Prebiotic Evolution

Molecular & Cell Biology

Lecture 1

Prof. David Fitch
Department of Biology
New York University
david.fitch@nyu.edu
http://www.nyu.edu/projects/fitch/

What is life?

- **Self-sustained heritability**
  - Functionality is *limited* by the **genome**
  - Life cannot be explained entirely by functionality ("design")
    - Why do snakes have femurs?
    - Why do humans have 5 fingers per hand?
    - Why do mitochondria and chloroplasts have their own rDNA and genetic codes?
  - Because of heredity, **history** is the explanation for **current patterns**
    - "Historical constraint" (phylogenetic or genomic constraint)
  - History of life = evolution

What is evolution?

- **Evolution is a *result*: heredity + variation**
  - "Descent with modification"
  - *Fact*: evolution has occurred
    - Began as a hypothesis
    - Not falsified, though falsifiable
  - *Theory*: evolution is a body of explanatory principles
    - **Embodies a mechanism**: natural selection, itself a *result* of:
      - Variation in fitness (ability to survive to reproduce)
      - Heredity (ability to pass alleles to progeny)
      - Reproduction / multiplication (ability for population to grow)
    - **Explains a variety of phenomena**
      - Adaptations
      - Shared nonadaptive features
- **"Origin of species" *vs*. origin of life**
  - Difficulty of historical reconstruction increases with elapsed time
  - Not possible to reconstruct actual events before heredity originated

Origin of Life

- **Precellular, Cellular**
- **Testing hypotheses**
  - *A priori* **assumptions** (often not stated) may themselves be tested
  - **Hypotheses** must make falsifiable "predictions"
    - "Null" hypothesis: chance alone is responsible
  - **Predictions** are confirmed (consistent with) or refuted by **data**
  - "**Parsimony**" distinguishes between alternative, unfalsified, working hypotheses
    - "Burden of proof" for less parsimonious (more complex) hypotheses
- **Alternative hypotheses**
  - Intelligent design
    - Least parsimonious (requires a creator)
    - Lack of evidence for "design" does not falsify existence of intelligence
  - Extraterrestrial origin
  - Spontaneous self-organization and natural selection
    - Also embodies many alternative hypotheses

Precellular evolution

- **"Prebiotic synthesis"**
  - **Hypothesis**: The molecules of life can be formed spontaneously under under "prebiotic conditions"
  - **Assumptions**: Prebiotic conditions, "uniformitarianism"
    - (NOT that humans should arise "continuously" from chimpanzees!)
  - **Predictions**
    - Amino acids and purines from prebiotic mixtures (Harold & Urey 1953)
    - Ribose from formaldehyde ("formose" reaction; Butlerow 1861)
    - Polymerization of "activated" nucleoside monomers without proteins
- **The "RNA World"**
  - **Hypothesis**: RNA preceded proteins and DNA as a primordial, information-bearing, catalytic molecule
  - **Assumptions**: Heredity, uniformitarianism
  - **Predictions**:
    - Conserved roles for RNA in fundamental machinery of life
    - At least some RNAs should demonstrate elementary catalytic activity
    - Role for RNA as precursor to DNA may be conserved

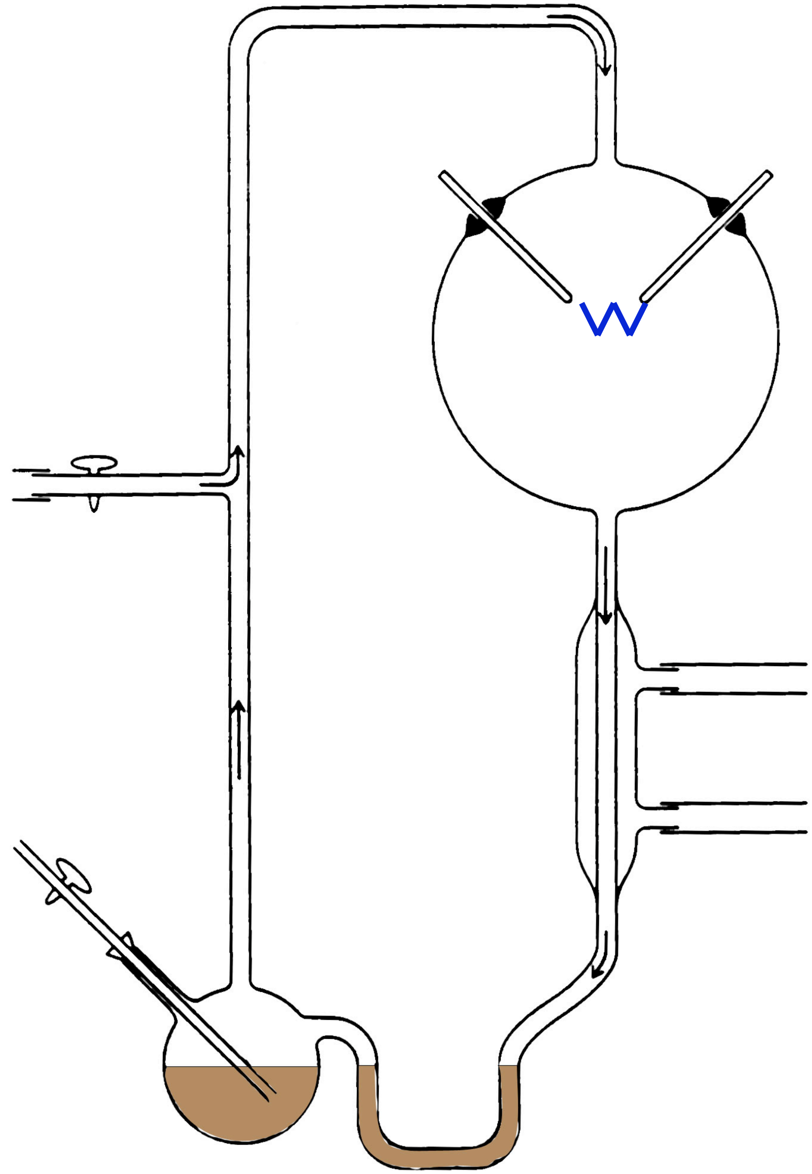
Prebiotic conditions?
(*a priori* assumptions)

- **Molecules** (potentially available from atmosphere, deep ocean rifts, benthic clays)
  - H_2, CH_4, NH_3, CO, H_2S
  - Mineral catalysts (phosphorus, pyrite, clays)
- **Energy**—potential sources (cal cm^{-2} yr^{-1})

– Solar radiation	2.6×10^5	Reasonable energy levels for organic synthesis
– UV at wavelengths:		
• 300-400 nm	3.4×10^3	
• 250-300 nm	5.6×10^2	
• 200-250 nm	4.1×10^1	
• <150 nm	1.7	
– Electrical discharges	4.0	
– Shock waves	1.1	
– Radiactivity	8×10^1	
– Volcanoes	1.3×10^{-1}	
– Cosmic rays	1.5×10^{-3}	

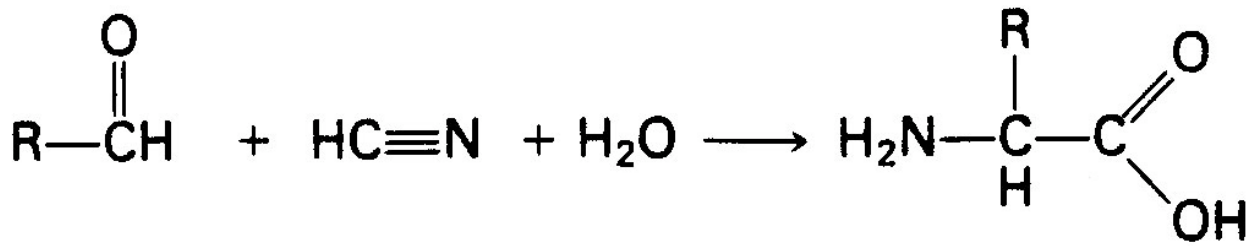
Oparin & Haldane (1920) hypothesis

- **Hypothesis:** The origin of life was prebiotic
- **Prediction:** Molecules of life should arise spontaneously in prebiotic conditions
- **Test** (Stanley Miller & Harold Urey 1953)
- **Results:**
  - >10% of C from CH_4 was in organic molecules
  - These included amino acids and precursors
    - Amino acids (G, A, D, V, L)
    - HCN and other cyano compounds
    - Aldehydes

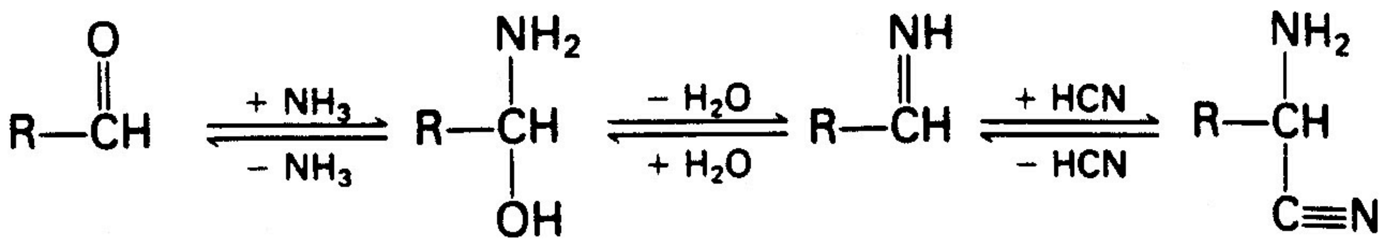


Precursors to amino acids

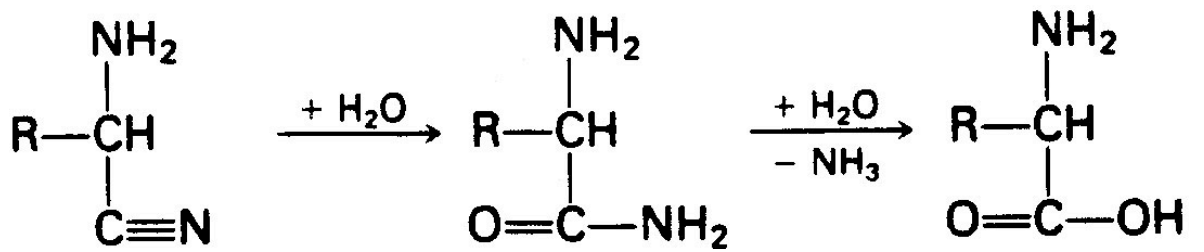
- **"Strecker synthesis"**
  - Overall reaction:



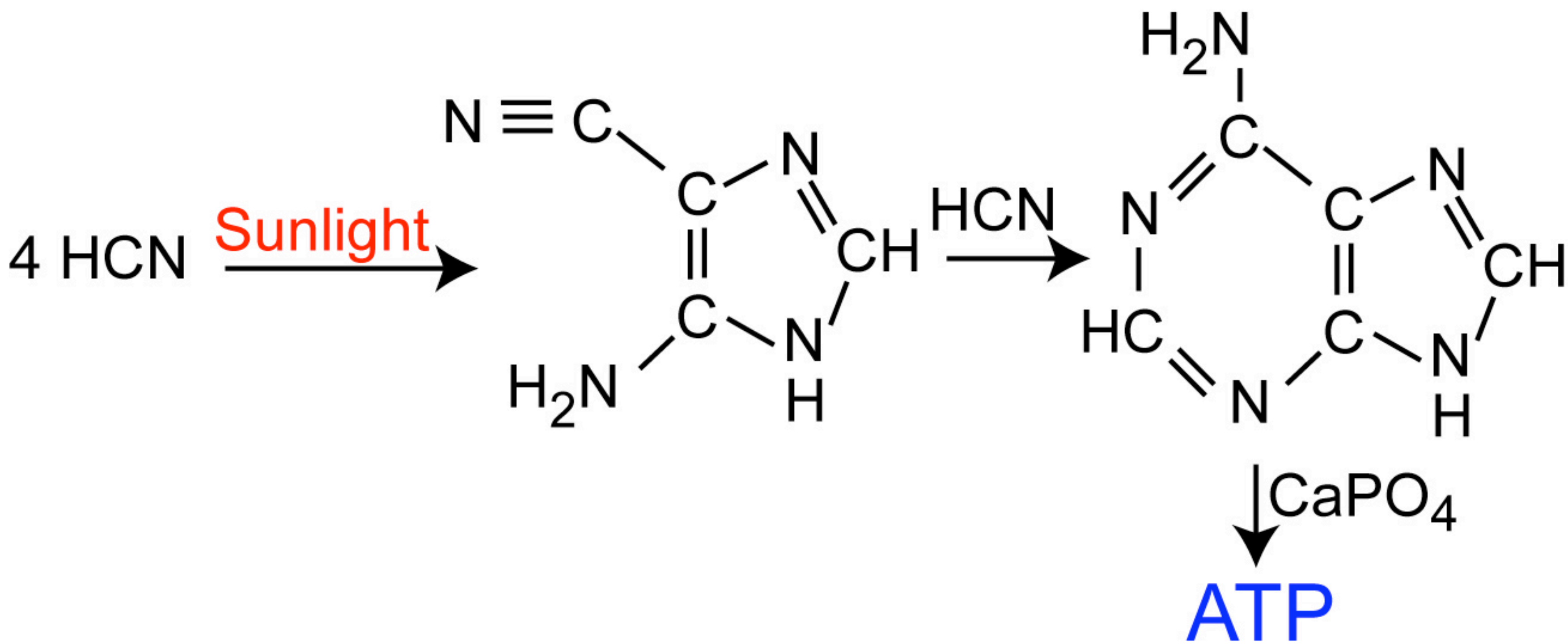
  - In atmosphere:



  - In ocean:

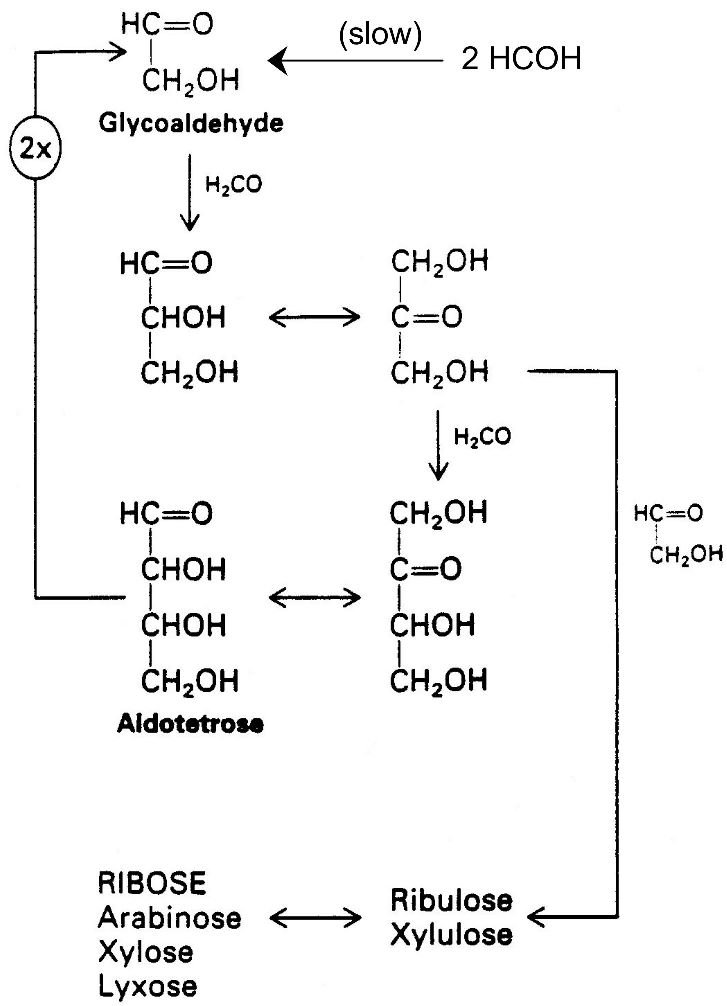


Precursors to purines



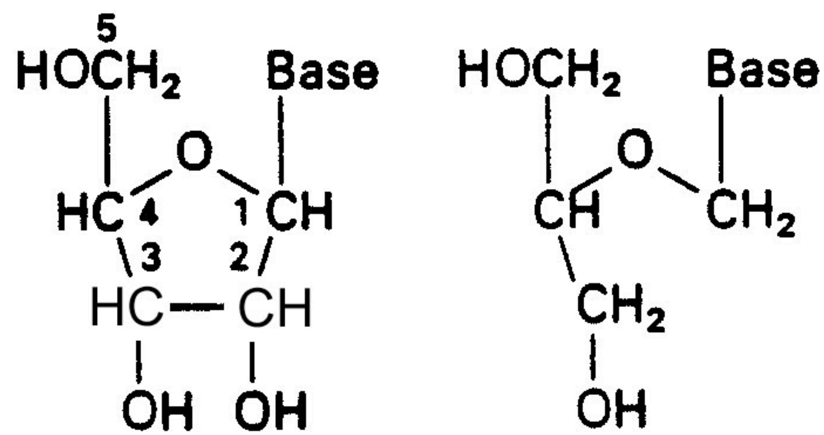
Precursors to ribose

- **"Formose" synthesis** (Butlerow 1861)
  - Series of condensations beginning with formaldehyde



Ribose-based nucleoside

Acyclo-nucleoside

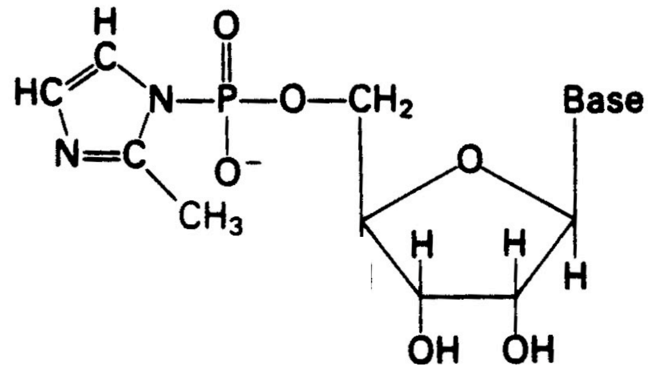


Phosphorylation would "activate" monomers

Origin of polymers

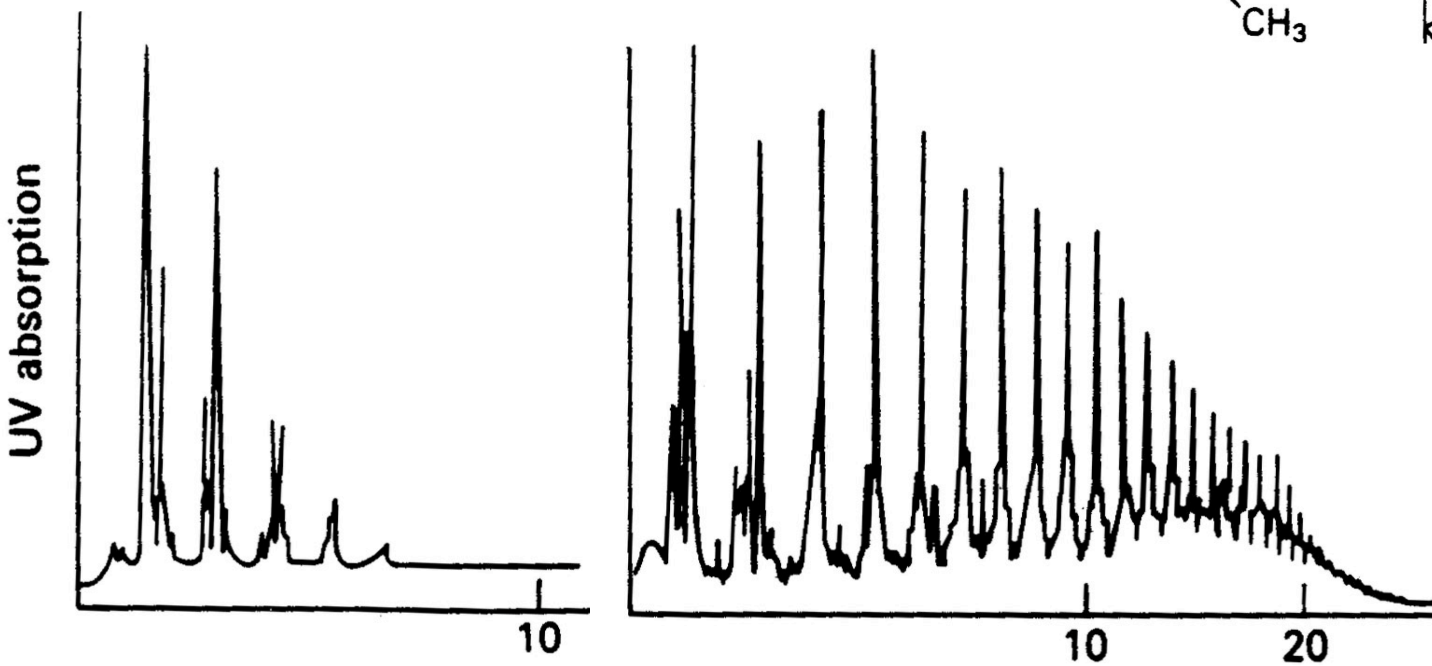
- **Model for formation of RNA by activated nucleosides**

e.g., methylated nucleoside 5'-phosphorimidazolide:



Without metals

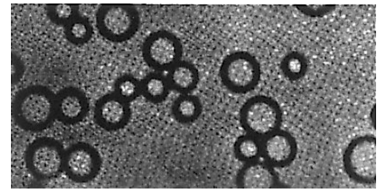
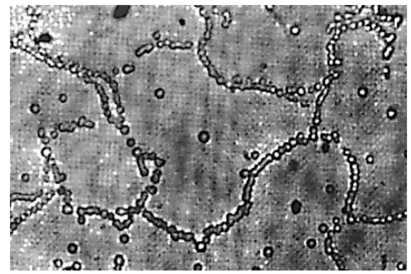
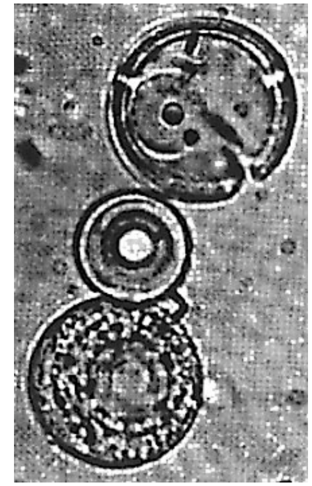
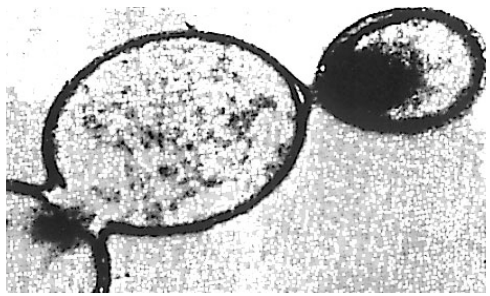
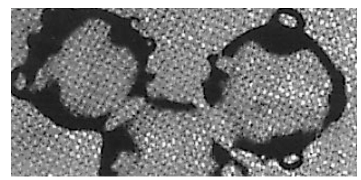
With Zn^{++}



Chain length (poly-G)

Self-assembly of macromolecules

- **Order and complexity result from self-assembly**
  - Proteinoid microspheres with internal structure
  - Multisphere assemblages
  - Membrane-like bilayers with "junctions"

- **Novel microenvironments allow:**
  - Selective permeability via lipid or proteinoid "membranes"
  - Novel (high) concentrations and enhancement of interactions
  - Chained reactions (concentrated products available as substrates)
  - Localized precipitation and organization (compartmentalization)
  - Entropy can *decrease* in subsystems (**not** a violation of the 2nd Law of Thermodynamics)

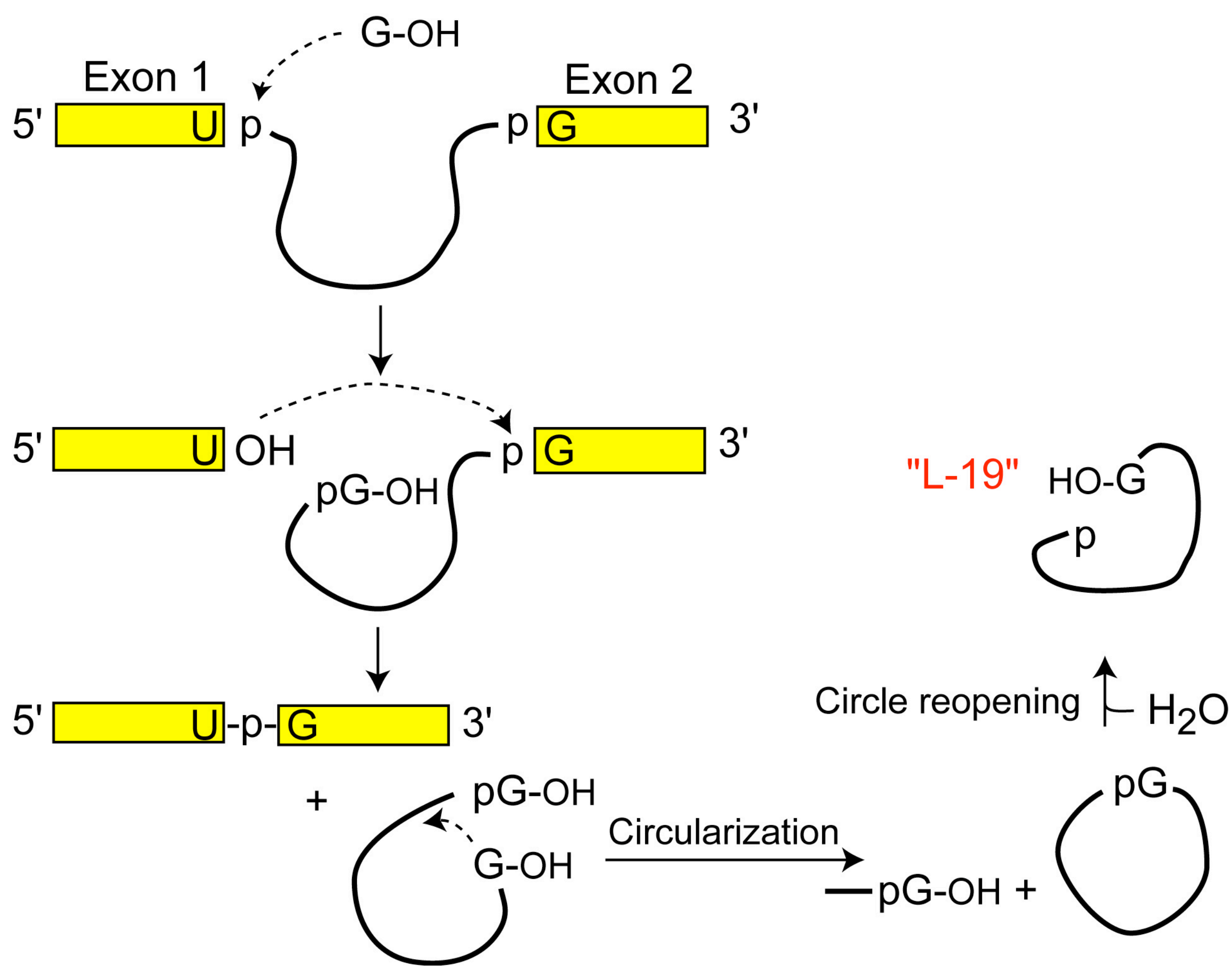
"RNA World" hypothesis

RNA preceded proteins & DNA (Orgel, Crick, Woese, 1960s)

- **RNA has the essential role in peptide assembly**
  - mRNA, tRNA, rRNA (which can promote translation even missing some proteins)
  - snRNAs (e.g., U1, U2, U4/6, U5)
- **RNA is required for DNA replication and synthesis**
  - Primer RNAs required for DNA replication
  - Telomerase RNA required for telomere synthesis
  - Deoxyribonucleotides are *derivatives* of ribonucleotides
  - Reverse transcriptase copies DNA from RNA template
- **RNAs are key cofactors**
  - 7S RNA (protein secretion), ATP, Coenzyme A
- **Some RNAs are catalytic**
  - Catalytic unit of RNase P (processes *E. coli* pre-$tRNA^{Tyr}$)
  - Self-splicing of introns
- **Some RNAs are regulatory**
  - miRNAs

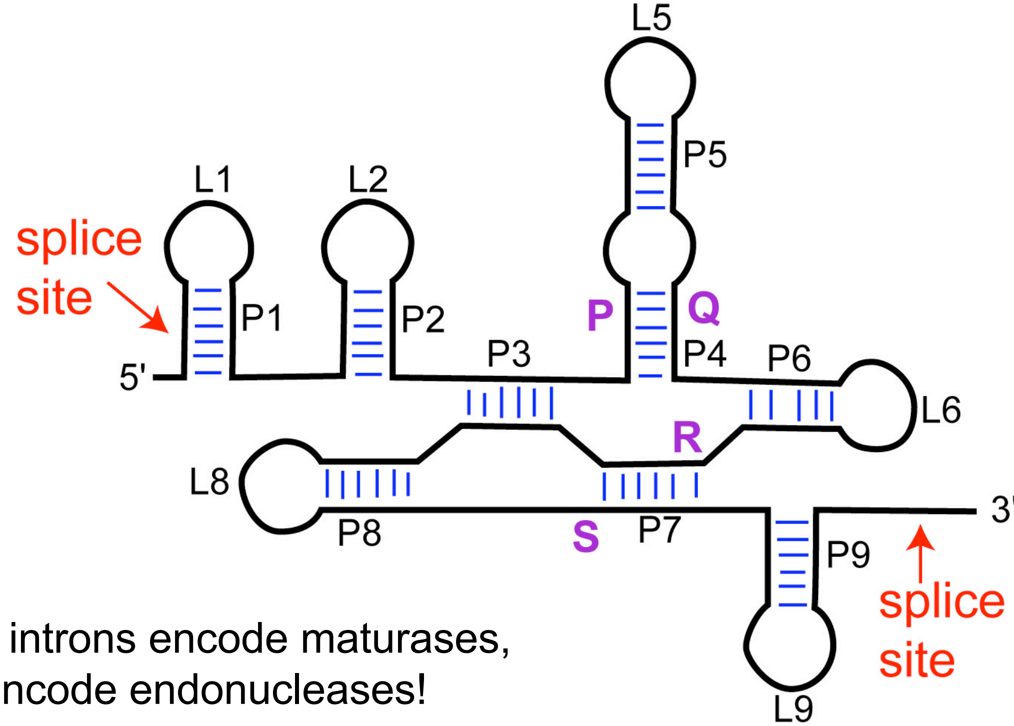
Group 1 self-splicing introns

- **Processing occurs as a series of transesterifications**



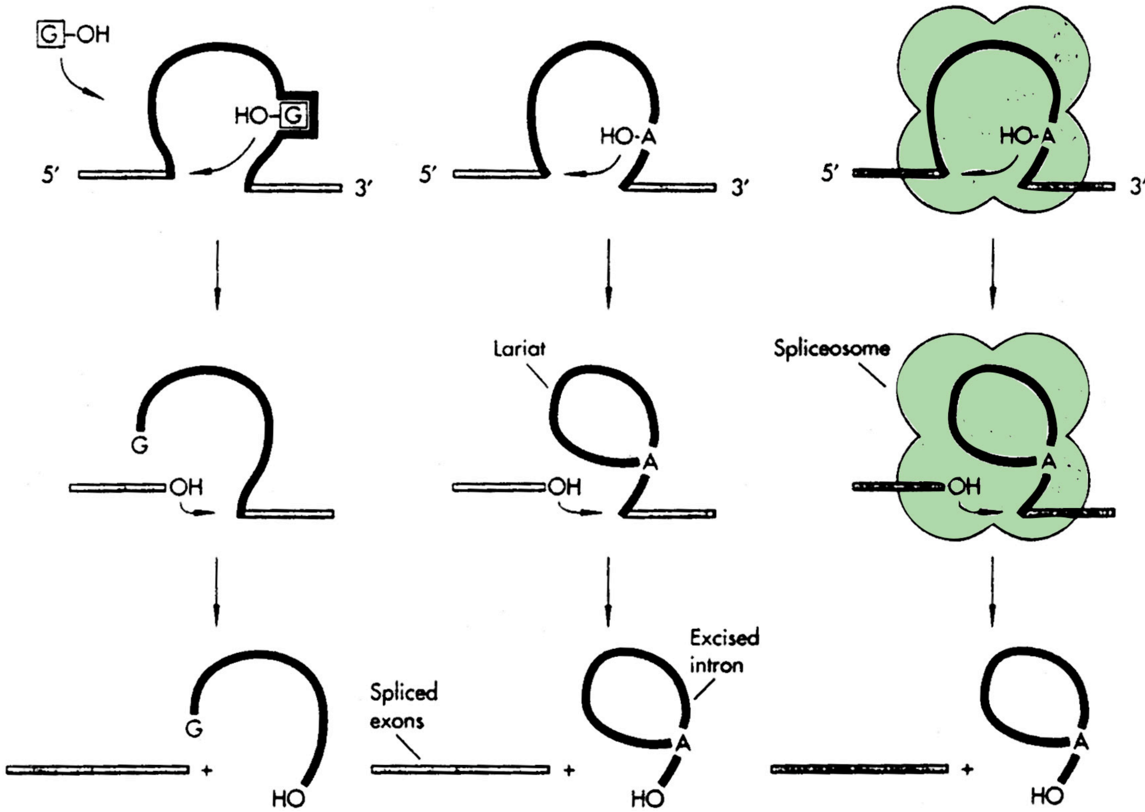
Conservation of Group 1 introns

- **Common ancestry allows comparative analysis of function**
  - Functionally important sequences/structures are often conserved
  - Group 1 intron structures are conserved:
    - in different genes
    - in different species (slime molds, mitochondria, chloroplasts, some bacteriophage



Some Gr. 1 introns encode maturases, and some encode endonucleases!

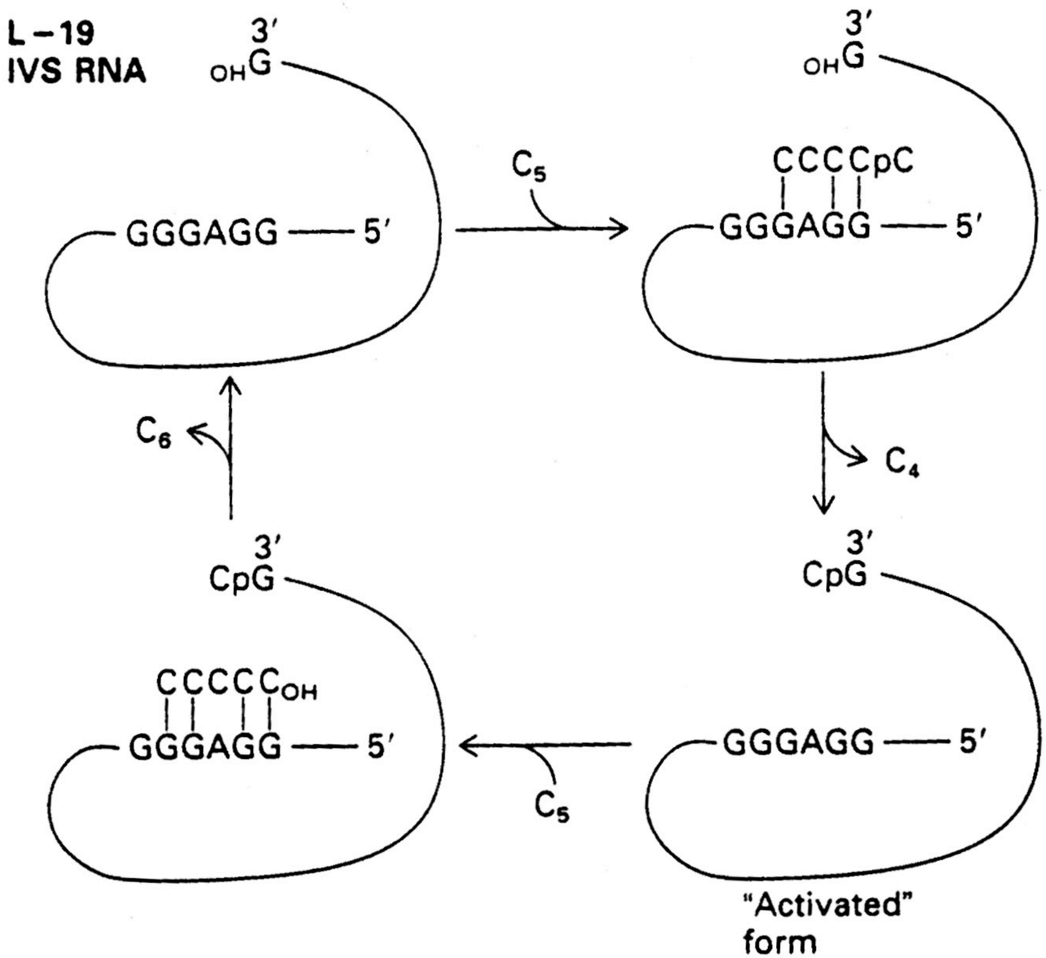
Other introns



- **Group 2 self-splicing a precursor to eukaryotic spliceosome?**
  - Note that Group 2 does not require a cofactor and makes a "lariat"
- **Trans-splicing also occurs** (all mRNAs of trypanosomes, many in *C. elegans*)
  - Important implications for "exon shuffling"

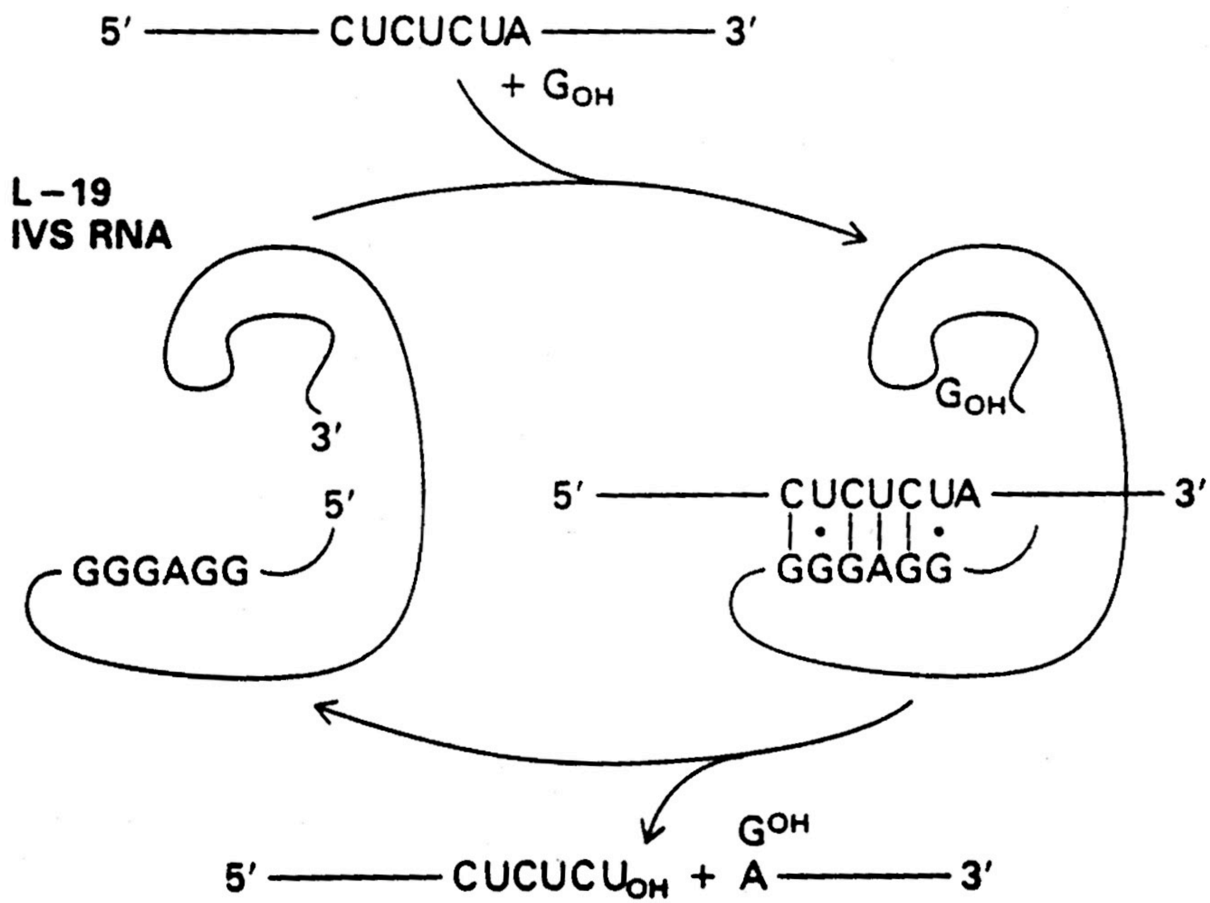
RNA as polymerase

- **"L-19" RNA can direct template-dependent extension**
  - Depends only on availability of (spontaneous) oligos

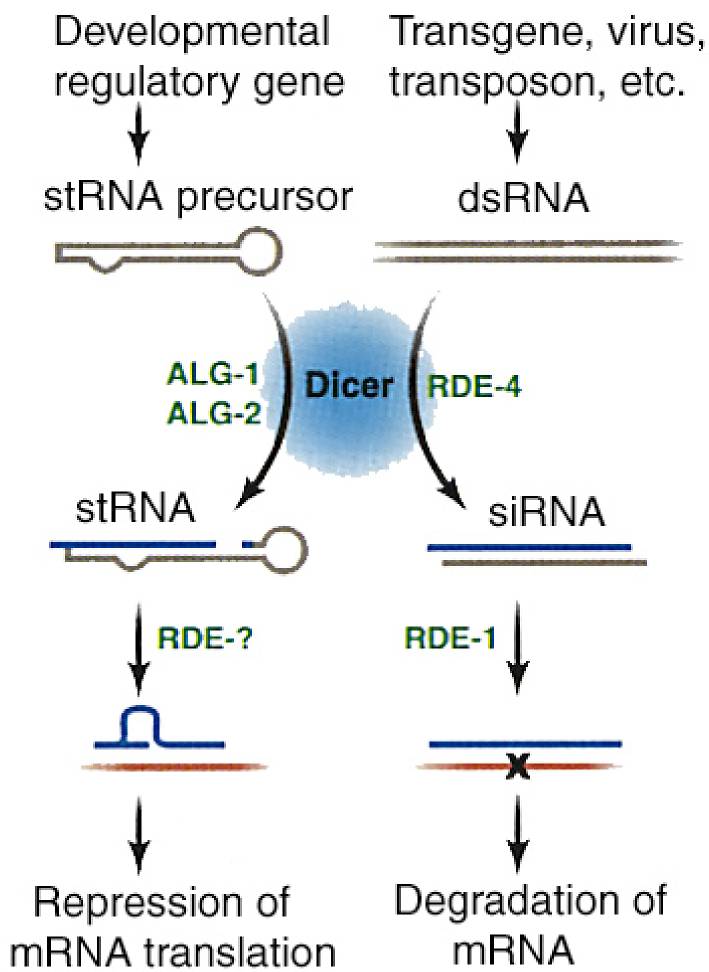


RNA as nuclease and ligase

- **Nuclease activity is similar to splicing, but site-specific**
- **Ligase activity is energetically the reverse**



RNA as regulator of gene expression



(miRNA)

Ambros, 2001

"RNA World" hypothesis

RNA preceded proteins & DNA (Orgel, Crick, Woese, 1960s)

- **RNA has the essential role in peptide assembly**
  - mRNA, tRNA, rRNA (which can promote translation even missing some proteins)
  - snRNAs (e.g., U1, U2, U4/6, U5)
- **RNA is required for DNA replication and synthesis**
  - Primer RNAs required for DNA replication
  - Telomerase RNA required for telomere synthesis
  - Deoxyribonucleotides are *derivatives* of ribonucleotides
  - Reverse transcriptase copies DNA from RNA template
- **RNAs are key cofactors**
  - 7S RNA (protein secretion), ATP, Coenzyme A
- **Some RNAs are catalytic**
  - Catalytic unit of RNase P (processes *E. coli* pre-$tRNA^{Tyr}$)
  - Self-splicing of introns
- **Some RNAs are regulatory**
  - miRNAs

RNA genome

- **Early RNA genome could probably self-replicate**
  - Template-dependent synthesis, ligation
- **Self-splicing would have allowed rapid evolution**
  - Different combinations of sequences and thus functions
- **Early protein synthesis was directed by RNAs**
  - tRNAs, rRNA
- **Early gene expression could be regulated by RNAs**
  - miRNAs
- **Evolution of the Genetic Code**
  - Once the codons began to be set up, and complexity of the code increased, it would be difficult to change (historical constraint)
  - Order to codon groupings suggests a possible stepwise adoption of codon assignments...

Genetic code

1st position	2nd position U	2nd position C	2nd position A	2nd position G	3rd position
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Hydrophobic Hydrophilic

Genetic code

- **The code is "degenerate"**
  - Third codon position is often completely **synonymous**
    - Perhaps the original machinery used only 1st & 2nd positions?
- **Second position determines hydrophobicity/hydrophilicity**
  - If pyrimidine, codon is hydrophobic; if purine, codon is hydrophilic
- **Easiest non-enzymatic RNA synthesis encodes protein order (β-sheets)**
  - RYR-YRY-RYR... : hydrophobic-hydrophilic-hydrophobic...
- **Heritable order is subject to Darwinian selection**
  - The stepwise process of selection will always lead to adaptations
  - Novel features arise by (duplication and) modification

Precellular evolution

- **"Prebiotic synthesis"**
  - **Hypothesis**: The molecules of life can be formed spontaneously under under "prebiotic conditions"
  - **Assumptions**: Prebiotic conditions, "uniformitarianism"
    - (NOT that humans should arise continuously from chimpanzees!)
  - **Predictions**
    - Amino acids and purines from prebiotic mixtures (Harold & Urey 1953)
    - Ribose from formaldehyde ("formose" reaction; Butlerow 1861)
    - Polymerization of "activated" nucleoside monomers without proteins
- **The "RNA World"**
  - **Hypothesis**: RNA preceded proteins and DNA as a primordial, information-bearing, catalytic molecule
  - **Assumptions**: Heredity, uniformitarianism
  - **Predictions**:
    - Conserved roles for RNA in fundamental machinery of life
    - At least some RNAs should demonstrate elementary catalytic activity
    - Role for RNA as precursor to DNA may be conserved

Next time

- **Cellular evolution**
  - **Molecular systematics**
  - **The evolution of plastids**
  - **Rooting the tree of life with gene duplications**
  - **The evolution of introns: recent or ancient?**
  - **Exon shuffling in the evolution of novel functions**