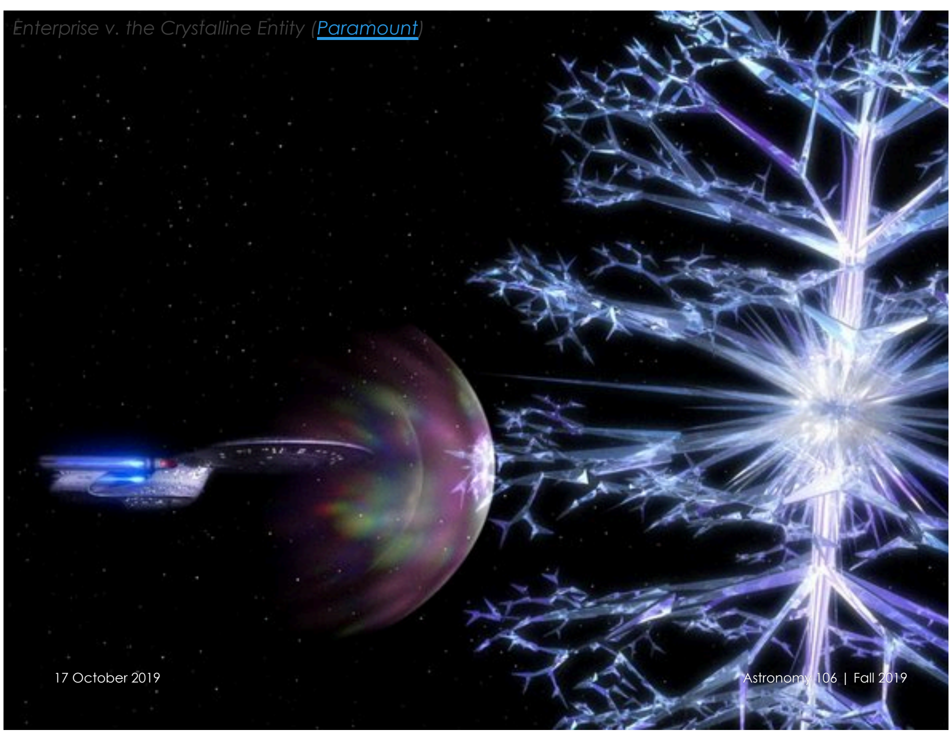
17 October 2019

THE LONG MOLECULES OF LIFE, FROM POLYMERS TO LIFE, & LIFE NOT AS WE KNOW IT

Homework #4 on WeBWorK due Monday at 7pm
TI surveys!



Enterprise v. the Crystalline Entity (Paramount)

The long molecules of life, from polymers to life, & life not as we know it

Replication or mass production of nucleic acids

Transcription

Codons

Translation: the current fashion in protein manufacture

The chicken-egg problem

- Protein-based primitive life?
- RNA world

Emergence of the genetic code

How long does all this take? The importance of baby steps and the heroism of time.

Attempt to shed the Earth bias and consider:

- Clays and crystals as catalysts for the first bio-macro-molecules
- Alternatives to carbon
- Alternatives to water
- Alternatives to atoms

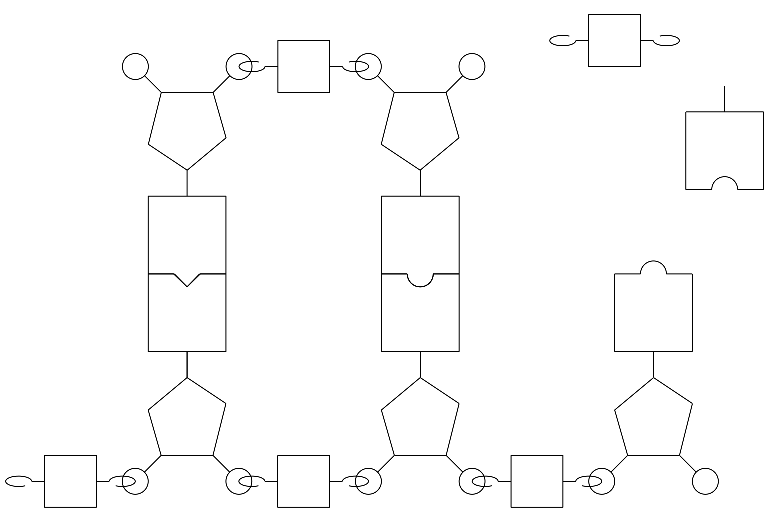
Replication of nucleic acid

Requirement (#3) of life is replication. Can proteins and nucleic acids replicate, or at least mass-produce?

- If they can, this could explain the emergence of classes of proteins and nucleic acids as components, the assembly of which could serve more complex, life-like roles.

Perhaps the easiest to envision is for short strands of RNA. Suppose a short RNA lies in a solution containing nucleobases, ribose, and phosphoric acid.

- A **complementary polymer** can form by hydrogen-bonding nucleotides onto the RNA's nucleobases, hooking up the phosphates and sugars, and then severing the hydrogen bonds.
- The complementary polymer can do the same, thus replicating the original.
- Hydrogen bonds are much weaker than covalent bonds. The "copies" can be stripped off without harming the original chain.
- Other polymers are capable of "unzipping" the H bonds in DNA. Once unzipped, both sides can bond to nucleotides in the solution just like RNA, and thus the DNA replicates: bases in same order as original.



Replication of RNA

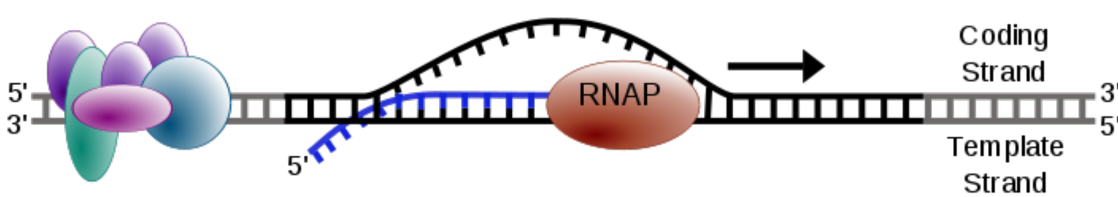
Transcription

Long chains of DNA can be partially unzipped, capture a complementary chain of nucleotides, have this new chain of RNA zip off, and have the DNA zip back up. This form of partial replication is called **transcription**.

Unzipping does not frequently happen spontaneously, except when the solution temperature experiences small increases.

- Certain polymers with the right molecules in the right spacing can do it, though.
- Among Earth life forms, the unzipping and rezipping of DNA is done with a protein we call **RNA polymerase**.
- This is an example of an **enzyme**: proteins that catalyze chemical reactions. This means that they do not get chemically changed in the process, like the dust grains that catalyze the formation of molecular hydrogen.
- Replications of DNA and RNA, and transcriptions, require – in the current age – a special class of proteins.

Animation of transcription



Replication of proteins

(Recall that a protein is a polymer of amino acids. The order of the amino acids primarily determines the type of protein.)

This has been envisaged in several different ways…

Long chains of amino acids can polymerize in solution without much help.

- Some configurations might be **self-organizing**: certain sequences have higher binding energies than others and will preferentially form.
- There will be lots of copies of these for this reason.

Alternative mass production or replication: Perhaps formation on a sequenced structure, like a crystal or another molecule, that preferentially attracts the monomers in a certain order?

Codons – The genetic code

Three-nucleotide sequences have a special significance, and thus their own name: **codon**.

Because there are four different nucleobases used in either nucleic acid, there are

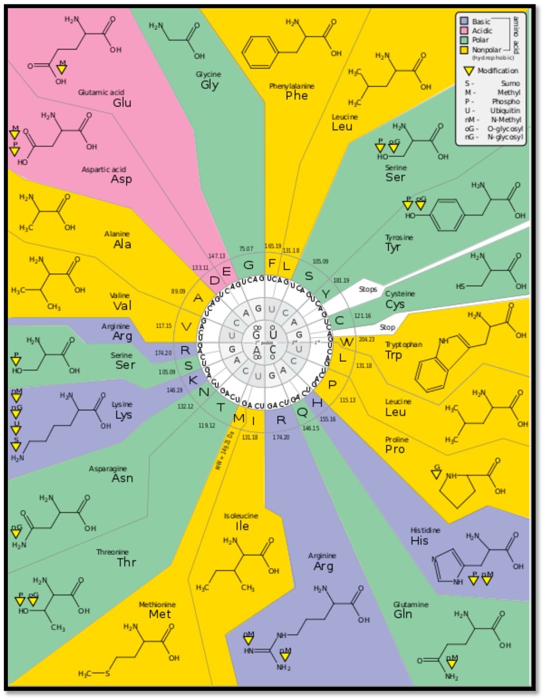
$$4 \times 4 \times 4 = 64$$

different codons.

These codons are used to encode the information used by modern organisms to build proteins.

- For example, there are 20 amino acids* used in human proteins, and the base sequences that indicate the beginnings and ends of codon sequences. 20 + 2 = 22 < 64, so three-base codons work.
- *Note: There are really 21 amino acids used in human proteins, but the 21st (selenocysteine) is relatively rare in proteins and uses a special signal in addition to the Stop codon.

Codons



Question

Instead of 20, suppose Earthly organisms used only 10 different amino acids. How many nucleotides could be used in a codon?

A. 1
B. 2
C. 3
D. 4
E. Any number greater than 1

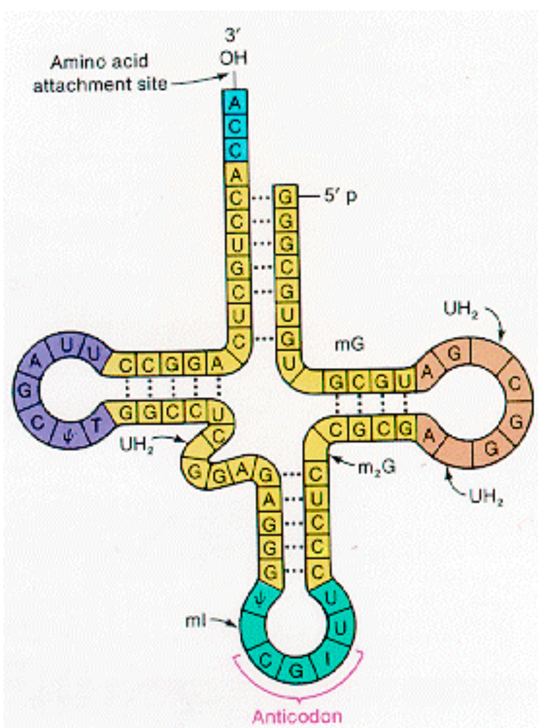
Translation

In modern lifeforms, this is quite complicated and intimately involves nucleic acids following a transcription.

Section of DNA or RNA transcribed onto a short length of RNA, called **messenger RNA (mRNA)**.

mRNA is transported to a **ribosome**, a collection of proteins and nucleic acids that will hydrogen-bond onto a certain codon of the mRNA.

The ribosome has shorter stretches of RNA around, called **transfer RNA (tRNA)**, that have exposed codons and, bonded elsewhere, an amino acid captured from the solution that is unique to the exposed codon.



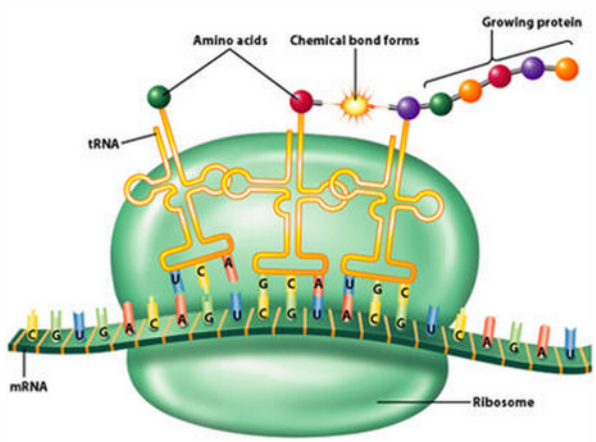
tRNA

Protein replication

The tRNA complementary to the first codon on the mRNA H-bonds to it, and the AA is the first element of the protein. The next element is brought in with the next tRNA, and so on.

The ribosome's (and tRNA's) job is called **translation**.

https://youtu.be/2BwWavExcFI?t=2m48s



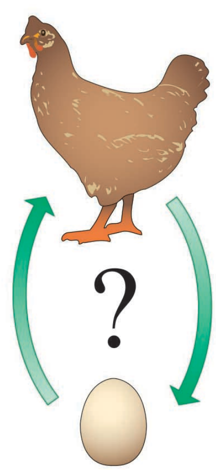
The original chicken and egg problem

In modern Earthly life, the replication or mass production of nucleic acids and protein are interdependent: each needs certain of the others in order to be produced in large quantities.

Which came first? Was the most primitive life originally just a protein or nucleic acid?

How did they come to be interdependent?

How did useful proteins and nucleic acids develop?



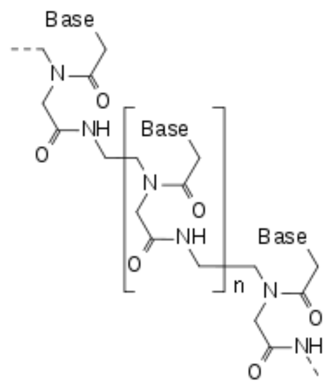
Protein-based primitive life?

There is no experimental evidence that protein-only life could exist, but there are theories and candidate substances.

Closest to modern life: peptide nucleic acids, which are proteins but which have nucleobases that allow replication and information storage à la RNA.

So far, peptide nucleic acids are a lab curiosity; there is no evidence in substructure of cells or chemistry of primitive monomers that it is a viable life form.

- They can hydrogen-bind to DNA, though.



Peptide nucleic acid with a N-(2-aminoethyl) glycine backbone

Protein-based primitive life?

Prions do not count: They reproduce by folding existing proteins into weird shapes, and in general seem to be things that developed after life was already on its way with both nucleic acids and proteins.

Before the genetic code was traced to DNA, Alexander Oparin (1924 – the originator of the idea behind the Miller-Urey experiment) noted the tendency of amino acids to combine and fatty acids (lipids) to form in solution, and postulated that this would lead to spherical **coacervate** droplets of these compounds in the "primordial pond."

- He thought that this would lead to cells: membranes made from lipids, most of the rest proteins.

Similarly, Fox later proposed **proteinoid** droplets (protein-like molecules which form abiotically from amino acids).

Protein-based primitive life?

Freeman Dyson (1982, 1999), a famous physicist turned evolutionary biologist, picked up on Oparin's scheme and conjured a model in which metabolism and proteins were primary, with nucleic acids and the genetic code developing much later.

- "Cells" first, enzymes next, nucleic acids much later.
- Cells provide sites for absorption of monomers.
  - He had acervates/proteinoids and amino acids in mind, but noted that the reasoning would work on crystals and nucleotides as well.
- Repeatable assembly of complex polymers through "order-disorder" transition of absorbed molecules.
- Dyson's model predicts that polymers with 100-200 monomers can be made repeatedly.
  - About the minimum required for a reproducing system, according to experiments on viruses (and thus RNA, not proteins).
- Works best if there are 8-10 different monomers; does not work at all if there are 3-4, so the mechanism prefers amino acids to nucleotides.
- Also achieves an error rate of about 1 in 4 if the total number of monomers present is of order 10,000 (fewer errors if more monomers).
- Hard to extend to nucleic acids, though, so this model has not found much favor.

Nucleic-acid based primitive life?

RNA can

- Replicate, as we have seen (transcription)
- Catalyze the formation of proteins (tRNA + RNA) and nucleic acids.
- Store information: the genetic code, so that specific proteins can be mass-produced.
- Transmit information: mRNA copies of short sections of a master RNA (translation).

This is the most consistent, and therefore popular, story of emergence of biology from chemistry: an RNA-dominated primitive form of life that preceded proteins and DNA.

This stage of evolution is generally called the **RNA world**.

RNA world

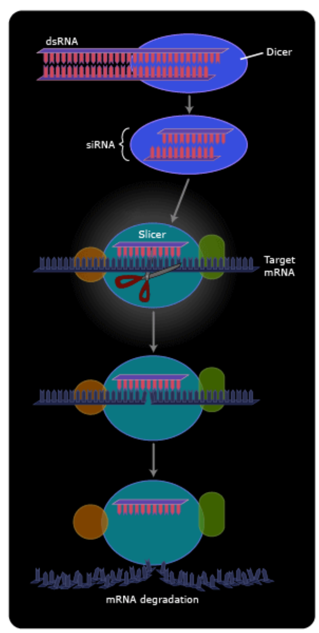
Scenario, due mainly to Eigen, Orgel, and Cech:

- In the beginning, there were short strands of bare, self-replicating RNA (Eigen et al. 1981). Eigen calls the varieties of the strands **quasispecies**.
- Most numerous and robust quasispecies predominate (appearance of natural selection).
- Quasispecies mutate and therefore diversify as time goes on.
- Truncated stretches of these RNAs catalyze reactions of other RNAs (Cech & Bass 1986).
  - Example: **192-base RNA that catalyzes the replication of an RNA** as long as itself with 98.9% accuracy.
  - At this point, the D-sugars have been "chosen."
- These RNA catalysts, called **ribozymes**, presage the function of modern ribosomes. (1989 Nobel Prize to Cech and Altman for discovering ribozymes.)
- Soon, the ribozymes begin to catalyze protein formation. L-amino acids "chosen" at this point.
- In turn, some of the proteins thus formed begin to influence RNA cleaving and hydrolysis; these are the first **enzymes**.
- This interchange of RNA and protein catalysts is termed a **hypercycle** by Eigen.

RNA world

Simulations and theories show that proteins made from other quasispecies would be likely to perform novel catalytic functions on a given quasispecies.

RNA can hydrogen-bond to complementary DNA. Eventually, DNA took over as the more robust and stable master copy of the base sequences that encode the formation of proteins, but RNA retained its role in transcription and translation.



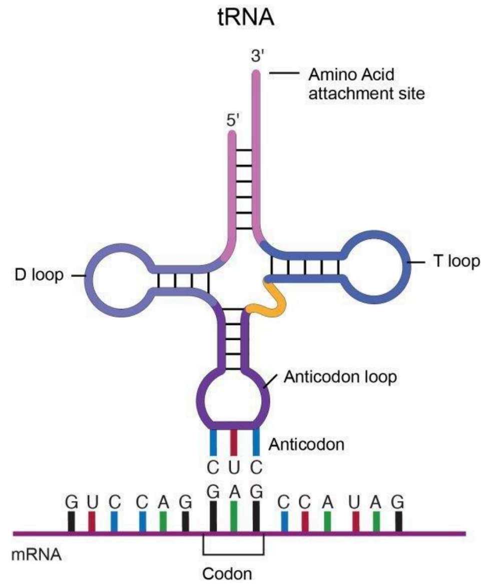
Ribozymes in action on RNA (Cech 2004)

The genetic code

One of the most striking adaptations of ribozymes is the development of **tRNA**: rather long molecules, each with a three-base sequence on one end which uniquely matches to an amino acid on the other.

As we have mentioned, life on Earth has 20* amino acids in use, which could not be specified by two bases (only 16 possible combinations) but can be with three (64 possible combinations).

This code must have been established very early in the development of life, as even cell-organelle DNA and RNA differ very little from nuclear DNA in the code they use.



		Nucleotide 2			
		U	C	A	G
Nucleotide 1 (5')	U	UUU Phenylalanine	UCU Serine	UAU Tyrosine	UGU Cysteine
		UUC Phenylalanine	UCC Serine	UAC Tyrosine	UGC Cysteine
		UUA Leucine	UCA Serine	UAA Ochre (Stop)	UGA Opal (Stop)
		UUG Leucine	UCG Serine	UAG Amber (Stop)	UGG Tryptophan
	C	CUU Leucine	CCU Proline	CAU Histidine	CGU Arginine
		CUC Leucine	CCC Proline	CAC Histidine	CGC Arginine
		CUA Leucine	CCA Proline	CAA Glutamine	CGA Arginine
		CUG Leucine	CCG Proline	CAG Glutamine	CGG Arginine
	A	AUU Isoleucine	ACU Threonine	AAU Asparagine	AGU Serine
		AUC Isoleucine	ACC Threonine	AAC Asparagine	AGC Serine
		AUA Isoleucine	ACA Threonine	AAA Lysine	AGA Arginine
		AUG Methionine	ACG Threonine	AAG Lysine	AGG Arginine
	G	GUU Valine	GCU Alanine	GAU Aspartic acid	GGU Glycine
		GUC Valine	GCC Alanine	GAC Aspartic acid	GGC Glycine
		GUA Valine	GCA Alanine	GAA Glutamic acid	GGA Glycine
		GUG Valine	GCG Alanine	GAG Glutamic acid	GGG Glycine

The genetic code

Amino acid	Codons
Alanine	GCU, GCC, GCA, GCG
Arginine	CGU, CGC, CGA, CGG, AGA, AGG
Asparagine	AAU, AAC
Aspartic acid	GAU, GAC
Cysteine	UGU, UGC
Glutamine	CAA, GAG
Glutamic acid	GAA, GAG
Glycine	GGU, GGC, GGA, GGG
Histidine	CAU, CAC
Isoleucine	AUU, AUC, AUA
Start	AUG
Leucine	UUA, UUG, CUU, CUC, CUA, CUG
Lysine	AAA, AAG
Methionine	AUG
Phenylalanine	UUU, UUC
Proline	CCU, CCC, CCA, CCG
Serine	UCU, UCC, UCA, UCG, AGU, AGC
Threonine	ACU, ACC, ACA, ACG
Tryptophan	UGG
Tyrosine	UAU, UAC
Valine	GUU, GUC, GUA, GUG
Stop	UAA, UGA, UAG

Also used to code seleno-cysteine

The genetic code

Baby steps

Every step in the development of life has involved the development, not just of molecules of a certain size and class, but of molecules that *do* something specific. That is a big restriction.

For example, take the minimum-length RNA-replication ribozyme mentioned earlier: 192 nucleotides.

There are four different nucleotides, so the number of combinations of 192 nucleotides is

$$n = \underbrace{4 \times 4 \times 4 \times K \times 4}_{\text{192 factors of 4}} = 4^{192} \cong 4{\times}10^{115}$$

If they form randomly, we may have to try $4{\times}10^{115}$ times for a good chance of getting the correct one. (Why? Well…)

Question! Quick education in probability…

Flip a coin. What is the probability of it coming up heads?

A. ½
B. 1/20
C. 1
D. 2/3
E. Cannot tell until you flip the coin.

Question!

What is the probability of the coin turning up heads in ten consecutive tosses?

A. $\frac{1}{2} = 0.5$
B. $\frac{1}{10} = 0.1$
C. $\left(\frac{1}{2}\right)^{10} = 9.8\times10^{-4}$
D. $\left(\frac{1}{2}\right)^{\frac{1}{10}} = 0.933$
E. $\left(\frac{1}{10}\right)^{2} = 0.01$
F. $\left(\frac{1}{10}\right)^{\frac{1}{2}} = 0.316$

So how many trials of ten coin tosses would you think you would have to make to have a good chance of ten consecutive "heads" at some point?

Question!

A. 2
B. 10
C. 100
D. $2^{10} = 1024$
E. $2^{100} = 1.3\times10^{30}$

In total, how many coin tosses would that be?

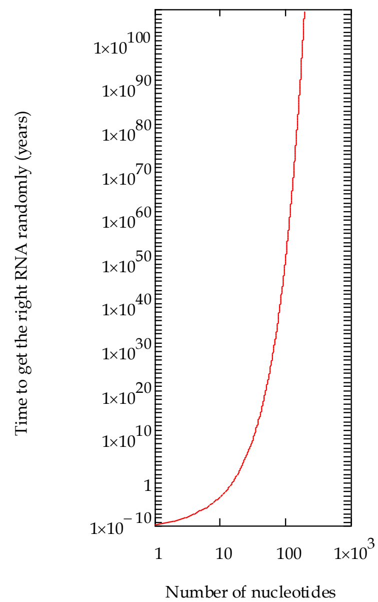
Question!

A. 20
B. 40
C. 10240
D. $1.3x10^3$
E. Cannot estimate from the given information.

Baby steps

If that many attempts were needed to have a reasonable chance of getting the correct molecule, **making them out of single nucleotides one by one**, it would take *much* longer than the age of the Universe.

But it takes much less time to make shorter RNAs. And that, plus natural selection, is the key: **make short RNAs first, then make longer RNAs out of these**.



Baby steps

For example, consider a two-step process.

Start with water at $T = 300$ K that has the Earth-crust abundance of carbon, 10% of it in nucleotides.

- Then it takes about 0.04 seconds to form an RNA with 24 nucleotides and 0.32 seconds to form one with 192, adding nucleotides one by one.
- To have a good chance to get a *specific* RNA with 24 nucleotides by this means, it takes $4^{24}\times0.04$ s $= 3.6\times10^5$ years, while it takes $4^{192}\times0.32$ s $= 4\times10^{107}$ years to get a specific 192-long RNA.

Clearly, 24-nucleotide RNAs are going to form the slow way much faster than the 192s; also, at least some of them can **self-replicate**, as we have already discussed.

How many different four-nucleotide-long RNA molecules are there?

Question

A. 4
B. 8
C. 16
D. 32
E. 64
F. 128
G. 256

About how many times would you have to construct 4-base RNAs at random to have a good chance of getting a specific one?

Question

A. 4
B. 8
C. 16
D. 32
E. 64
F. 128
G. 256

In the worst case, you would have to do this serially (one RNA at a time). If it takes 0.05 s to synthesize a 4-base RNA, how long does it take to have a good chance of synthesizing a specific one?

Question

Baby steps

Suppose that once certain 24-long RNAs formed, **self-replication led quickly ($\ll 10^5$ years) to the incorporation of all the nucleotides around into chains of 24.**

- It gives about 1.6 s to make a chain of 192 out of chains of 24, given the above conditions.

Furthermore, let us assume that there are 24 different ones that have risen to prominence by replication.

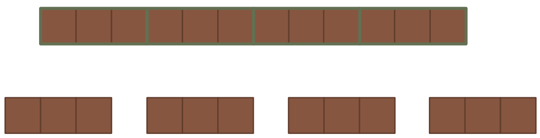
Then it will take $24^{\left(\frac{192}{24}\right)}\times 1.6\ \mathrm{s} = 5000$ years to have a good chance of making a *specific* 192-long RNA like our "RNA replicase" ribozyme. The total time is therefore less than a million years, instead of greater than 10^{107} years.

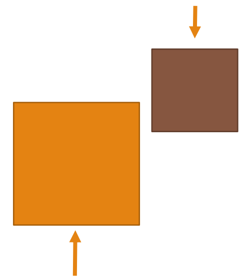
- And thus much less than the age of the Solar System.
- Goes even faster with more, shorter steps, as we will see.

Probability math simplified

Two pieces of information to figure out:

- Length of molecule (number of blanks)
- How many options you have for each blank





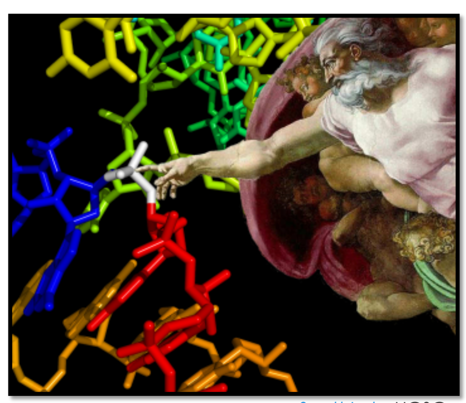
Baby steps

It would be nice to start with primordial soup and see the evolution of ribozymes produced in the lab in this manner, but unfortunately there are few graduate students willing to wait 5000 years to get their PhDs.

- Existing ribozymes have been "intelligently designed" by the likes of Cech and Altman, rather than evolved from a mixture of prebiotic chemicals.

Meanwhile, there are **computer simulations** based on lab data:

- Ma et al. (2007, 2010a): starting with a prebiotic mixture and allowing the basic sequences to be 5-10 nucleotides long (instead of the 24 we used a minute ago), "RNA replicase" is generated in only 52 years.



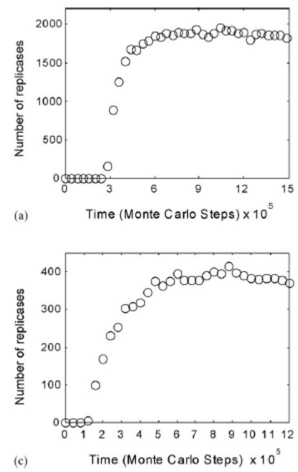
Scott Lab, UCSC

Baby steps

Also in the simulations: once the first kind of RNA replicase is created, more efficient, *mutated* replicases begin to arise.

In similar computer simulations, Ma et al. (2010b) have also shown that ribozymes provided with a protective cell-like membrane soon (within only a few years!) produce ribozymes which can replicate membrane-component molecules.

- Emergence of *cell* self-replication?



Computer simulation of RNA-replicase evolution, using 10-nucleotide (upper) and 6-nucleotide (lower) basic sequences. A "Monte Carlo step" is approximately 0.38 hours (Ma et al. 2007).

Time is still the hero

A sophisticated biochemical function like the "RNA replicase" ribozyme takes less than one million years to produce from scratch, *via* chemistry and biological **natural selection**.

Of course, there are lots of functions to evolve, but all the functions of the molecules of life did not need to develop serially.

Thousands of millions of years (i.e. Gyr), very roughly, seem reasonable for making many functions of this level of sophistication. Complex simulations bear this out.

Still a long time, but it is much shorter than the time available and is consistent with the fossil record of primitive life on Earth.

If length-12 RNA nucleotides are created, and during this time, 10 length-12 quasispecies develop the ability to self-replicate, how many length-144 RNA nucleotides would we need to construct from length-12 RNA nucleotides so that we have a good chance of creating one with the desired arrangement?

Review question!

Panspermia

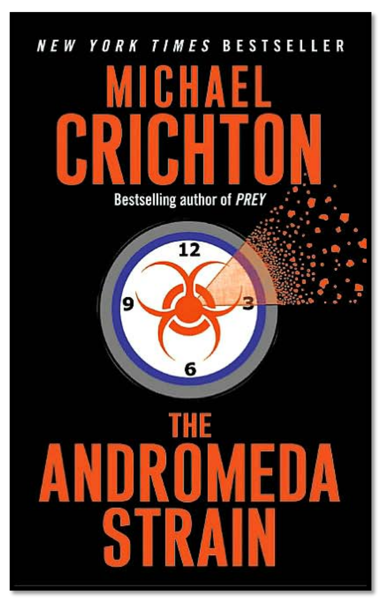
Could life be brought to the Solar System from somewhere else, frozen on the surface of dust grains?

No. There is no evidence of biomacromolecules in meteorites or IDPs, let alone viruses or bacteria.

Hoyle's famous suggestion that the 1918 flu epidemic originated in IDPs from Comet Halley's tail also does not pass medical tests.

Also begs the question of how *that* life began.

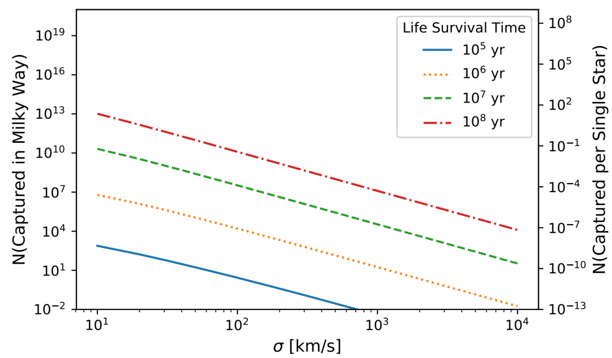
Nevertheless, this still a common suggestion.



Panspermia

With the discovery of 'Oumuamua being of interstellar origin, the concept of Galactic pansermia has seen a resurgence.

- Benefit – If life evolves on one habitable planet, it can "easily" spread to all habitable planets in the Galaxy.
- If this were the case, then we would expect to find evidence of bacteria, viruses, etc. in meteorites in the Solar System. Which we do not.



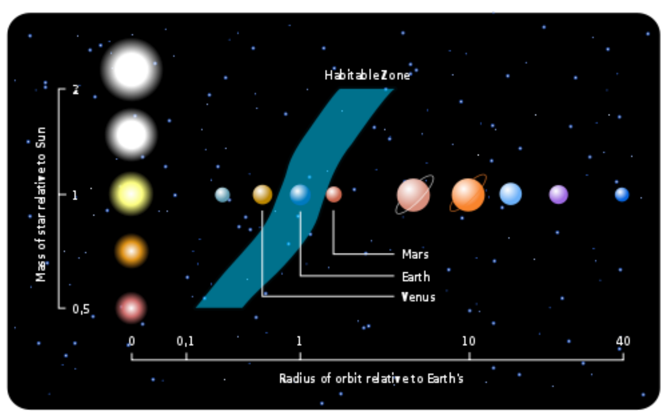
Capture rate of object as a function of velocity dispersion (Ginsburg et al. 2018)

Not quite panspermia…

Lately, it has been even more frequently suggested that primitive life may have come to Earth from Mars because Mars solidified first.

- If early Mars were habitable, and
- If life could form quickly enough, say within 500-600 Myr,
- Then primitive life like the RNA world could have evolved on Mars while Earth was still molten,
- And transferred to Earth by meteorites.

This seems unlikely, though, based on the likely habitable-zone history.



Why molecules? Clay & crystals as catalysts

Certain solid surfaces are good at trapping molecules and therefore have the potential to catalyze the polymerization of amino acids and nucleotides.

Any trapping surface will suffice from this viewpoint with the Dyson (1999) disorder-order-transition model of polymerization, discussed last time.

And some of the available solid surfaces are orderly: even periodic over large dimensions compared to amino acids or nucleotides.

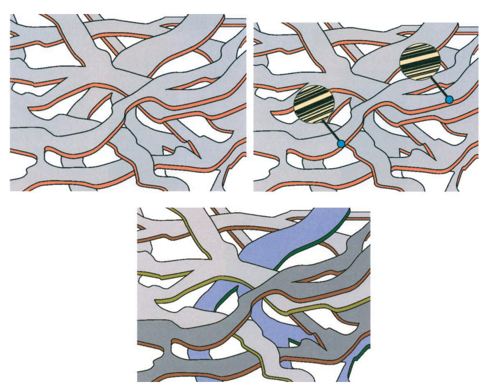
We know that crystalline solids form very early in the life of protoplanetary disks and are present throughout the development of life. Can they serve as templates for the first biopolymers?

Clay & crystals as catalysts

One interesting setting: **clays**, which consist of very thin layers of crystalline silicate (SiO_x-containing) minerals, with liquids and/or finer grains **intercalated** between the layers.

- Molecules in the interlayer would have many sites in which to be trapped, and there would be patterns of such traps at the spacing of the crystal lattices of the two layers and at the sum and difference of the two spacings.
- It is conceivable that such varieties of spacings will promote the assembly of proteins (Cairns-Smith 1968, 1985).

This is plausible and could be provided rather early…



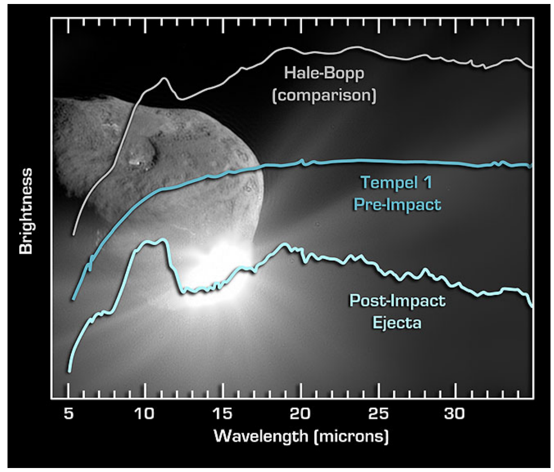
Schematic diagram of crystal planes in a clay (Cairns-Smith 2008)

Clay & crystals as catalysts

We normally think of clays as sedimentary deposits caused by chemical weathering of rocks, but they are seen in extraterrestrial objects as well, such as in Comet 9P/Tempel by the Deep Impact experiment and the Spitzer Space Telescope.

Post-impact Temple emission features include some from smectite clay.

Lisse et al. (2006)

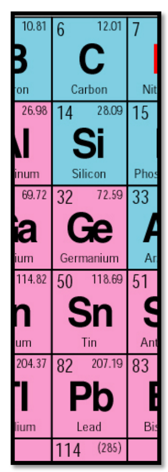


Why carbon? Other tetravalent atoms

To lead to life – especially the intelligent sort – the chemical basis needs to be capable of holding a great deal of information, which means that if it is molecular, it needs to be able to make complex molecules like carbon.

Carbon's complexity is due to its tetravalence, meaning that the atom has four electrons that can make covalent bonds. These elements are all in the same column of the periodic table. Atoms in other columns make fewer bonds.

By far, the most abundant of these elements besides carbon is silicon. Silicon life is a favorite of science-fiction writers when considering alternative forms of life.



Other tetravalent atoms

Unfortunately, Si-containing molecules do not seem to be very promising as originators of a dominant life form.

- C-C bonds are particularly strong, twice as strong as Si-Si.
- Even Si-H and Si-O bonds are stronger than Si-Si bonds; this makes long chains of Si very unstable.
- Si does not easily make double or triple bonds, which with carbon are very important in chemistry (e.g. the COOH group in amino acids).
- Silicon's oxides, carbides, and nitrides have high melting points and are insoluble, so they are difficult to keep in gas and solutions and thus do not react much under "lifelike" conditions.
- And Ge, Sn, and Pb are even worse.

Other tetravalent atoms

In the case of carbon v. Si, Ge, etc., "Earth chauvinism" is probably not a factor for the *origin* of life.

- In Earth's crust, silicon is hundreds of times more abundant than carbon, yet life on Earth wound up based on carbon.
- If silicon life could not beat carbon life with that much of an edge, how would it do so under cosmically-normal conditions, with carbon a factor of ten more abundant than silicon?

Keep in mind the possibility of non-primitive silicon life, like sentient computers, which are based upon crystalline silicon. Maybe we are just an evolutionary stage on the way to silicon life…

Vita in silico

Large single crystals of pure silicon do not occur in nature. However, as you know, *intelligent* life forms – like humans, c. 1960 – can create such crystals.

They can also create **transistors** (solid-state switches) in the crystals – far easier in silicon than in any other material.

Transistors can currently be made at a density of about one billion per square centimeter of silicon surface.

With this many transistors, one can make very powerful computers.

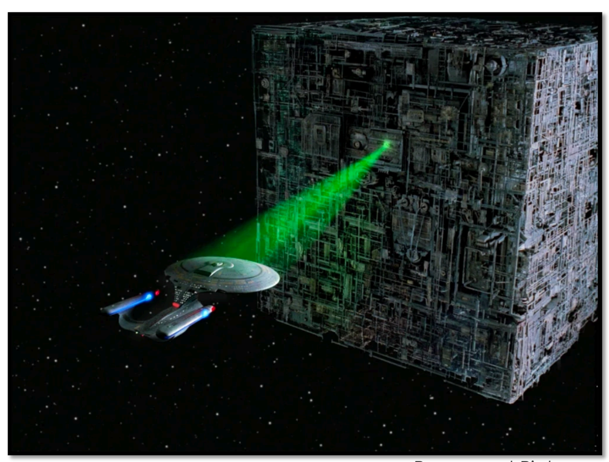
Orion Pictures

Vita in silico

And, though it cannot be currently proven, many computer scientists think that powerful software on powerful computers could become independent, self-aware intelligences.

Not all life becomes intelligent, but surely intelligence will develop all of the characteristics of life.

So silicon life may be possible in this form, but it would require other intelligent life forms first, with origins in media other than silicon: does not affect Drake's f_l.



Paramount Pictures

Why water? Other solvents

Lots of other molecules that make good solvents are abundant in the ISM and in protoplanetary disks, and are thus delivered **together** to planets in large quantities. The table includes some samples, along with their freezing and boiling points at Earth-surface atmospheric pressure:

Solvent	Freezing point [K]	Boiling point [K]
Water (H_2O)	273	373
Ammonia (NH_3)	195	240
Formaldehyde (H_2CO)	181	252
Methanol (CH_3OH)	179	338
Ethanol (CH_3CH_2OH)	159	352
Methane (CH_4)	91	109
Ethane (C_2H_6)	90	184

Other solvents

Cosmically speaking, water is the most abundant of these.

All but methane and ethane are **polar solvents** (like water) – one end of the molecule tends to always have a positive electric charge, and the other negative.

So they tend to dissolve things the same way water does, as well as promote the formation of organic zwitterions.

Water can dissolve significantly more than any other molecule.

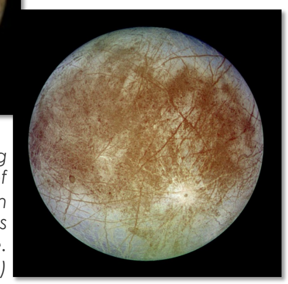
Solvent found in meteorites and IDPs	Number per 1000 water molecules
Water (H_2O)	1000
Ammonia (NH_3)	9
Formaldehyde (H_2CO)	10
Methanol (CH_3OH)	20
Ethanol (CH_3CH_2OH)	10
Methane (CH_4)	7
Ethane (C_2H_6)	4

Other solvents

Water also has a larger heat capacity and is a more efficient evaporative cooler than the others, which is valuable in temperature regulation.

So wherever it is liquid, water is probably life's solvent.

On the other hand, the others are liquid at lower temperatures than water: they could support life outside the usual definition of the habitable zone.



Leading (top left) and trailing (bottom right) hemispheres of [Europa](), which has a ~160 km deep water ocean under its ~10 km water-ice surface. (Galileo/JPL/NASA)

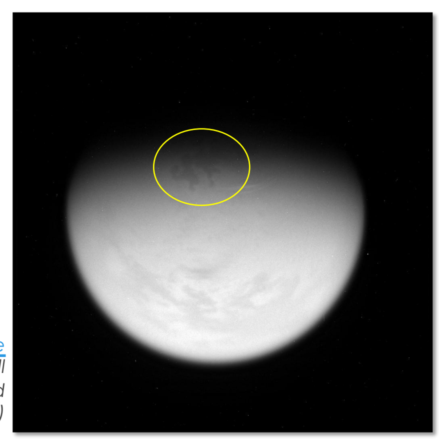
Nonpolar solvents

Methane and ethane are **nonpolar** solvents.

Some molecules (e.g. oils) dissolve more easily in nonpolar solvents than in polar ones.

No zwitterions are made in the nonpolar solvents!

So these solvents would support very different biochemistries, about which we can only speculate.



Titan's North-Pole Great Lake ("Kraken Mare"), which is full of liquid methane and ethane. (Cassini/JPL/NASA)

Biochemistry in liquid methane

One life-like step demonstrated (theoretically) in liquid methane is the plausibility of **membranes** that would play the role of Oparin's coacervate droplets as possible progenitors of cell membranes.

The mostly-nitrogen atmosphere of Titan is fairly rich in cyanoacetylene (HC_3N), delivered with this moon's original ingredients. This reacts with other molecules to produce acrylonitrile (CH_2CHCN).

Name	Structure	Abundance (ppm)
HCN	N≡C—H	200
Cyanoacetylene	N≡C—C≡CH	40
Acrylonitrile	N≡C—CH=CH_2	10

Stevenson, Lunine, & Clancy (2015)

We know from NASA *Cassini* observations that these molecules condense into aerosols and rain down onto the surface and into the methane/ethane lakes.

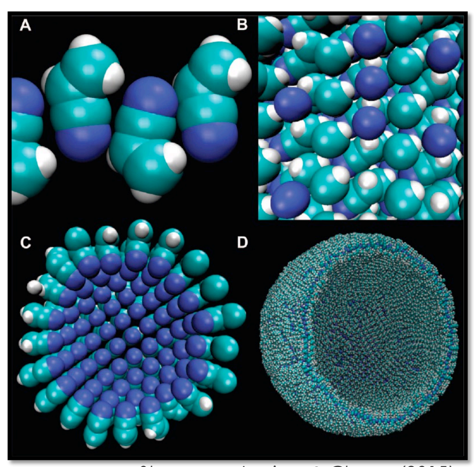
Biochemistry in liquid methane

Acrylonitrile molecules, in turn, have a tendency to hydrogen-bond with each other with the nitrogen on one end matching up with the hydrogens on the other end.

In a methane solution, simulations show that acrylonitrile then self-assembles into spherical shells 9 nm in diameter.

- Called **azotosomes**, in analogy to liposomes.

Mechanical simulations of azotosomes show that they are stretchy, despite the low T.



Stevenson, Lunine, & Clancy (2015)

Biochemistry of liquid membranes

This structure is backwards from the way lipids form into liposomes in water:

- Liposomes present a polar "face" to the surrounding water, which itself is polar.
- Azotosomes present a nonpolar face which goes with the nonpolar solvent, methane.

Thus at least one of the steps of Life is possible in liquid methane, but the stiffness of long C chains at low temperatures is still a problem.

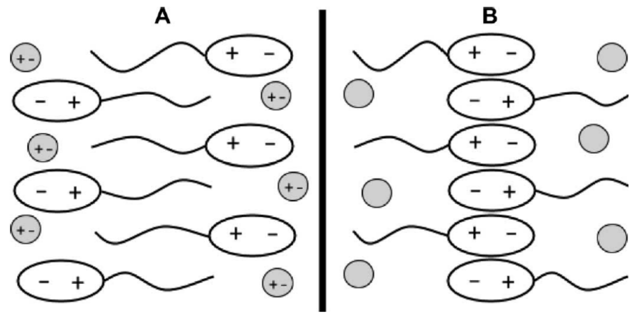


Fig. 1. Liposomes and azotosomes. (**A**) Liposome in polar solvent Polar heads are braced by nonpolar lipid tails. (**B**) Azotosome in non-polar solvent. Nonpolar tails are braced by polar nitrogen–rich heads.

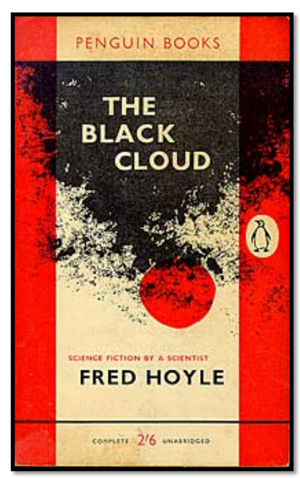
Stevenson, Lunine, & Clancy (2015)

Why atomic matter at all?

It would take too long to discuss all the possible non-matter-based life forms raised by science-fiction writers; here are a few motifs.

"Pure energy." Epitomized by Hoyle's *The Black Cloud*, in which a molecular cloud in the interstellar medium is imagined to come alive, its mind present in electric and magnetic fields and molecules. This particular scheme can be safely ignored: interstellar clouds and all their electromagnetic fields are observed to live for only tens of Myr, and being so low in density, they would need much more time to organize, not less.

- This is not to say that "pure energy" life would never become important; just not in the beginning. But if you have a looooooooong time to evolve…
- If the Universe is open (as indeed it seems to be), then after thousands of Gyr nothing will be left by black holes, photons, and gravity. Believe it or not, it is possible to conceive even intelligent life under these circumstances.



Why atomic matter at all?

Star life. The remains of dead stars – white dwarfs and neutron stars – become much more orderly as they cool off through eternity; they even crystallize as they do so (atomic crystals in a WD, nuclear crystals in a NS). And they are certainly dense and long lived…

- Hard to evaluate this as it is unsafe to get too close to these objects (even if we could); only objections are that the temperature stays high for many Gyr, and extremely strong gravity may restrict such forms of life to two dimensions. (See *The search for life in the Universe* by Don Goldsmith and Toby Owen, 2002.)

Quantum life. Perhaps reenergized by the new field of quantum computing, some have imagined that subatomic systems, or others that operate quantum mechanically, could be alive and intelligent.

- Upside: huge numbers of internal rearrangements are possible, and time does not really enter the problem.
- Downside: the arrangements do not seem to be deterministic and are instead ruled by probability. As they interact with their surroundings, they would spontaneously switch between alive and dead. Memory would be difficult to imagine.

Summary

It looks as if carbon/water-based life has a few promising ways by which it could have arisen spontaneously, and one – RNA world – that seems as if it could get going very quickly.

If RNA developed first – rather than proteins – a "genetic takeover" would not have had to occur to produce the nucleic-acid version of life we now know.

No other set of molecules seems nearly as promising… and that is not just geocentrism.

In temperature ranges over which water is not liquid, there may be other ways in which life could develop. We do not know, chemically, another "guaranteed" way.