The Transition to Life

The Transition to Life

Chemical Evolution —?→ 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

$\Rightarrow$ Protolife?

Protolife

1. “Virus” Free living but equivalent in complexity

   Protein + Nucleic Acid + Supply by Environment

   Genetic Code

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

   Amino Acids + Dry Heat $\longrightarrow$ Proteinoids

   (Hot Tidepool?)

   $\downarrow$ H_2O (Tide)

   Microspheres

Protocells

Protolife? (Look like life)

Can Add Proteinoid

- Split
- Bud
- Form Chains

Grow

Divide } "Reproduce"

Bud }

Like Bacteria

But "Reproduction" not exact

Later incorporate Nucleic Acids

Proteinoid $\longrightarrow$ Cells $\longrightarrow$ Genes

<u>Problem:</u> How to incorporate Nucleic acids?

Picture of Proteinoid Microspheres

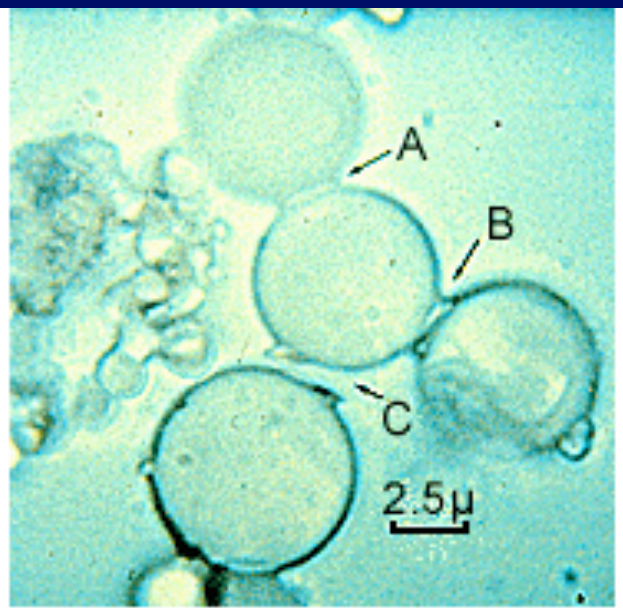


FIGURE 5.15 — Photograph of proteinoid microspheres produced by repeated energizing and dehydrating the primordial soup. The main features of this figure can be simulated by shaking a mixture of oil and water and watching the globs of oil cluster on the surface of the water. Seen here through a microscope, 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 $\longrightarrow$ Genes $\longrightarrow$ Protein $\longrightarrow$ 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)

$\Rightarrow$ 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)

5. Use lipids to make protocells

6. Competition leads to biological evolution

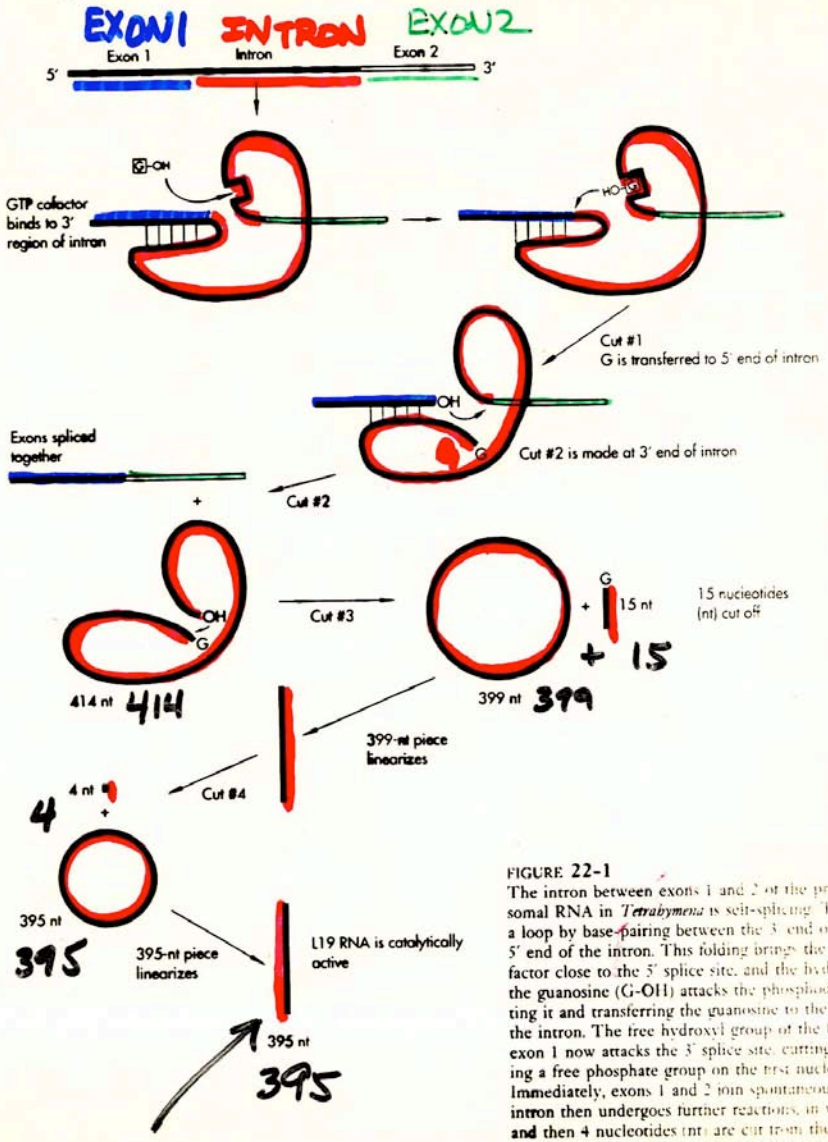


FIGURE 22-1
The intron between exons 1 and 2 of the pre- somal RNA in *Tetrahymena* is self-splicing. The a loop by base-pairing between the 3′ end of e 5′ end of the intron. This folding brings the gu factor close to the 5′ splice site, and the hydro the guanosine (G-OH) attacks the phosphate ting it and transferring the guanosine to the 5′ the intron. The free hydroxyl group of the last 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 3′ nal product of these reactions, L19 RNA (so c 19 nucleotides have been removed) is catalytic It can act as both a polymerase and a nuclease on pH and the concentrations of the substrates

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

Problems with Nucleic Acid First Scenario

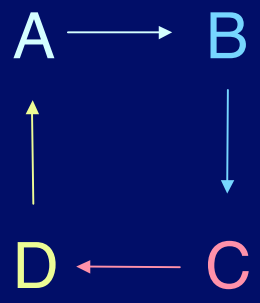
1. Hard to get monomers
2. Unlikely to link correctly
3. Need existing proteins and lipids
4. Hypercycles subject to instabilities

N = size of molecular population

If N small → Population Collapse

If N large → Selfish RNA

If N large → Short Circuit



If B ⟶ D Short Circuit

⇒ Only narrow range of sizes works

The Origin of the Genetic Code

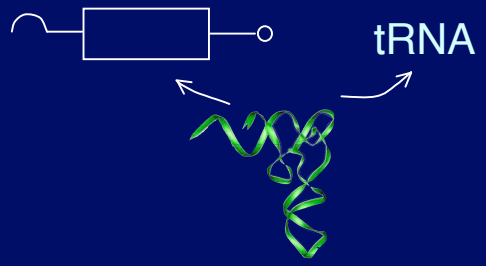
- We need more than **either** protein or RNA protolife
- Need interaction via genetic code
- Need **translation**
- Consider first a scenario by R. Shapiro

Shapiro's Fable

The case for the "chicken"

Protein first ⇒ replication problem

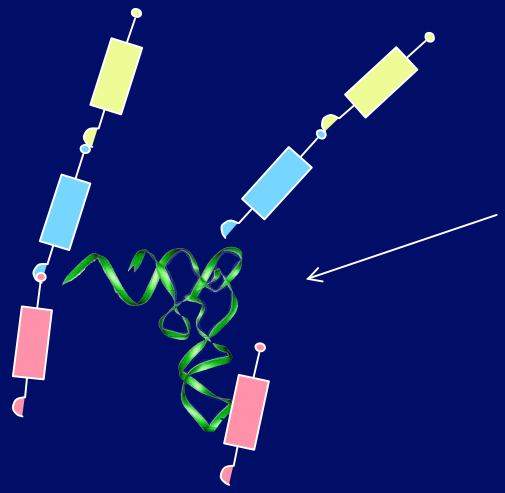
"interpreters" aminoacyl tRNA synthetases



Match tRNA &

Amino acids

Could an earlier version have copied proteins directly?



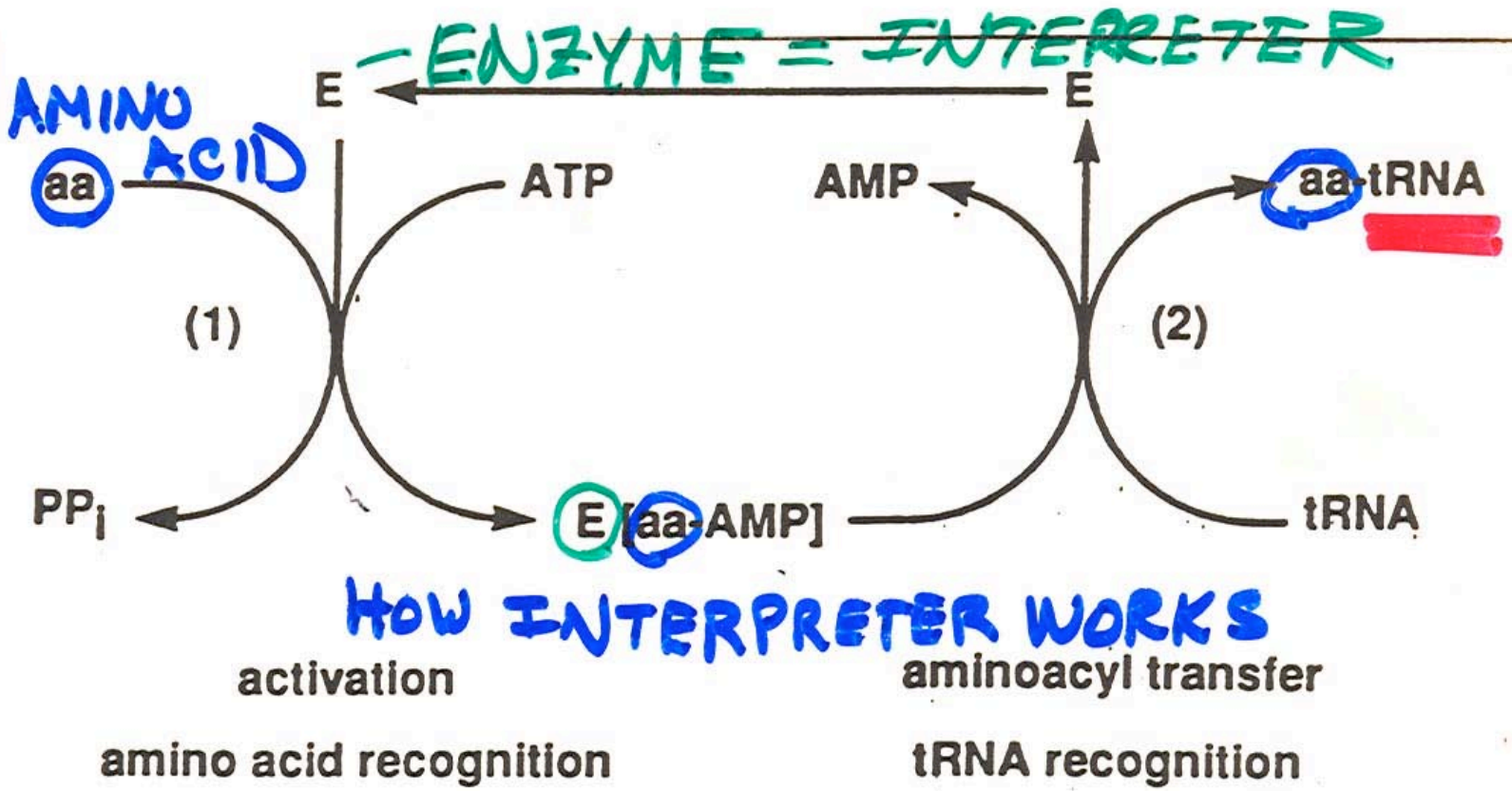


Fig. 1 The two steps of the reaction catalysed by aminoacyl-tRNA synthetases. The tRNA is recognized in the second step (through the features designated paracodon in this article) by the enzyme carrying a bound aminoacyl-AMP intermediate. Participation of the aminoacyl group in the recognition process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

2. 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

Origin of the Genetic Code

Crucial step in any theory

Allows communication

Nucleic Acids $\longleftrightarrow$ Proteins

Early versions probably coded fewer amino acids - less specific

Some evidence for RNY and G - C more stable

(Purine, Either, Pyrimidine)

$\Rightarrow$ 4 codons

GGC	glycine
GCC	alanine
GAC	aspartic acid
GUC	valine

Common in Miller-Urey and Meteorites

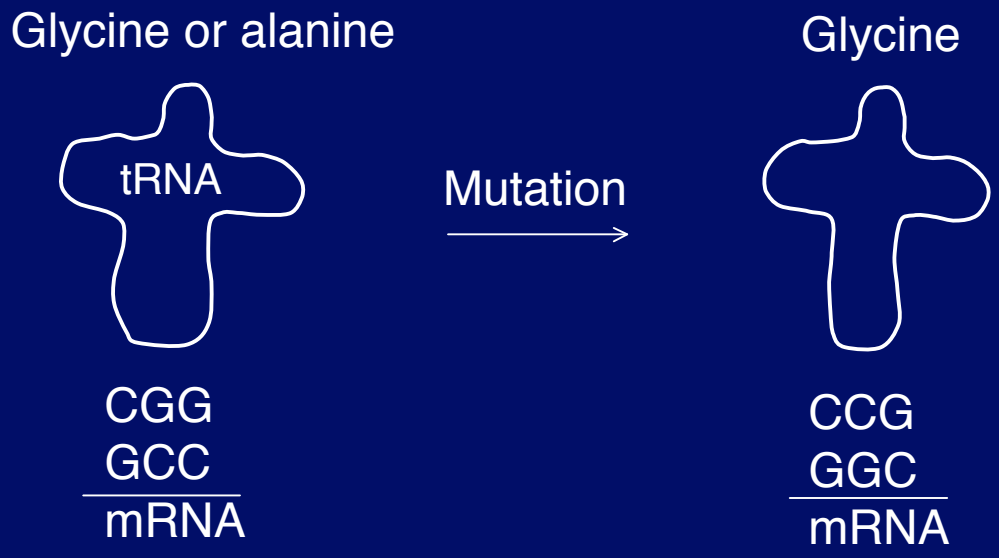
Others added later

Evolution of Genetic Code

Gaining specificity

If early tRNAs carries 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

$\Rightarrow$ Natural Selection

Still Evolving

Some organisms have slightly different codes in mitochondria or in nucleus

Other Ideas

- Neither the chicken nor the egg came first
- Transitional forms that were later discarded

Or was it the “egkin”?

Some experiments with peptide nucleic acid (PNA).

PNA: Peptide backbone with bases

Can act as template for polymerization of RNA

From activated nucleotides

(Böhler, et al., *Nature*, **376**, 578 & comments by Piccirilli, pg. 548) 17 Aug. 1995

PNA could be simpler to form under prebiotic conditions

Main point is that a simpler thing (not necessarily PNA) could have preceded RNA

Membranes

- Membranes provide enclosure
  - Also fundamental for metabolism
- Membranes never arise from scratch
  - Always passed down and added to
  - All derived from ancestral cell
- T. Cavalier-Smith proposes membranes
  - Plus nucleic acid formed “ob-cell”
  - Merger of 2 ob-cells formed first cell

Thioester World

1. Need precursor to RNA world

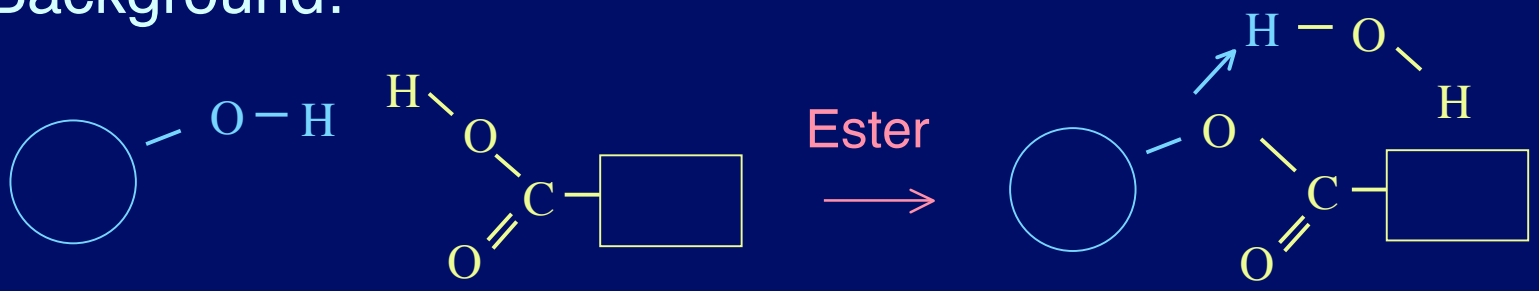
C. deDuve

2. Need energy conversion

In Vital Dust

Protometabolism

Background:



Hydroxyl + Carboxyl

Thiol + Carboxyl

Thiols involved in metabolism, particularly in ancient pathways

Also can catalyze ester formation by group transfer Reactions
e.g. peptide bonds

Catalytic Multimers

C. de Duve

"Multimer" short peptides and esters

(NH_2) (OH)

of amino acids and hydroxy acids

Will form from thioesters. Assume some catalytic ability, lead to protometabolism

Energy Sources

Basic need is hydrogen atoms
(or electrons in excited states)

In pure water $\underline{H^+ + OH^-}$ more if acidic
H_2O

$e^- + H^+ \longrightarrow H^*$ excited H

Now chlorophyll + sunlight

Then?

1. UV light + Fe^{+2} $\longrightarrow$ Fe^{+3} + e^-

   e^- + H^+ $\longrightarrow$ H^* $\longrightarrow$ H for reactions

2. H_2S in H_2O $\longrightarrow$ SH^-

   $2SH^- \longrightarrow S_2^{-2}$ + H_2

   Fe^{+2} + S_2^{-2} + FeS_2 iron pyrite

Transition to Phosphate

Energy currency in life now is ATP

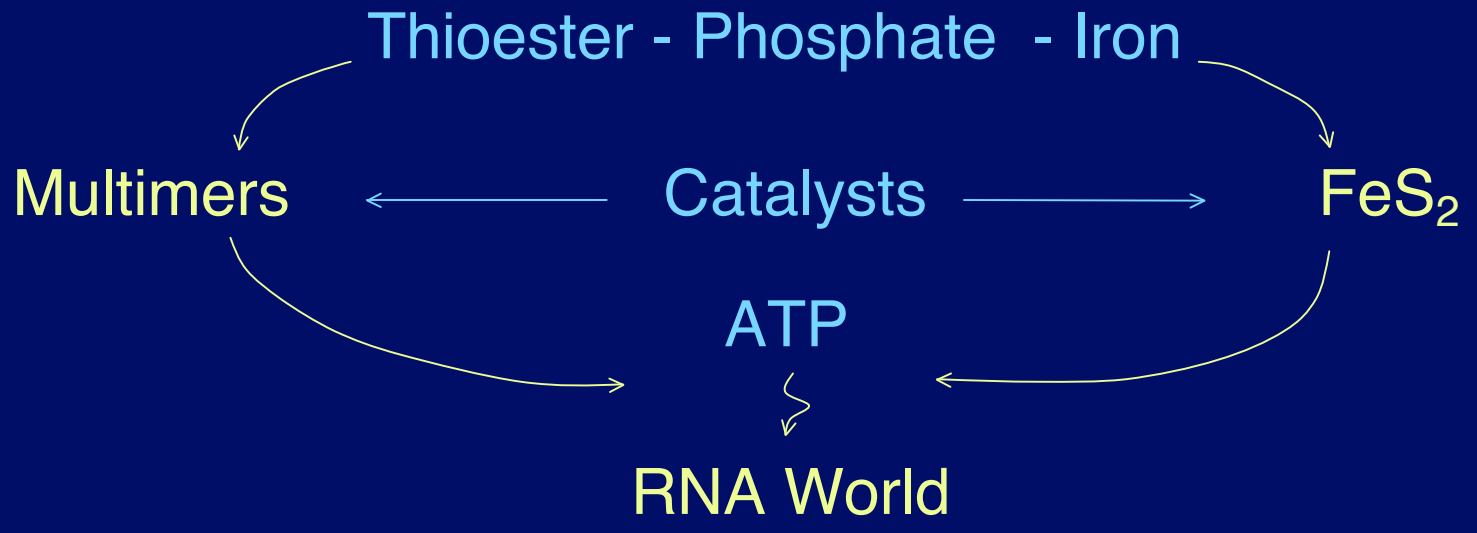
Adenosine Triphosphate

used to make bonds, remove H_2O

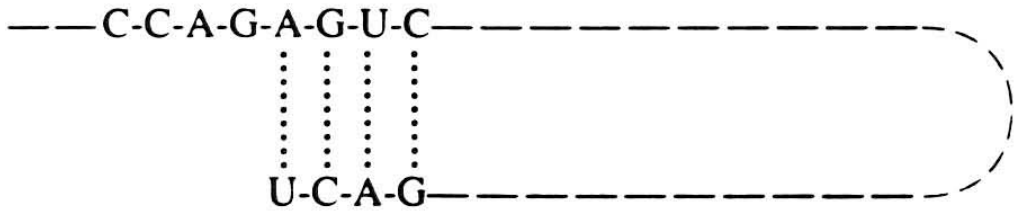
Earlier, inorganic phosphate

p–p diphosphate or polyphosphate

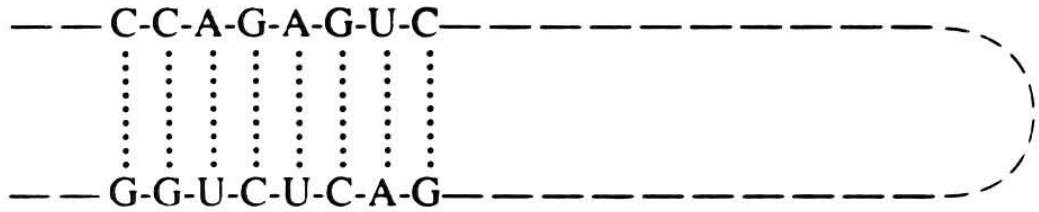
still involved in ATP reactions



age—will be followed by GUC. This AGUC sequence is complementary to the terminal sequence written in antiparallel fashion, and will cause the chain to double up as follows:



Assume now that this folded chain is subject to elongation, by the addition of new nucleotides, from right to left, to the U end. The presence of G next to the A paired with the terminal U is likely to favor the addition of a complementary C over that of the other three possible nucleotides. Repeat the process and you get U added opposite A, G opposite C, G again opposite the next C, and so on. What you get is the formation of a stretch complementary over all its length to the other end of the molecule:



Summary of Proto-Life Development

Stage	**Proteins**	**Halfway # 1** Peptide Nucleic Acids	**Halfway # 2** RNA Ribozyme	**Nucleic Acids**
Monomers	Amino Acids	Bases Amino Acids	Ribose Sugars Bases Phosphates Amino Acids	Ribose Sugars Bases Phosphate
Polymerization	Proteinoids	Short strands of PNA's	Short strands of RNA + amino acids	Short strands of RNA
Replication	?	Affinity for complementary bases + ease of peptide bonding	Affinity for complementary bases	Affinity for complementary bases
Pre-life	Proteinoids + RNA?	Separation of proteins and nucleic acids	Separation of nucleic acids and protein parts	RNA adapts proteinoids as needed
Life	Proteins	Disappears	Disappears	DNA and RNA

Update: Mimivirus

- A very large virus was discovered in 2003
- Both RNA and DNA
- More DNA than some bacteria
- Genes for translation, DNA repair enzymes
- Leading to reevaluation of viruses
- May be ancient lineages
  - Precursors to bacteria, etc.
  - Controversial

Image of Mimivirus