The RNA world

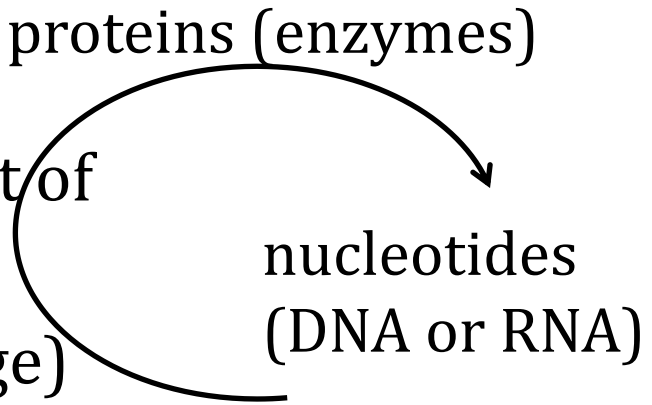
Definitions of life

Evidence for RNA world

Problems with RNA world
Alternatives (maybe there was no RNA world)

Andrew Torda, April 2013

Today versus history

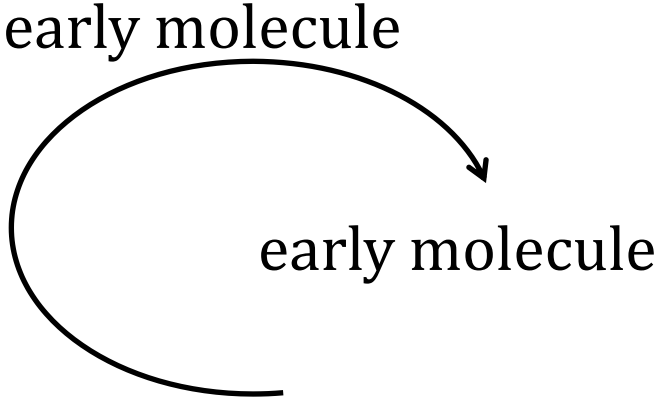


Picture today

- implies simultaneous development of
  - proteins (copying)
  - nucleotides (information storage)

Suggestion

- one molecule
  - self copying
    - possibilities
    1. protein like
    2. nucleotide like
    3. something else



This is templated

- later remove this requirement

What is life ? Practical – not philosphical

Practical – not philosophical

- people, trees, …
- bacteria
- viruses ?
- infectious DNA / RNA ?

Some concepts

- life consumes energy – better formulated
- life avoids equilibrium, needs energy, consumes entropy

Equilibrium

Reaction A + B ↔ C + D $\Delta G = RT \ln \frac{[C][D]}{[A][B]}$

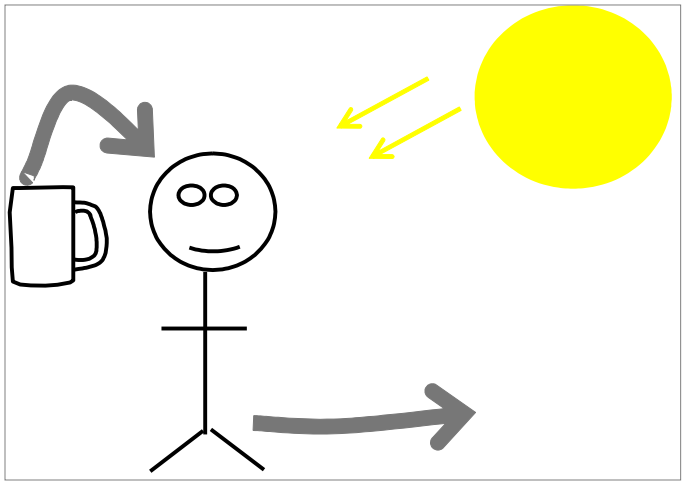
Decay A ↔ B + C, then $\Delta G = RT \ln \frac{[B][C]}{[A]}$

In a closed system, if $\ln \frac{[B][C]}{[A]} = 0$ you are dead

Consequence

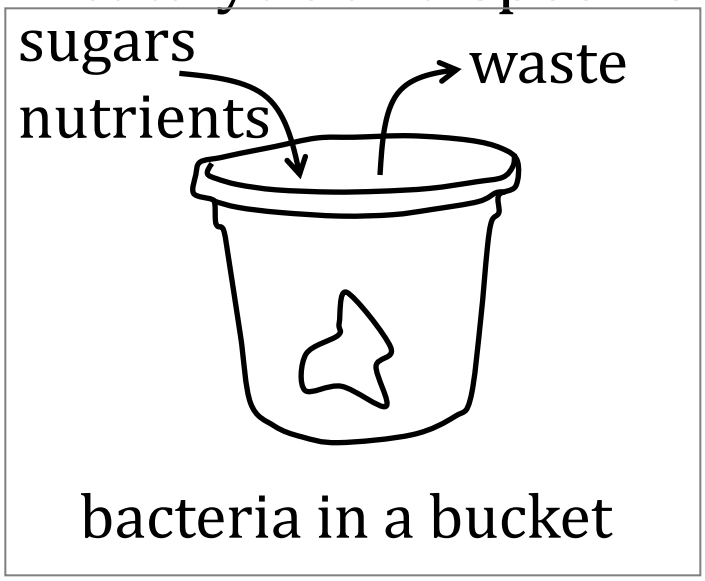
- life is in "steady state"

Steady state systems

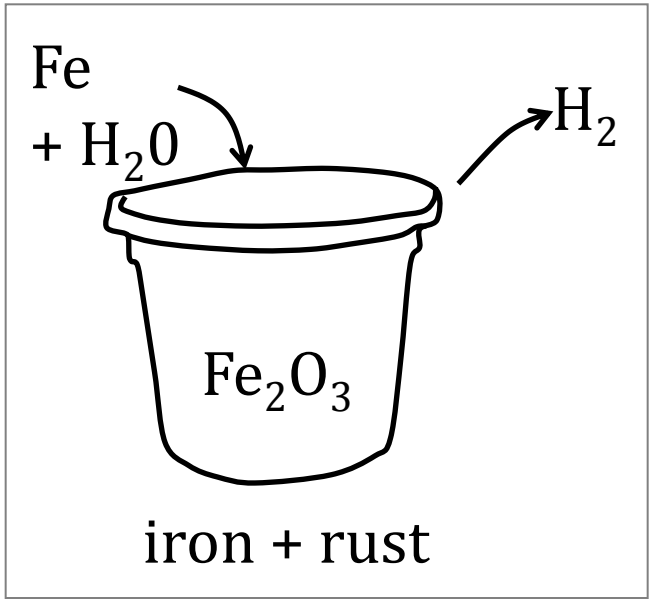


Input of energy

- maintenance of order
- grows
- catalytic and specific



bacteria in a bucket



iron + rust

- bacteria and rust
  - grow, eat nutrients, catalyse their own copying

Rust

Why is rust not life

- low information
- no ability to change and evolve

information / entropy

Entropy is easy to define

- N_{states} equal probability $S = k \ln N_{states}$

- or with different probabilities $S = -k \sum_{i=1}^{N_{states}} p_i \ln p_i$

- life has information, but what is it ?

Information

- pretend a genome is a string amongst possible genomes
- *E. coli* ?

  ≈ 5 million base pairs (5×10^6) ... $4^{5\times10^6} \approx 10^{3000000}$
- how many states could e. coli's genome have ?
  - of these possibilities, very few are used
  - "information" per genome is big

- genome of rust ? information in rust ?
  - alphabet is 1 ? length is 1 ?

Claim

- evolution is information increase via selection

Complexity

Smallest genomes

- viruses – few proteins – parasitic

Free living ?

- a few hundred proteins

Is there a minimum complexity for life ?

- no answer, but rust is very simple

Life

Rust can

- catalyse the production of rust, does not adapt

"life" can

- general copying machinery
- copy sequence_1 or sequence_2
- templated copying

This flexibility necessary for evolution

Summary of life

- not at equilibrium / consuming energy
- catalytic
- creating information
- copying with possibility of change / selection

- minimum complexity ? no evidence yet

RNA world properties

- replication of RNA (directed / templated)
- Watson-Crick base pairing (not at start)
- no protein catalysis

  - did it exist ?

Why believe in an RNA world ?

1. both phenotype and genotype
2. information
3. roles of nucleotides
4. Selex
5. biosynthesis
6. ribosome

In turn..

Why believe in an RNA world ?

You have to carry information between generations and perform chemistry..

1. Information
- proteins rarely code for other proteins

2. Both phenotype and genotype
- simplicity (parsimony) – one type of molecule

Why believe in an RNA world ?

3. Roles of nucleotides

All AMP/ADP/ATP chemistry (or GMP)

- basically nucleotides

Lots of classic biochemistry

- CoA (AMP-phosphopantetheine)
- NAD (nicotinamide adenine dinucleotide)

Why believe in an RNA world ?

4. Selex (topic in biochem lectures, later this course)
select for desired activity from random molecules

- from a soup of RNA molecules, one can select desired activities
- activities were there

- start of life – just a big selection experiment

Why believe in an RNA world ?

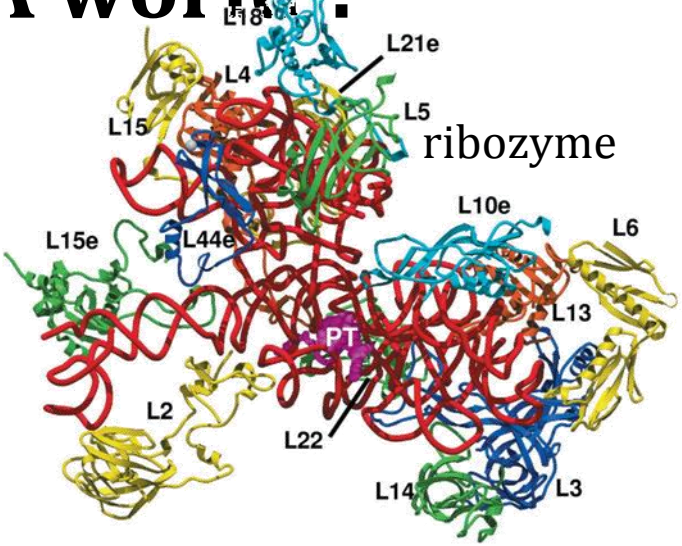
5. Biosynthesis

- much machinery devoted to RNA biosynthesis
  many enzymatic steps
- DNA is just a modification afterwards
- looks as if RNA is the older molecule

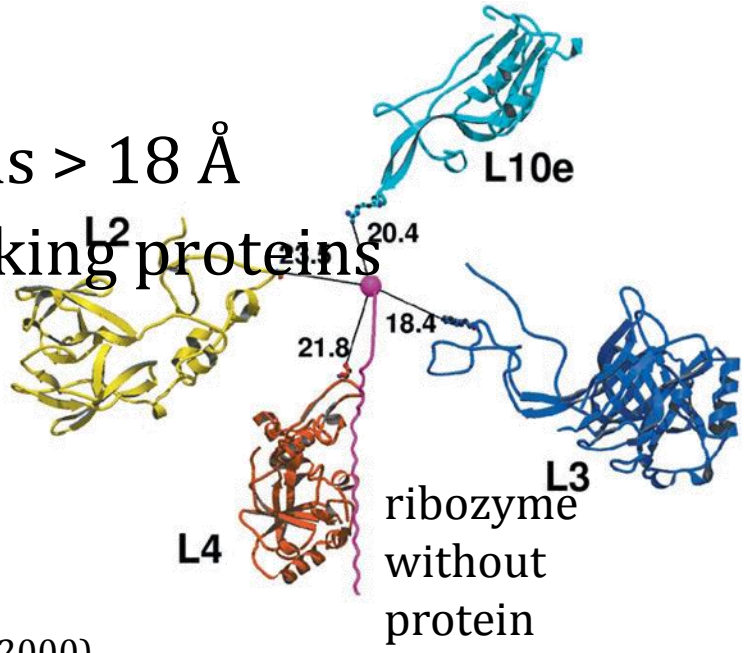
Why believe in an RNA world ?

6. ribosome

- incredibly conserved
- part of ribosome near active site



- remove all the RNA
- the nearest protein to active site is > 18 Å
- the fundamental operation of making proteins from a template –
  carried out by a ribozyme



Nissen P, Hansen J, Ban N, Moore PB, Steitz TA., Science 289, 920-930 (2000)

RNA World – requirements

Source of basic requirements

- ribose
- bases (A, C, G, U + more T, I, X, …)

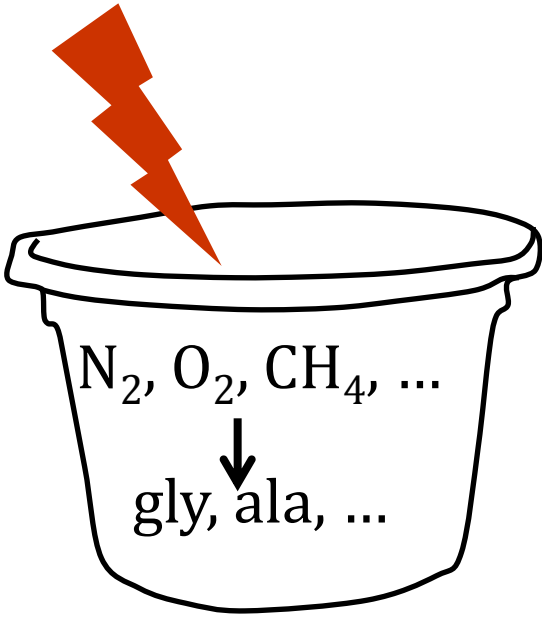
Vague source

- Miller experiments from 1950's

Can one make nucleosides ? nucleotides ?

- polynucleotides ?

Lots of problems…



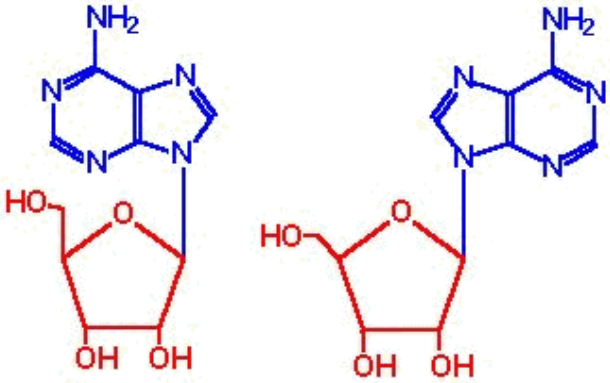
Specificity - sugars

Make sugar in lab

- condensation from smaller molecules
- result ?
  - mixture of 5 member sugars (ribose, pyranose, …)
  - ribose is not dominant

Enantiomers, isomers, ..

- details of linkages different, but only one is used in modern world
  - syn / anti, L / D



syn / anti

How to make nucleotides ?

N-sugar + xxx-PO_4 →N-sugar-PO_4 (+H_20 + xxx)

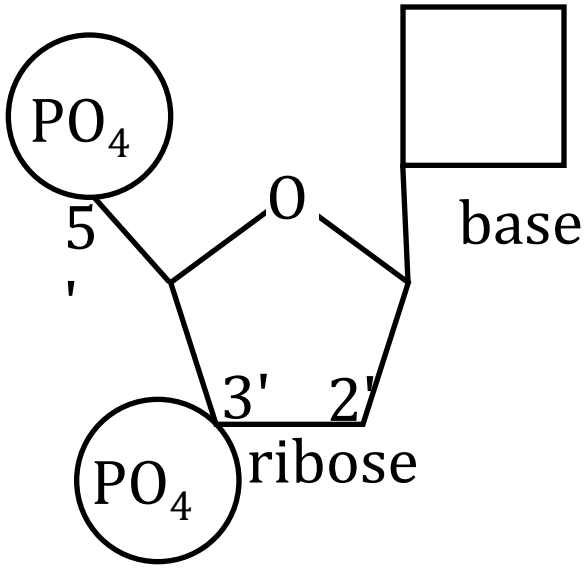
ribozymes have been made for related reactions

- quite plausible
  - no really good candidates yet

Abiotic ?

- many examples of catalysis exist
  - Pb^{2+}, BO_3^{3-}, …

Joining monomers (problems)



Modern chemistry always 5' to 3'

Nucleotides (NMP)

- 3 reactive groups
  - 5' PO_4, 3' OH, 2' OH

Soup of 5' NMPs and condense

- mixture of
  - 5', 5' pyrophosphate
  - 2', 5' PO_4 diester
  - 3', 5' desired diester

RNA replicase

- One model – we have one replicase
- Basic requirement – replicase should
  - act on itself (or similar copies)
  - should produce
    - itself or
    - complementary copies

Length constraints

- define fidelity q = probability that one residue is correctly added
- probability of copying chain length n correctly = q^n
- no mistakes – no evolution

q	n	perfect copies
0.9	4	0.66
0.9	10	0.35
0.95	10	0.65
0.95	20	0.36

Replicase Quality

- Is there are magic q ?
- Must we wait for some chemicals with correct q ?

- No ! Evolution helps

Joyce / Orgel – first replicase

How likely are we to take a random soup of nucleotides

- ribozyme of 40 bases
- $q = 0.9$
  - not very likely, but if
- a replicase starts
  - copies related molecules better than unrelated
- if it copies better / faster it will be selected for and evolve
- could this happen ?
  - copying by other catalysts using RNA as template

Alternative Genetic Systems

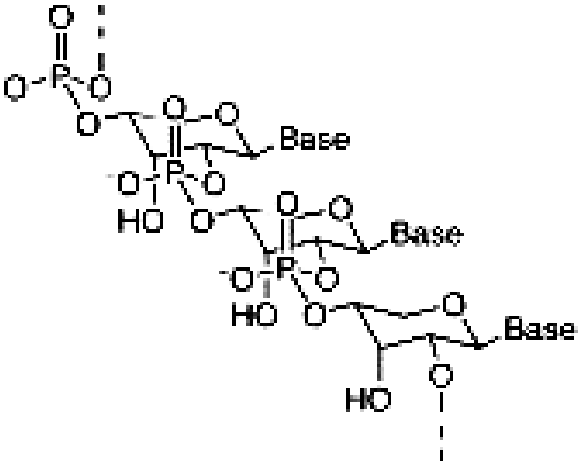
Must we start with RNA ?

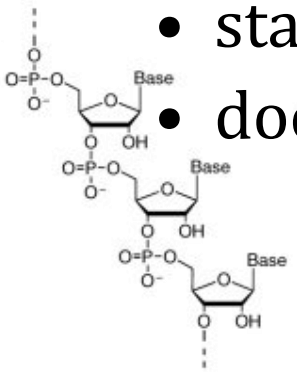
If not, bias is towards a system

- can pair specifically with RNA sequences
  - XYZW pairs to ACGU so we can have template directed RNA synthesis
- should form an open (helical) structure

Examples

- replace ribose with pyranose (p-RNA)
  - stable, helical
  - does NOT form paired dimers with RNA





from Joyce, GF & Orgel, LE in The RNA World, (eds Gesteland, RF, Cech, TR, Atkins GF) Cold Spring Harbor Lab Press 2006

different sugars RNA

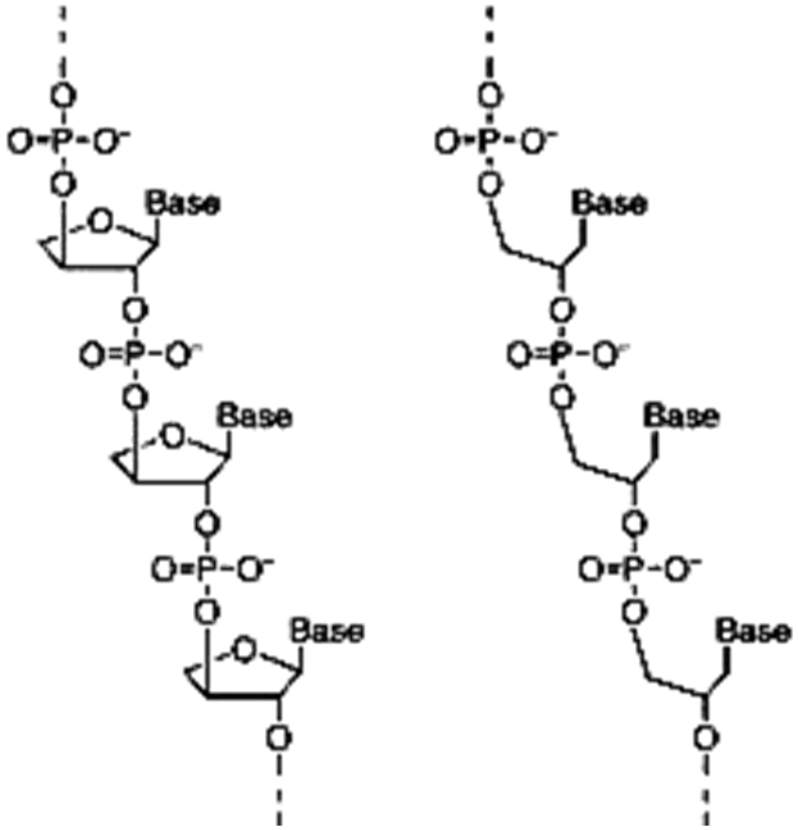
Use threose (left) TNA

- forms stable double helix
- threose may be easier to make

Use glycol (right) GNA

- also forms double helices

- Other examples possible



Main point:

- There may have been something before RNA

Complete change of philosophy

maybe we do not need an RNA world

Do we need this general templating ?

So far – search for general replicase, polymerase

- Can one build a living system from less general components ?

Examples

- peptides made without ribosomes
  - antamanide
  - glutathione
- the "information" is stored in enzyme structures

reference: Kauffman, SA, The Origins of Order, Oxford University Press, NY 1993

Requirements for RNA/DNA/Protein world

1. RNA can catalyse formation and cleavage of internucleotide bonds
2. abiotic formation of the monomers
3. solutions must be concentrated (small volume)
4. anabolic flux (making larger polymers)
5. catalytic closure
   formation of each member of set is catalysed by some other member

We could apply these rules to proteins or nucleotides

- change nature of monomer
- consider the first four problems

Some prerequisites are easy

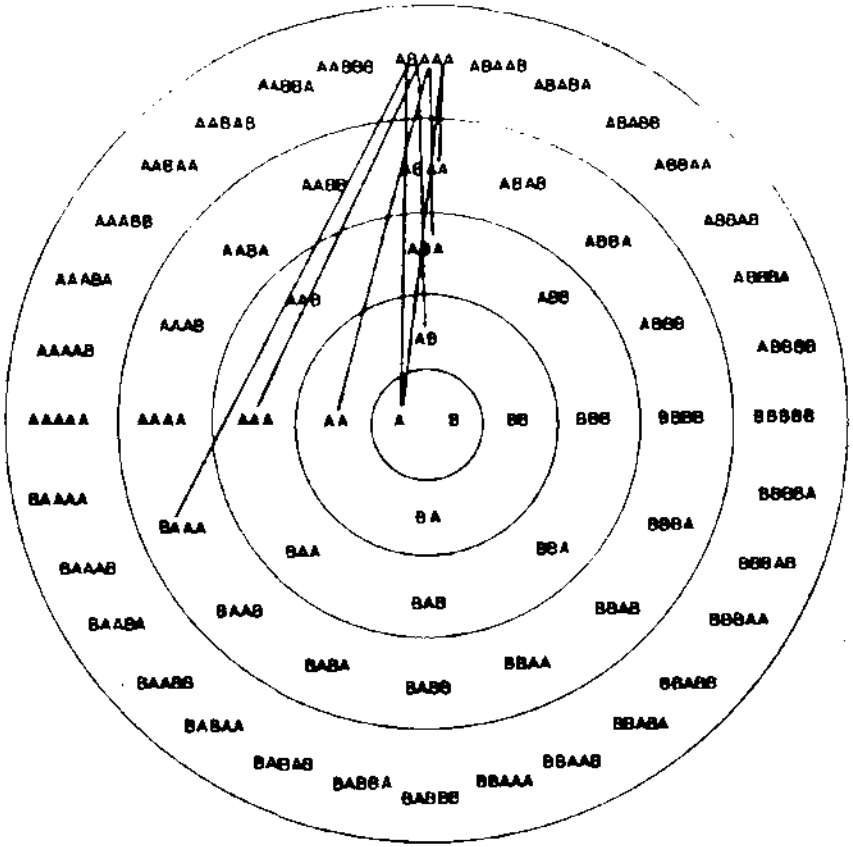
1. proteolytic enzymes or ribozymes
2. tolerate a very imperfect soup of molecules, complex peptides or mixed 3', 5' + 2', 5' nucleotides
3. confinement – drops, minerals, agglomerations
4. most reactions are $A + B \leftrightarrow AB + H_20$
   - removing water drives equilibrium to right
5. catalytic closure – not by simple templating
   - radical philosophy – exam questions, take notes

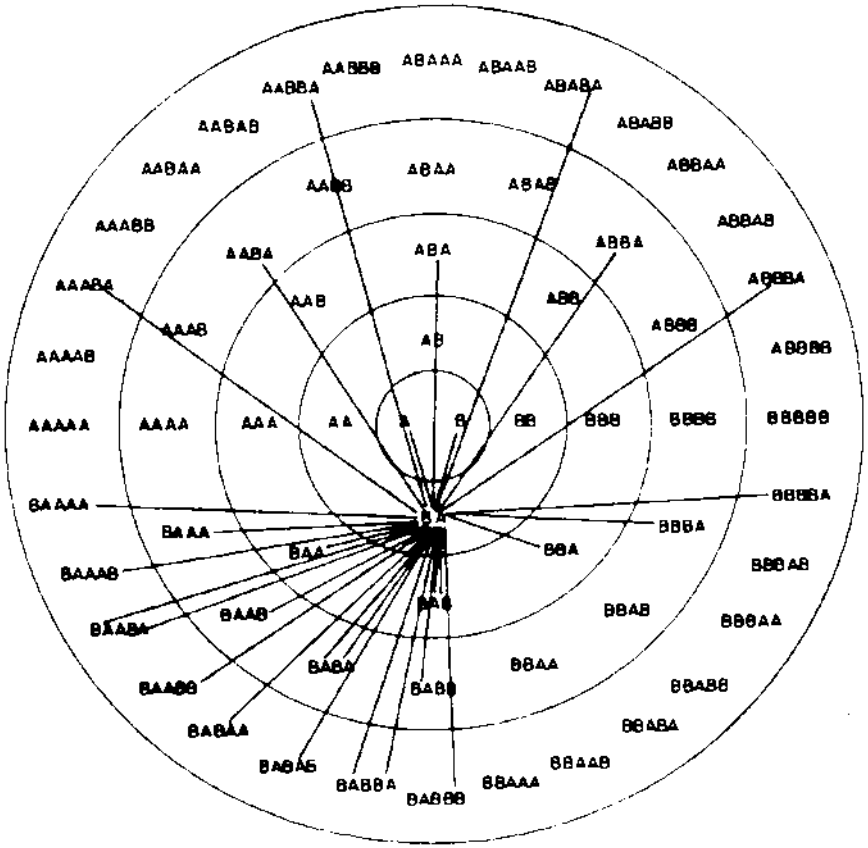
Catalytic closure

Imagine a soup of polymers with conversions

- cleavage or ligation ABCDE↔ ABC + DE

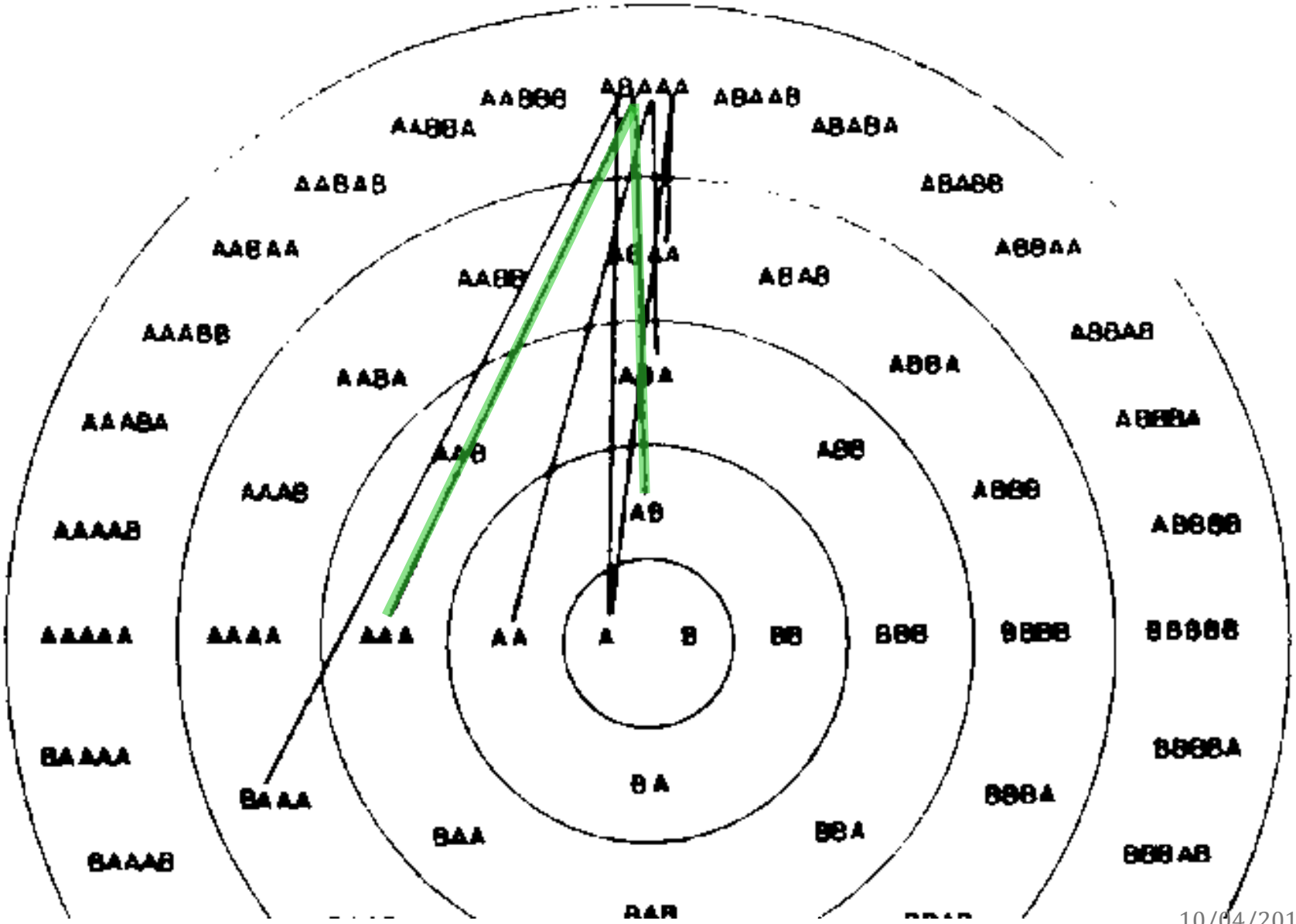
How many ways can we form a 5-mer ? or 2-mer ?





Kauffman, SA, The Origins of Order, Oxford University Press, NY 1993

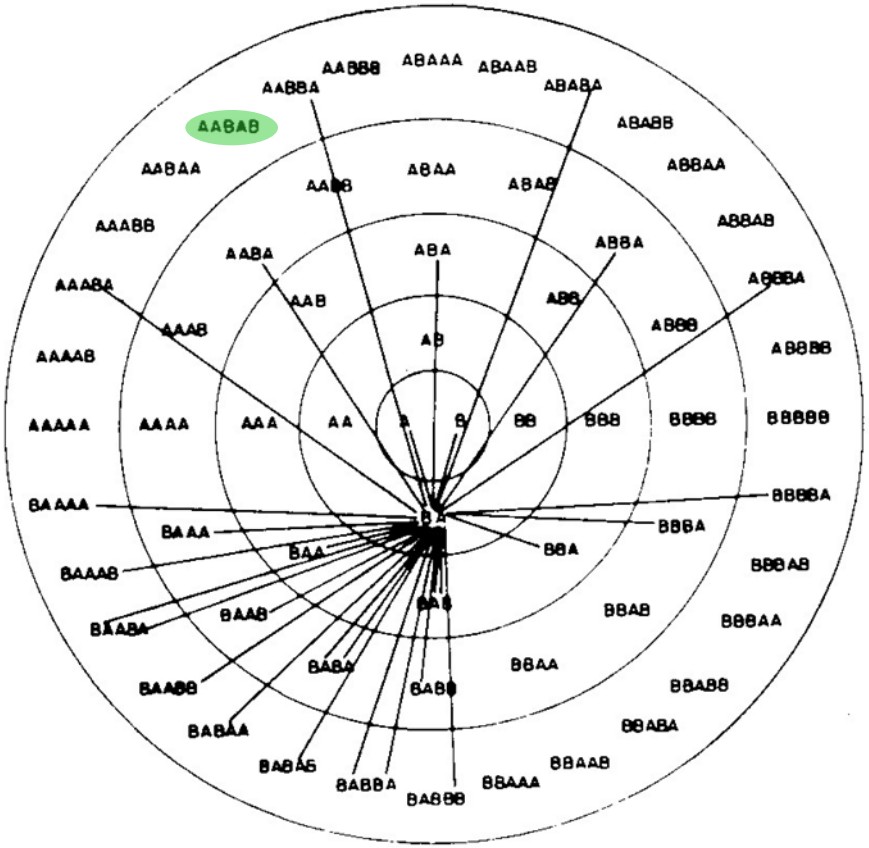
AB + AAA form ABAAA



Catalytic subset

Within set of polymers some are enzymatic for joining / breaking units

- for RNA 4×4=16 X-Y types
- pick a polymer
- with probability *p* pick a react it catalyses
- imagine green sequence catalyses all AB bonds
  - leads to huge number of edges
- go to next sequence, maybe assign a reaction

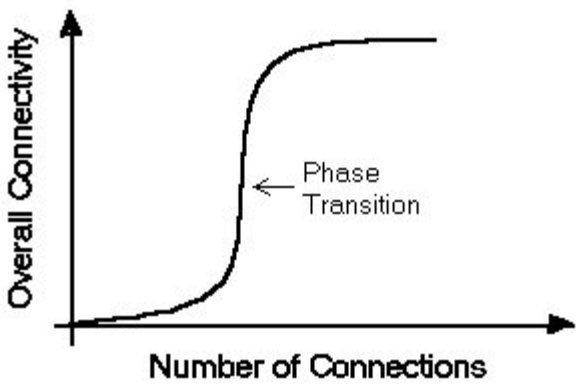


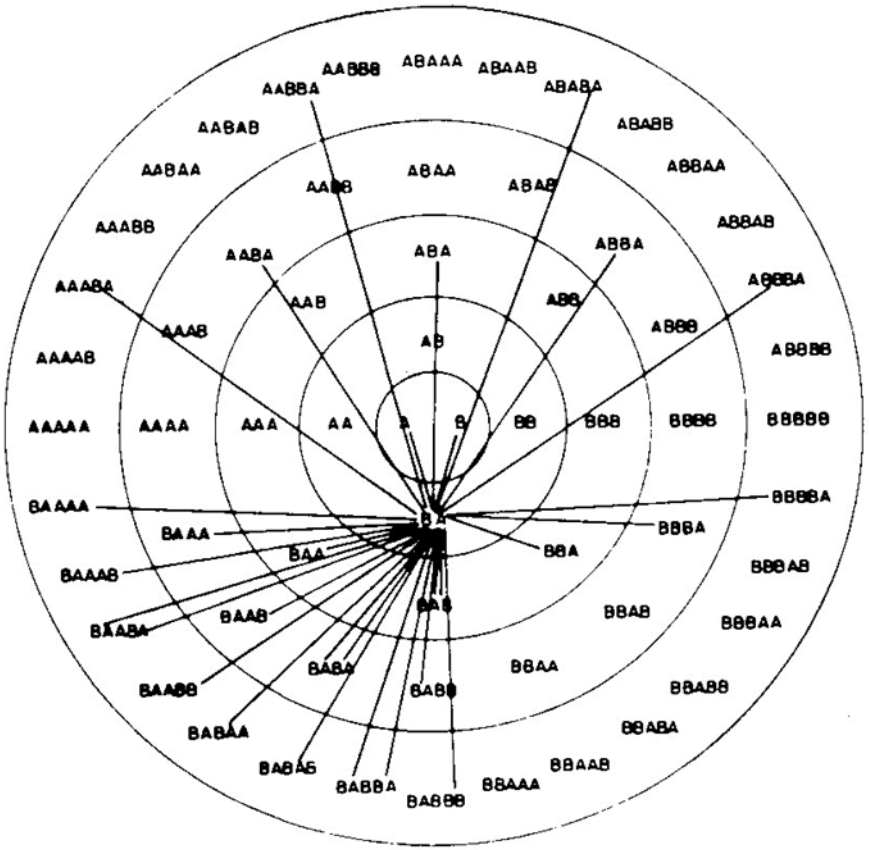
Catalytic subset

How many real enzymes and edges do we need ?

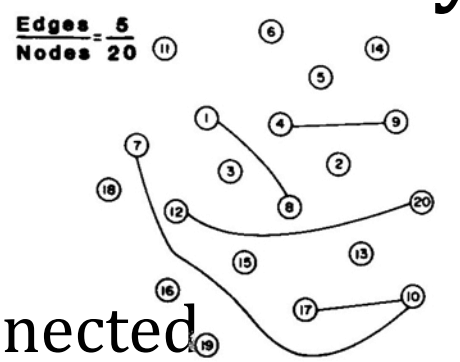
- I do not have to be able to synthesise everything

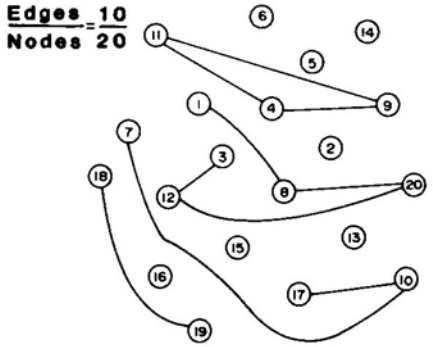
- Behaviour with random graphs ?





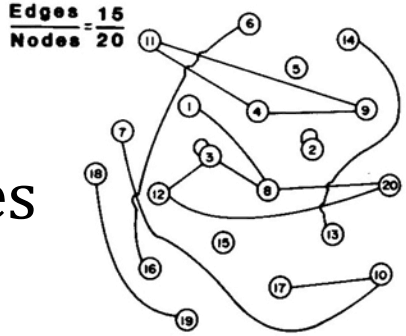
edges and connectivity

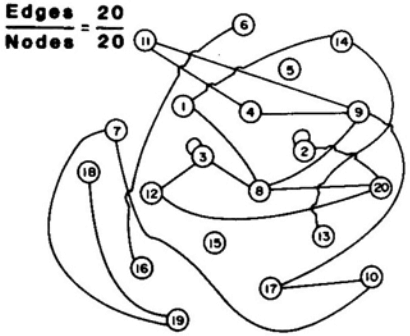


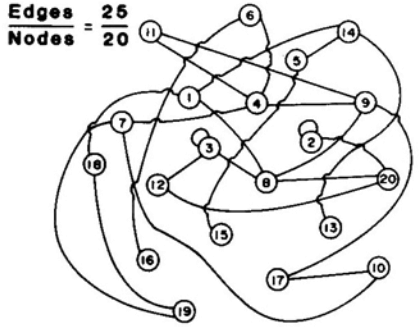


Standard results

- as edges ≈ nodes/2
  - most components are connected







When edges≈nodes

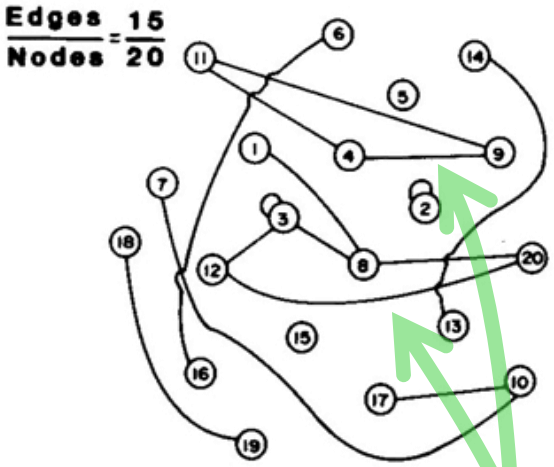
- cycles appear

Those nodes in cycles

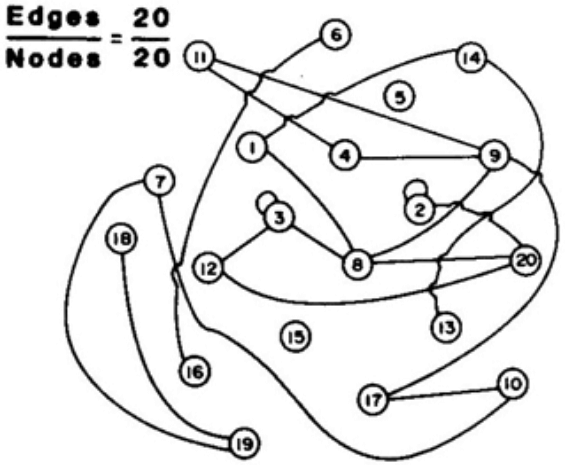
- can be synthesised using only other components in the cycle
- probability of cycles is near 1

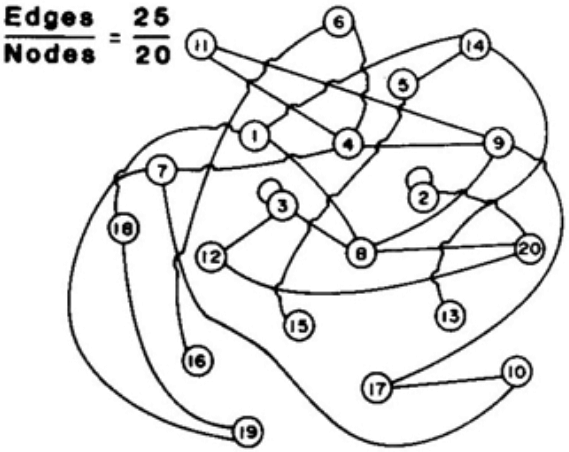
Connectivity

- As soon as I have a cycle..
  - Self-reproducing system… Life ?



here ?





Catalytic cycles

Gross simplifications

- no specificity
- one enzyme does all XY bonds regardless of context
- all rates the same…

Reasoning valid for 4 bases (RNA) or 20 residues (protein)

Auto-catalytic model

Without real "information" system

- self reproducing
- minimum complexity
- may have errors, tolerance of errors = evolution

- life may emerge suddenly
- order appears suddenly (Entropy disappears ..OK ?)

Autocatalytic model consequence

Anti-evolution

- what are the chances of molecules coming together to form a 200 residue protein ?
- what are the chances of a hurricane blowing bricks and building a house ?

This model

- the hurricane does not have to re-assemble a house
- any self-sustaining network will do
- our world is just one outcome
- whatever chemistry is most successful…

Experimental evidence

- not like ribosomes (difficult to explain without an RNA world)
- artificial systems.. example

RNA example of cooperating cycles

- ribozyme with four regions, ABCD
- four autocatalytic reactions
  - A+BCD →ABCD
  - AB+CD →ABCD
  - ABC+D →ABCD
  - ABC+D →ABCD
- ABCD is a better catalyst than the parts
- recognition / pairing site can be varied
- possibility of cooperation
  - $A_1 + B_1\ C_1D_1 \xrightarrow[A_2B_2C_2D_2]{} A_1B_1C_1D_1$

Vaidya,N Manapat, ML. Chen, IA Xulvi-Brunet,IR, Hayden, EJ, Lehman, N, Nature, 2012, 491, 72

RNA example of cooperating cycles

Throw all ingredients into bucket

- A, AB, ABC, BCD, CD, D × sequence variants

48 possible products (comes from joining and sequence)

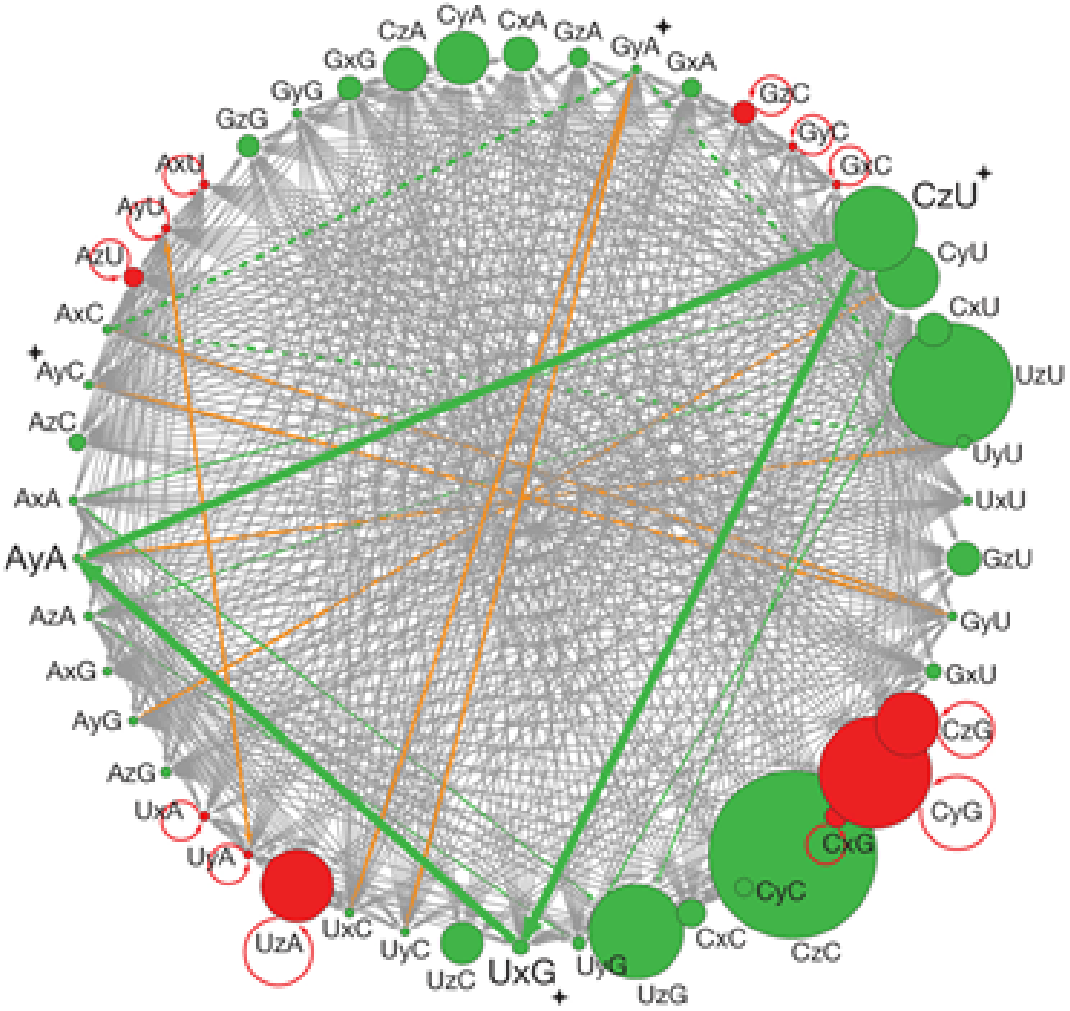
Results ?

48 products

size: how much of product after 8 hr

red: autocatalysts

green: cooperators



Claim:

cooperators are winners

Proof ? No – nice example of feasibility

Vaidya,N Manapat, ML. Chen, IA, Xulvi-Brunet, IR, Hayden, EJ, Lehman, N, Nature, 2012, 491, 72

For an Exam

- characteristics of life
- evidence for RNA world
- problems with RNA world
- auto-catalytic models

Summary

- life, entropy, information
- evolution, errors and tolerance of errors
- RNA world
  - ribosome – strong evidence
  - search for (possibly indirect) template directed replication
  - difficult to specify exact reactions producing
    - activated monomers
    - polymers
- search for simple template-directed replication may not be necessary
- self reproducing system may spontaneously form