# Origin of Life: I
# Monomers to Polymers

# Synthesis of Monomers

Life arose early on Earth

1. Conditions
    1. Liquid Water
    2. Reducing or Neutral Atmosphere
    3. Energy Sources

2. Originally thought atmosphere was

$NH_3$, $CH_4$, $H_2O$, $H_2$

Miller-Urey Experiment

Now Believe $CO_2$, $H_2O$, $N_2$

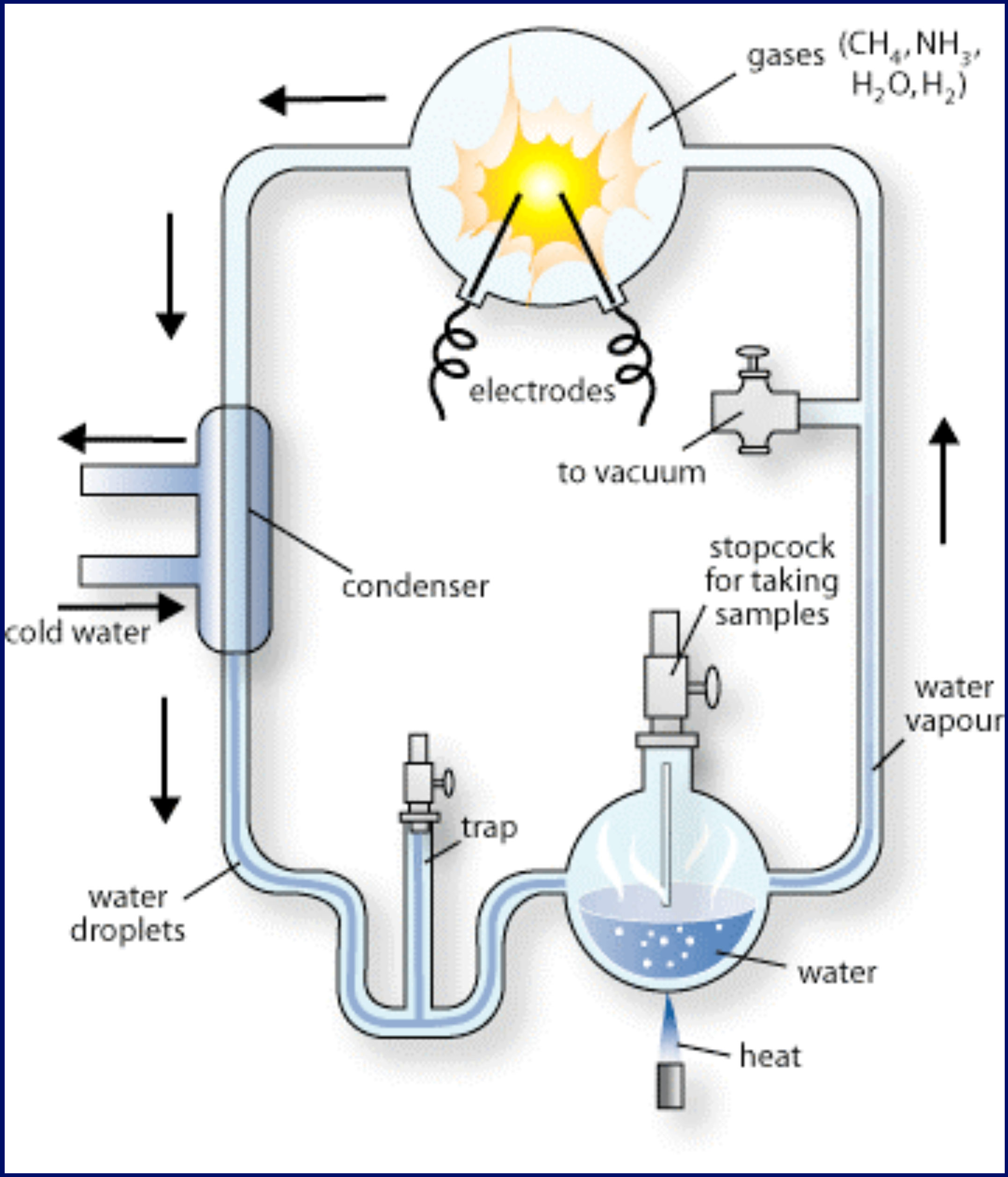
3. Energy Sources

Ultraviolet Light (No Ozone)

Lightning

Geothermal (Lava, Hot Springs, Vents, …)

# Miller-Urey Experiment



| COMPOUND | Relative Yield |
|---|---|
| **Glycine** | 270 |
| Sarcosine | 21 |
| **Alanine** | 145 |
| N-methylalanine | 4 |
| Beta-alanine | 64 |
| Alpha-amino-n-butyric acid | 21 |
| Alpha-aminoisobutyric acid | 0.4 |
| **Aspartic acid** | 2 |
| **Glutamic acid** | 2 |
| Iminodiacetic acid | 66 |
| Iminoacetic-propionic acid | 6 |
| Lactic acid | 133 |
| Formic acid | 1000 |
| Acetic acid | 64 |
| Propionic acid | 56 |
| Alpha-hydroxybutyric acid | 21 |
| Succinic acid | 17 |
| Urea | 8 |
| N-methyl urea | 6 |

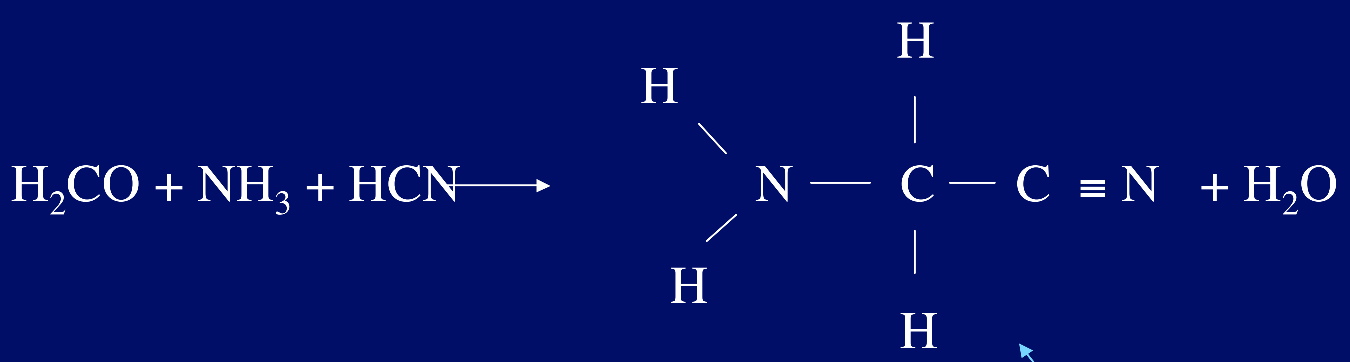
# How did Amino Acids form in Miller-Urey Experiment?

## Strecker Synthesis

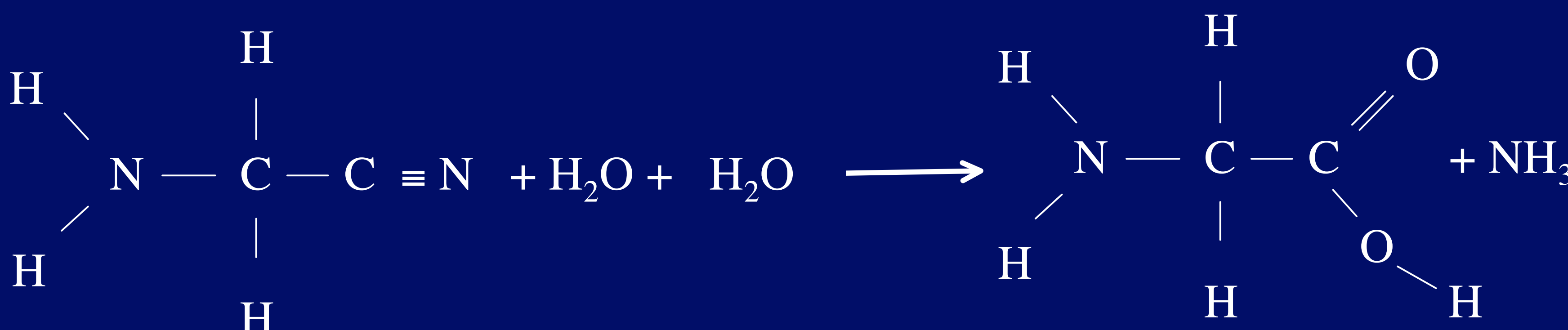
$CH_4$, $H_2$, $NH_3$ + Energy → $H_2CO$, $HCN$, $HC_3N$, Urea ($H_2NCONH_2$)

e.g. Glycine Synthesis

Reactive

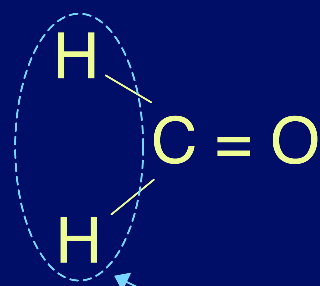



Aminoacetonitrile



glycine

$H_2CO$



form Aldehyde

More complex group - other aldehydes

↓

more complex amino acids

Lower yield if atmosphere was $N_2$, $CO_2$, $H_2O$

(If $H_2/CO_2 > 2$, get good yield)

# Problems with Miller-Urey

Atmosphere was $N_2$, $CO_2$, $H_2O$

$NH_3$, $CH_4$ would react $\longrightarrow$ $N_2$, $CO_2$

Try $N_2$, $CO_2$, $H_2O$ in Miller-Urey simulation

Only get trace amounts of glycine
Need $CH_4$ to get more complex amino acids

Need $H_2/CO_2 > 2$ to get much of any amino acid

## Miller-Urey with Cosmic Rays

A group in Japan has obtained good yields of amino acids from slightly reducing gases ($CO_2$, CO, $N_2$, $H_2O$)

When they used high energy protons (simulate cosmic rays)

Apparently not Strecker Synthesis (Low abundance of aminoacetonitrile)

# Building Blocks of Nucleic Acids

Not formed in Miller-Urey, but some intermediates were

1. Ribose Sugar:

   $5 \ H_2CO + \text{Heat} \longrightarrow H_{10}C_5O_5$

   [Clay Catalyst]

2. Bases

   a) Purines $5 \ HCN \rightarrow H_5C_5N_5$ (Adenine)

   b) Pyrimidines

   $HC_3N + \text{Urea} \longrightarrow H_5C_4N_3O$ (Cytosine)

   (1995) Cyanoacetaldehyde + Urea $\longrightarrow$ Uracil

3. Phosphate

Rock Erosion

Less understood than amino acids

**Other Possibilities for building blocks:**

Seafloor Vents

Interstellar Molecules

Comets

# Alternative Delivery

Molecular clouds - strongly reducing, contain many molecules used in Miller-Urey ($H_2$, $NH_3$, $H_2O$, $CH_4$) and intermediates (HCN, $H_2CO$, $HC_3N$) and possibly glycine

Problem:  These would not have survived in part of disk where Earth formed

But interstellar ices → comets

Evidence from similar molecules

(e.g.  $C_2H_2$, $CH_4$, HNC, …)

Clearly indicates interstellar chemistry

Cratering record on moon, …

⇒ heavy bombardment early in history

Comets and their debris could have brought large amounts of "organic" matter to Earth

(and maybe some of the oceans)

Some evidence for non-biological amino acids in layer deposited after asteroid impact 65 million years ago

# Sources of Organic Molecules

Quantitative comparison by Chyba & Sagan, Nature 1992, Vol. 355, p. 125

Currently, Earth accretes ~ $3.2 \times 10^6$ kg $y^{-1}$ from interplanetary dust particles (IDP)

~ 10% organic carbon $\Rightarrow 3.2 \times 10^5$ kg $y^{-1}$

~ $10^3$ kg $y^{-1}$ comets

~ 10 kg $y^{-1}$ meteorites

~ $10^3 \times$ more at $4.5 \times 10^9$ yr ago (?)

(cratering record)

UV + reducing atmosphere $2 \times 10^{11}$ kg $y^{-1}$

But if $H_2/CO \lesssim 0.1$ IDP's dominant source

# So if atmosphere very neutral, IDP's may have been important

Most of mass in IDP's in range of size ~ 100 μm

mass ~ $10^{-5}$ g

Complex structure - composites of smaller grains some carbon rich

Enhanced deuterium implies low T

Deuterium enhancement also found in interstellar molecules

May imply connection back to interstellar chemistry

2 kinds (mass ranges) can supply organic matter

1. Interplanetary dust particles ($m \lesssim 10^{-5}$ g)

2. Smaller meteorites ($m \lesssim 10^{8}$ g)

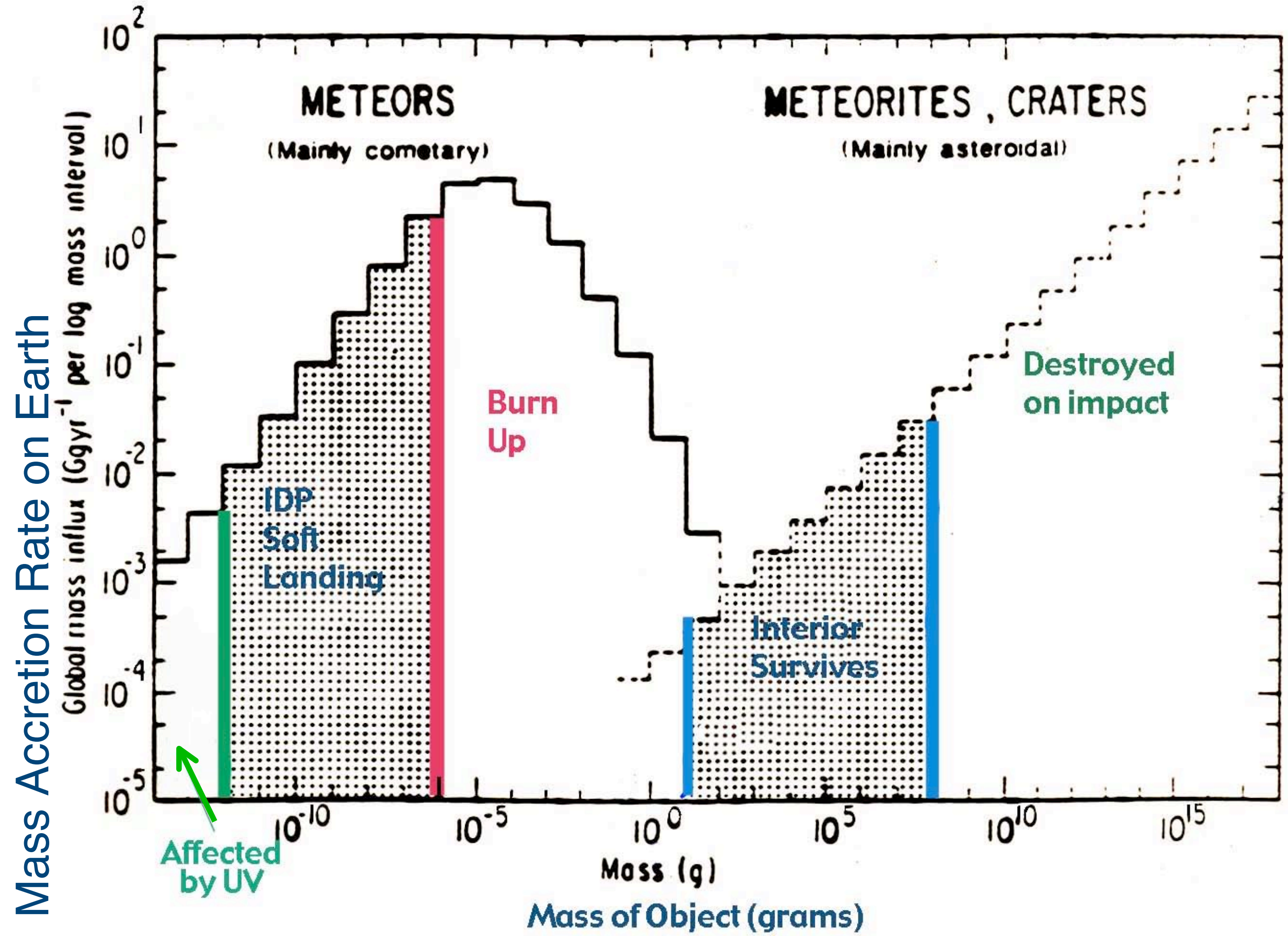



FIG. 1 Infall rate of meteoritic matter on Earth (adapted from ref. 5). Intervals where organic matter can survive passage through atmosphere are shaded. The curve on the right is based on the relation[5] $N = 0.54\, r^{-2.1}$ ($N$ = number of impacts per Myr, $r$ = radius in km), for an assumed density of 3 g cm$^{-3}$. The corresponding mass accretion rate (Gg yr$^{-1}$) between $r_1$ and $r_2$ is 15.83 $(r_2^{0.9} - r_1^{0.9})$.

E. Anders (1989) Nature, 342, 255

# Alternative Sites

Locally reducing environments

1. Ocean vents

   Sources of $CH_4$ and $H_2S$

   Current Vents have ecosystems based on energy from chemicals - not photosynthesis

   $H_2S \longrightarrow$ Bacteria $\longrightarrow$ Clams, Tube Worms

   Pre-biotic amino acid synthesis?

2. Inside Earth

Many bacteria now known to live deep (~ 2 miles) in Earth. Energy from chemicals, adapted to high temperature.

Genetic makeup is very ancient.

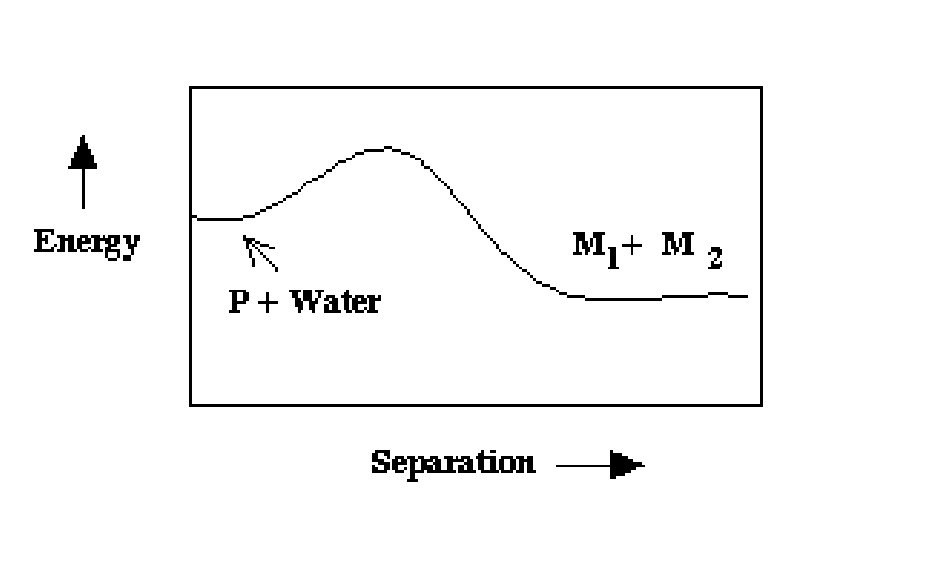
3. Hot Springs

Bacteria may be important in precipitating minerals. Adapted to high T and ancient.

# Synthesis of Polymers

$M_1 + M_2 \longrightarrow P + H_2O$

$\longleftarrow$ more likely in liquid $H_2O$



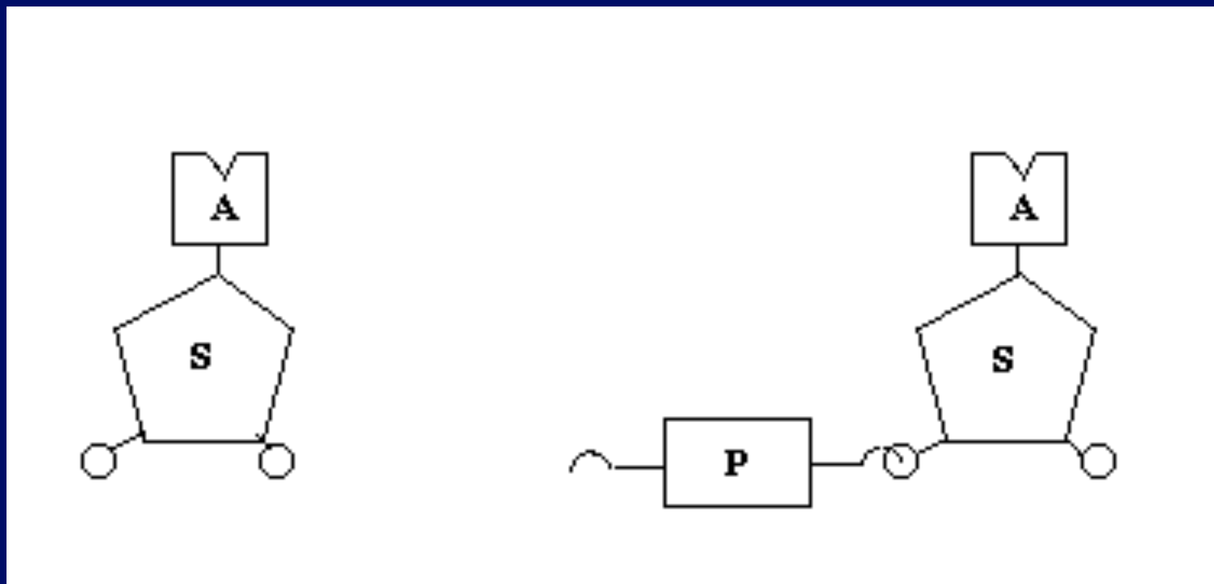

Solutions Remove $H_2O$ (Drying, Heat)

Sydney Fox $\longrightarrow$ Proteinoids

Energy Releasing Reactions ($H_2NCN$ or $HC_3N$)

Catalysts: Clays

Problem is worse for Nucleic acids because more complex




nucleoside

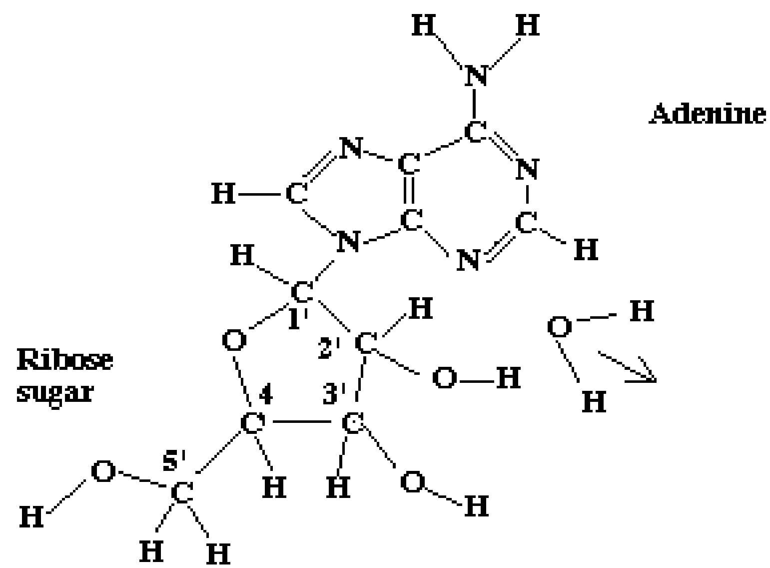
nucleotide

↑

Monomers of nucleic acids

# Synthesis of Adenosine

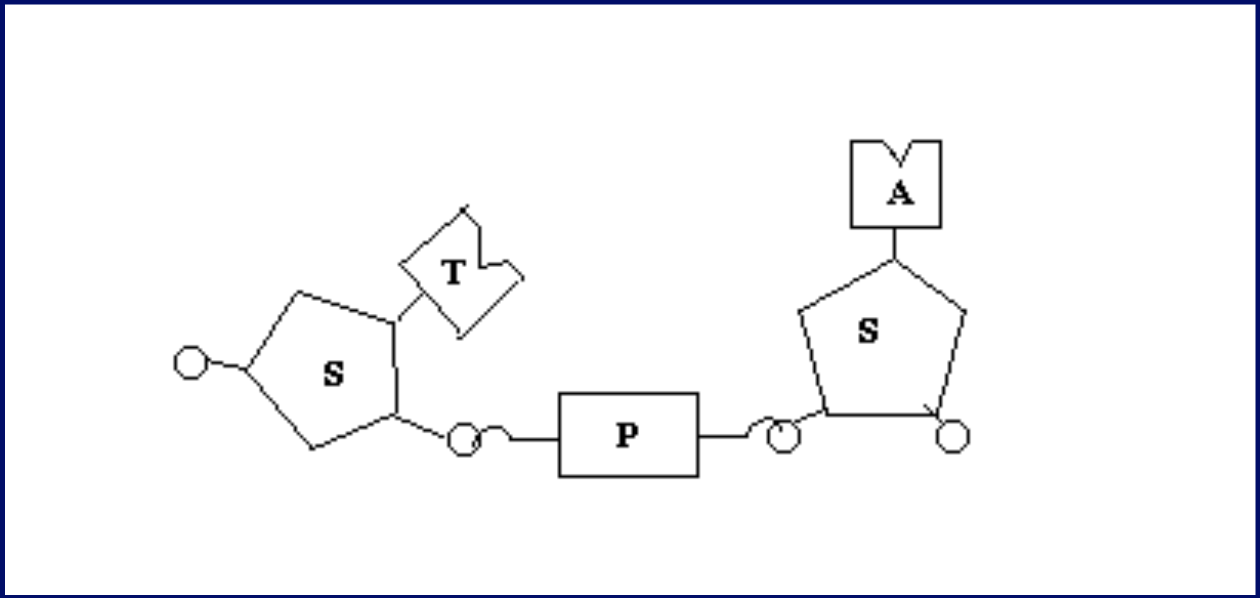
Base on 1’ Carbon (Why?)




Adenine + ribose sugar ⟶ adenosine + $H_2O$

Also phosphates 3’ & 5’ carbons

Otherwise, you are likely to get Misalignment




Leslie Orgel has had some success in getting high percentage of correct linkages, in presence of Zinc ions.

## Experimental Results

Sugar + base + heat  yield some nucleosides

Activated nucleosides + phosphoric acid + $Zn^{+2}$

Get polymers up to 50 nucleotides in length

linkages (mostly) correct

# The Odds

- We need to get an “interesting” polymer
  - Enzyme
  - Self replicator
- Properties of polymer depend on
  - Order in which monomers combine
- If we combine monomers at random,
  - How likely to get something interesting?

# Statistics

- What unlikely event happened just before the Super Bowl?

# Statistics of an unlikely event

Random reactions in primordial soup?

Unlikely event versus many trials

Probability Primer: Consider tossing 10 coins

Probability of all heads = product of prob.

$$P = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) \cdots \left(\frac{1}{2}\right)^{10} = \frac{1}{1024}$$

Probability of getting 10 amino acids $\longrightarrow$ protein

Chosen from 20 in a particular <u>order</u>

$$\left(\frac{1}{20}\right)^{10} = \frac{1}{1 \times 10^{13}}$$

Based on discussion by R. Shapiro

## But if you try many times, the chance of success is higher

$$P(r) = \frac{n!}{r!\,(n - r)!}\; p^r\,(1 - p)^{n-r}$$

r = # of successes    p = prob. of success on each trial

n = # of trials

n! = n (n-1) (n-2) … 1

e.g. make $n = \frac{1}{p}$    (flip all 10 coins 1024 times)

$$P(1) = \frac{n!}{1!\,(n-1)!}\left(\frac{1}{n}\right)\left(1-\frac{1}{n}\right)^{n-1} = 0.37$$

Chance of one or more successes = 0.63

For reasonable chance need $n \sim \frac{1}{p}$

## How many do we have to get right?

1. How many atoms?

| | |
|---|---|
| Lipids | $10^2 - 10^3$ |
| Enzymes, RNA | $10^3 - 10^5$ |
| Bacterial DNA | $10^8 - 10^9$ |
| Bacterium | $10^{11} - 10^{12}$ |
| Human Being | $10^{27} - 10^{28}$ |

If we choose from H,C, N, O (ignore S,P)

probability of right choice 1/4

So for enzyme: $(\frac{1}{4})^{10^3} \sim 10^{-600}$

# of trials:  R. Shapiro computes

$N = 2.5 \times 10^{51}$ (surely an overestimate)

$n << \frac{1}{p}$ for simple enzyme

2. What if we start with amino acids?
   Need ~ $10^{13}$ trials to get 10 amino acid protein

To get 200 amino acids in right order

$$\left(\frac{1}{20}\right)^{200} = 10^{-260}$$ Hopeless!

Need something besides random combinations

Selection (Natural?)

# Improving the Odds

Many proteins composed of interchangeable segments (Domains)

10 to 250 amino acids

One domain found in ~ 70 different proteins

Intermediate building blocks?

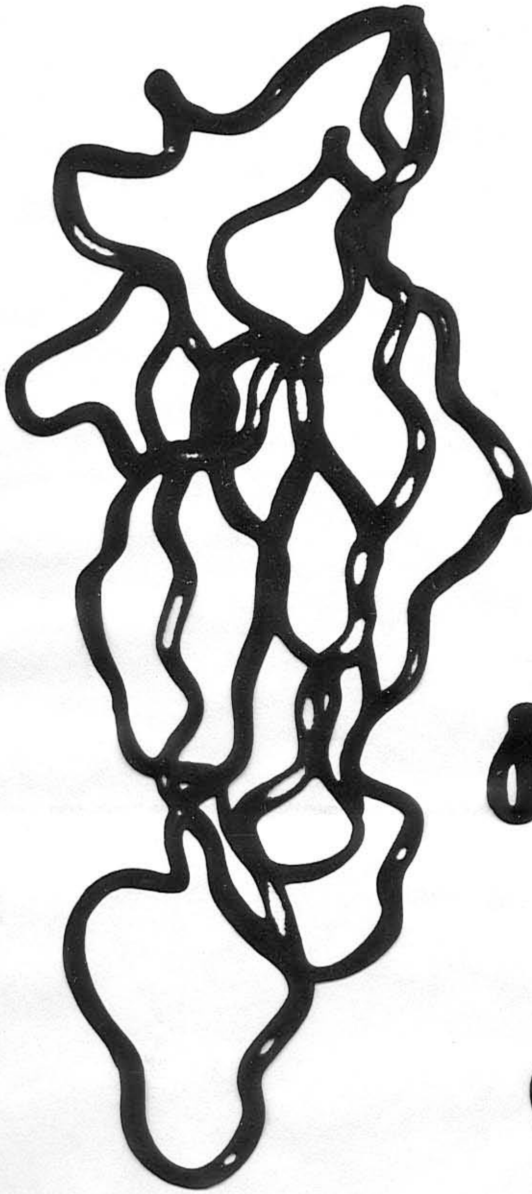
If so, may only need to get enough amino acids in right order for a domain e.g. 18 amino acid domain

$$P = \left(\frac{1}{20}\right)^{18} = 10^{-23}$$

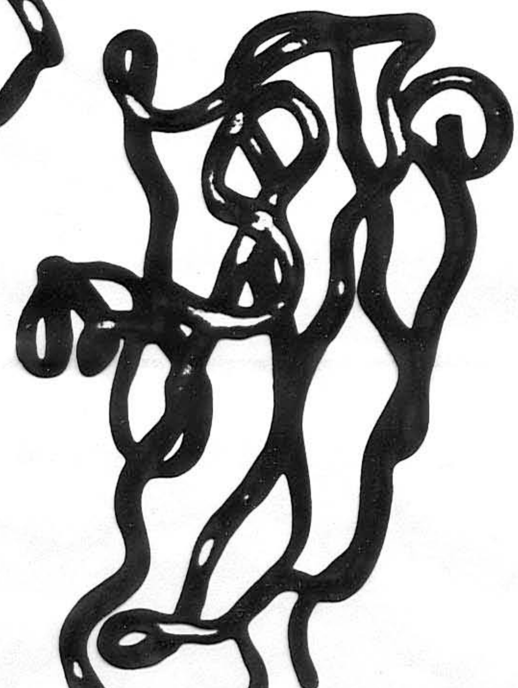
Also, many variations in amino acids don’t destroy function

and many different sequences may be interesting

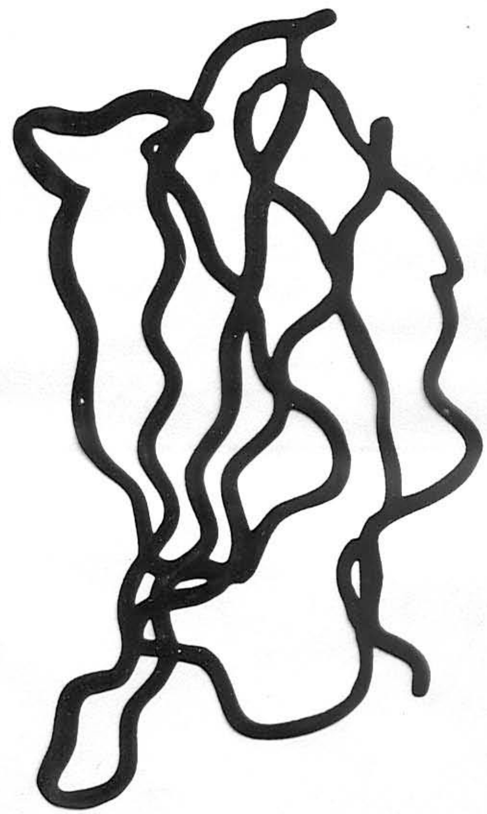
**GENETICALLY MOBILE MODULES** have been found in many proteins. Two types of these modules, or domains, are shown here. The Fn3 and the GHR domains are examples of fibronectin type III modules. The PapD and CD2 domains are immunoglobulin domains. These modules are linear sequences of amino acids that can fold themselves into consistent, recognizable structures with specific biochemical properties. During evolution, these domains can move as discrete units from one protein to another, which helps new types of proteins to appear.



PapD DOMAIN 1



GHR DOMAIN 1



Fn3 (10TH DOMAIN)

SCIENTIFIC AMERICAN *October* 1993

Scientific American Doolittle & Bork
Oct. 1993, pg. 50

Proteins made of domains, assembled in various ways
10-250 amino acids for ones containing disulfide bonds

18 - 100 for those without

Of all amino acids available

$\left(\frac{1}{20}\right)^{40}$ or $\left(\frac{1}{20}\right)^{18}$

$\log_{10} = 40 \log 20$     $-18 \log 20$

$= -52$     $= -23.4$

so $10^{-52}$     $10^{-23.4}$

# Interesting fact on how the improbable happens

1st winner of Texas Lotto lottery

Picked all 6 numbers correctly in the <u>same</u> order as they were drawn.

Each number runs from 1 to 50, and once chosen, cannot be repeated (balls are taken from a box).

So the odds against getting them in order is

$$\left(\frac{1}{50}\right)\left(\frac{1}{49}\right)\left(\frac{1}{48}\right)\left(\frac{1}{47}\right)\left(\frac{1}{46}\right)\left(\frac{1}{45}\right) = \frac{1}{11{,}441{,}304{,}000}$$

You don't need to get them in the same order to win - odds against winning include any combination, so 1 in 16 million