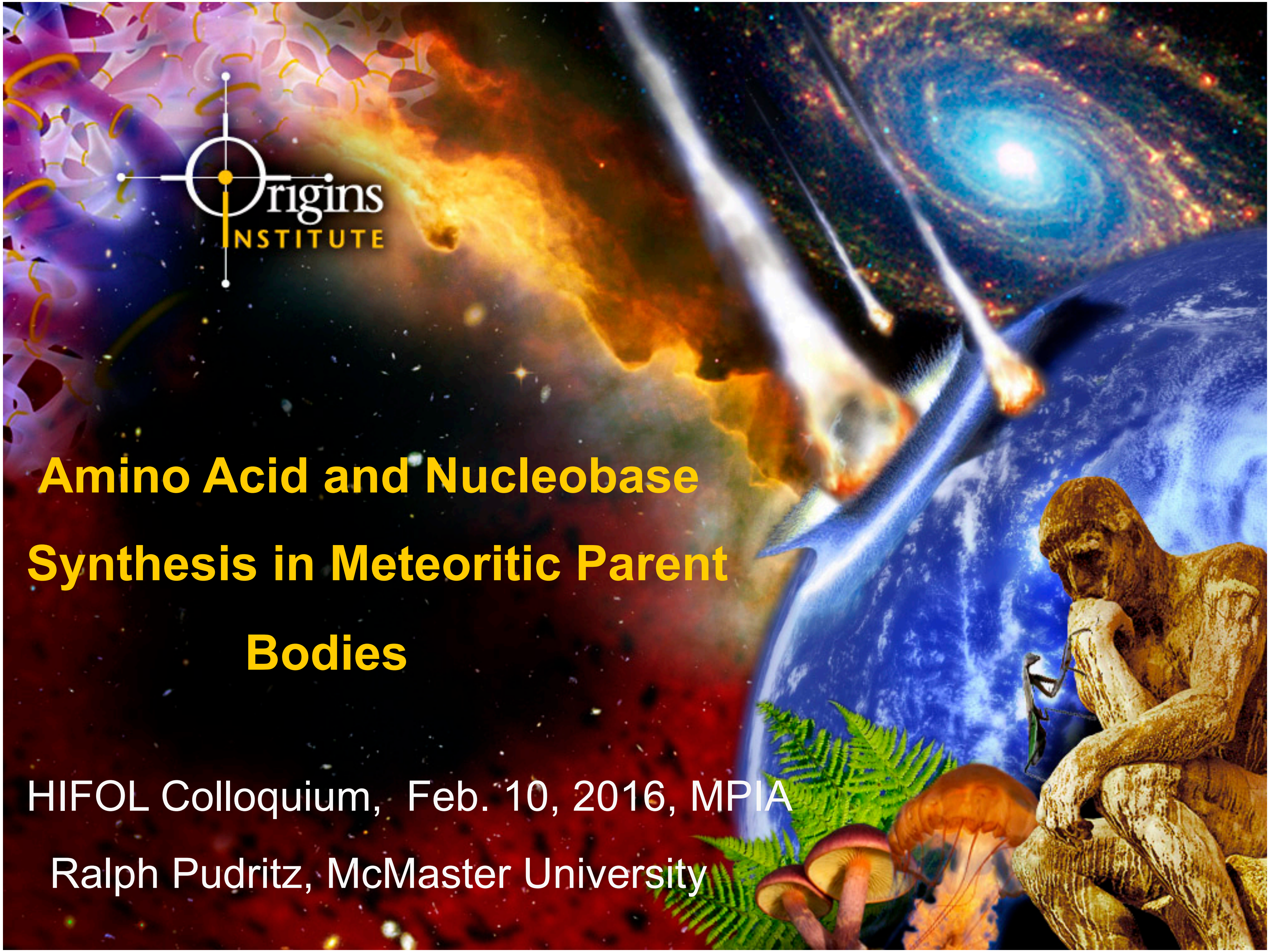


Origins
INSTITUTE

Amino Acid and Nucleobase Synthesis in Meteoritic Parent Bodies

HIFOL Colloquium, Feb. 10, 2016, MPIA

Ralph Pudritz, McMaster University

Thanks to my students and collaborators:

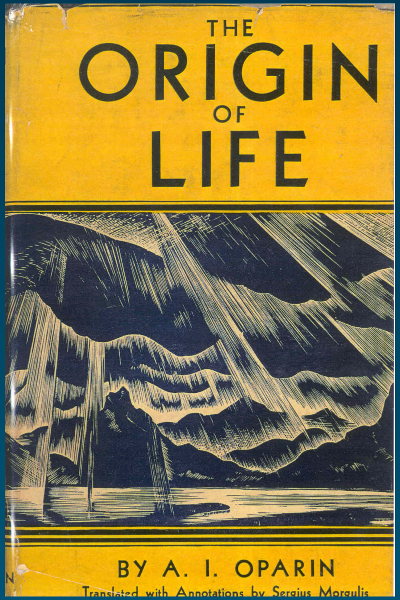
Paul Higgs: Depts. Physics and Astronomy, & Biochemistry
Alyssa Cobb (M.Sc. McMaster, Northup-Grumman),
Ben Pearce (M.Sc. McMaster)
Paul Ayers (Theoretical chemistry group – McMaster)

Undergrads:
Jeff Emberson: NSERC USRA (U. of T. astrophysics)
Darren Fernandes: NSERC USRA (U of T., biophysics)

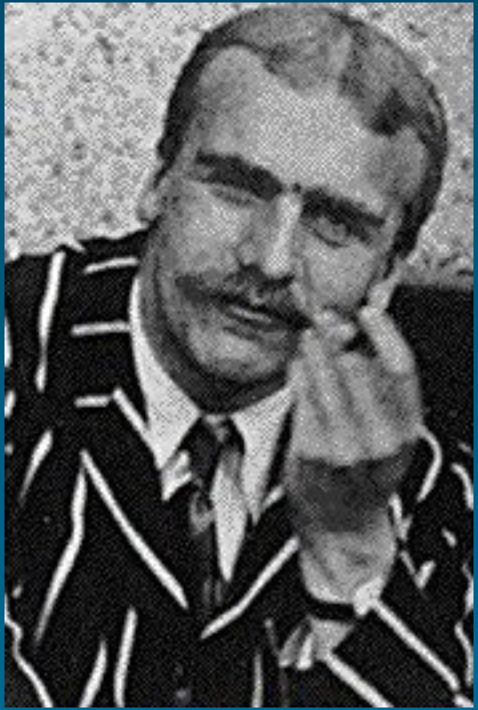
Research Support:

NSERC CREATE: Canadian Astrobiology Training Program (CATP)
NSERC, Origins Institute

I Origins of Life



A.I. Oparin (1894-1981)



J.B.S. Haldane

(1892-1964)

Oparin-Haldane Hypothesis (1920's):

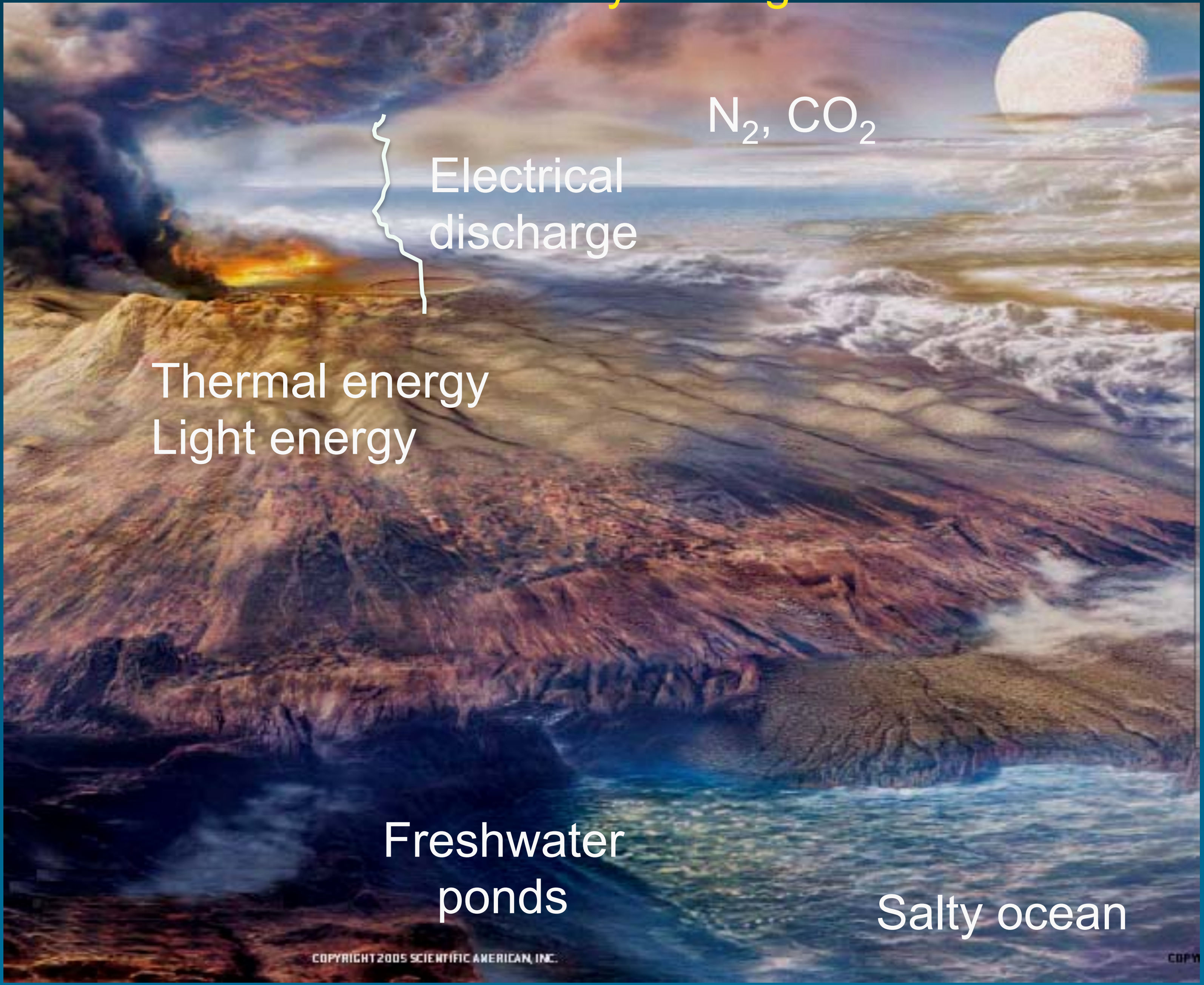
- life formed only once on Earth in "hot dilute soup"

- Origin follows basic laws of chemistry and physics + Darwin's law of evolution

Oparin: metabolism first

Haldane: genes first

Four billion years ago….



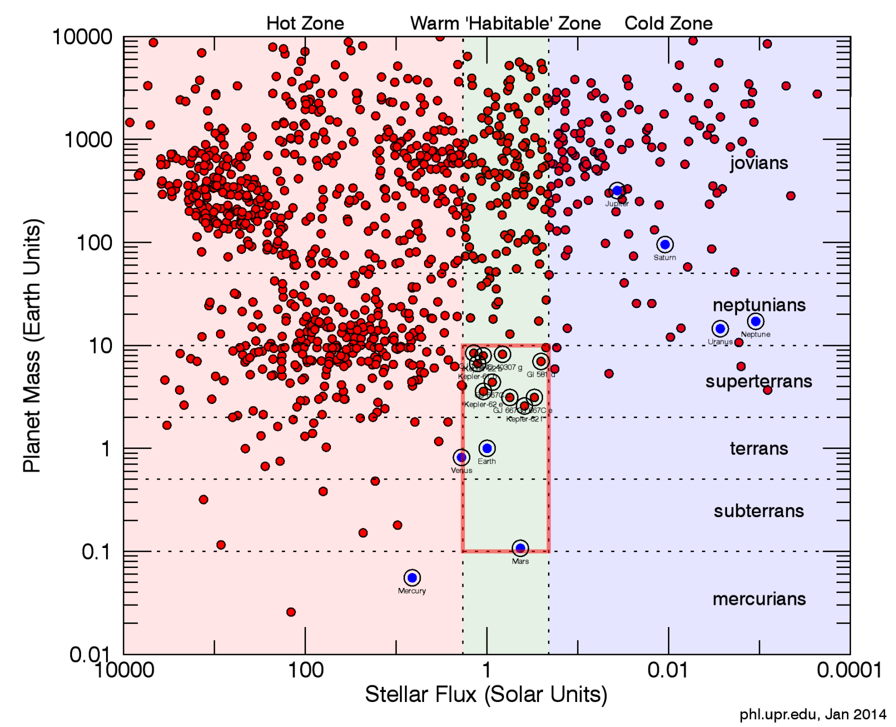
David Deamer

What is needed for life as we know it?

- “Habitable” rocky planets
- Energy source (chemical, the Sun,...)
- Water
- Biomolecules

Exoplanet Observations: Habitable Rocky Planets

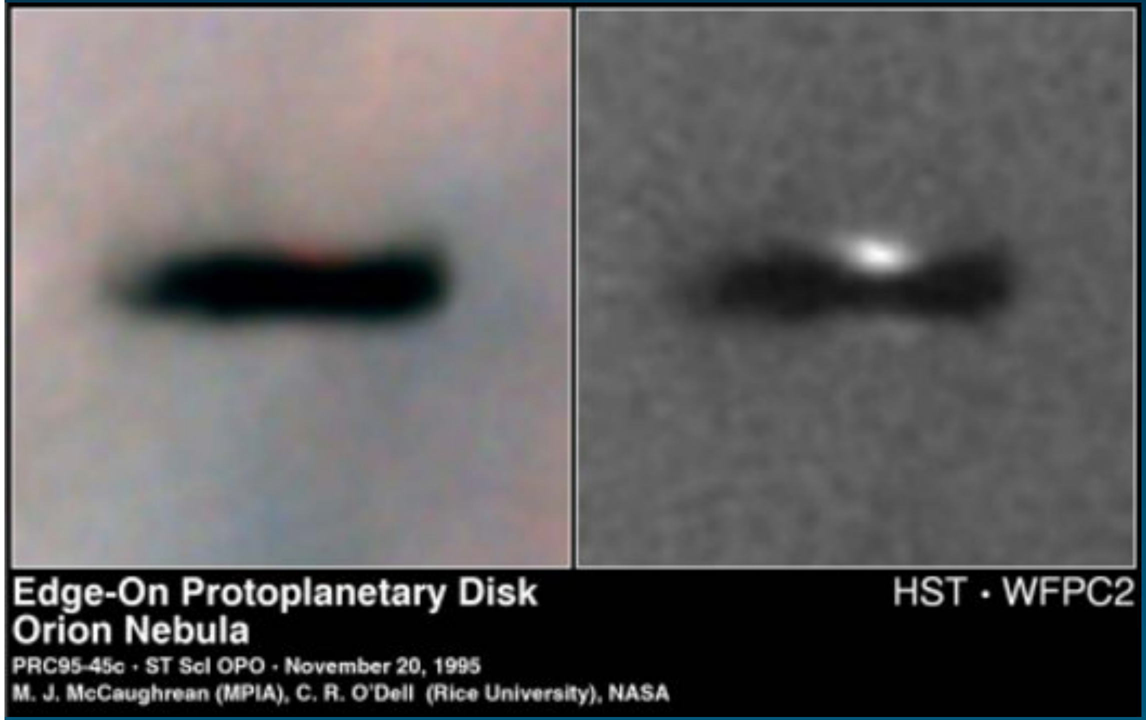
- Jupiters at orbit of Mercury (“hot Jupiters”)
- Pile up of massive Jupiters at 1 AU
- SuperEarths: 1-10 M_E dominant population
- Nearly 2000 confirmed planets
- About a dozen known SEs in habitable zones



Kepler data release 2015

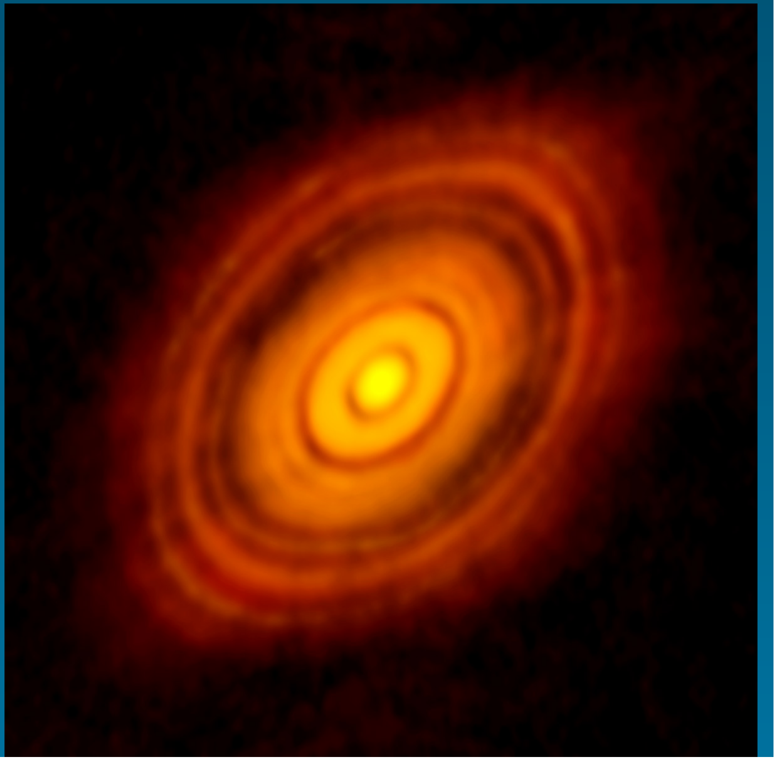
Planet formation and chemistry in disks around stars:

- most of star’s mass accretes from disk
- rocky planets made from the dust and ices

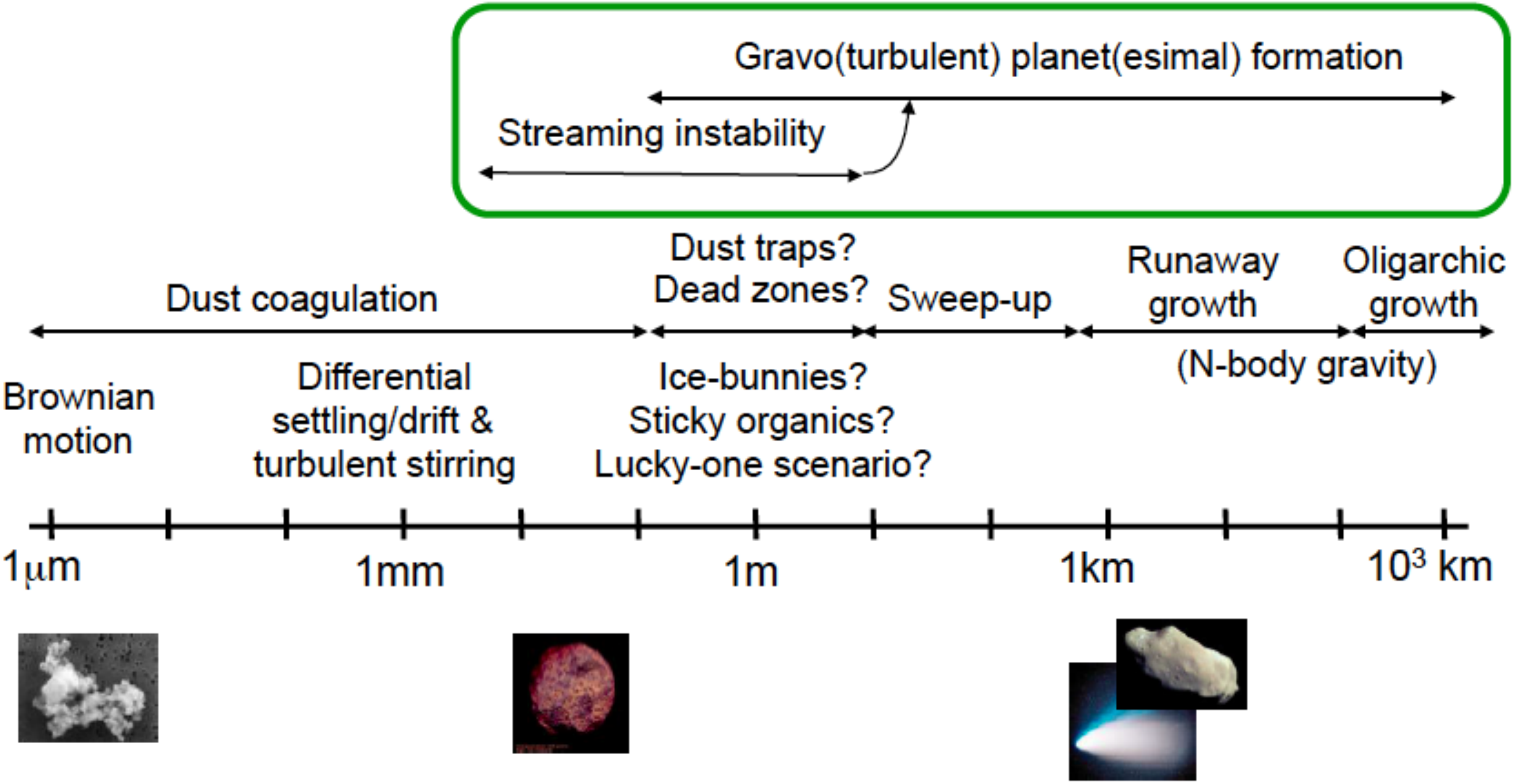


Top: gas/dust protoplanetary disks around young stars in Orion Nebula (Hubble Space Telescope image)

Right: 1 Myr old disk around a young star, HL Tau (ALMA mm image)



Making Rocky Planets, Step 1: Dust to Planetesimals



Gravo(turbulent) planet(esimal) formation

Streaming instability

Dust coagulation

Dust traps?
Dead zones?

Sweep-up

Runaway
growth

Oligarchic
growth

(N-body gravity)

Brownian
motion

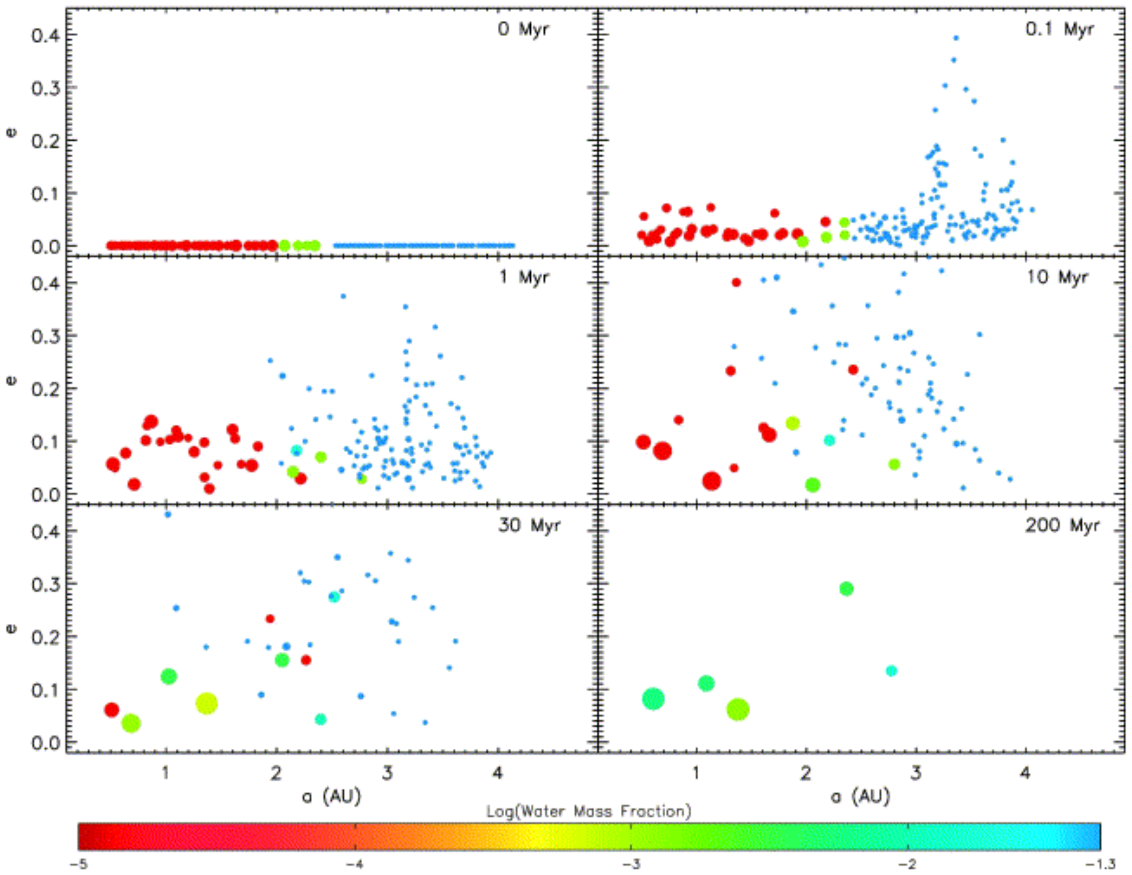
Differential
settling/drift &
turbulent stirring

Ice-bunnies?
Sticky organics?
Lucky-one scenario?

Dullemond, 2012

Step 2: Planetesimals to Watery Planets

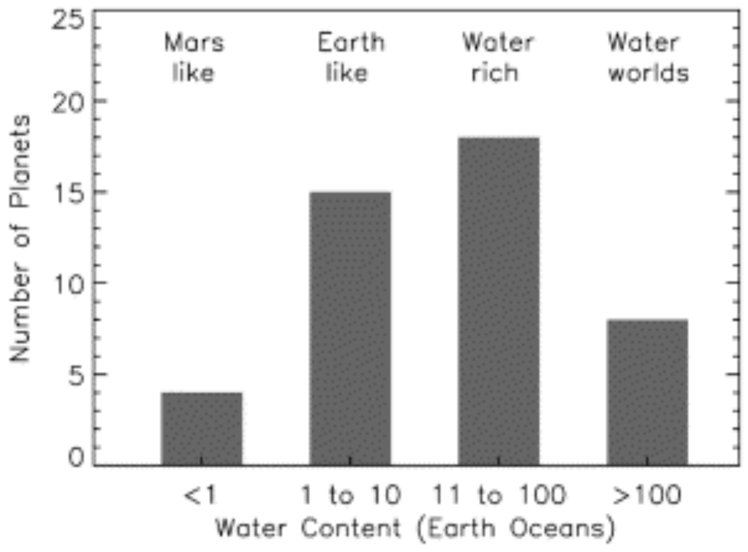
(Raymond et al, 2004)



N- body simulations of planetesimals being perturbed by Jupiter

Water laden asteroids beyond “snow line” (2.5 AU) predominant carriers of water

- Water abundant!



Sources of Biomolecules

1. Meteorite parent bodies - planetesimals

Eg. Murchison's meteorite
– impacted Australia (1969): carbonaceous chondrites

Organics ~ 1.5 % of total mass

More than 70 amino acids

Source of sugars, alcohols, sugar acids, and 3 nucleobases

(Cooper et al. 2001, Nature)

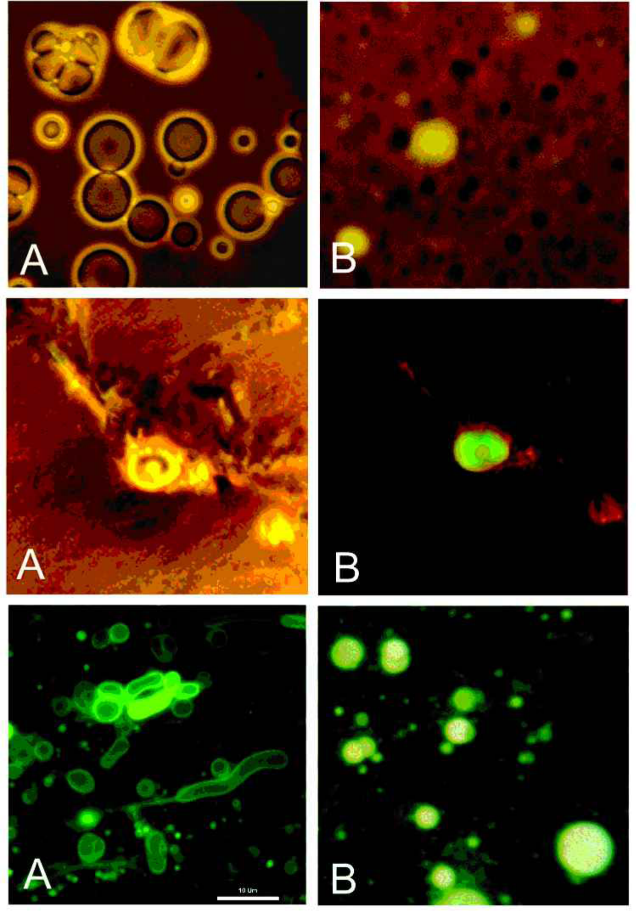
Source of amphiles – lipids that make membranes

Steps towards the first cells:

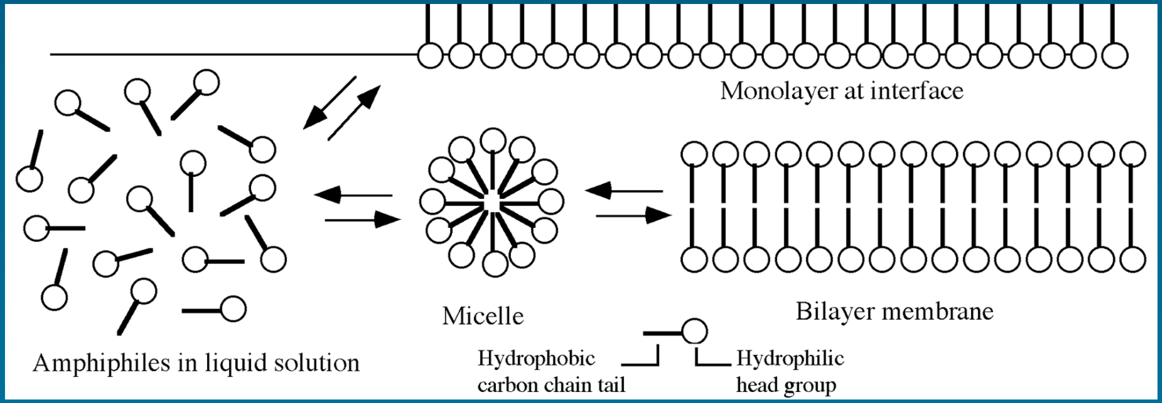
Grind up Murchison (top)… suspend in water … get bags (vesicles) – primitive membranes.

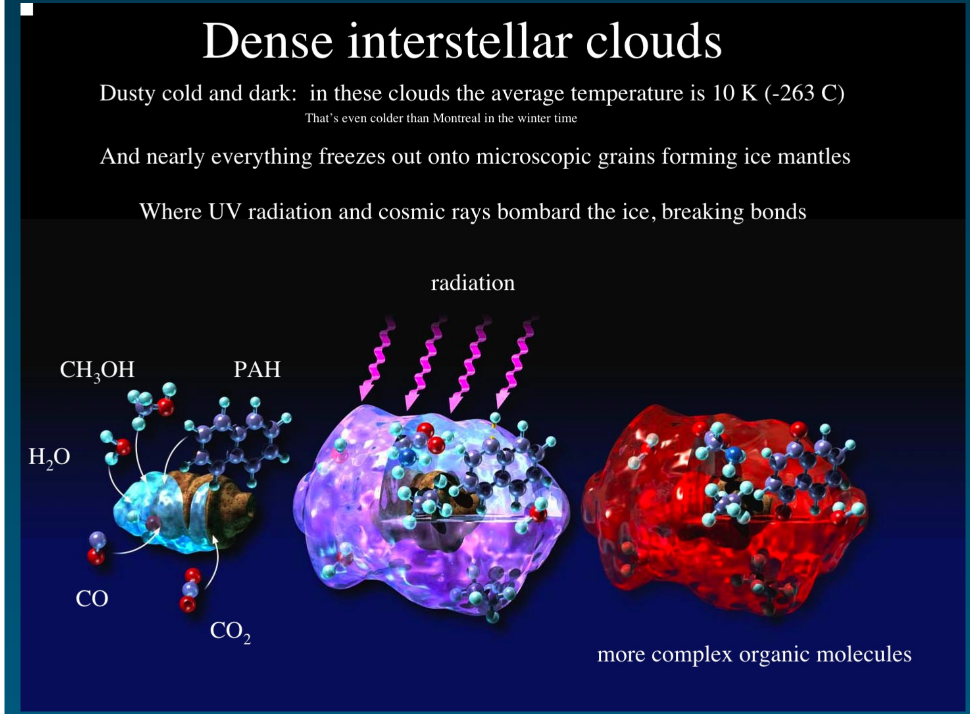
Synthesized analogue (middle)

Synthesized vesicles encapsulate large molecules - such as DNA (lower right)



micrographs: Deamer, 2002)





Dense interstellar clouds

Dusty cold and dark: in these clouds the average temperature is 10 K (-263 C)

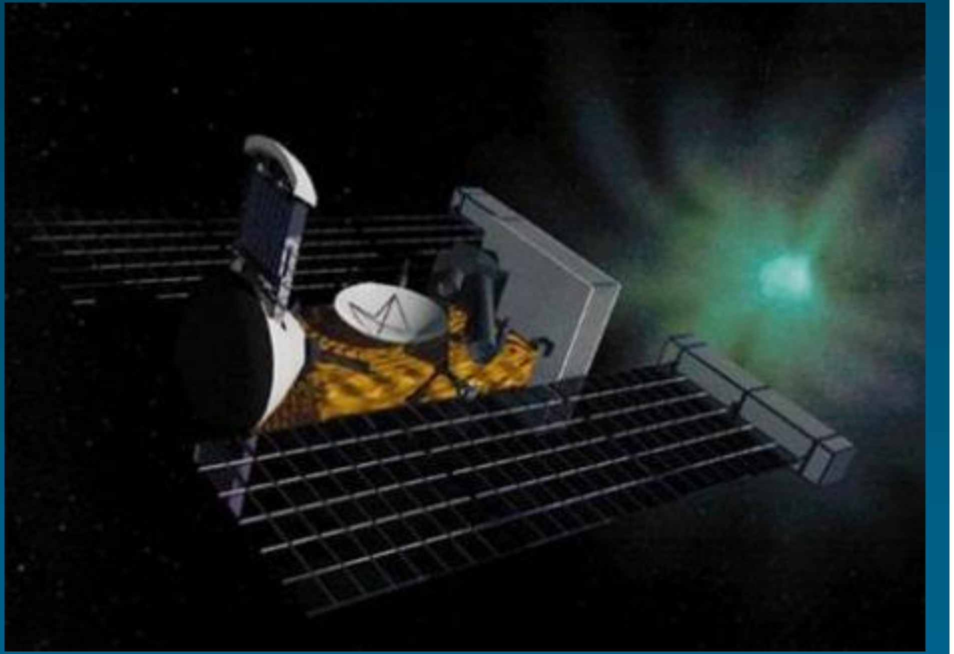
That's even colder than Montreal in the winter time

And nearly everything freezes out onto microscopic grains forming ice mantles

Where UV radiation and cosmic rays bombard the ice, breaking bonds

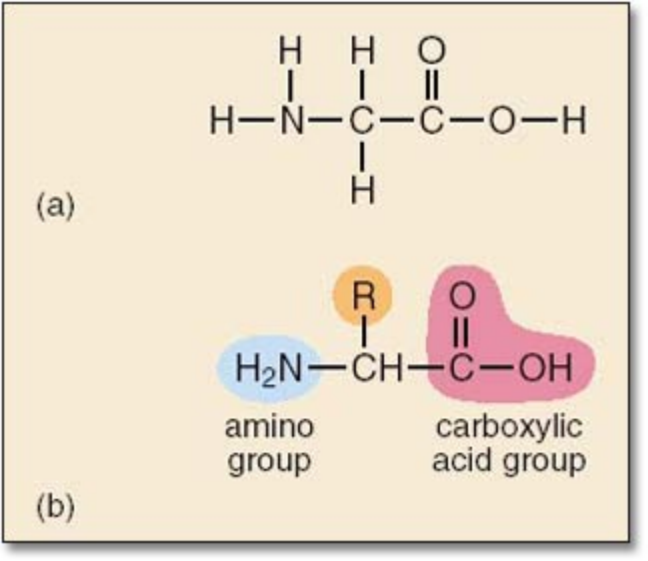
…. or comets?

Bernstein et al (Nature, 2002) – experiments produce glycine



Stardust observatory collects icy grains from tail of comet 81P/Wild 2 - Glycine detected

(Elsila et al 2009)



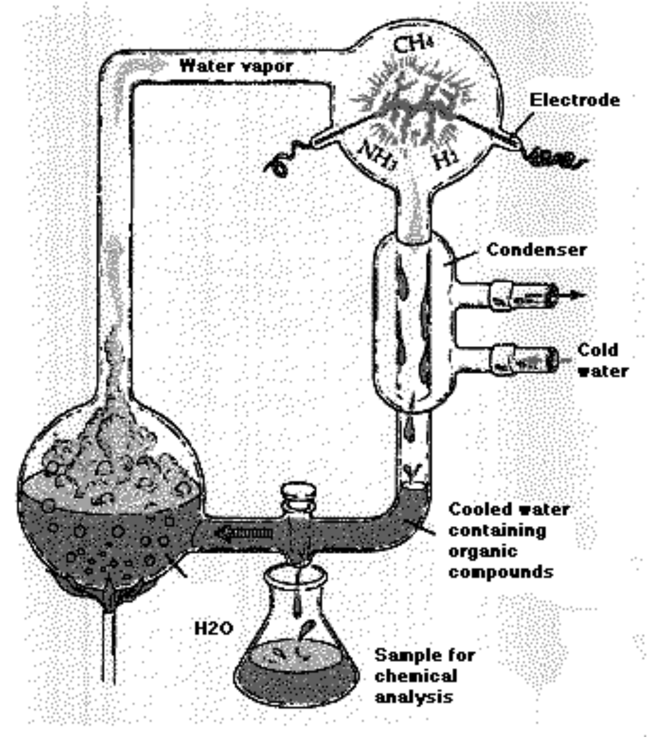
Glycine

2. Planetary atmospheres?

Stanley Miller 1953

Miller- Urey (1953) experiment:

- Jupiter-like: reducing atmosphere of hydrogen, methane, ammonia.
- Presence of water (early ocean)
- Energy source (“lightning”).



Modern studies: H is released from volcanoes as H_2O, rather than H_2:

- most C as CO_2 rather than CO or CH_4

3. Hydrothermal vents in oceans?

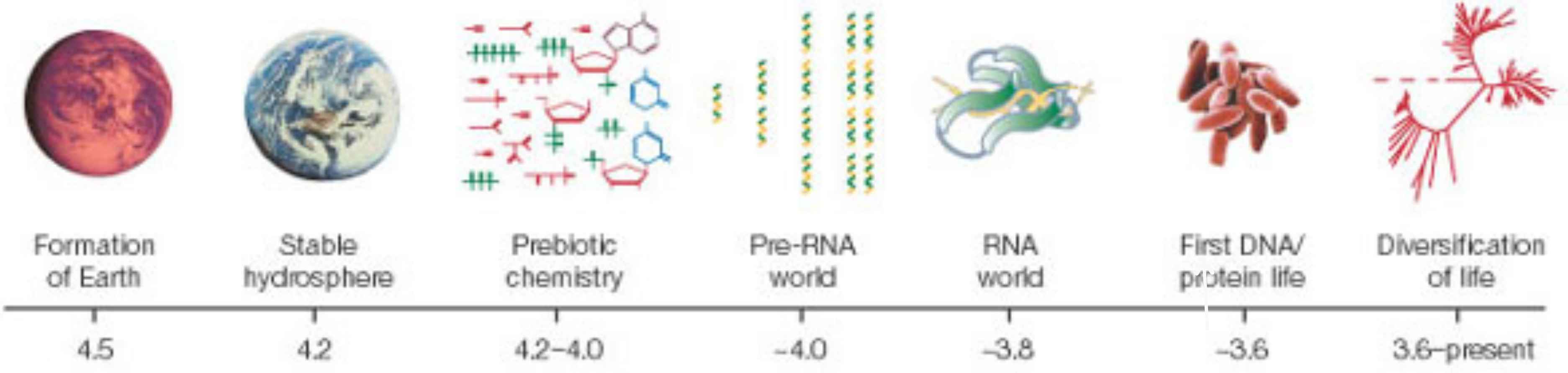
- ocean floor spreading zones
- hyperthermophile micro-organisms – seemingly at root of tree of life.

Temperatures: 270-380 C.
Water with hydrogen sulfide spews out of cracks in Earth's crust

Early history of life on Earth

Are there general properties of prebiotic soups?

Do these shape resulting genetic codes?



Dating of rocks and meteorites

Last ocean-vaporizing impact. Lunar craters

RNA World

Chemical evolution

Isotopic evidence for life (3.8)

Origin of Genetic Code

Microfossil evidence? (3.5)

Joyce 2002

II. Thermodynamic constraints on early amino acids – and genetic codes

(Higgs & Pudritz, *Astrobiology,* 2009)

1. Observed frequencies of amino acids synthesized non-biologically

- M1, M2, M3 = Meteorites (e.g Murchison meteorite)
- I1 = Icy dust grains in space
- A1, A2, A3 = Atmospheric chemistry experiments (e.g. Miller-Urey)
- H1, H2 = Hydrothermal synthesis
- S1, S2, S3 = Other chemical synthesis experiments

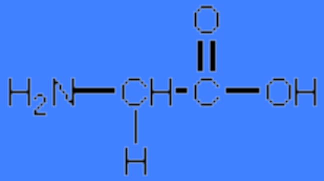
2. Rank amino acids in order of decreasing frequency in these 12 experimental observations. Derive mean ranking R_{obs}.

3. Two distinct groups of amino acids:

EARLY: found non-biologically

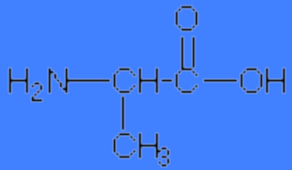
LATE: found biologically

Early Amino Acids: simpler and thermodynamically less costly


$H_2N-\underset{H}{CH}-\overset{O}{\overset{\|}{C}}-OH$

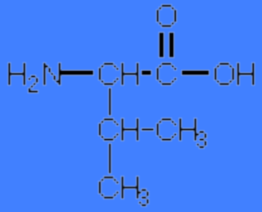
Glycine

Gly G


$H_2N-\underset{CH_3}{CH}-\overset{O}{\overset{\|}{C}}-OH$

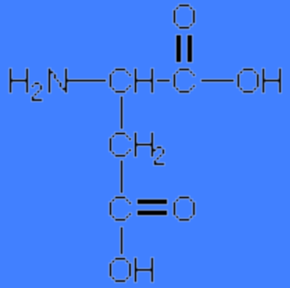
Alanine

Ala A



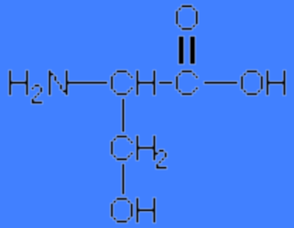
Valine

Val V



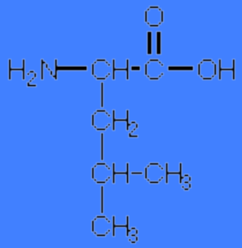
Aspartic acid

Asp D



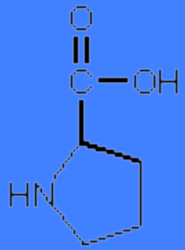
Serine

Ser S



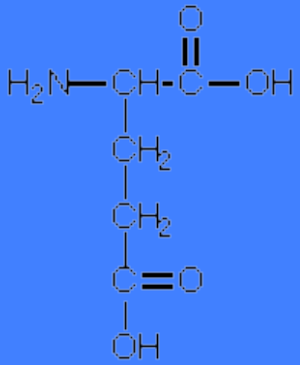
Leucine

Leu L



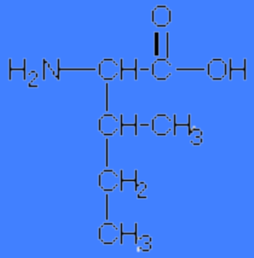
Proline

pro P



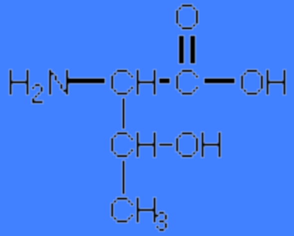
Glutamic acid

Glu E



Isoleucine

Ile I

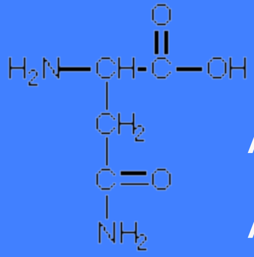


Threonine

Thr T

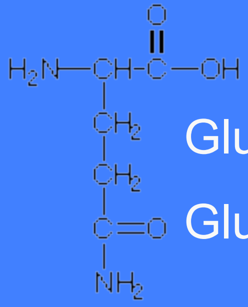
Hydrophobic ↔ Hydrophilic Acidic

Late Amino Acids: more complex and thermodynamically costly



Asparagine

Asn N



Glutamine

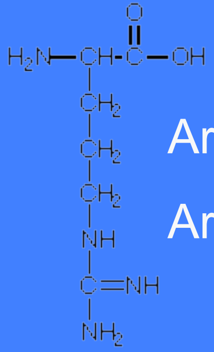
Glu Q

Amines

Hydrophilic

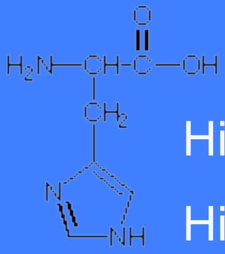
Lysine

Lys K



Arginine

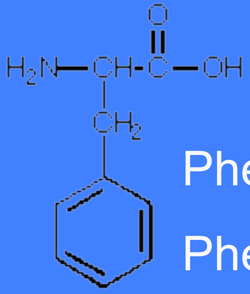
Arg R



Histidine

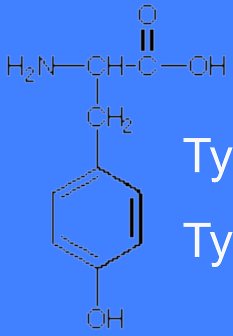
His H

Basic



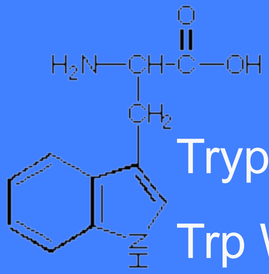
Phenylalanine

Phe F



Tyrosine

Tyr Y

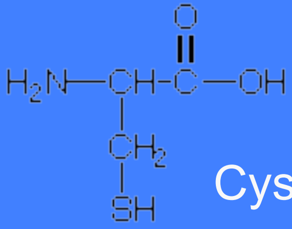


Tryptophan

Trp W

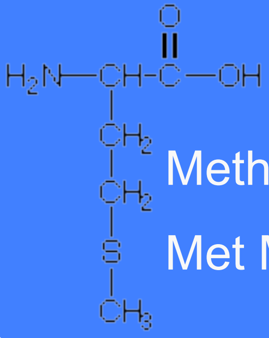
Aromatic

Hydrophobic



Cysteine

Cys C



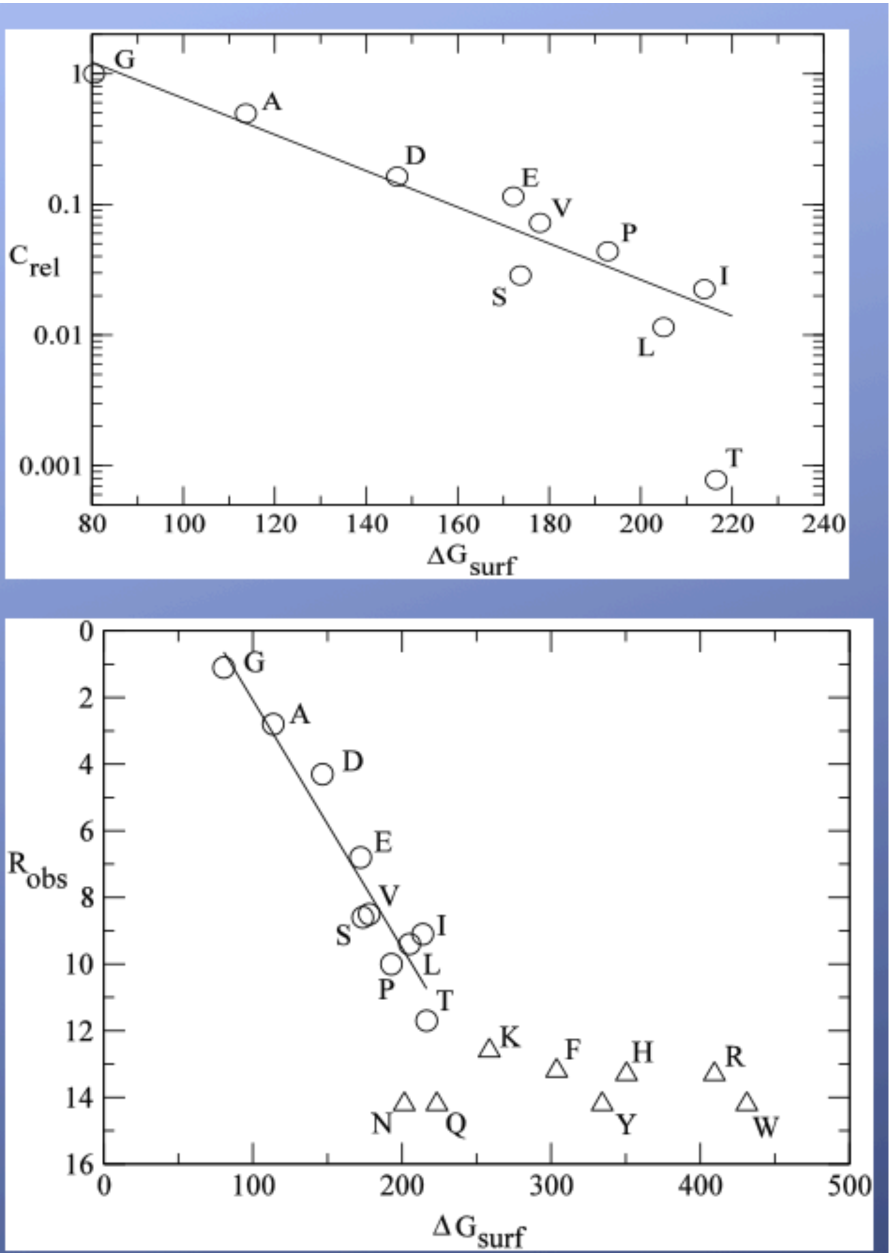
Methionine

Met M

Sulphur containing

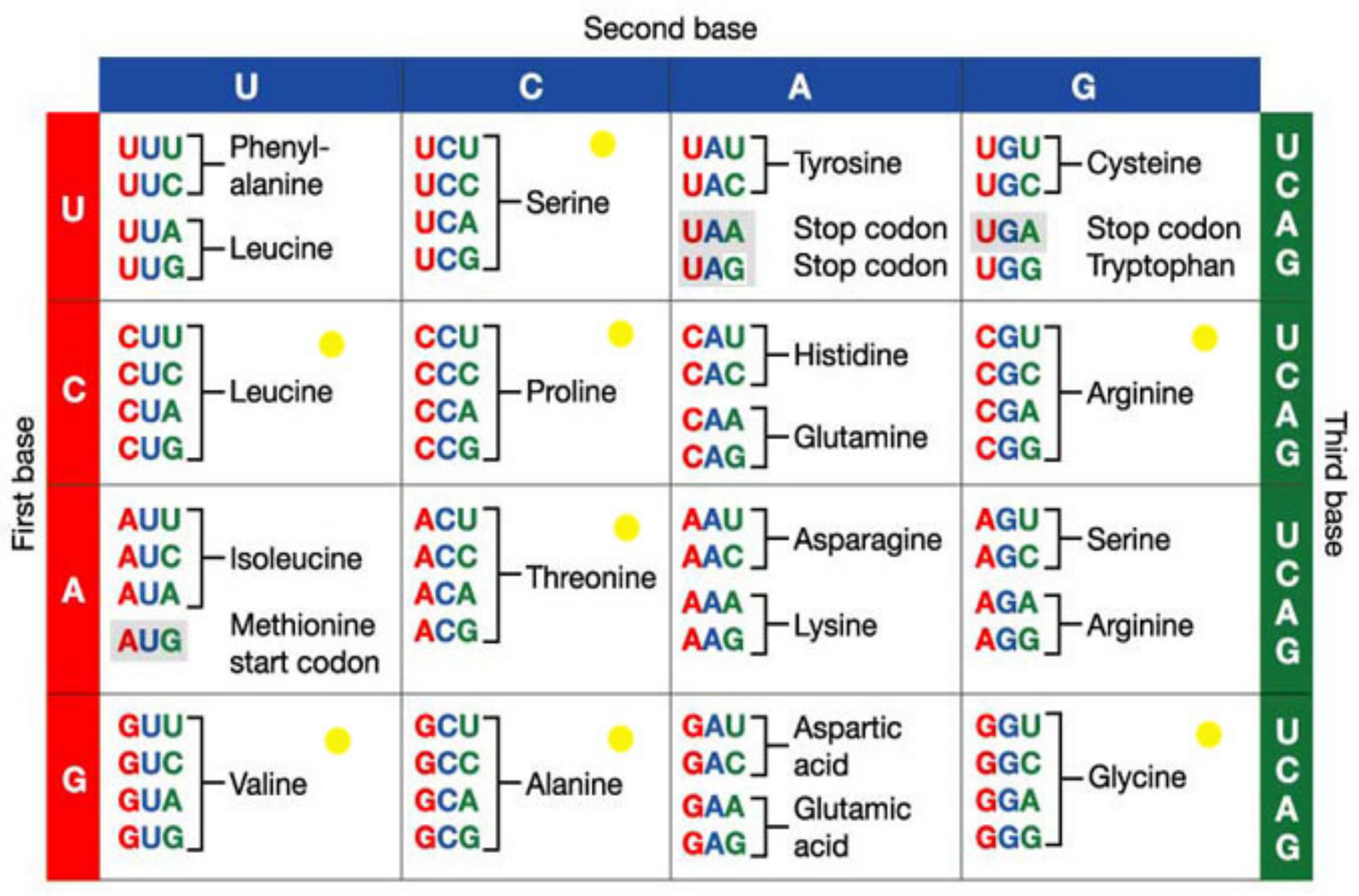
B. How amino acids influence codes:

- Compare results to Higgs & Pudritz (2009) survey of amino acid data - most common amino acids are easiest to make: role of thermodynamics
- Amino acids in life: 10 “early” - prebiotic; and 10 “late” biotic.
- “Early” code simple - based on content of organic soup
- Later code - designed by life - adds new capability with 10 more specially designed (Wong 2005)



Gibbs data from Amand & Shock 1992

Modern Genetic Code: ie, mapping between codons (combination of 3 bases) and amino acids



	Second base								
First base	**U**		**C**		**A**		**G**		**Third base**
U	UUU	Phenylalanine	UCU	Serine ●	UAU	Tyrosine	UGU	Cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA	Leucine	UCA		UAA	Stop codon	UGA	Stop codon	A
	UUG		UCG		UAG	Stop codon	UGG	Tryptophan	G
C	CUU	Leucine ●	CCU	Proline ●	CAU	Histidine	CGU	Arginine ●	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Isoleucine	ACU	Threonine ●	AAU	Asparagine	AGU	Serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lysine	AGA	Arginine	A
	AUG	Methionine start codon	ACG		AAG		AGG		G
G	GUU	Valine ●	GCU	Alanine ●	GAU	Aspartic acid	GGU	Glycine ●	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glutamic acid	GGA		A
	GUG		GCG		GAG		GGG		G

The Canonical Genetic Code was established ~3.5 Billion years ago

AAs in same column cluster in physical property space (Urbina+ 2006)

Amino acids in the genetic code: random or natural selection?

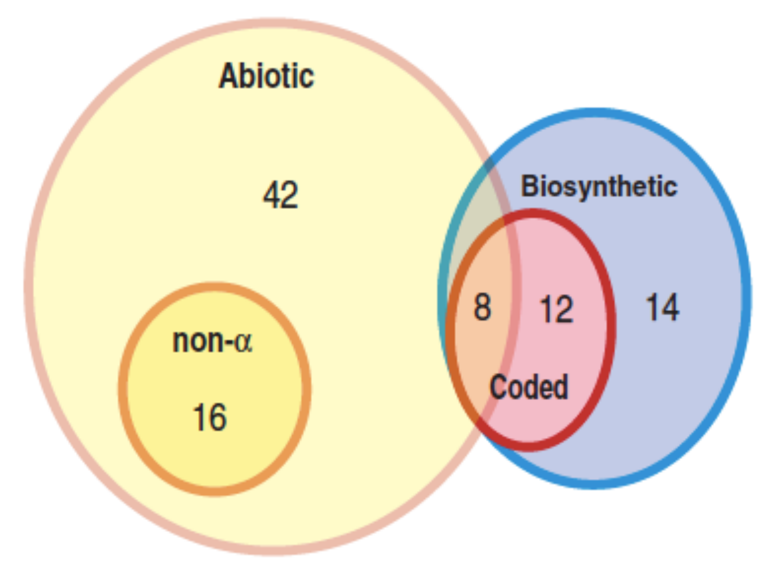
Abiotic = 66 AAs in Murchison; 50 of which could build proteins

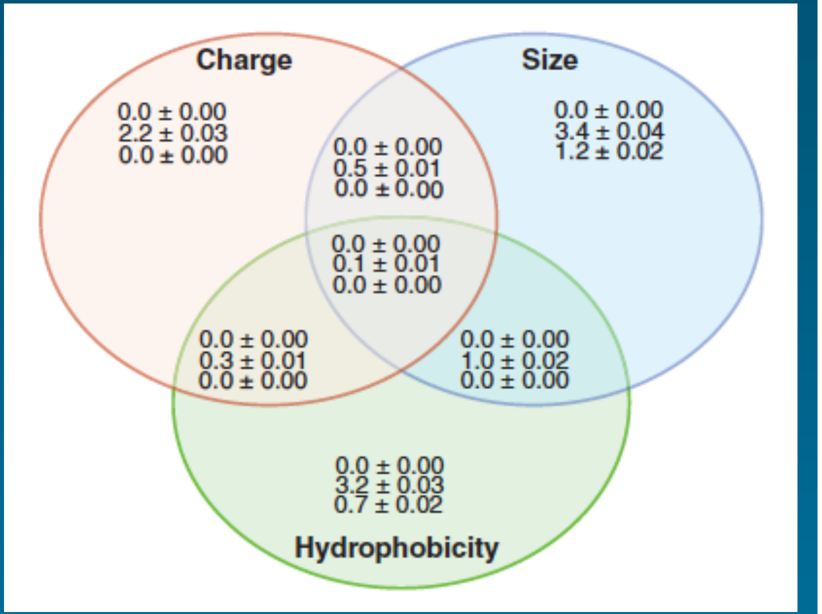
Biosynthetic: 12 AAs in code not found in Murchison built by life + 14 intermediaries

Bottom: Quantify a 20 sequence “alphabet” in terms of eveness of spread and breadth:

(mean and 95% confidence levels: iii is for 20 out of 76)

RESULT: Only 0.03% of random sets of 20 cover chemical space better - evolution?





Philip & Freeland, *Astrobiology*, 2011
(Lu & Freeland, 2009)

Implications for general evolution of genetic code?

1. Thermodynamics:
   provides natural frequency of amino acids for first code.

2. Earliest code used smaller repertoire of amino acids – each with larger no. of codons – stripped down version of ours.

- Lowest "cost" amino acids (eg. G) found in most highly expressed proteins (eg. Akashi & Gojobori (2002)

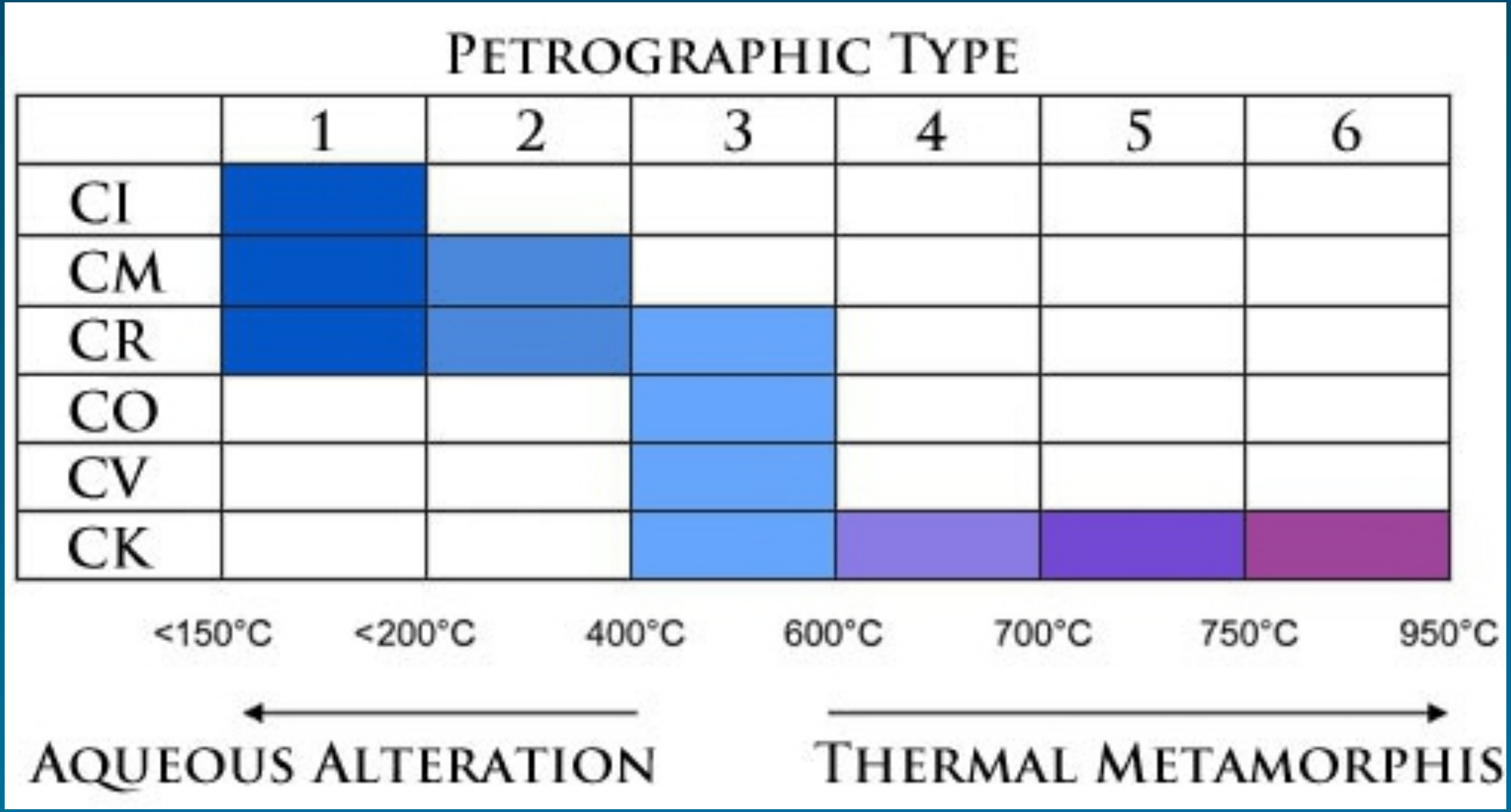
3. As more amino acids added, proteins ever more useful
   - finally DNA/protein code takes over (eg. Wong 2005)

4. Thermodynamics + Darwinian selection may produce early codes with similar attributes

III. Amino Acids

1. AA Meteoritic Data

Carbonaceous chondrites: classified by chemical composition, and secondarily, by amount of aqueous and thermal processing.



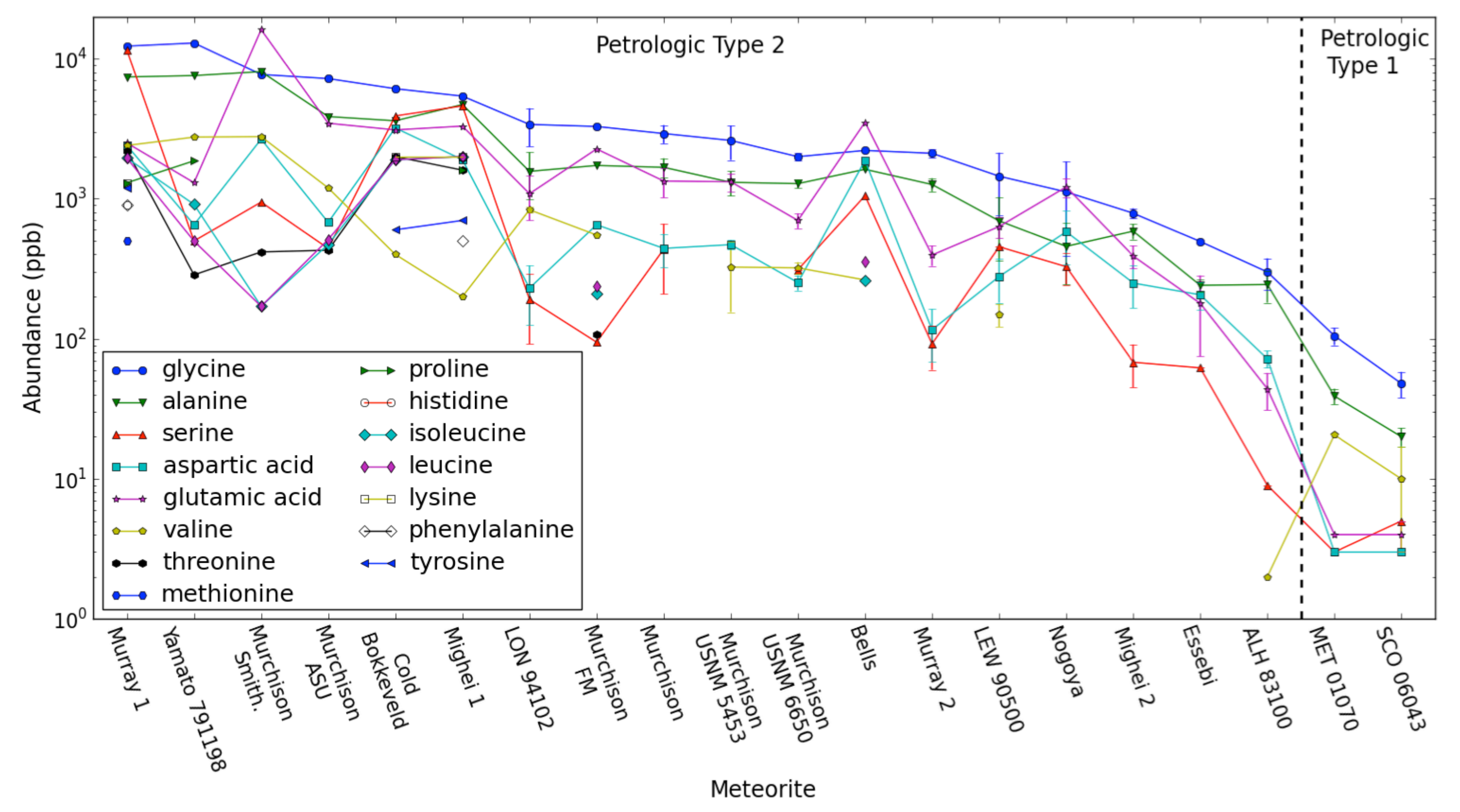
PETROGRAPHIC TYPE

	1	2	3	4	5	6
CI	🟦					
CM	🟦	🟦				
CR	🟦	🟦	🟦			
CO			🟦			
CV			🟦			
CK			🟦	🟪	🟪	🟪

Sephton (2002)

Amino acid abundances in CM meteorites

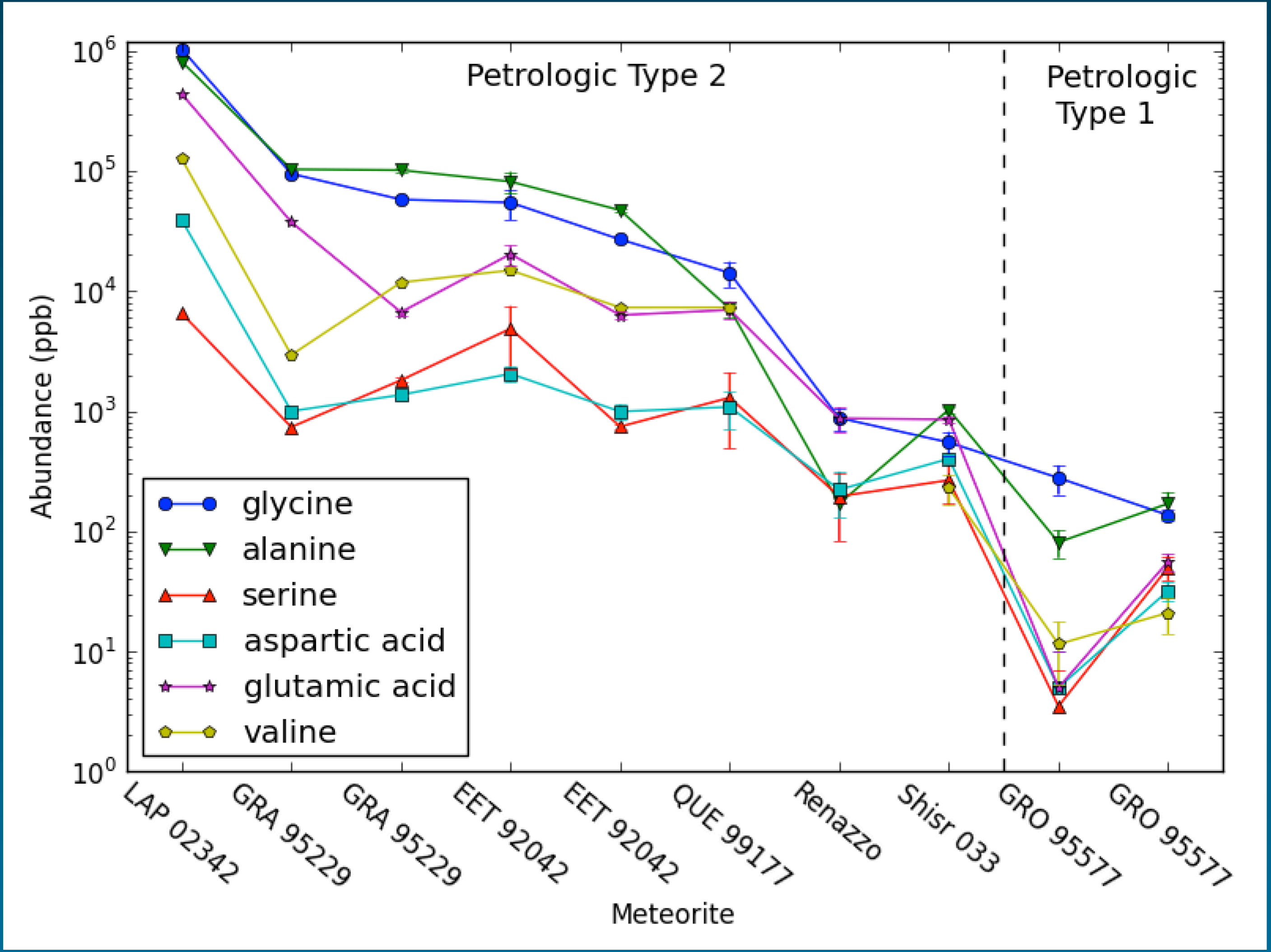
(order meteorites in monotonic sequence with glycine)



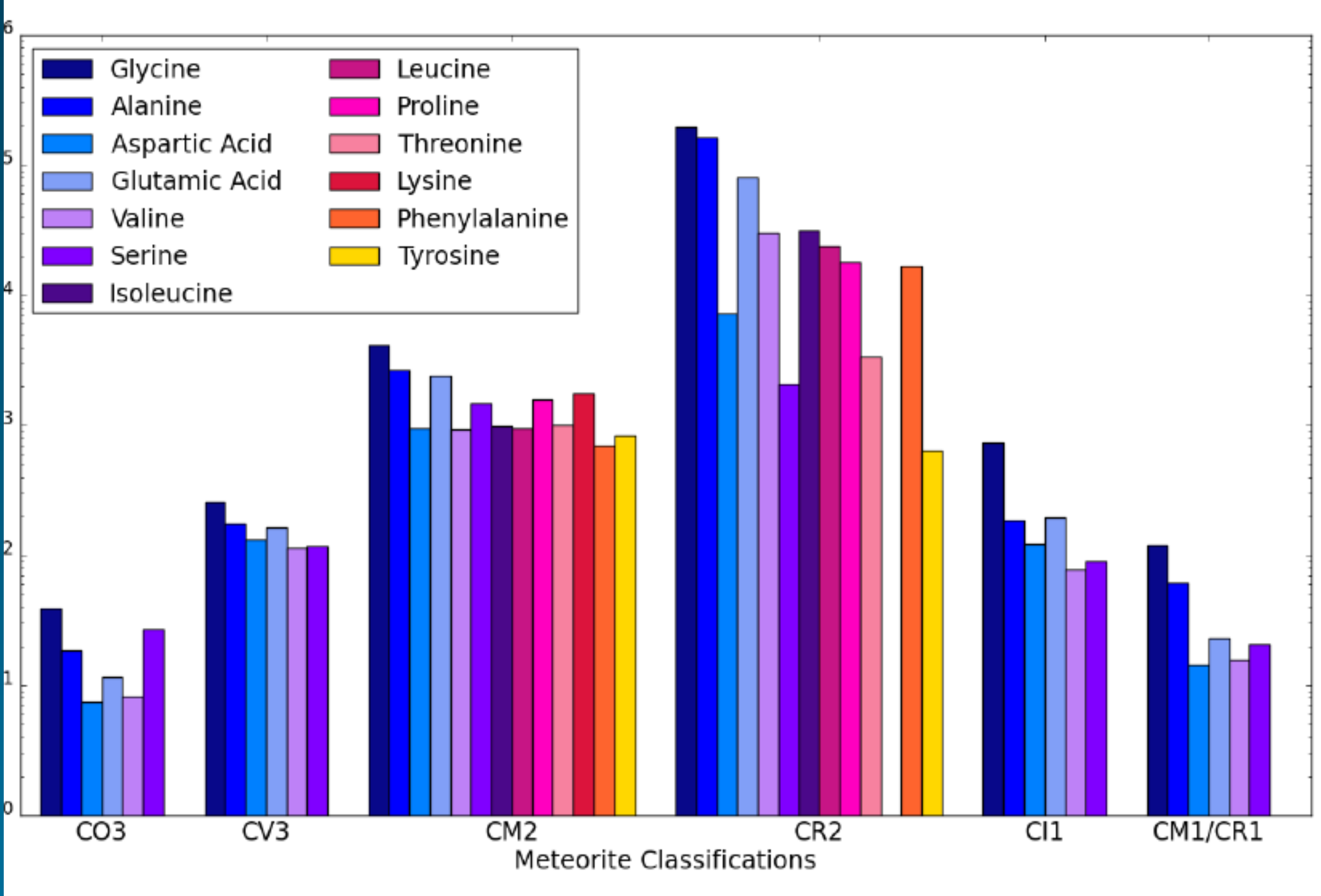
Cobb & Pudritz 2014, Ap J

Data from: Glavin et al (2006, 2011); Botta et al (2002); Peltzer et al (1984); Cronin & Pizzarello (1983); Shimoyama et al (1985); …

Amino acid sequences for CR meteorites

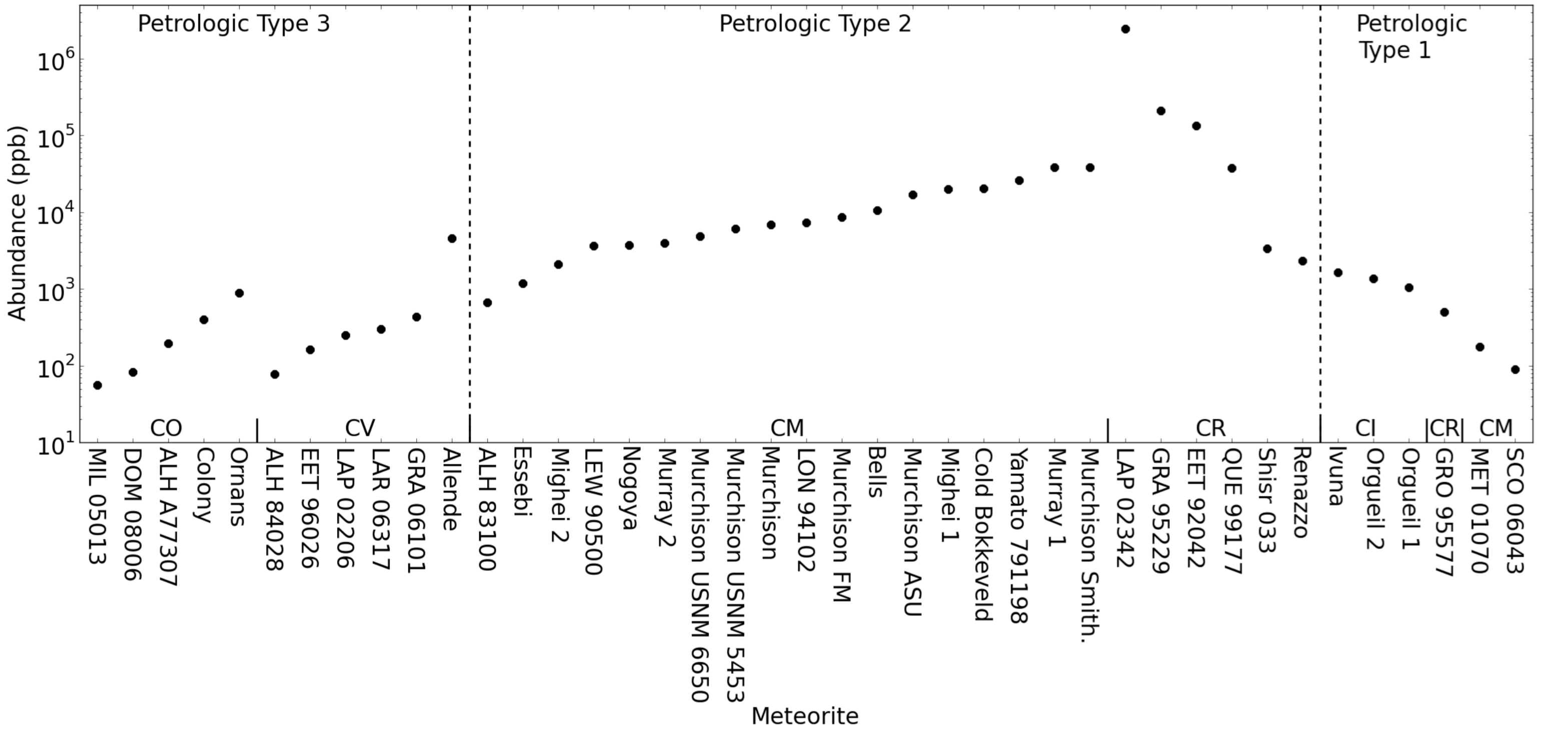


Amino acids in carbonaceous chondrites
- AA ordering follows Higgs & Pudritz 2014



Cobb & Pudritz 2014, ApJ

Overall trends for total amino acids across classifications.



Most abundant in CM and CR (eg. Glavin + 2011, Burton + 2012)

- Average concentrations 10^4 ppb, and 10^5 ppb…

Amino acid trends:

- Optimal range for temperatures T ~ 0 – 100 ° C
- Type 3 – most aqueous altered, no peak in abundance, T 50-150?
- Type 2 – cooler temperatures T ~ 0 – 100 C, most abundant AA in all petrographic groups
- Type 1 – minimally aqueous alteration.

Explanations:

- Onion shell model for parent body? (Weiss & Elkins-Tanton 2013).
- Planetesimal formation in chemically differentiated protostellar disks? (Cobb, Pudritz, & Pearce 2015)

2. Amino Acid Synthesis in Planetesimals: Theory

(Cobb, Pudritz, & Pearce 2015, ApJ)

Planetesimals are natural biomolecule factories!

Aqueous interiors (for a few Myr) - heating by radiogenic elements

Routes to amino acids

1. Aqueous alteration of PAHs (eg. Shock & Shulte 1990)
2. Strecker synthesis in aqueous solution (in presence of NH_3)

aldehyde (eg. formaldehyde) + HCN + H_2O -> amino acid (eg. glycine))

Equilibrium chemistry:

minimization of Gibbs energy for reactions:

$$\Delta G_r = \Sigma G_f^{products} - \Sigma G_f^{reactants}.$$

Structure of planetesimals

3D models simulations of 50km radius body

- rock 80%, ice 20% by volume
- heating by short lived radionuclide ^{26}Al, heating for several Myr
- hydrothermal convection is found

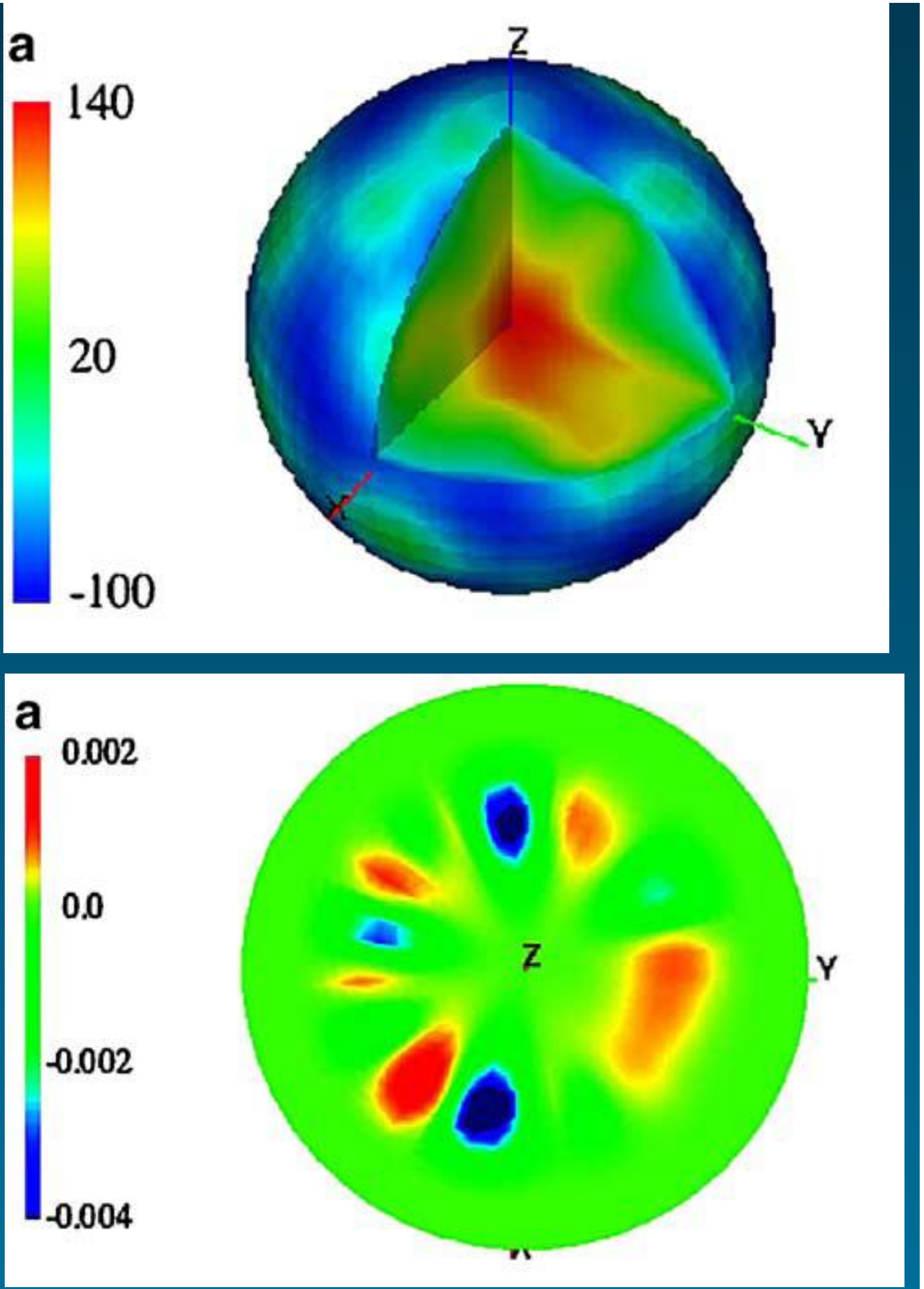
Top: Temperature structure (^{o}C)

Bottom: Velocity structure, equatorial slice:

green – stationary

blue – motion radially outwards

red – motion radially inwards



(Travis & Shubert 2005)

Numerical data for our models.

- Chemistry depends upon local T, P (eg. Gibbs free energy)

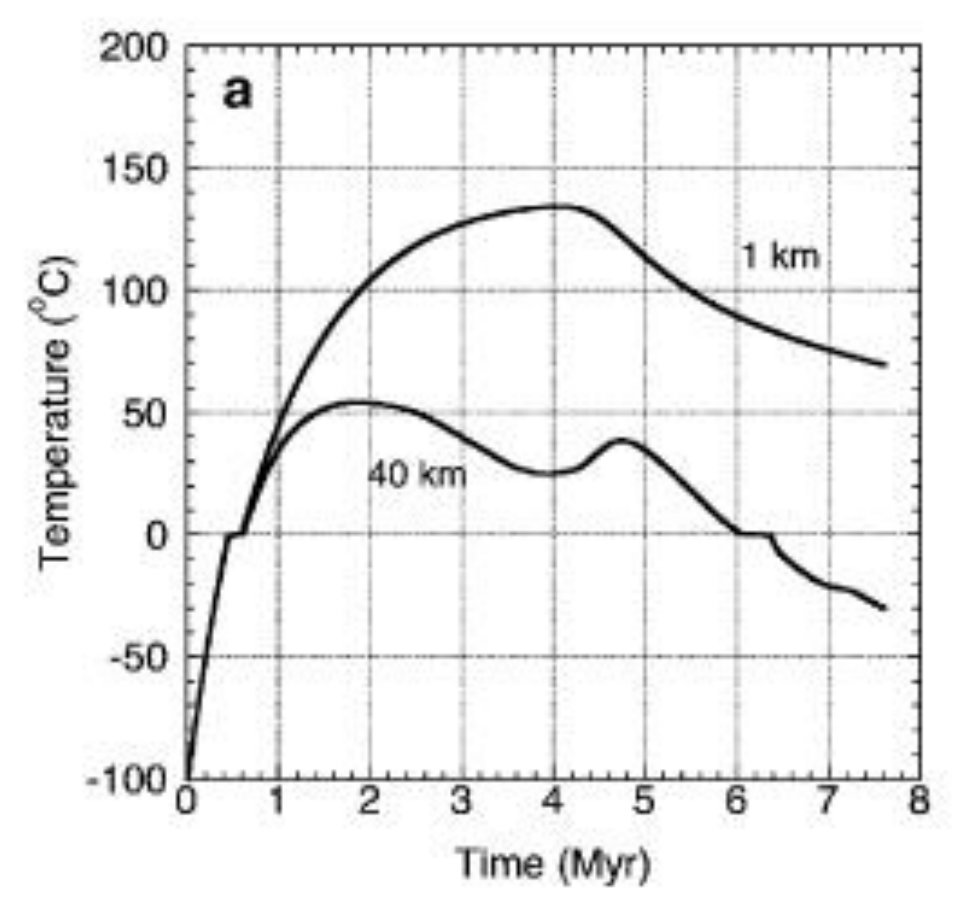
- T data from model, also compute pressures

  P = 8.7 bar (1 km)

  = 3.2 bar (40 km)

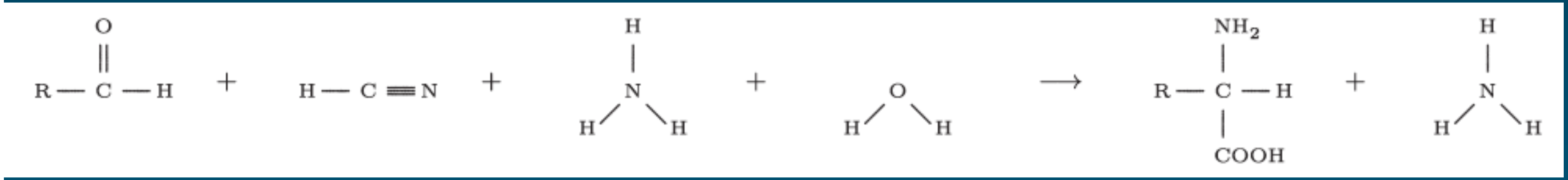
- Liquid water appears after 0.6 Myr, and lasts for 5Myr near surface, and until end of simulation at depth (1km)

LOTS OF TIME AVAILABLE TO COME TO EQUILIBRIUM



Temperature at 1 km and 40 km from centre of body

Pathways:



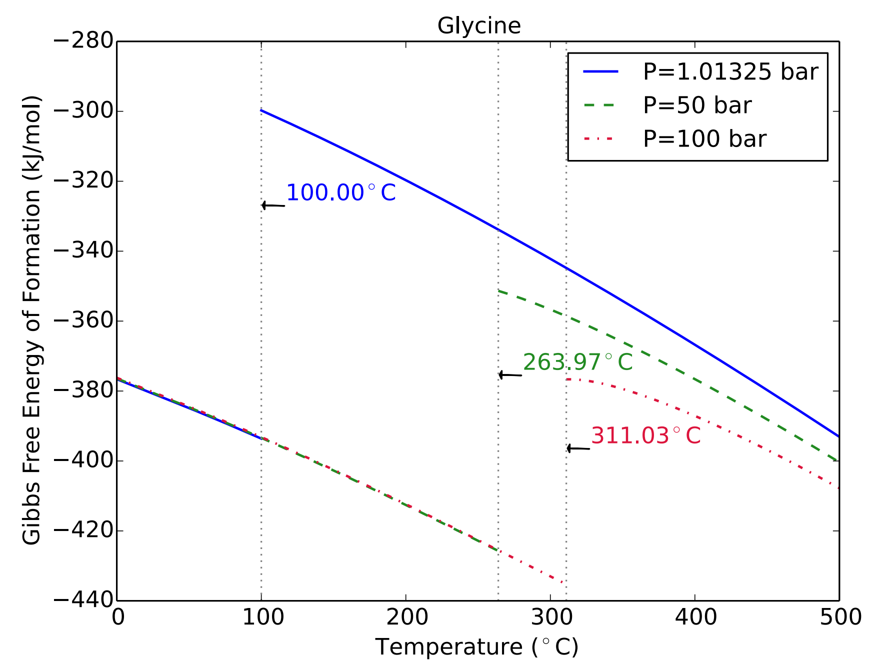
Strecker synthesis for amino acids:

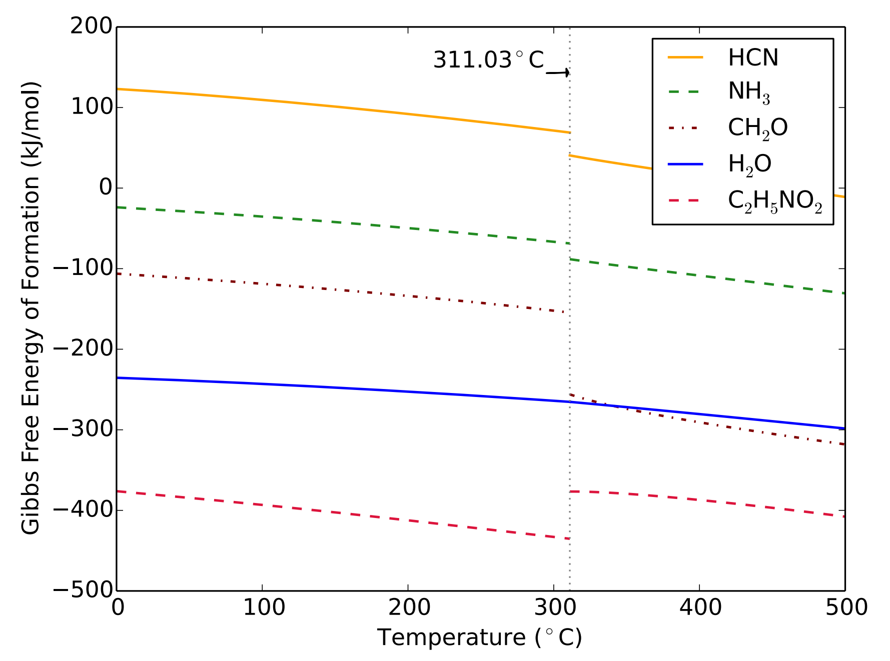
aldehyde + HCN + H_2O in presence of NH_3 -> Amino acid + NH_3

Eg. Formaldehye <-> Glycine
Acetaldehyde <-> Alanine
Glycoladehyde <-> Serine

Initial abundances: from cometary data (pristine material)

Eg. Glycine synthesis (gory details…)

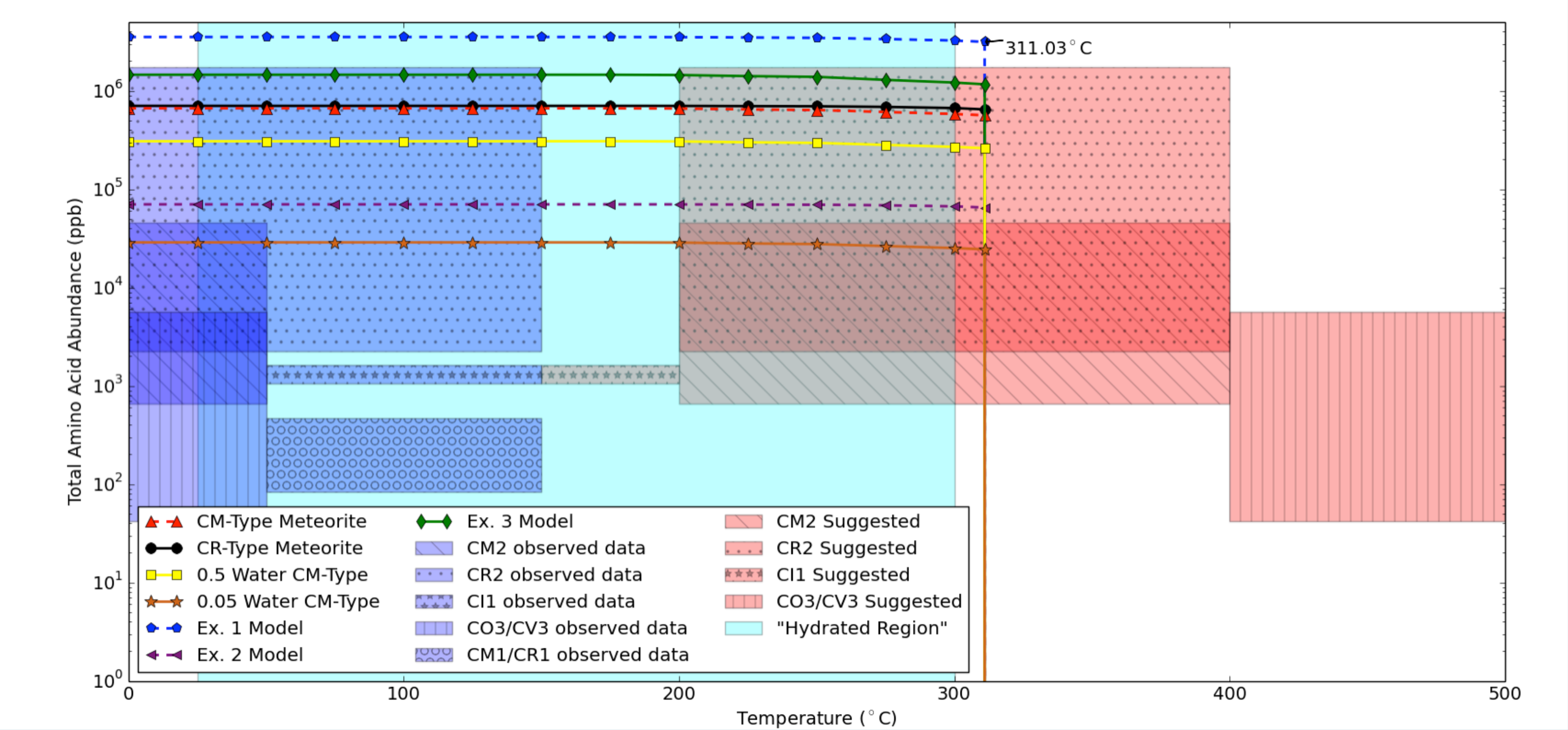




G(T, P) for glycine. In aqueous phase – no influence of pressure (water is incompressible fluid)

Gibbs free energy of formation for reactants @ 100 bar. Vertical line is boiling point of water at 100 bar.

Amino acid yields (ppb) compared to data constraints



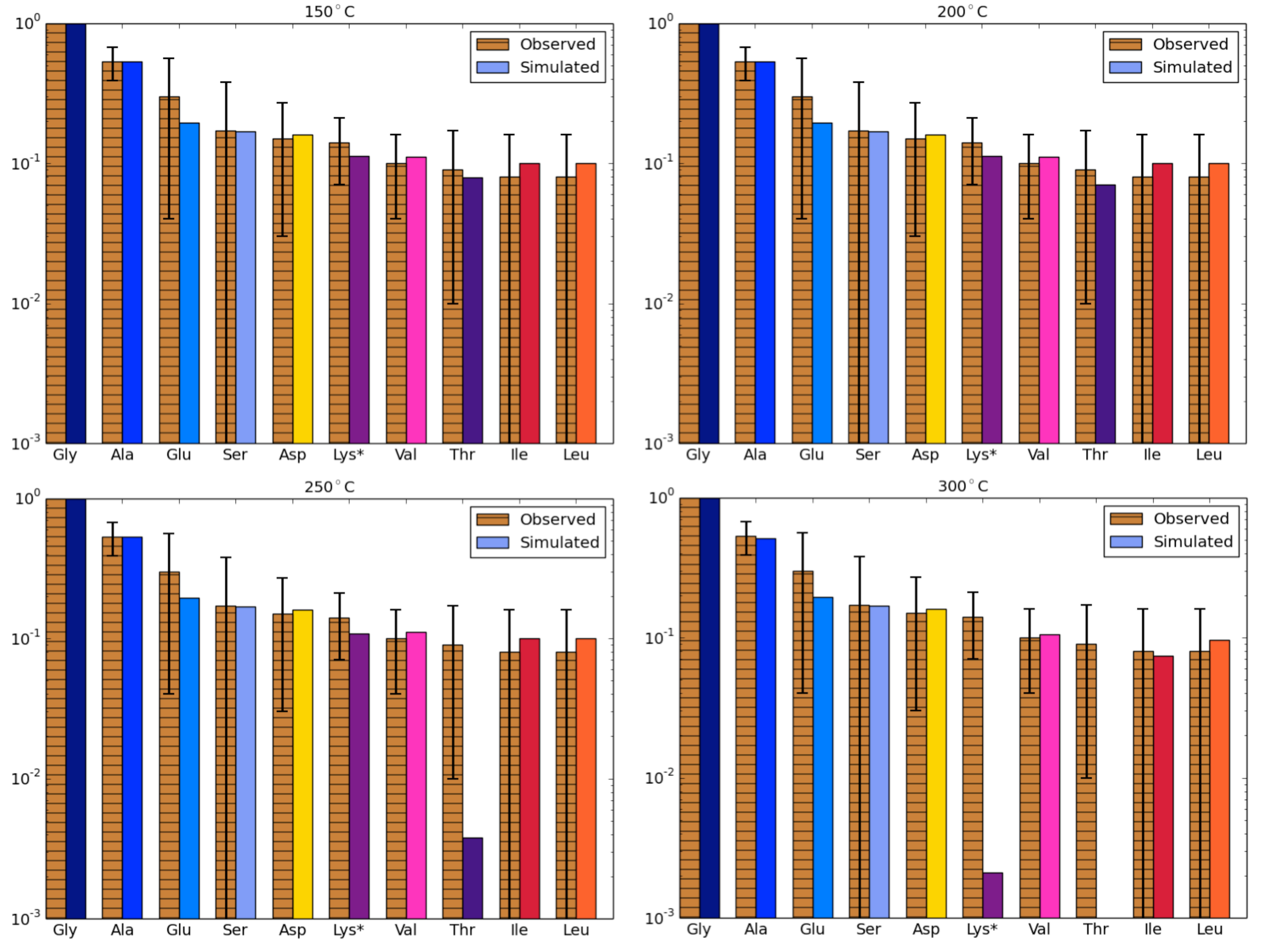
Total amino acids (gly, ala, glu, val, thr, leu, ile, lys). Data constraints from earch carbonaceous chondrite subclass. Red shaded areas suggested constraints by Sephton 2002. Ex 1-3 models for varying X/ H_2O where X is aldehyde concentration.

Amino acid frequencies for CM2 meteorites..

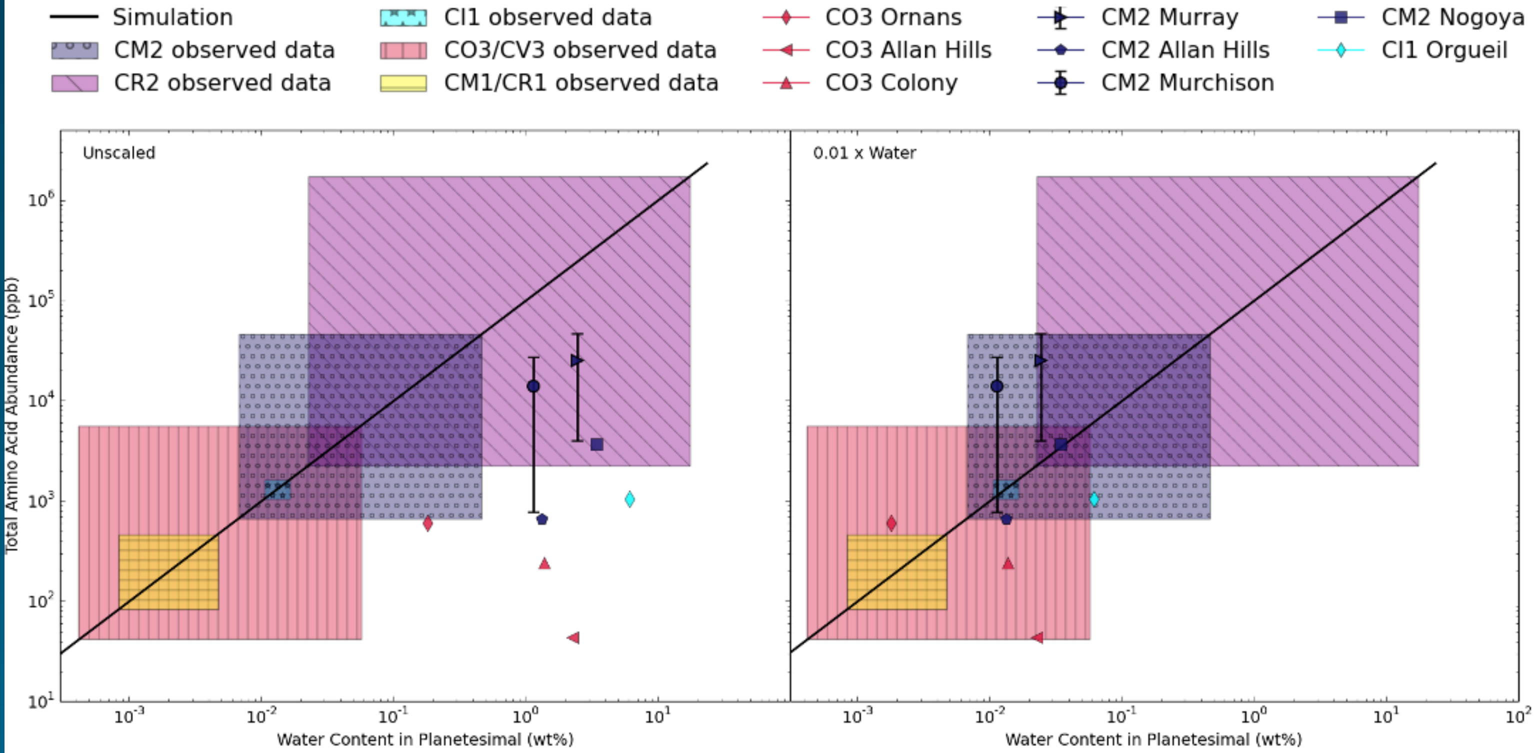
Theoretical vs observed AA frequencies.

Beyond 200 C, see breaks with the data.

Good agreement with planetesimal model.



Effects of varying water content… planetesimals from different regions of a disk



Coloured squares: measured % uncombined water content. Weathering effects affect this (water absorbed on Earth)

Right panel; correcting for weathering by factor of 1/100.

Our model – AAs reflect planetesimals formed in different chemical regions of the disk…

- Water content reflects differentiation with respect to water ice line

- Hydroxy acid (OH instead of NH_2_ & formed by Strecker pathway) to amino acid ratio changes:

  - low NH_3 (inner disk) -> lower AA -> CM class

- Overall, loss of AAs beyond 200 C reflects max inner temperature of planetesimals.

IV. Nucleobases

- purines (guanine + adenine; G & A)
- pyrimidines (cytosine, thymine, and uracil; C, T, U)

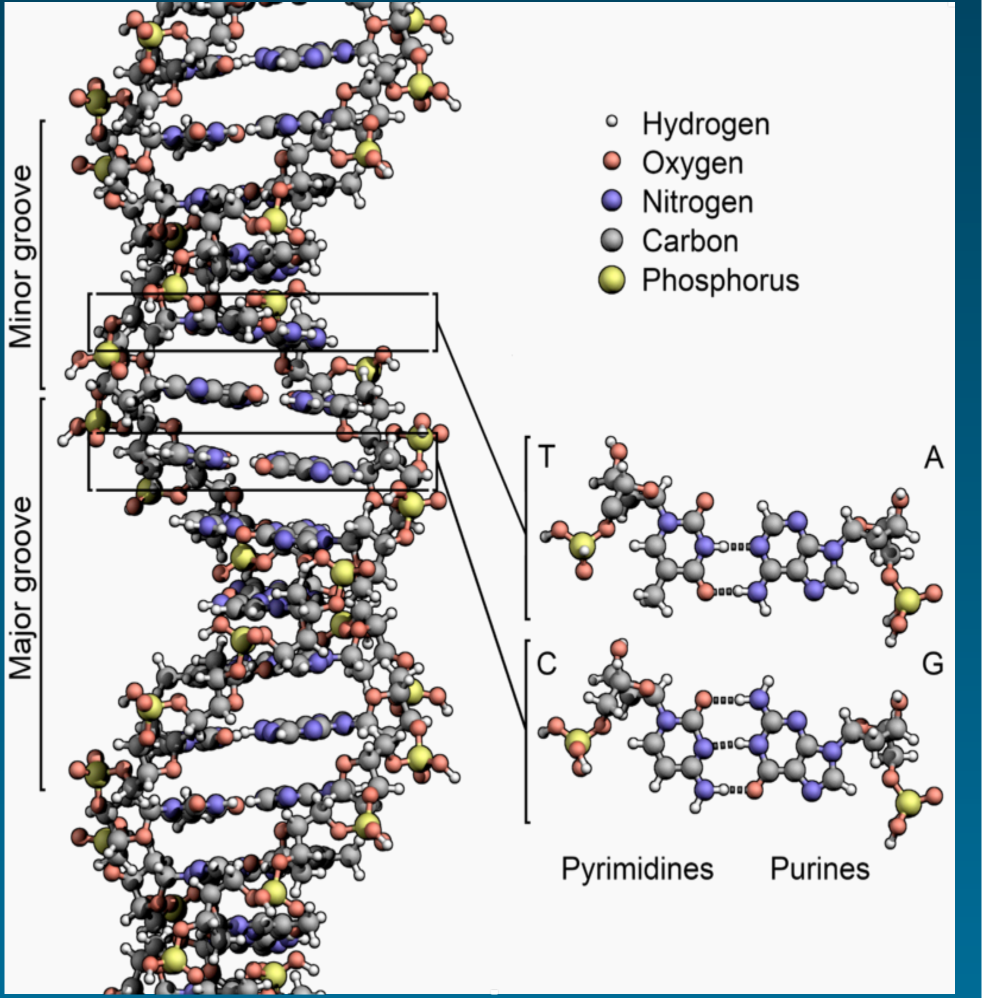
Base pairs:

RNA, G-C and A-U;

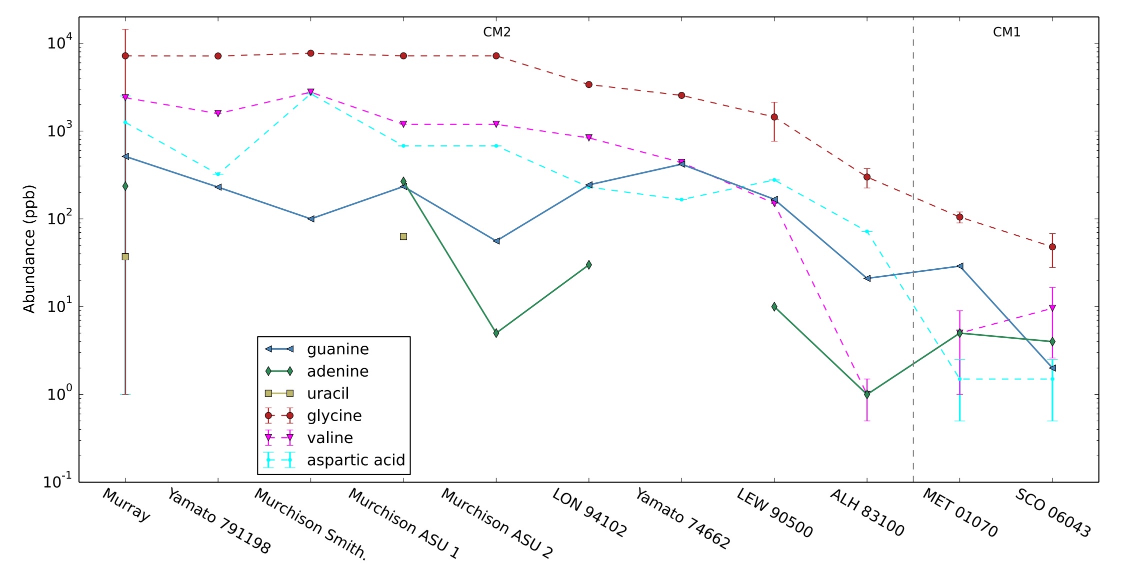
DNA, G-C and A-T.

Nucleotides: (base + sugar + phosphate) – hard to synthesize (eg. Powner + 2009)

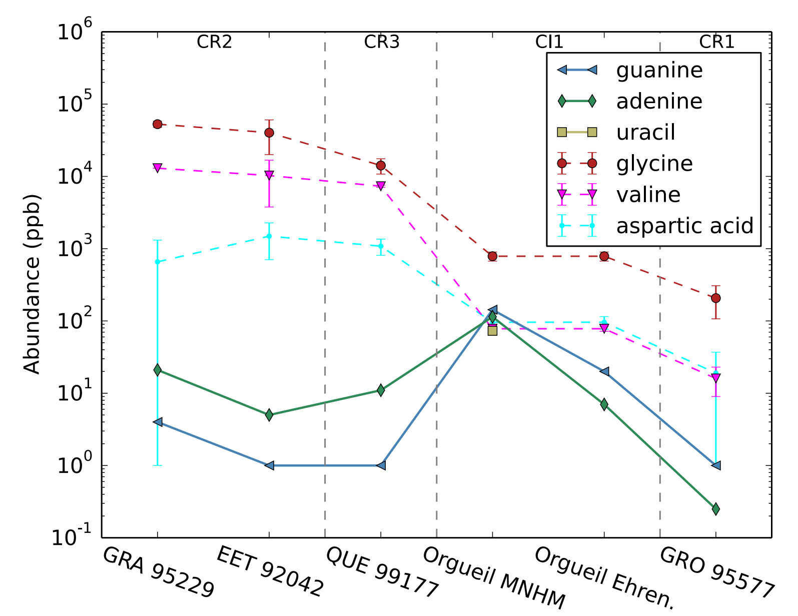
Meteorites: C and T Absent! WHY?



1. Nucleobase data (Pearce & Pudritz 2015, ApJ)

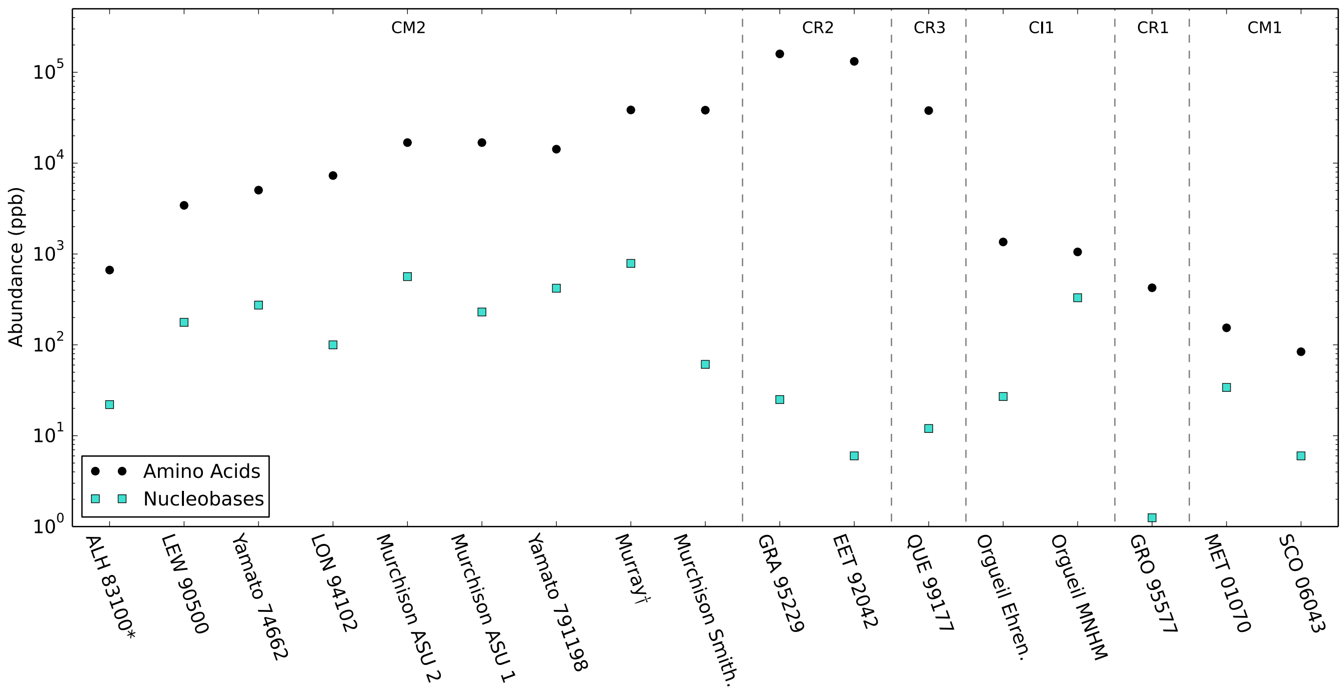


CMs



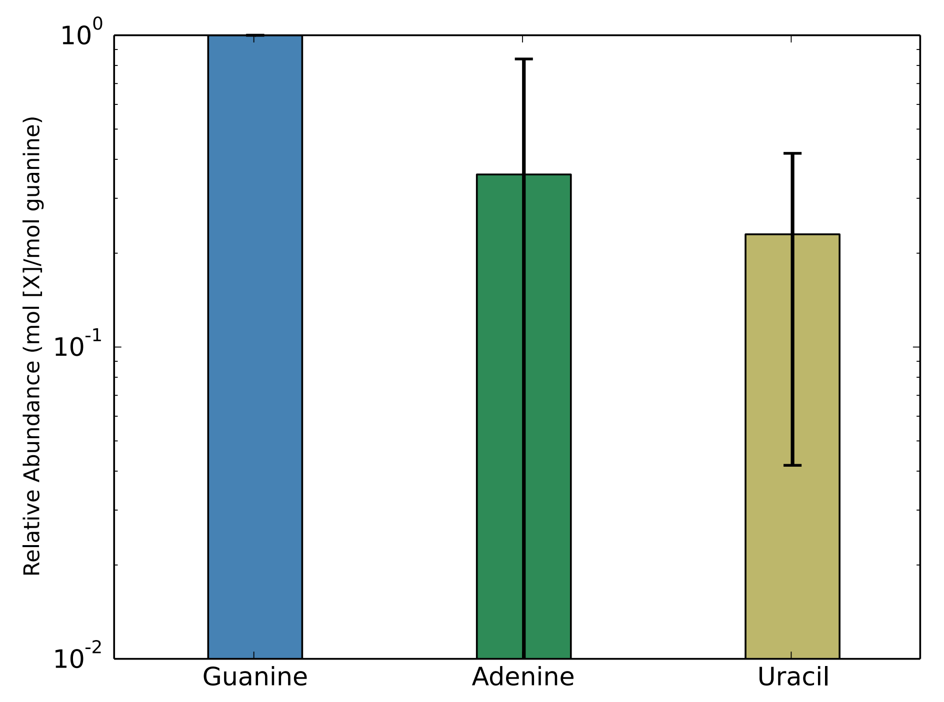
CRs and CIs

Comparing nucleobase with AA patterns



Total nucleobases from 17 meteorite samples – bottom curve. Total Aas for same samples - top

Relative frequencies of nucleobases in CM2s



2. Nucleobase synthesis in planetesimals

(Pearce & Pudritz 2016, submitted *Astrobiology*)

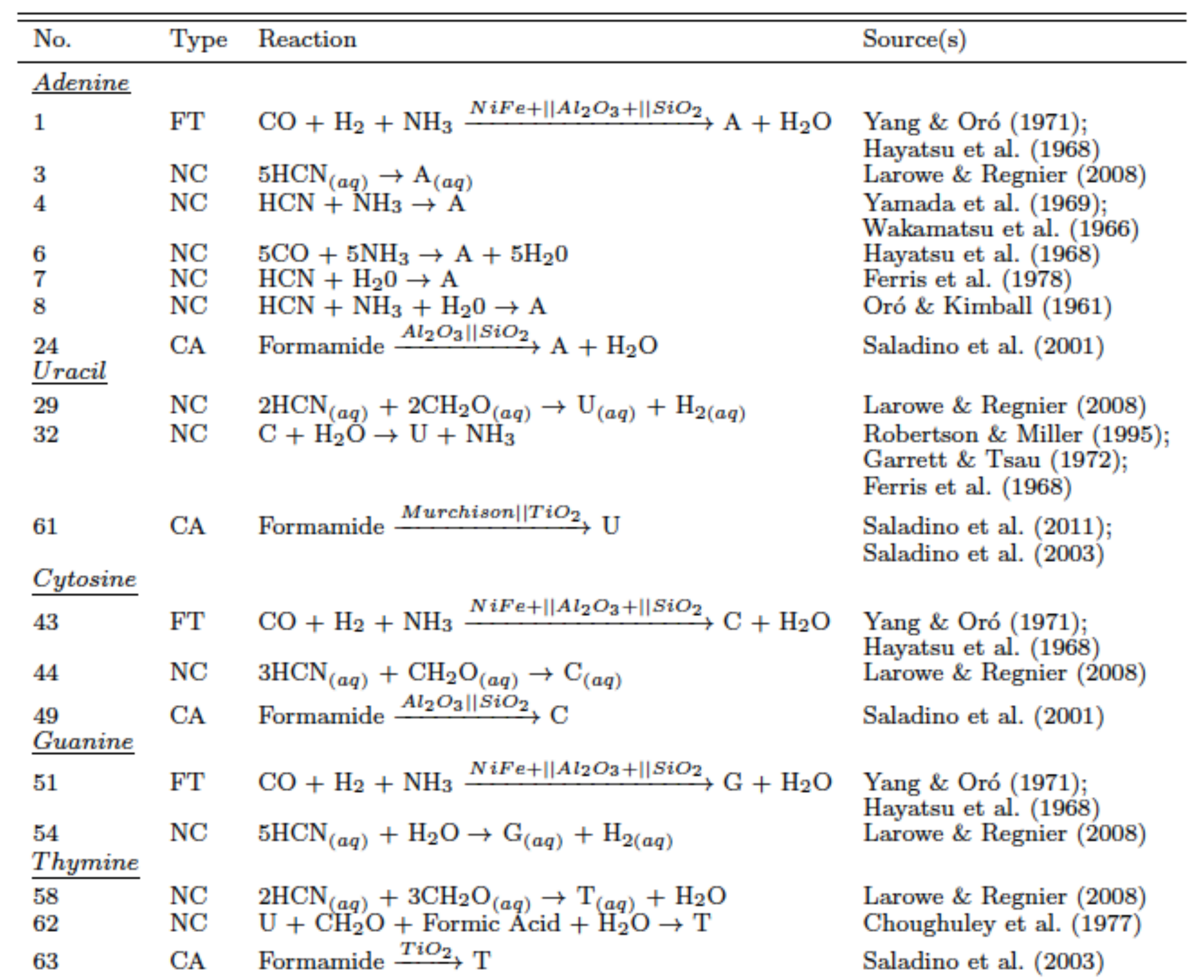
- 3 Classes of reactions (we gathered 68 from the literature)

- Fischer-Tropsch (FT): H_2, CO, NH_3 in presence of a catalyst (eg. nickel-iron alloy) can make all 5

- Non-catalytic (NC): reactants, usually involving HCN heated and cooled, no catalyst. Can make all 5

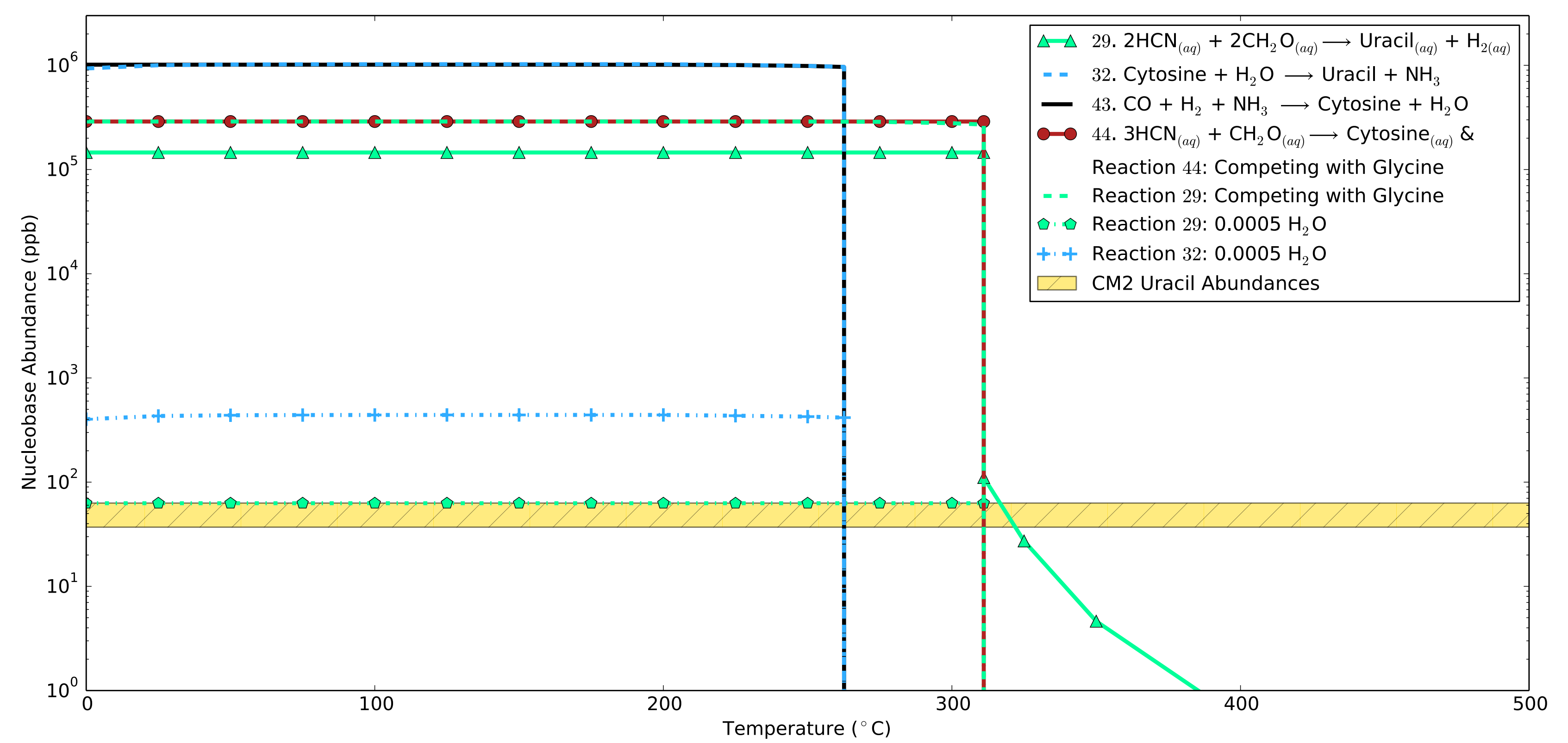
- Catalytic (CA): typically use formamide as sole reactant in presence of catalysts

- Total of ~ 18 reactions relevant for planetesimal conditions (eg. range in T, no UV, …)



No.	Type	Reaction	Source(s)
Adenine			
1	FT	$CO + H_2 + NH_3 \xrightarrow{NiFe+\|\|Al_2O_3+\|\|SiO_2} A + H_2O$	Yang & Oró (1971); Hayatsu et al. (1968)
3	NC	$5HCN_{(aq)} \rightarrow A_{(aq)}$	Larowe & Regnier (2008)
4	NC	$HCN + NH_3 \rightarrow A$	Yamada et al. (1969); Wakamatsu et al. (1966)
6	NC	$5CO + 5NH_3 \rightarrow A + 5H_20$	Hayatsu et al. (1968)
7	NC	$HCN + H_20 \rightarrow A$	Ferris et al. (1978)
8	NC	$HCN + NH_3 + H_20 \rightarrow A$	Oró & Kimball (1961)
24	CA	$\text{Formamide} \xrightarrow{Al_2O_3\|\|SiO_2} A + H_2O$	Saladino et al. (2001)
Uracil			
29	NC	$2HCN_{(aq)} + 2CH_2O_{(aq)} \rightarrow U_{(aq)} + H_{2(aq)}$	Larowe & Regnier (2008)
32	NC	$C + H_2O \rightarrow U + NH_3$	Robertson & Miller (1995); Garrett & Tsau (1972); Ferris et al. (1968)
61	CA	$\text{Formamide} \xrightarrow{Murchison\|\|TiO_2} U$	Saladino et al. (2011); Saladino et al. (2003)
Cytosine			
43	FT	$CO + H_2 + NH_3 \xrightarrow{NiFe+\|\|Al_2O_3+\|\|SiO_2} C + H_2O$	Yang & Oró (1971); Hayatsu et al. (1968)
44	NC	$3HCN_{(aq)} + CH_2O_{(aq)} \rightarrow C_{(aq)}$	Larowe & Regnier (2008)
49	CA	$\text{Formamide} \xrightarrow{Al_2O_3\|\|SiO_2} C$	Saladino et al. (2001)
Guanine			
51	FT	$CO + H_2 + NH_3 \xrightarrow{NiFe+\|\|Al_2O_3+\|\|SiO_2} G + H_2O$	Yang & Oró (1971); Hayatsu et al. (1968)
54	NC	$5HCN_{(aq)} + H_2O \rightarrow G_{(aq)} + H_{2(aq)}$	Larowe & Regnier (2008)
Thymine			
58	NC	$2HCN_{(aq)} + 3CH_2O_{(aq)} \rightarrow T_{(aq)} + H_2O$	Larowe & Regnier (2008)
62	NC	$U + CH_2O + \text{Formic Acid} + H_2O \rightarrow T$	Choughuley et al. (1977)
63	CA	$\text{Formamide} \xrightarrow{TiO_2} T$	Saladino et al. (2003)

Cytosine synthesized – but quickly decomposes in water by deamination to Uracil and NH_3



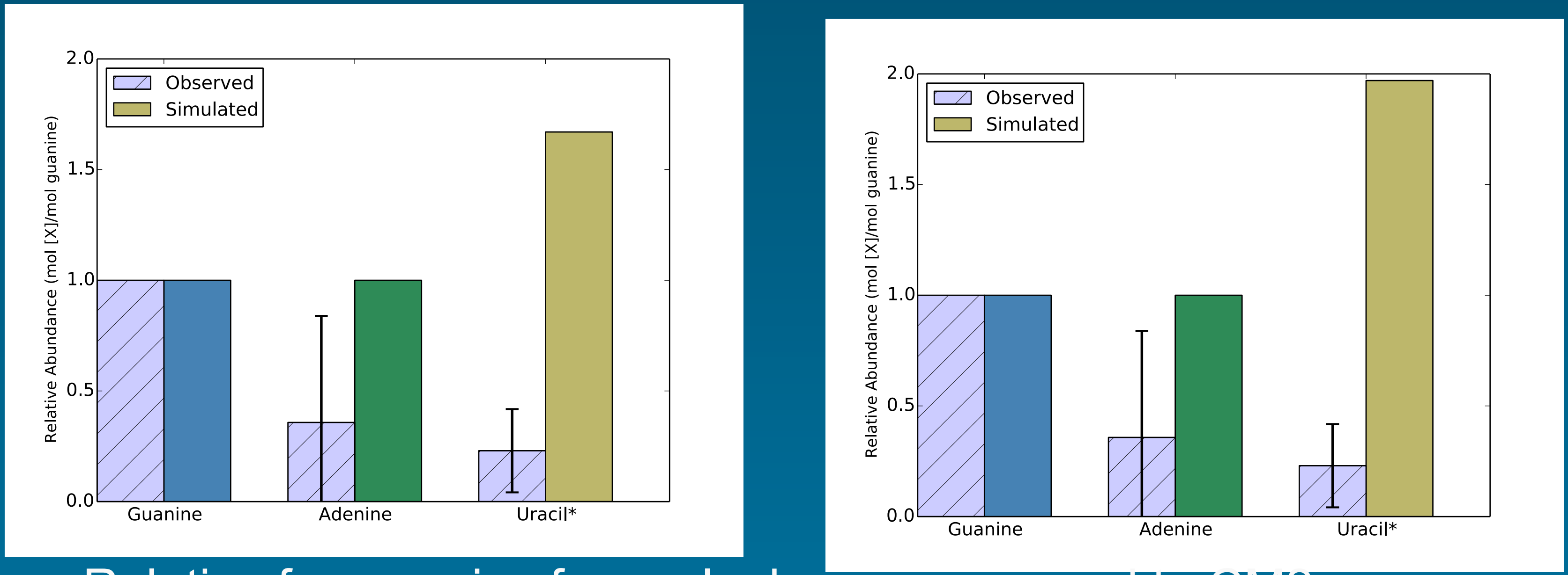
Deamination in 17,000 yr (Levy & Miller 1998)

Results:

- Uracil over produced by destruction of cytosine through deamination

- Thymine is produced readily by NC reaction from U, formic acid, and formaldehyde (eg reaction 62)
  BUT - molecule is quickly destroyed by H_2O_2 at 120^o C (Shadyro + 2008)
  Hydrogen peroxide is observed in comets…

- Most favourable reactions: FT or NC, involving simple molecules
- - water, ammonia, carbon monoxide, hydrogen cyanide.

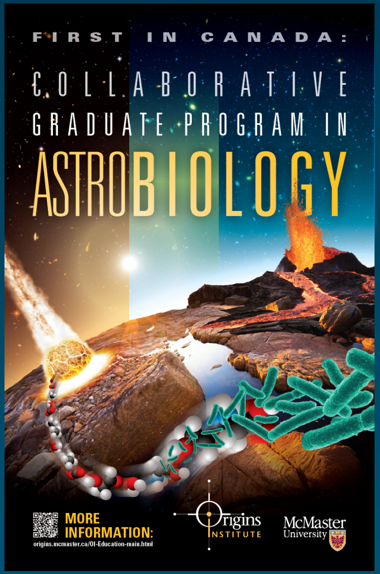


Relative frequencies for nucleobases compared to CM2s:
Left - FT reactions Right – NC reactions

Summary

- Thermodynamic equilibrium in planetesimals provides good description of organics seen in meteorites
- Initial amino acid frequencies <-> thermodynamics conditions early DNA/protein codes
- Universality : planets equipped with similar biomolecular complements -> implications for first codes everywhere?
- Nucleobases C, T formed in other ways?
  - ice grains, and then into comets?
  - direct synthesis of nucleotides on planet surfaces (Powner + 2009) - utilizing UV activation
    – nucleobases need not apply?
- Are there general arguments about nucleotide formation as for AAs?

Origins Institute programs on Origins of Life

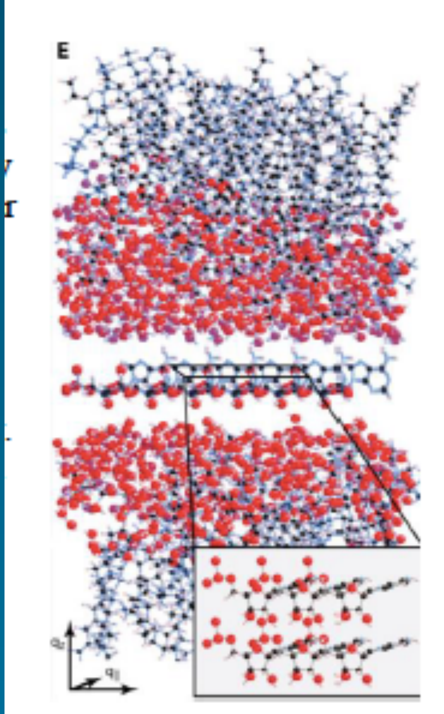


1. Canada's first collaborative graduate program in Astrobiology (launched 1.01.2013):

- Origins Institute + 5 collaborating depts

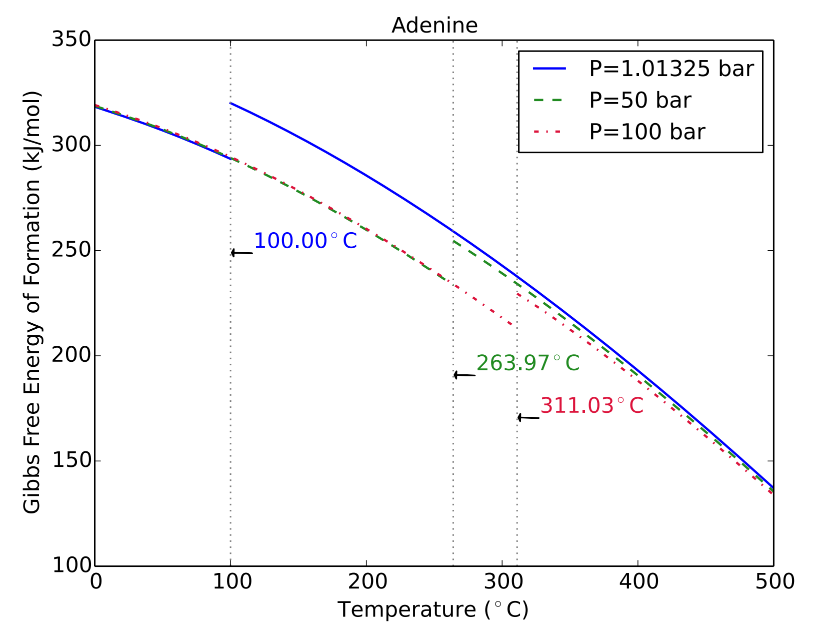
- M.Sc. And Ph.D. program
- "Home Dept" degree + Astrobiology specialization

2. Origins of Life Laboratory funded. Simulation of early planet conditions -> RNA polymerization, RNA world
(M. Rheinstadter (PI), R. Pudritz, Y-F Li)

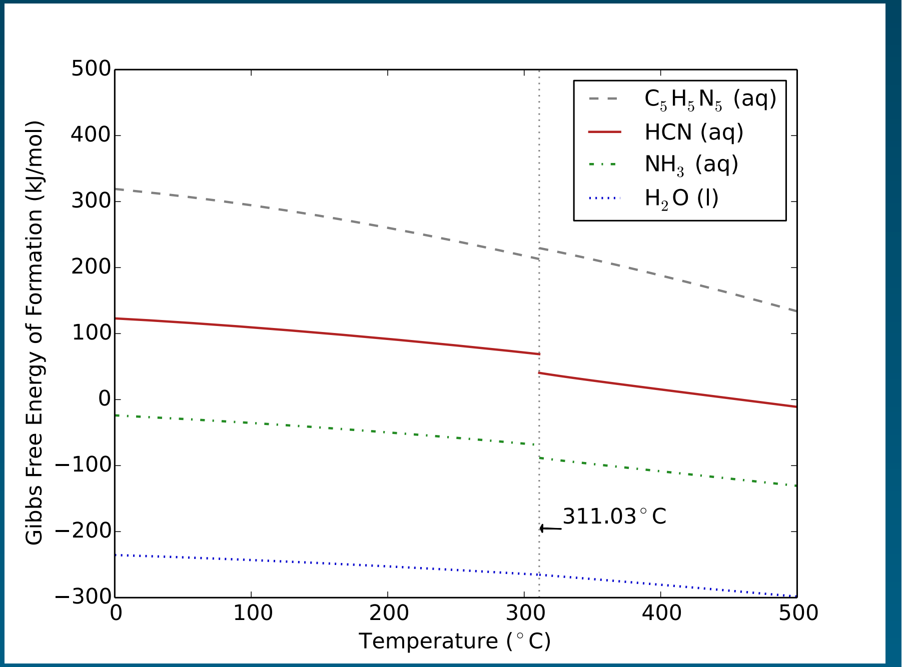


Toppozini + 2013

Eg. Synthesis of Adenine



Gibbs free energy dependence on T and P - Adenine



NC synthesis of adenine – reaction 8