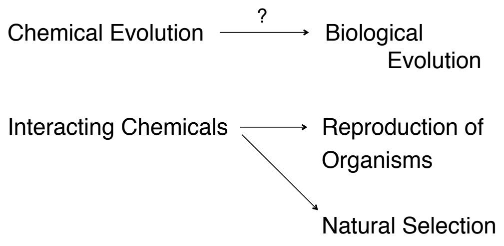
The Transition to Life

The Transition to Life



Chemical Evolution $\xrightarrow{?}$ Biological Evolution

Interacting Chemicals → Reproduction of Organisms

Interacting Chemicals → Natural Selection

Based on Simplest Life Now:

Need:

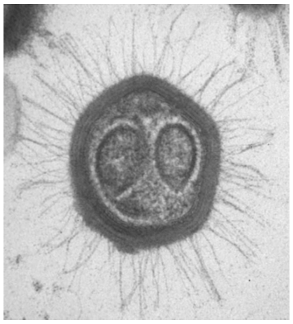
1. Nucleic Acids	Replicable Information
2. Proteins	Enzymes (Catalysts)
3. Lipids	Membranes (Enclosure)
4. Carbohydrates	Energy Storage
(Pigments)	(Energy Conversion)

Too much to ask of chemical evolution
⇒ Protolife?

Update: Mimivirus

- A very large virus was discovered in 2003
- Both RNA and DNA
- More DNA than some bacteria (> 1000 genes)
- Genes for translation, DNA repair enzymes
- Leading to reevaluation of viruses
- May be ancient lineages
  - – Precursors to bacteria, or eukaryotes
  - – Controversial

Image of Mimivirus



Protolife

1. “Virus” Free living but equivalent in complexity
   $\underbrace{\text{Protein + Nucleic}}_{\text{Genetic Code}}$ Acid + Supply by Environment
2. Protein Protolife
   Protein ⟶ Self Replication?
3. Nucleic Acid Protolife
   RNA ⟶ Self Catalysis?

---

4. Something Else
   Minerals
   Clay Layers
   Mineral - Molecule
   Pyrite
   Thioesters

Genetic Takeover
? ⟶ RNA ⟶ DNA

Protein-Based Protolife

1. Proteinoid microspheres - Sidney Fox
   $\text{Amino Acids + Dry Heat} \xrightarrow[\text{(Hot Tidepool?)}]{} \text{Proteinoids}$
   ↓ H_2O (Tide)
   Microspheres

Protocells		
Protolife?	Bacteria	
Can Add Proteinoid	Grow	
Split	Divide	“Reproduce”
Bud	Bud	
Form Chains	Form chains	

But “Reproduction” not exact
Later incorporate Nucleic Acids
Proteinoid ⟶ Cells ⟶ Genes

Problem: How to incorporate Nucleic acids?

Picture of Proteinoid Microspheres

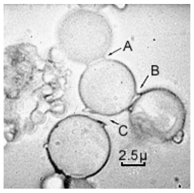


FIGURE 5.15 — Photograph of proteinoid microspheres produced by repeated energizing and dehydrating the primordial soup. The main features of this figure can be simulated by shaking a mixture of oil and water and watching the globs of oil cluster on the surface of the water. Seen here through a microscope, each microsphere contains a large\ concentration of amino acids. (The scale shown, 2.5 microns, equals 2.5x10-4 cm.) (Sidney Fox)

Nucleic Acid Based Protolife

RNA → Genes → Protein → Cells
Self-replicating RNA molecules
Experiment by Sol Spiegelman
RNA from Q_β Virus - parasite on bacteria
Injects RNA - Bacterium makes replicase
↑
Enzyme to Replicate RNA

RNA multiplies, using activated nucleotides in bacterium to copy RNA and make new viruses

In Test Tube: Template RNA, Replicase, Activated Nucleotides (ATP, CTP, GTP, UTP)

⇒ RNA copied **without** machinery of cell

Variation: **No** template RNA

Replicase made RNA from nucleotides
↑
Protein

Manfred Eigen - further experiments with RNA in test tube:

Mutant RNA strands compete
Degrade to smallest (~ 200 nucleotides)
RNA that replicase could recognize
(Monster - Selfish RNA)

RNA can do self-catalysis in some cases
Could this have led to self replication?

Eigen scenario

1. A replicating RNA molecule forms by chance (random replicator - not a gene) ribozyme (catalyst, made of RNA)
2. Family of **similar** RNA's develops (quasispecies)
3. Connection to proteins
   (quasispecies specialize to make parts of protein)

---

4. Complex interactions (hypercycles)

6. Use lipids to make protocells

6. Competition leads to biological evolution

---

434 | CHAPTER 22

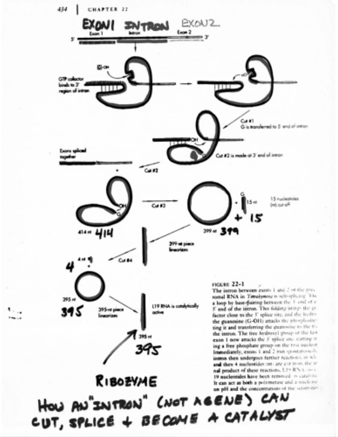


FIGURE 22-1
The intron between exons 1 and 2 of the pre
somal RNA in *Tetrahymena* is self-splicing. Th
a loop by base-pairing between the 3' end of e
5' end of the intron. This folding brings the G
factor close to the 5' splice site, and the hydro
the guanosine (G-OH) attacks the phosphodie
ting it and transferring the guanosine to the
the intron. The free hydroxyl group of the las
exon 1 now attacks the 3' splice site, cutting it
ing a free phosphate group on the first nucleot
Immediately, exons 1 and 2 join spontaneously
intron then undergoes further reactions, in wh
and then 4 nucleotides (nt) are cut from the
nal product of these reactions, L19 RNA
19 nucleotides have been removed, is catalytic
It can act as both a polymerase and a nuclease
on pH and the concentrations of the substrate

HOW AN "INTRON" (NOT A GENE) CAN CUT, SPLICE + BECOME A CATALYST

Problems with Nucleic Acid First Scenario

1. Hard to get monomers
2. Unlikely to link correctly
3. Need existing proteins and lipids
4. Hypercycles subject to instabilities

N = size of molecular population

If N small → Population Collapse

If N large → Selfish RNA, Short Circuit

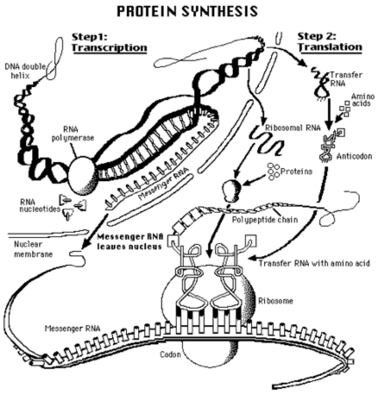
$$\begin{tikzcd} A \arrow[r] & B \arrow[d] \\ D \arrow[u] & C \arrow[l] \end{tikzcd}$$

If B → D Short Circuit

⇒ Only narrow range of sizes works

The Origin of the Genetic Code

- We need more than **either** protein or RNA protolife
- Need interaction via genetic code
- Need **translation**
- Let’s recall what is needed for translation…

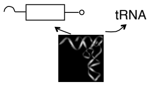


Shapiro’s Fable

The case for the “chicken”

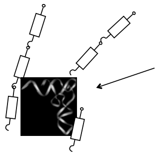
Protein first ⇒ replication problem

“interpreters” aminoacyl tRNA synthetases



Match tRNA &
Amino acids

Could an earlier version have copied proteins directly?



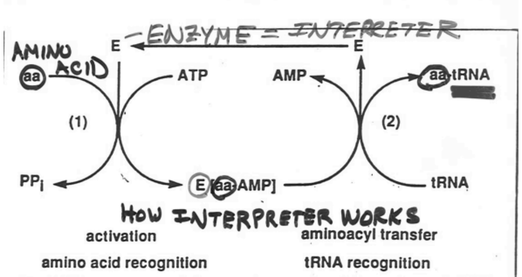


Fig. 1 The two steps of the reaction catalysed by aminoacyl-tRNA synthetases. The tRNA is recognized in the second step (through the features designated paracodon in this article) by the enzyme carrying a bound aminoacyl-AMP intermediate. Participation of the aminoacyl group in the recognition process is thus an attractive possibility.

1. Early Evolution: Start with 4-6 amino acid types, gradually add more
   enzymes increase in size and catalytic power
2. First use of phosphate as energy? (ATP)
   or sugar-phosphate chains for construction
   (Teichoic acids in membranes of some bacteria)
   (partial Q_β replicase)
3. Bases added for structure
   Support for protein synthesis → ribosome

---

4. Begin to copy RNA (Full Q_β replicase)
   Natural selection leads to better ribosome
5. Specialized, Short RNA aided attachment of amino acids to proteins; became tRNA
6. Then mRNA - to align tRNA's
   now a separate genetic system that evolves
7. DNA developed from RNA

Shapiro dates last step to prokaryote -eukaryote split (different ways of storing DNA info)

Tests:

1. Synthesize in lab? Not possible yet.
3. Molecular archaeology - vestigial ability of interpreters to recognize amino acids in proteins
3. Survivors of protein era? prions?

Support for the "chicken"

1. 1988 discovery that interpreter does not use tRNA codon to recognize correct tRNA
   (in some cases) ~ 1/2
   - instead a single base pair at the other end of tRNA
   ⇒ simpler, older code
   second genetic code

   ⇒ connection of interpreter and tRNA
   more primitive than current code

2. Dyson modeling of molecular "populations"

Transition from disorder to order
(non-life) (life)

Finds number of monomer types likely to be
9 - 11 (ok if used ~ 1/2 of modern proteins)
But nucleotides (only 4) - not enough

Favors protein first

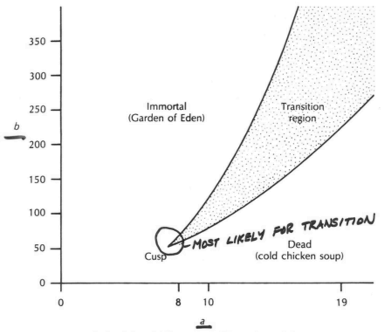


FIGURE 2.7 *Summary of Dyson's model*

1+a = # OF TYPES OF MONOMERS
b = DISCRIMINATION FACTOR OF CATALYST

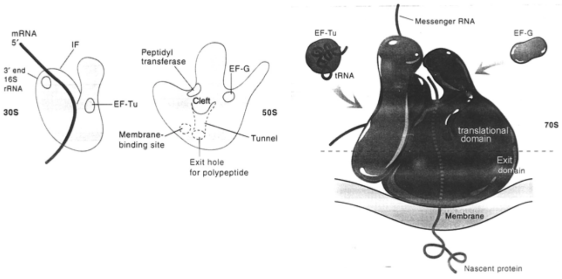
The Egg Strikes Back

Other work shows some RNA can catalyze
Non-RNA reactions

1. RNA in ribosome appears to be what catalyzes peptide bond formation
   Noller, et al. 1992, *Science*, **256**, 1416

2. RNA "ribozyme" catalyzes reactions between amino acids and tRNAs
   First "interpreter" may have been RNA
   Piccirilli, et al. 1992, *Science*, **256**, 1420

Translation



Origin of the Genetic Code

Crucial step in any theory

Allows communication
Nucleic Acids ⟷ Proteins

Early versions probably coded fewer amino acids - less specific

---

For mRNA Genetic Code

First RNA Base	U	C	A	G	Third RNA BASE
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	Aspartic Acid	Glycine	U
	Valine	Alanine	Aspartic Acid	Glycine	C
	Valine	Alanine	Glutamic Acid	Glycine	A
	Valine	Alanine	Glutamic Acid	Glycine	G

Amino Acids

Purine Either
Some evidence for RNY and G - C more stable
Pyrimidine

⇒ 4 codons	GGC	glycine	Common in Miller-Urey and Meteorites
	GCC	alanine	
	GAC	aspartic acid	
	GUC	valine	

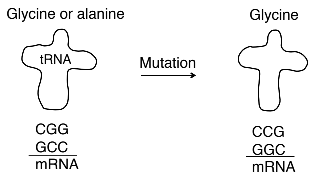
Others added later

Evolution of Genetic Code

Gaining specificity

If early tRNAs carried more than 1 kind of amino acid

e.g.



Evidence that code has evolved
Freeland, et al. Tested 10^6 other codes

Only one better at minimizing bad effects of mutations
⇒ Natural Selection

Still Evolving
Some organisms have slightly different codes in mitochondria or in nucleus

Summary

- Transition to life is poorly understood
- Need to consider "protolife"
- Can we get by with only one polymer?
  - If so, protein or RNA
  - If so, how do we get genetic code going?