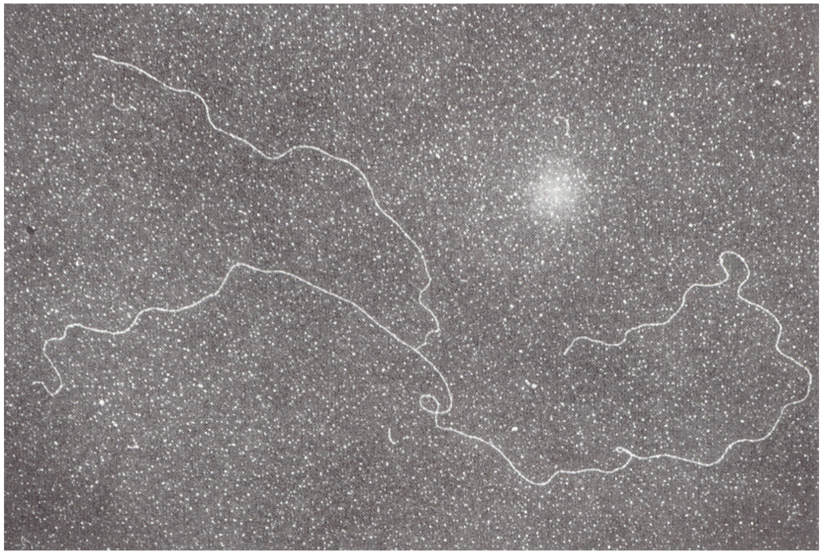
- each side contains all the required information to replicate the missing half
- each single strand can code for amino-acid chains (RNA)



Role of DNA in replication and protein synthesis:

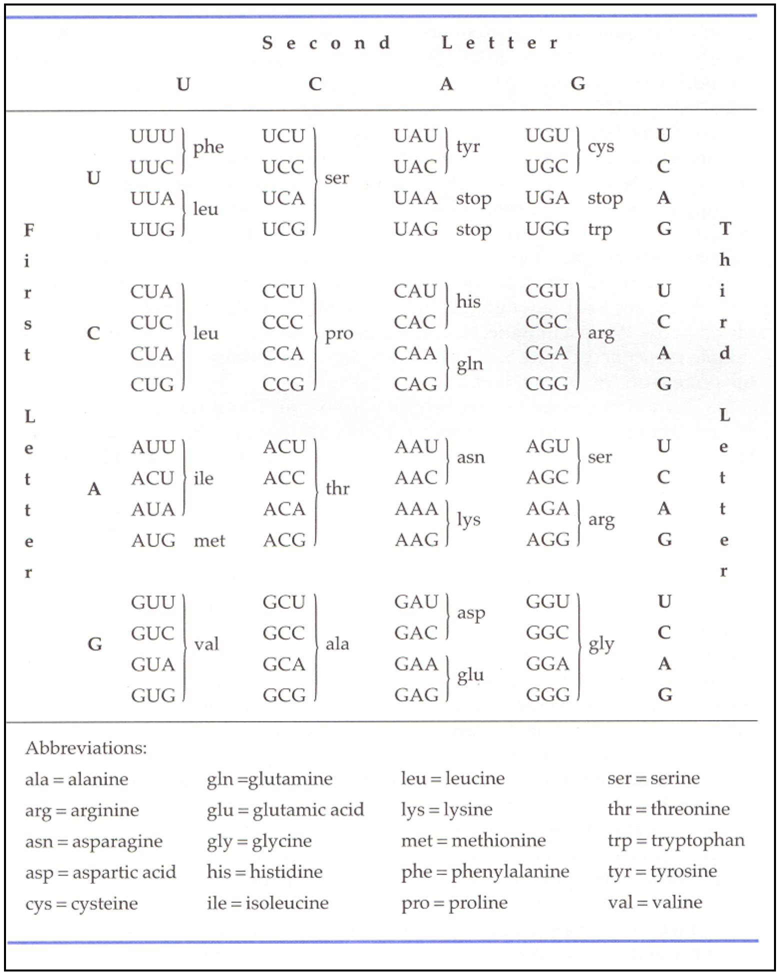


Replication through splitting now seen directly in electron microscopy



Electron micrograph of DNA in process of replication

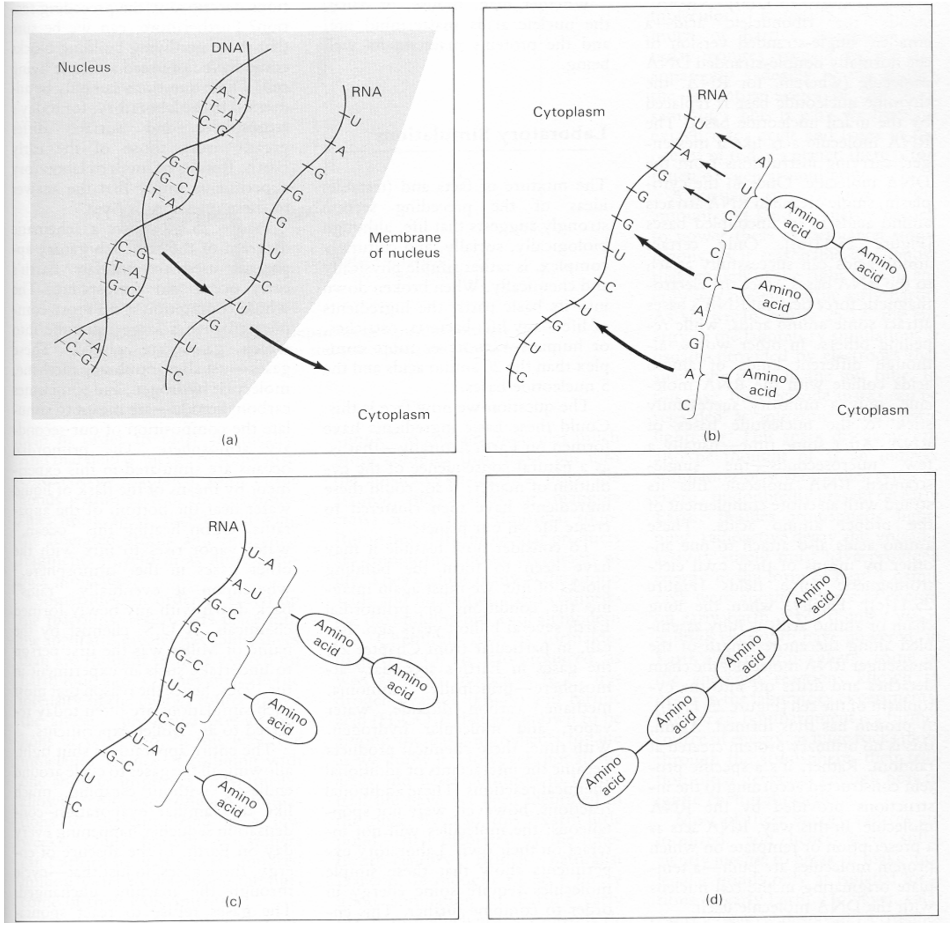
Triplets of bases on RNA code for individual aminoacids in protein chain (with some redundancy)



First Letter	Second Letter: U		C		A		G		Third Letter
U	UUU	phe	UCU	ser	UAU	tyr	UGU	cys	U
	UUC		UCC		UAC		UGC		C
	UUA	leu	UCA		UAA	stop	UGA	stop	A
	UUG		UCG		UAG	stop	UGG	trp	G
C	CUU	leu	CCU	pro	CAU	his	CGU	arg	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	gln	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	ile	ACU	thr	AAU	asn	AGU	ser	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	lys	AGA	arg	A
	AUG	met	ACG		AAG		AGG		G
G	GUU	val	GCU	ala	GAU	asp	GGU	gly	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	glu	GGA		A
	GUG		GCG		GAG		GGG		G

Abbreviations:

ala = alanine
arg = arginine
asn = asparagine
asp = aspartic acid
cys = cysteine
gln = glutamine
glu = glutamic acid
gly = glycine
his = histidine
ile = isoleucine
leu = leucine
lys = lysine
met = methionine
phe = phenylalanine
pro = proline
ser = serine
thr = threonine
trp = tryptophan
tyr = tyrosine
val = valine



Note: 4 bases give $4^3 = 64$ triplets (> 20 amino acids + 1 “stop”) but only $4^2 = 16$ pairs.
2 bases would give only 8 triplets.

A key feature of replication is that it is not perfect

Inaccuracies in replication of DNA (mutations) due to:

- Radiation (cosmic rays or radioactivity)
- Chemical agents
- Rare errors of DNA copying machinery

Differential propagation of the resulting “trials”

- “natural selection” in one form or another

Charles Darwin

Thomas Malthus

Gregor Mendeleev

Energy for Life:

- Chemical from surroundings (very small amounts)
- Geothermal
- Sunlight
- Other life products (sugars)

Origin of Life on Earth:

0. Earth's evolving atmosphere:

- Initially highly reducing (NH_3, H_20, CH_4)
- Early: photo-dissociation by solar ultraviolet and H loss → less reducing atmosphere (CO, CO_2, N_2 and H_20)
- Later: Production of O_2 by Life (photosynthesis) → oxidizing atmosphere (N_2, O_2)

0.1 Panspermia?

- Long duration interstellar travel hazardous (radiation, CR)
- "Martian" meteorites suggest interplanetary travel not impossible
- Begs the question of the "origin"
- But: complex organic molecules exist in space (see later)

1. Chemical synthesis of organic chemicals:

- Requires reducing atmosphere and absence of free oxygen
- Miller-Urey experiments (highly reducing atmosphere) make amino-acids, sugars, nucleotide bases etc: e.g. glycine

$$NH_3 + 2\,CH_4 + 2\,H_2O + \text{energy} \rightarrow C_2H_5O_2N + 5\,H_2$$

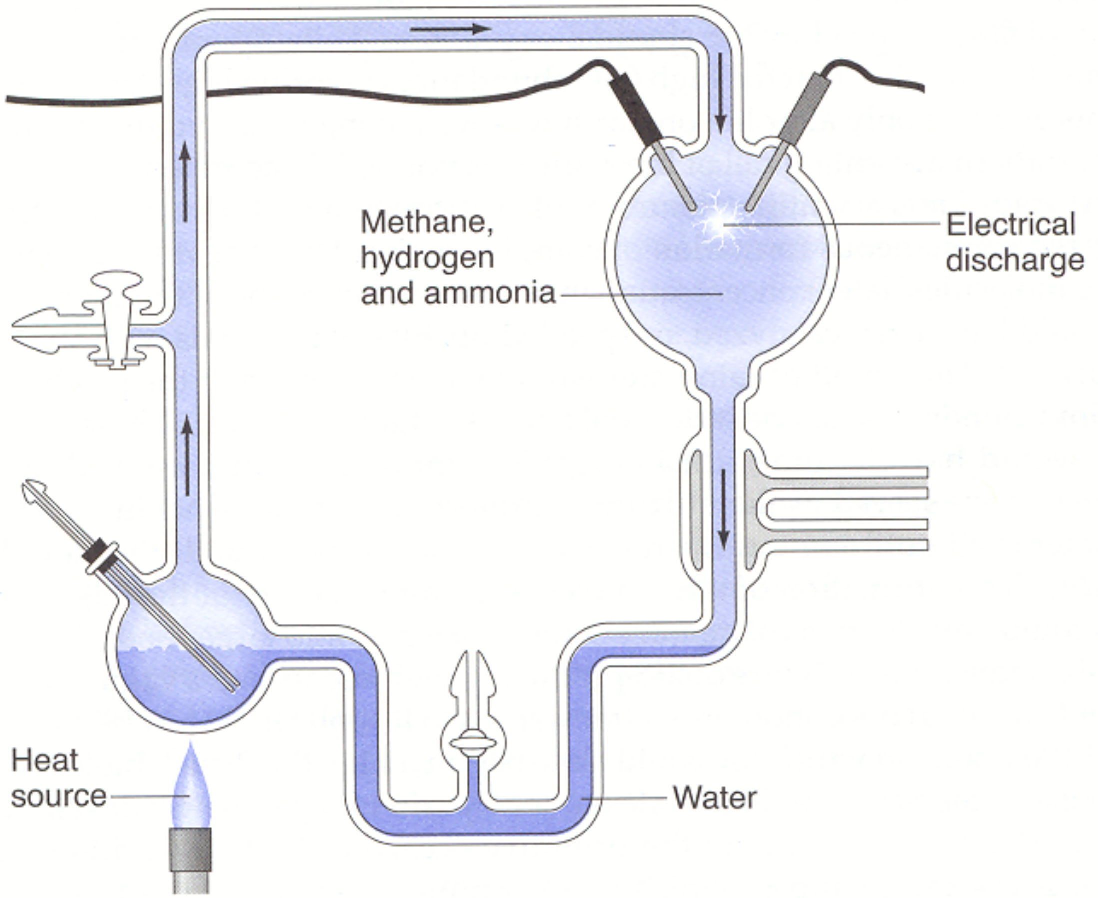
- Experiments with less reducing atmospheres also work:

$$3\,HCN + 2\,H_2O + \text{energy} \rightarrow C_2H_5O_2N + CN_2H_2$$

- Reactions in small surface pools are attractive: preserve reagents, utilize solar energy etc. c.f. undersea vents which have short duration, severe dilution, extreme conditions.
- Role of impacts? External origin of organic molecules? – e.g. amino-acids in meteorites. Renewed source of reducing atmosphere by replenishing volatile materials.

e.g. Murchison meteorite: 14,000 organic compounds including 70 amino acids, small excess of L- over R- chirality.

Miller-Urey-type experiments



Miller-Urey experiment (1953)
One of several similar conducted at the time.

Debate over precise composition of atmosphere (e.g. effect of volcanoes, CO_2 vs CH_4, H-content etc) but basic conclusions remain.

Analysis of sealed vials indicate results better than earlier analysis indicated: >25 amino acids produced. Most of the natural amino acids, hydroxyacids, purines, pyrimidines, and sugars have been made in variants of the experiment.

Bottom line: it is relatively easy to make basic organic building blocks of complex molecules in the kind of reducing atmosphere(s) that are expected on early Earth.

Amino acids in Miller-Urey, Murchison and used in proteins

Amino acid	Miller-Urey	Murchison meteorite	Used in proteins
glycine	****	****	✔
alanine	****	****	✔
α-amino-N-butyric acid	***	****	✗
α-aminoisobutyric acid	****	**	✗
valine	***	**	✔
norvaline	***	***	✗
isovaline	**	**	✗
proline	***	*	✔
pipecolic acid	*	*	✗
aspartic acid	***	***	✔
glutamic acid	***	***	✔
β-alanine	**	**	✗
β-amino-N-butyric acid	**	**	✗
β-aminoisobutyric acid	*	*	✗
γ-amino-butyric acid	*	**	✗
sarcosine	**	***	✗
N-ethylglycine	**	**	✗
N-methylalanine	**	**	✗

2. Polymerization (how to make the initial strand of DNA)

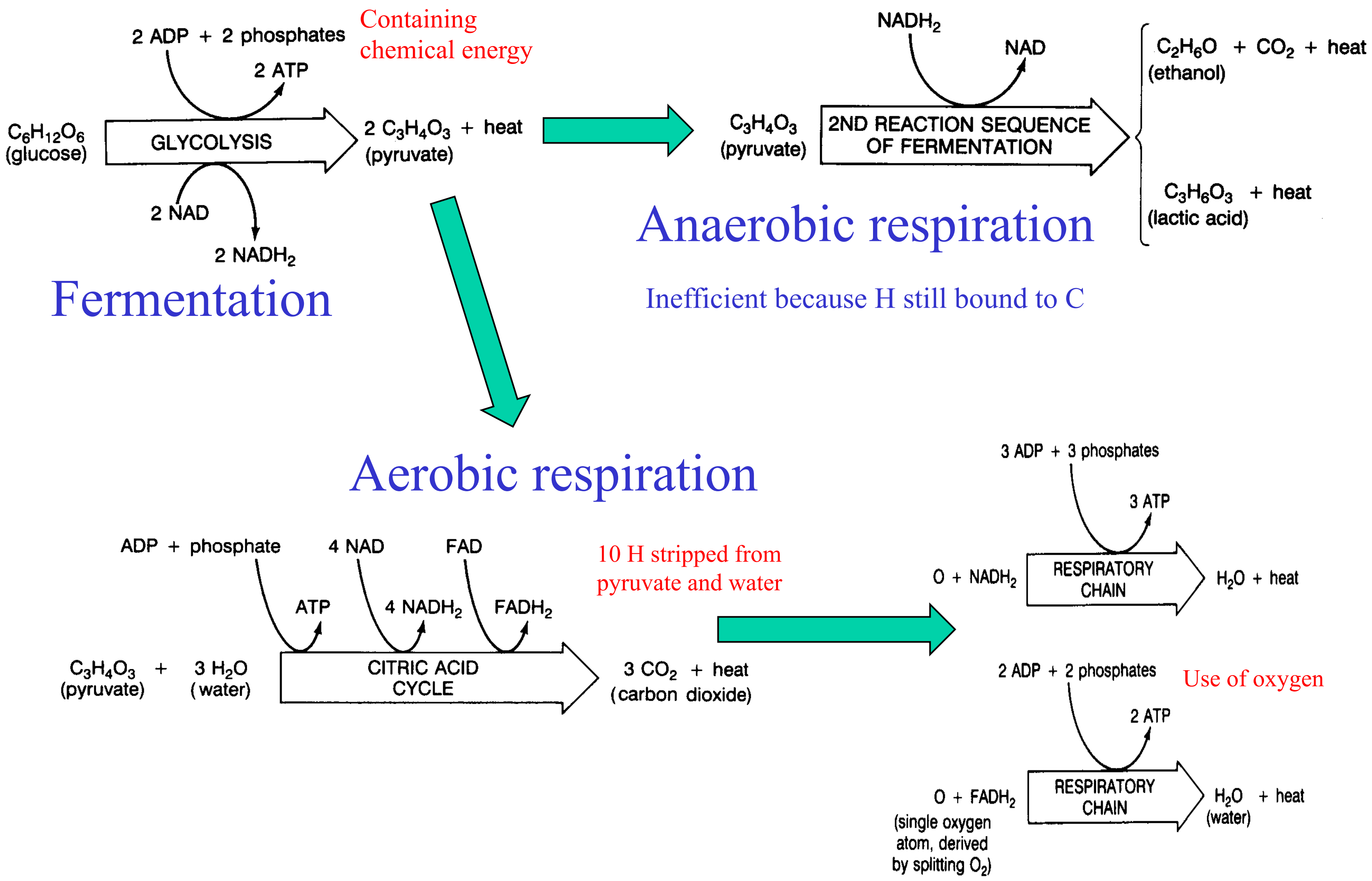
- Today, the production of polymers (e.g. proteins) from DNA (i.e. RNA) itself depends on catalysis by other proteins made from RNA. Chicken and egg?
- Some sequences of RNA can perform catalysis function → "RNA-World": how to make the RNA?
- Interest in clays as structures to promote polymerization
    - Regular structures with high surface area
    - Organic compounds attracted to surface
    - Water is absorbed
    - Located near standing water

- Possibility of earlier, simpler, but now disappeared, non-DNA "Life" as a kind of "scaffolding"?

3. Emergence of prokaryotic cells (~3.5 Gyr ago)

- Proto-cells as droplets in solution?
- Simplest cells: Prokaryotes (without nucleus): Archaea and Bacteria (differ on basis of RNA). Single long strand of DNA (several 1000 genes).
- Archaea likely developed first – still found as extremophiles (extreme conditions) and anaerobic systems. Depend on (small amounts of) local chemical energy.
- Source of free organic compounds gradually reduced as atmosphere became less reducing
- Some prokaryotes develop photosynthesis (earliest form may have used H_2S) – producing O_2 and sugars (e.g. glucose) out of sunlight (2-3 billion years ago). Enabled more complex Life using more efficient aerobic respiration. First solid fossil evidence.

Energy for cells: fermentation and respiration



2 ADP + 2 phosphates

2 ATP

Containing chemical energy

$C_6H_{12}O_6$ (glucose)

GLYCOLYSIS

$2\ C_3H_4O_3$ + heat (pyruvate)

2 NAD

2 $NADH_2$

Fermentation

$NADH_2$

NAD

$C_3H_4O_3$ (pyruvate)

2ND REACTION SEQUENCE OF FERMENTATION

C_2H_6O + CO_2 + heat (ethanol)

$C_3H_6O_3$ + heat (lactic acid)

Anaerobic respiration

Inefficient because H still bound to C

Aerobic respiration

ADP + phosphate

ATP

4 NAD

4 $NADH_2$

FAD

$FADH_2$

$C_3H_4O_3$ (pyruvate) + 3 H_2O (water)

CITRIC ACID CYCLE

3 CO_2 + heat (carbon dioxide)

10 H stripped from pyruvate and water

3 ADP + 3 phosphates

3 ATP

O + $NADH_2$

RESPIRATORY CHAIN

H_2O + heat

2 ADP + 2 phosphates

2 ATP

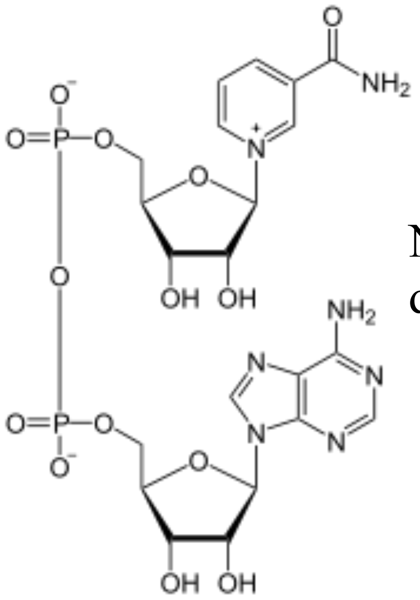
Use of oxygen

O + $FADH_2$ (single oxygen atom, derived by splitting O_2)

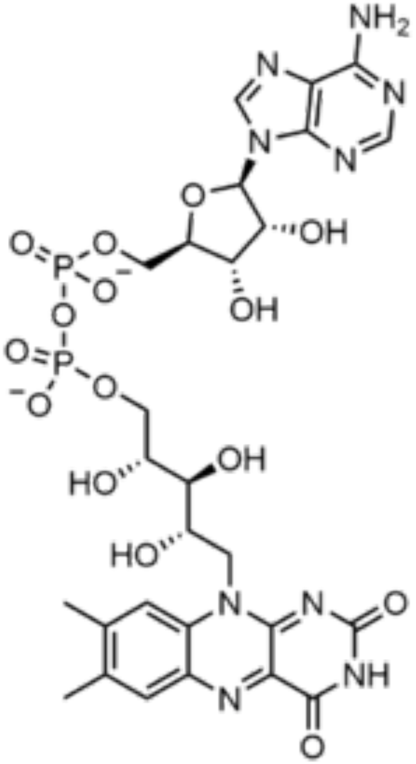
RESPIRATORY CHAIN

H_2O + heat (water)

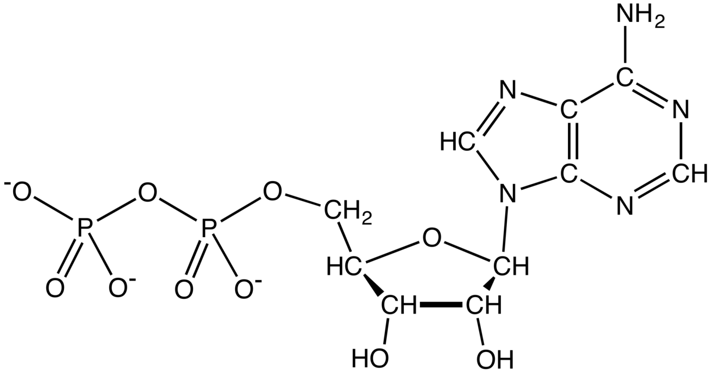
All very complicated...



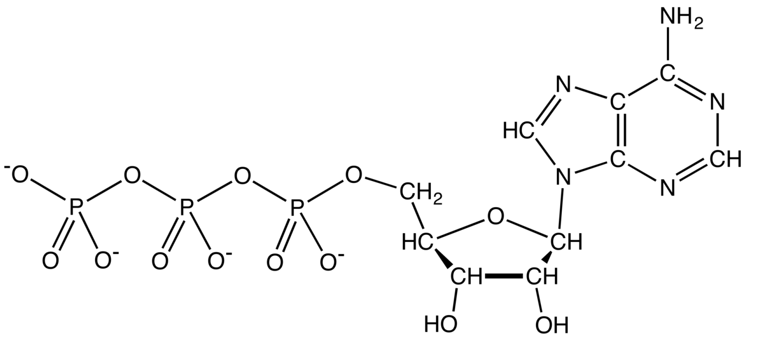
Nicotinamide adenine dinucleotide (NAD)



Flavin adenine dinucleotide (FAD)

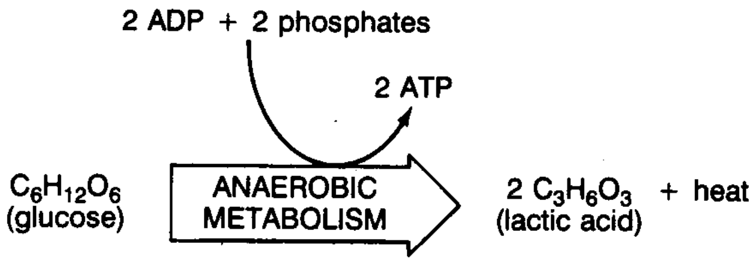


Adenosine diphosphate (ADP)

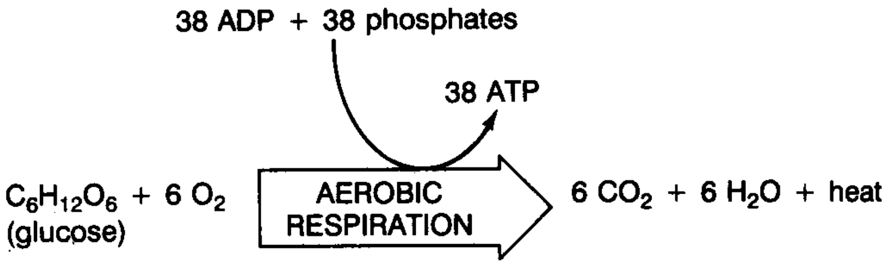


Adenosine triphosphate (ATP)

Net effect of “burning” glucose in controlled way



2% energy “efficiency”
(energy in *useful* form)



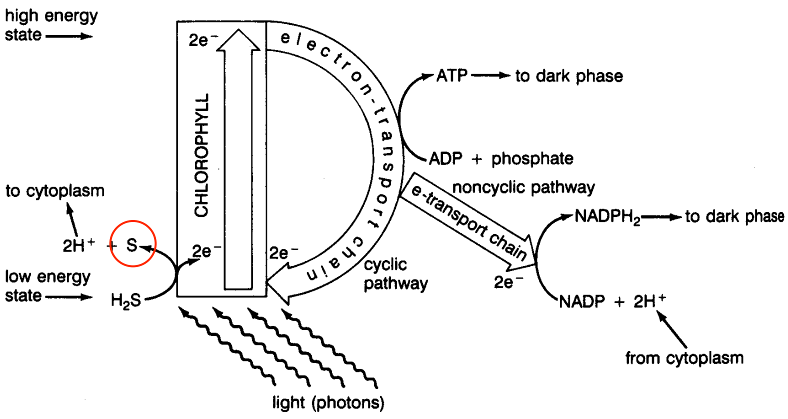
40% energy “efficiency”

(c.f. 0% in a wood fire – burning uncontrollably – and 30% in combustion engine)

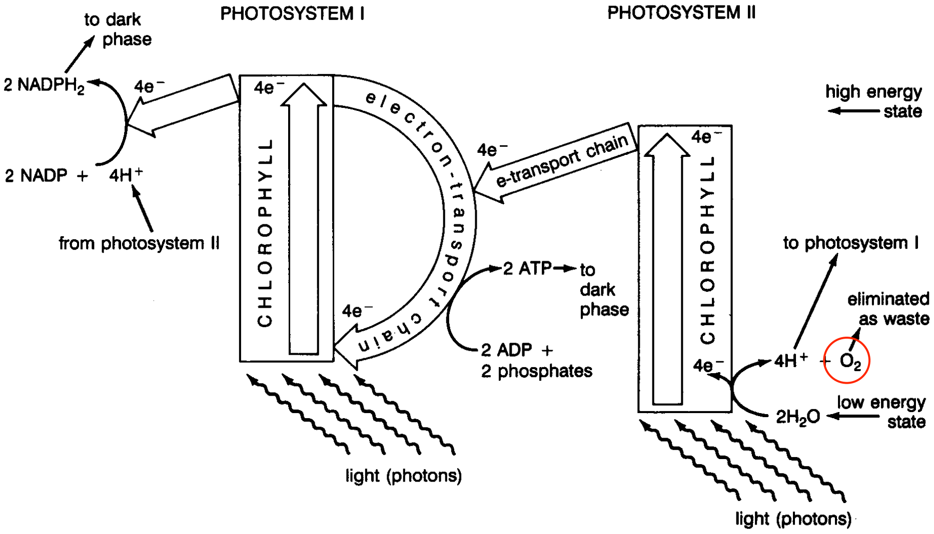
Supply of oxygen allows vastly more efficient extraction of energy

Photosynthesis: extraction of energy from sunlight to make glucose

Two forms: both produce glucose to be used in fermentation and/or respiration. The most efficient form also liberates O_2 as waste.



Anaerobic photosynthesis (light phase) liberating e.g. sulphur from H_2S



Aerobic photosynthesis (light phase) liberating oxygen from H_20.

Added new Photosystem II producing oxygen to Photosystem I from anaerobic.

Common dark phase: ATP, $NADPH_2$ synthesize glucose out of CO_2

High energy photons needed to split (abundant) water

Photosynthesis:

(some bacteria use S not O)

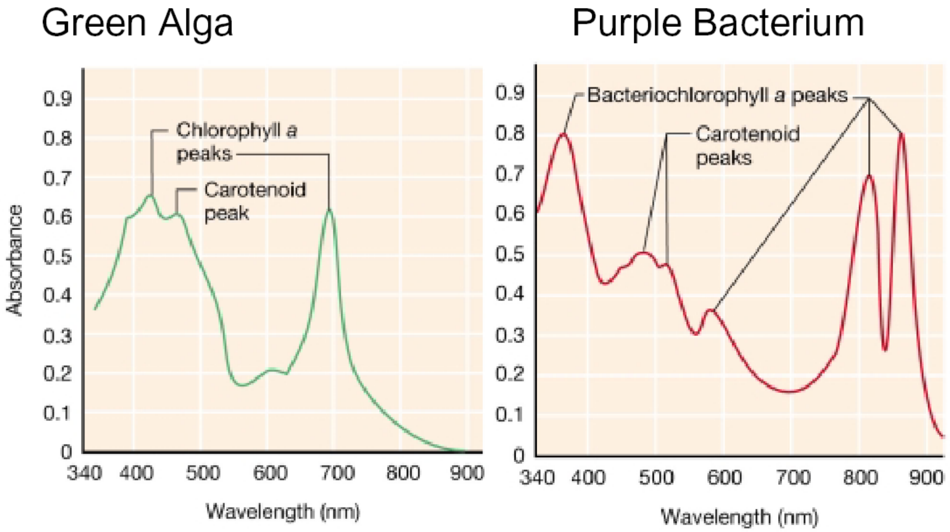
$$\boxed{6\ CO_2 + 6\ H_2 0 + \text{light} \rightarrow C_6H_{12}O_6 + 6\ O_2}$$

Extremely complex process in detail

Absorption spectrum of chlorophyll peaks

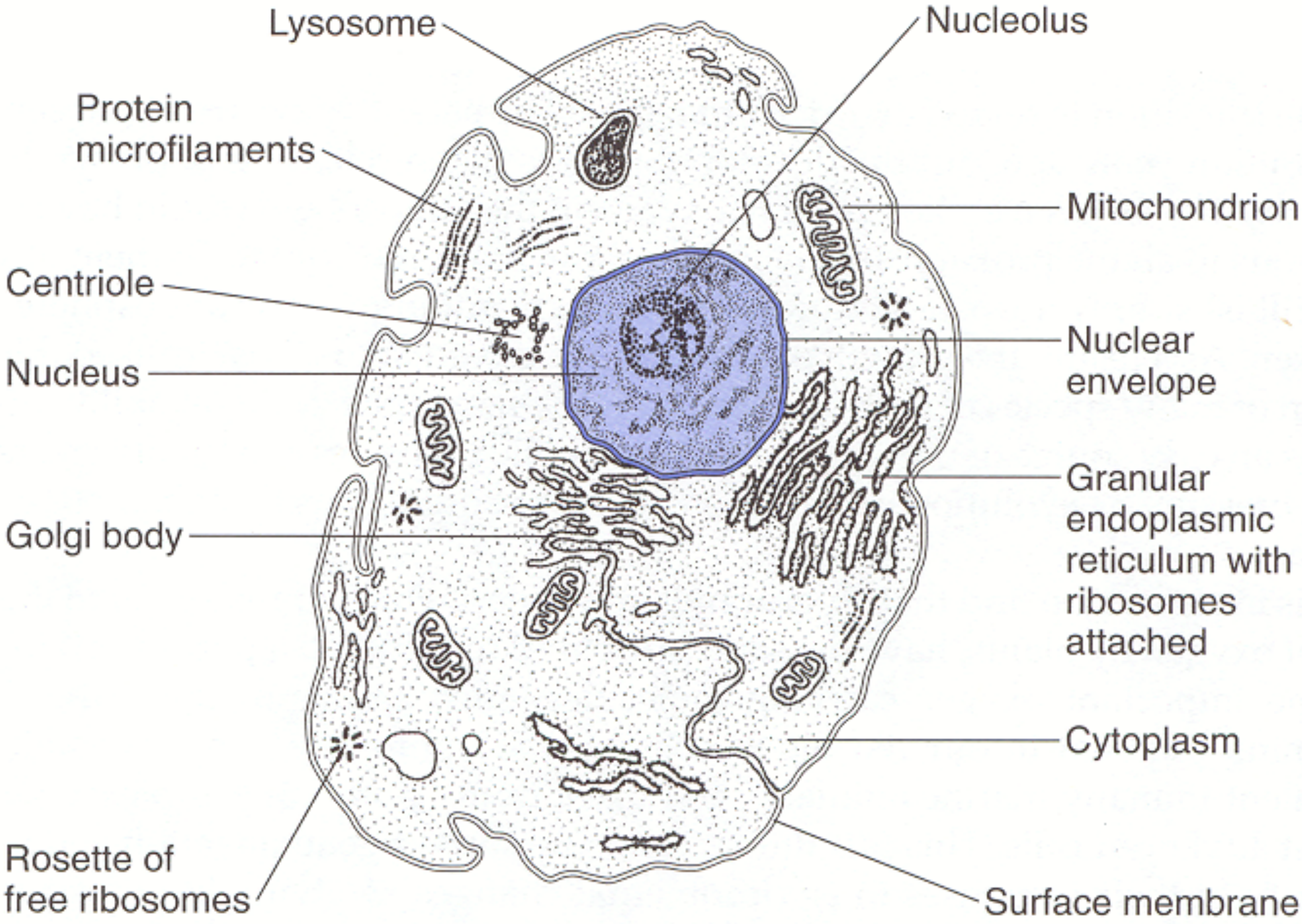
Oxygenic

Anoxygenic



Origin of Life on Earth (continued)

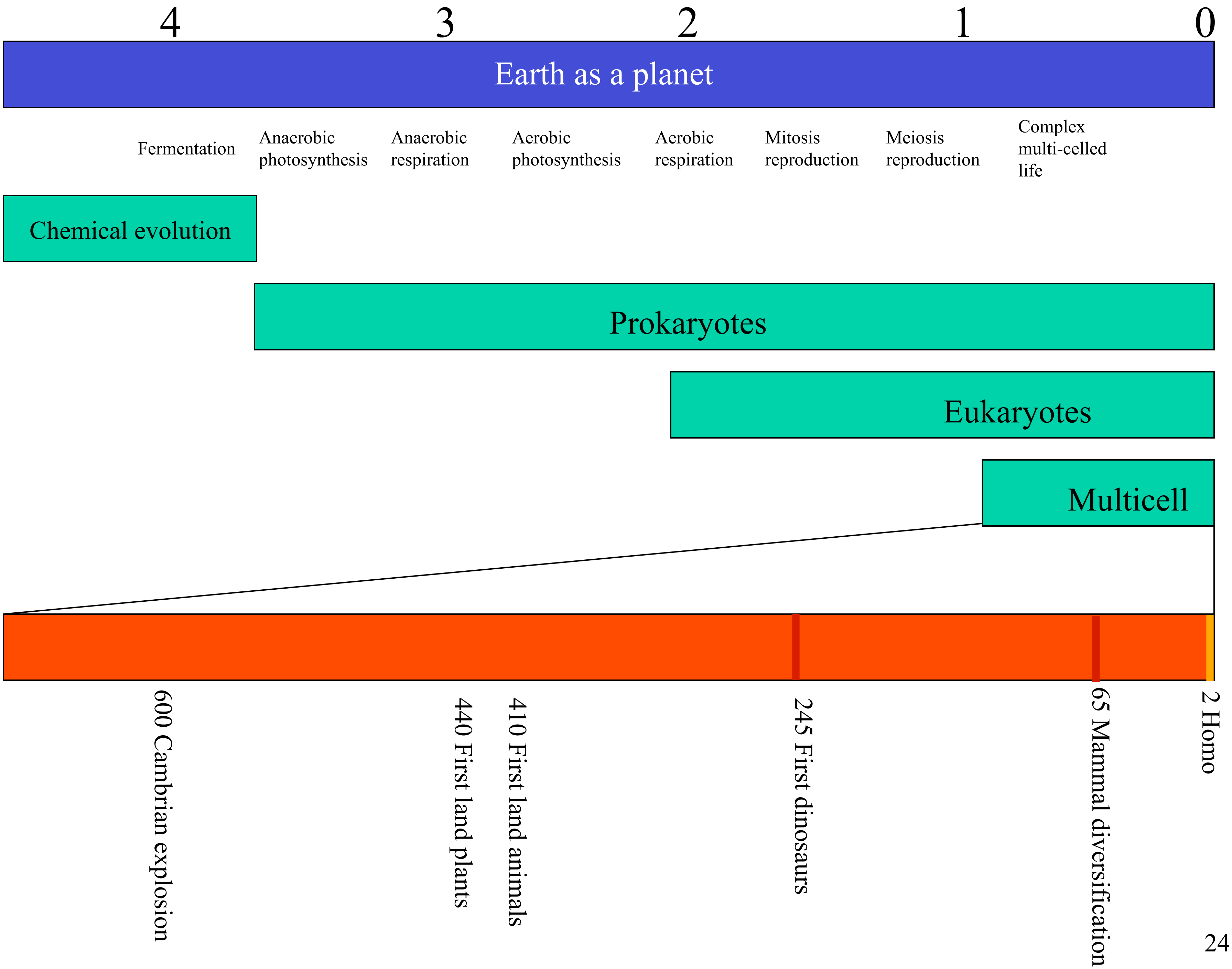
4. Emergence of Eukaryotic cells (~2.7 Gyr ago)
    - Cells with distinct cell nucleus, plus specialized organelles (mitochondria, chloroplasts etc.)
    - 10-1000 times more DNA than in bacteria
    - Incorporation of prokaryotes into proto-Eukaryote (similar gene systems to bacteria)?
    - Development of sexual reproduction (meiosis), producing vastly increased genetic sampling (1 Gyr ago?)
    - Development of multi-cellular organisms (multiple development)
        - → e.g. Jellyfish (700 Myr ago)


Lysosome
Nucleolus
Protein microfilaments
Mitochondrion
Centriole
Nuclear envelope
Nucleus
Granular endoplasmic reticulum with ribosomes attached
Golgi body
Cytoplasm
Rosette of free ribosomes
Surface membrane

Origin of Life on Earth (continued)

5. Emergence of advanced animals

- Cambrian explosion (600 Myr ago) → anatomically modern animals (two eyes, legs, shells/carapaces/skeletons, etc.)
- Invasion of land by plants and arthropods (440 Myr ago)
- Appearance on land of vertebrates, amphibians, insects, seed plants (410 Myr ago)
- Diversification of reptiles (285 Myr ago)
- Emergence of dinosaurs and reptiles (245 Myr) after a mass extinction
- Diversification of mammals (65 Myr) after another mass extinction
- Appearance of Homo (2 Myr)



4 3 2 1 0

Earth as a planet

Fermentation

Anaerobic photosynthesis

Anaerobic respiration

Aerobic photosynthesis

Aerobic respiration

Mitosis reproduction

Meiosis reproduction

Complex multi-celled life

Chemical evolution

Prokaryotes

Eukaryotes

Multicell

600 Cambrian explosion

440 First land plants

410 First land animals

245 First dinosaurs

65 Mammal diversification

2 Homo

What if anything can be inferred from the rapid emergence of the most basic life on Earth?

Life is common in the Universe? Not necessarily!

Interesting aside on conditional probabilities….

Suppose we have a sequence of "steps" in a process, each of which has a certain chance p_i of happening per unit time. Generally, we would expect to have to wait a time $t_i \sim p_i^{-1}$ for a given step to occur. Easy steps have $p_i > t_{avail}^{-1}$ and are likely to happen, difficult steps have $p_i < t_{avail}^{-1}$ and are unlikely to happen.

Clearly, if we have one or more difficult steps in the process, the chance that all the steps take place is small.

But, if we look only at those systems that were, by great good luck, able to complete all the steps, what is the distribution of the time spent waiting for each step?

Perhaps counter-intuitive answer:

- For easy steps, $<t_i> \sim p_i^{-1}$ (as expected)
- For difficult steps, $<t_i>$ is independent of the difficulty p_i, and is instead given just by the number N of difficult steps $<t_i> \sim (N+1)^{-1}$.

Example: Run a simulation with a time interval of 10^6 units, and eight "steps" with individual probabilities of occurring per unit time of 10^{-1}, 10^{-2}, 10^{-3} …. 10^{-8}.

The chance of them all happening in 10^6 time steps is very small, of order $10^{-2} \times 10^{-1} \times 1 \sim 10^{-3}$ (actually $\sim 3.3 \times 10^{-4}$). The distribution of $<t_i>$ is

"easy" steps, $<t_i> \sim p_i^{-1}$ | "difficult" steps, $<t_i> \sim N^{-1}$

Step p_i^{-1}	10	100	1000	10^4	10^5	10^6	10^7	10^8	left
$<t_i>$	9.8	95.8	989	9,810	68,385	207,920	257,860	254,910	220,015
$\sigma(t_i)$	9.3	99.5	837	11,236	60,212	164,590	179,680	189,212	161,118

Note that the dispersion in the individual step lengths $\sigma(t_i)$ is large, $\sigma(t_i) \sim <t_i>$

The explanation is actually straightforward:

If we know that all the hard steps completed, then the probability that we have a particular configuration of t_i is given by the product that each step completed, which is $p_i t_i$ (for $p_i t_i << 1$)

$$\boxed{P = \prod_i t_i p_i = \prod_i p_i \prod_i t_i}$$

The Πp_i is constant, so P is <u>maximized</u> when Πt_i is maximized subject to the constraint that $\Sigma t_i + t_{end} = t_{tot}$, where t_{end} is the time after the last step completed before the end of the experiment.

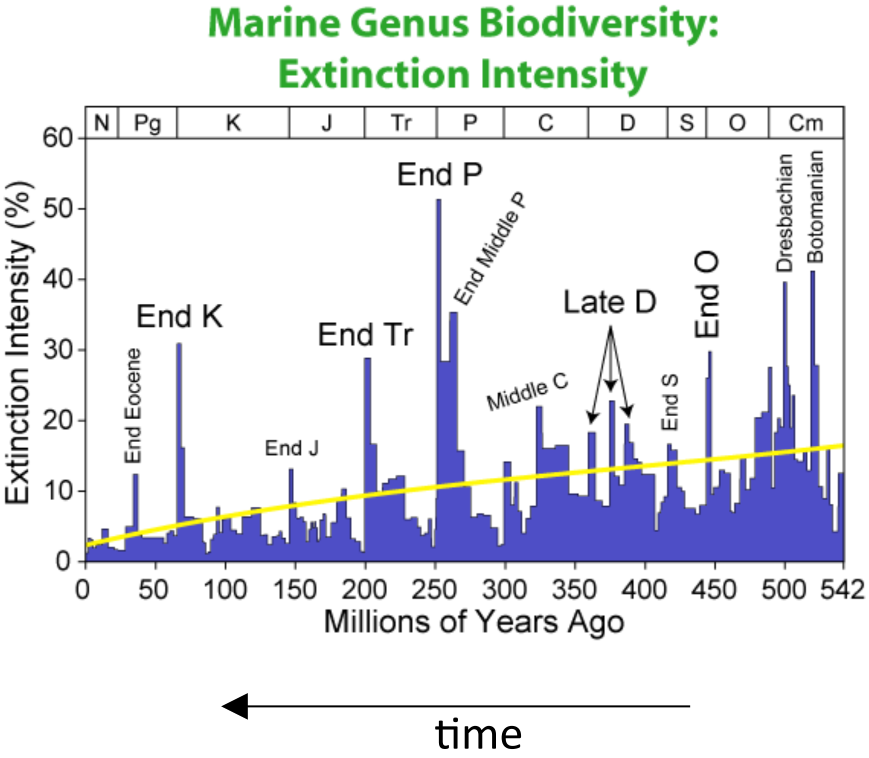
It is easy enough to show that P is maximized when all the t_i and t_{end} are all equal, i.e.

$$\boxed{t_i = \frac{t_{tot}}{(N+1)}}$$

We think on Earth that $t_1/t_{tot} \sim 0.1$. This may tell us nothing about p_1. It is just as consistent with all steps being very unlikely (p_i being very small) and $N \sim 10$.

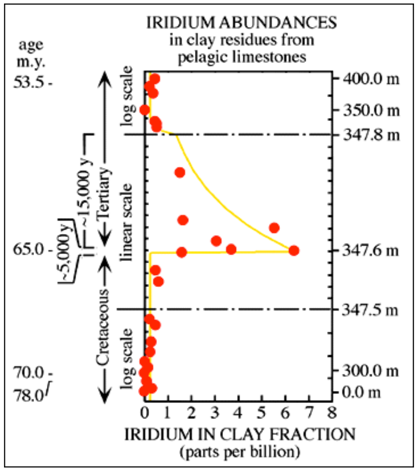
Mass extinctions

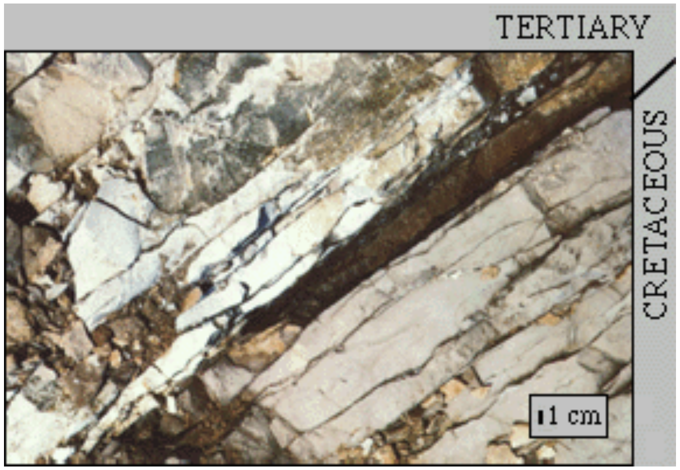
- The rate at which species (families etc) become extinct shows wide variations, leading to “sudden” loss of large numbers (e.g. Permian/ Triassic 50% of marine families = 95% of marine species, 70% of land species)
- Easiest to see in marine fossil record, but the effect is distributed across environments and across the planet (synchronous?)
- Almost impossible to estimate the duration if less than 10^4-10^5 years due to resolution of fossil record.


Marine Genus Biodiversity: Extinction Intensity

Causes are debated but there is good evidence that the K/T extinction 65 million years ago (dinosaur extinction) was associated with major impact (10-20 km asteroid). The KT boundary has:

- Iridium anomaly (Alvarez et al 1980) --> estimate of 10km asteroid
- Shocked quartz beads, especially near Caribbean
- Isotope anomalies in Iridium and Chromium that are similar to meteoritic material
- Major impact crater Chicxulub (Yucatan) accurately dated to KT event (200 km diameter)





Key Point: Large impacts do happen

There will be a broad spectrum of impacts: size (energy) vs. frequency

e.g. 2013 Chelyabinsk meteorite (~20m diameter, 500kT energy) was likely the largest impact since Tunguska in 1908 (~100m, 3-5 MT energy).

Resulting impact craters are produced by release of gas (vapourized rock) that is produced by the impactor tunneling through rock

Effects are expected to be *global* for impacts of objects with diameters above about 1 km. These occur of order once every million years.

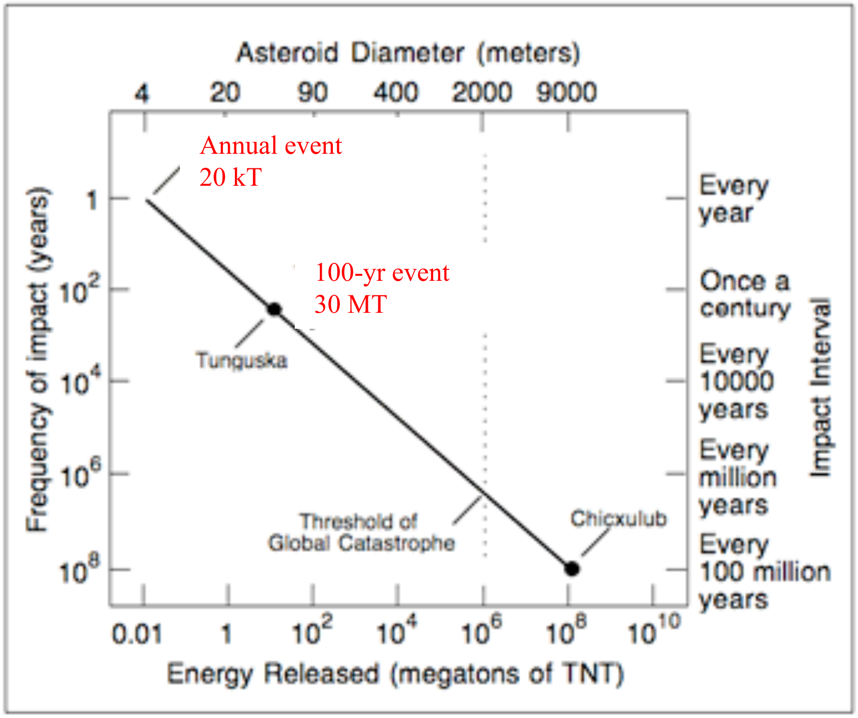
Expected effects:

- Dust in atmosphere blocks out sunlight
- Tsunamis
- Blasts of hot gas
- Fall-out of hot rock

Note: 10km object is equivalent material to 2mm deposited over entire surface of Earth

$$\frac{\Delta t}{yr} \sim \left(\frac{E}{0.01\,\mathrm{MT}}\right)^{0.8}$$

Interesting question: what is the worst impact witnessed by "historical humans" in 10,000 years (1000 MT) and by Homo in 1 Myr ($10^{5.5}$ MT)



How strange could life be?

Terrestrial Life is based on organic molecules (C-chains) and water as a solvent.

- Carbon's ability to form long hydrogenated chains: e.g. oxygen (H_2O, H_2O_2), nitrogen (NH_3, N_2H_2) c.f. carbon up to ($C_{90}H_{84}$)

Could well imagine different details (e.g. other handedness, use of other amino acids, etc.).

But what about other elements and/or solvents?

- e.g. Silicon-based instead of Carbon-based?
- e.g. NH_3 or CH_3OH instead of water as solvent?

Silicon-based Life (next element below C in Periodic Table)?

- Si-Si bond has half the strength of C-C
- Si-H and Si-O are stronger than Si-Si (c.f. C-O ~ C-H ~ C-C)
- Si_mH_n is unstable for $m > 3$
- SiO polymers exist (silicones) but are essentially inert
- Si has <u>extreme affinity</u> for O: SiH_4 only exists for T > 1000 K
  - e.g. Jupiter: NH_3, PH_3, AsH_3, CH_4, GeH_4, H_2S, H_20 in atmosphere. No SiH_4 (all in the form of SiO_2)
- Si found with other elements as Silicates (e.g. rock on Earth!)

CO_2	SiO_2	CH_4	SiH_4
Gas at > 200K	Gas at > 2300K	Exists in oxidising atmosphere (e.g. 10 yr on Earth)	Bursts spontaneously into flame
Soluble in H_2O	Insoluble (except HFl)		
→ C + O in uv	Stable (quartz)		

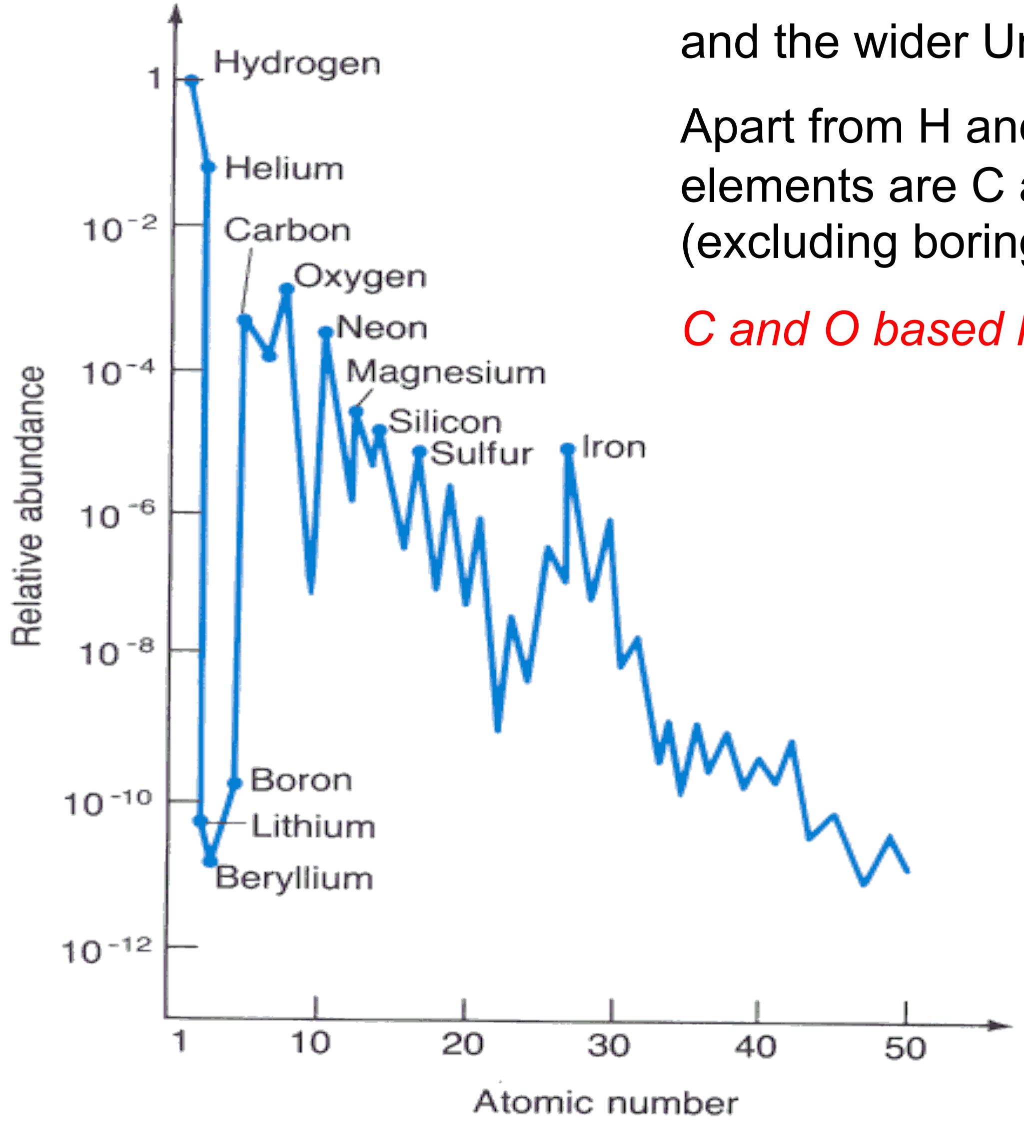
Alternative solvents?

Ideal solvent:

- Solvent should be liquid (better than gas for concentration reasons) over a range of temperature
  - Temperature high enough for chemical reactions to occur
  - Temperature low enough for survival of complex molecules
  - Range of T because of likely changes (e.g. planets)
- Should have high heat capacities for active temperature control

	H_2O	NH_3	CH_3OH
Liquid	0-100 C	-78 – -33 C	-94 – +65 C
SHC	1	1.2	0.6
LHV	600	300	290
Surface tension	1	0.5	0.33
“Solvent”	1	0.5	0.5

H_20 is best solvent, but NH_3 and CH_3OH are not bad



Relative abundances (by weight) of elements in the Sun (typical of the Galaxy and the wider Universe):

Apart from H and He, most abundant elements are C and O with N next (excluding boring Ne).

C and O based life is natural and easy!

Key points

- Hierarchy of structure and function
- Complexity gives extreme efficiencies for energy conversion
- Organic molecules were naturally produced on and around young Earth
- Multiple steps in the emergence of advanced Life on Earth
- Little can be said from the early emergence of Life on Earth
- Interaction of Life with the extraterrestrial environment via impacts etc.
- C-life is extraordinarily much easier than Si-life.
- Other solvents may be possible, but why not H_2O?