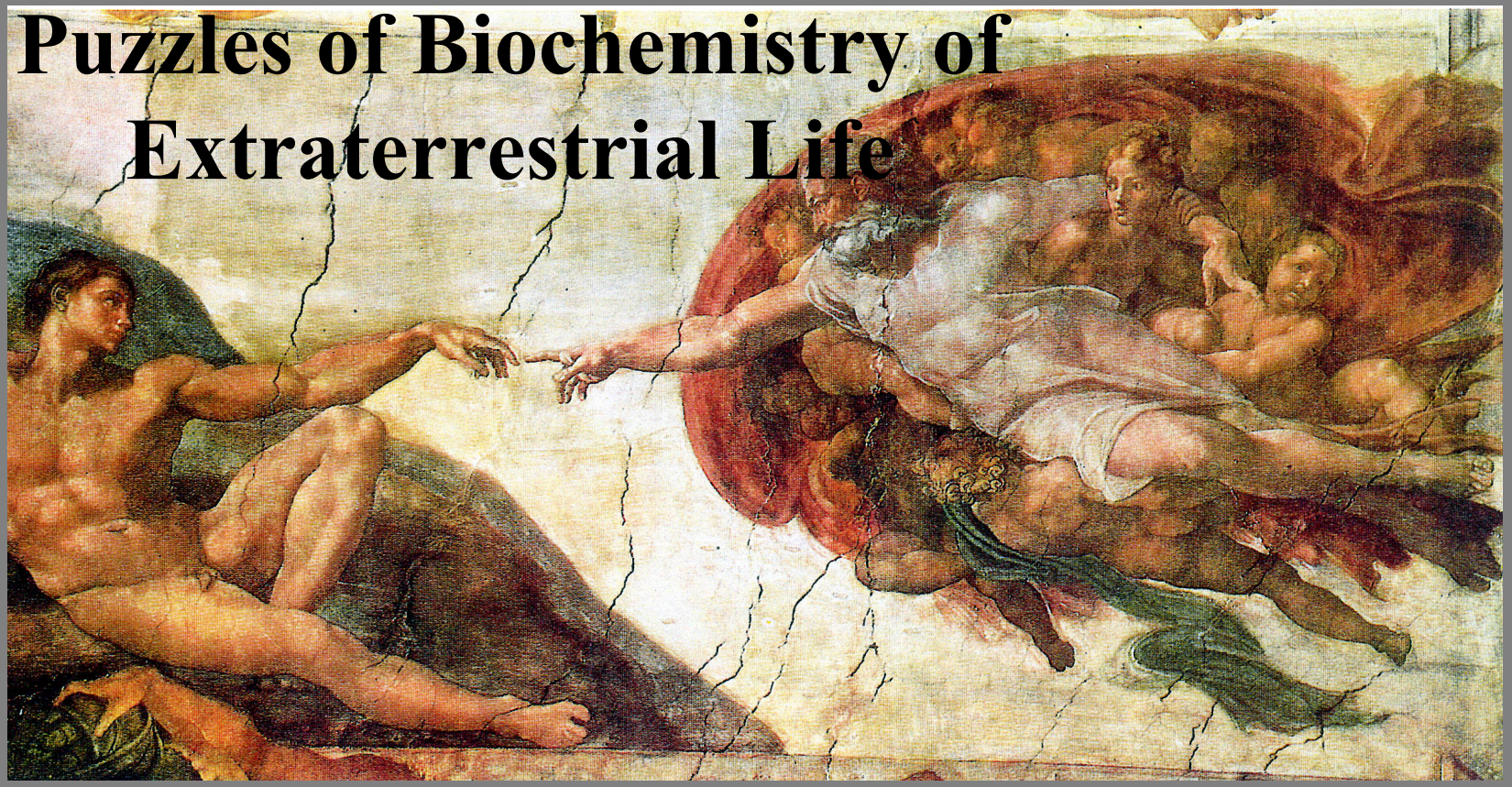
Puzzles of Biochemistry of Extraterrestrial Life



Tairo Oshima

Institute of Environmental Microbiology, Kyowa-kako Co., Machida, Tokyo, Japan

Terrestrial life is an informational machine consisting of nucleic acids as software and proteins as hardware

	Software	Hardware
Cell constituents	Genes	Enzymes Membranes
Chemical entity	Nucleic acids	Proteins Lipids
Main components	Bases, Sugars, Phospahte	Aminoaicds, Fatty acids, Glycerol
Elements	C,H,O,N,P	C,H,O,N

Life is a chemical machine

- However, elements and molecules are stricltly limited=Biochemical Exclusive Principle

- Only 4 elements are used as major components and they are identical to 4 major “active” elemnts of Universe

- 10 major elements of life is identical to top 10 elements of ocean

Biochemical exclusive principle is essence of life and key to ratiocination of extraterrestrial life

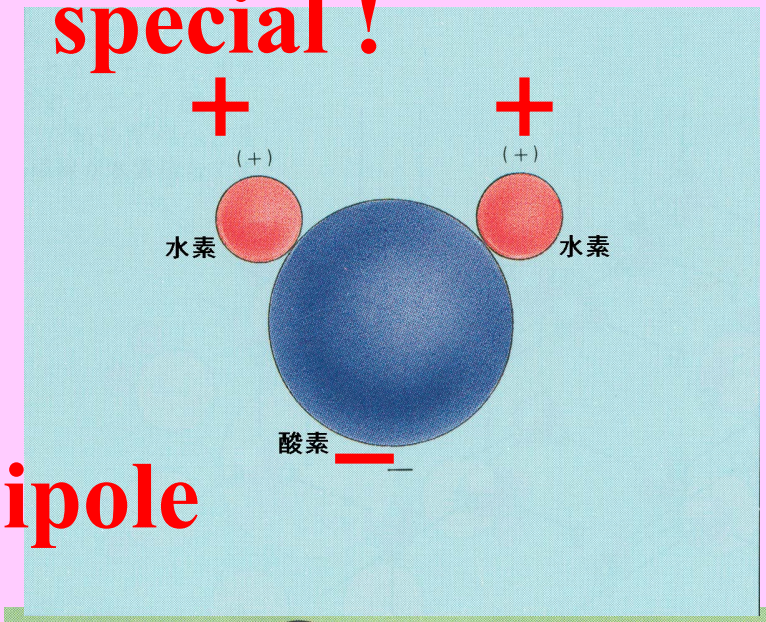
Elemental compositons of life

	Universe	Earth crust	Ocean	Animals	Plants
1	H	O	H	H	H
2	He	Si	O	O	O
3	O	H	Cl	C	C
4	C	Al	Na	N	N
5	N	Na	Mg	Ca	P
6	Ne	Fe	S	P	Ca
7	Mg	Ca	Ca	S	K
8	Si	Mg	K	Na	Mg
9	Al	K	C	K	S
10	Fe	Ti	N	Cl	Cl
11	S	P		Mg	Na

Why H and O ?

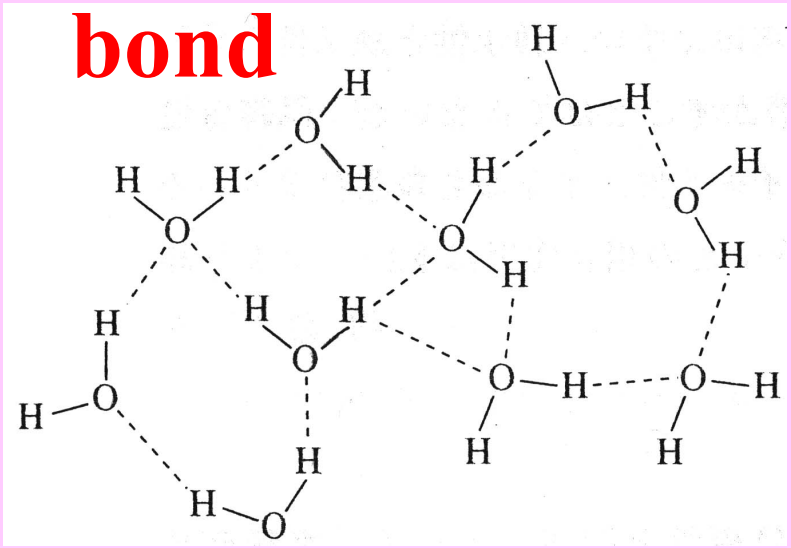
- Water

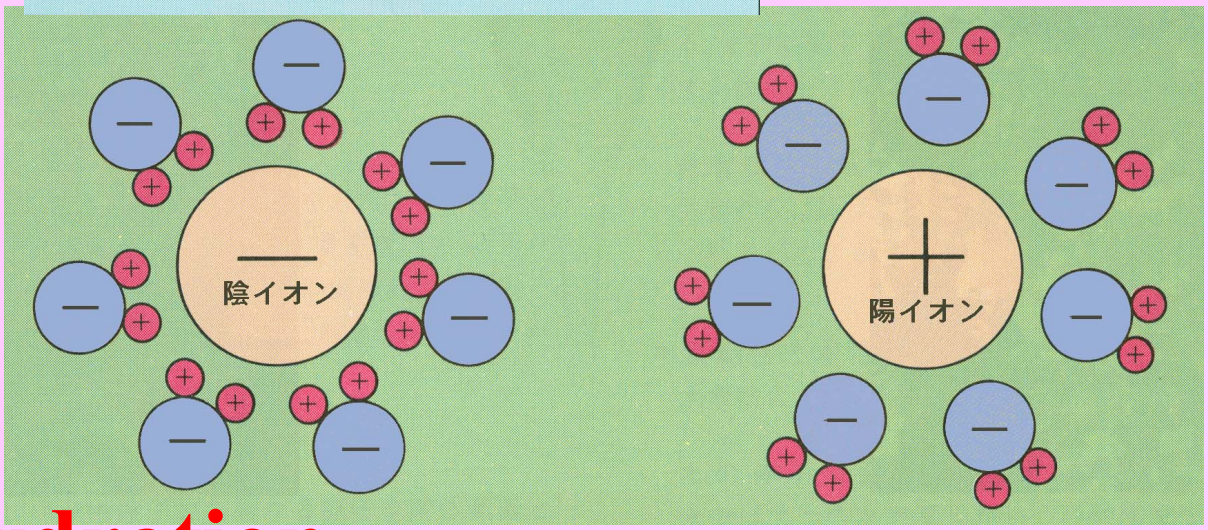
Water is special !



Dipole

Hydrogen bond





Hydration

Liquid at room temperature

Large sp heat=homeostasis

Hydrophobic bond

No water, no life

2003年6月、インドのグジャラート州は、過去数十年間で最も厳しい
干ばつに見舞われた。焼け付くような暑さの中、人々はバケツを持ち、
水の出る泉を求めて何kmも歩く。

Why Carbon ?

- Chemical valency is 4 → create a variety of compounds

- C-C bond is stable, Si-Si bond is unstable → cannot create long fattay acids

Speculations

- H, O, C, N would be major elements for Ets
  Si cannot act as a substitute for C

- P would be an important element for Ets

- H2O would be solvent of Ets' cells

Biochemical exclusive principle

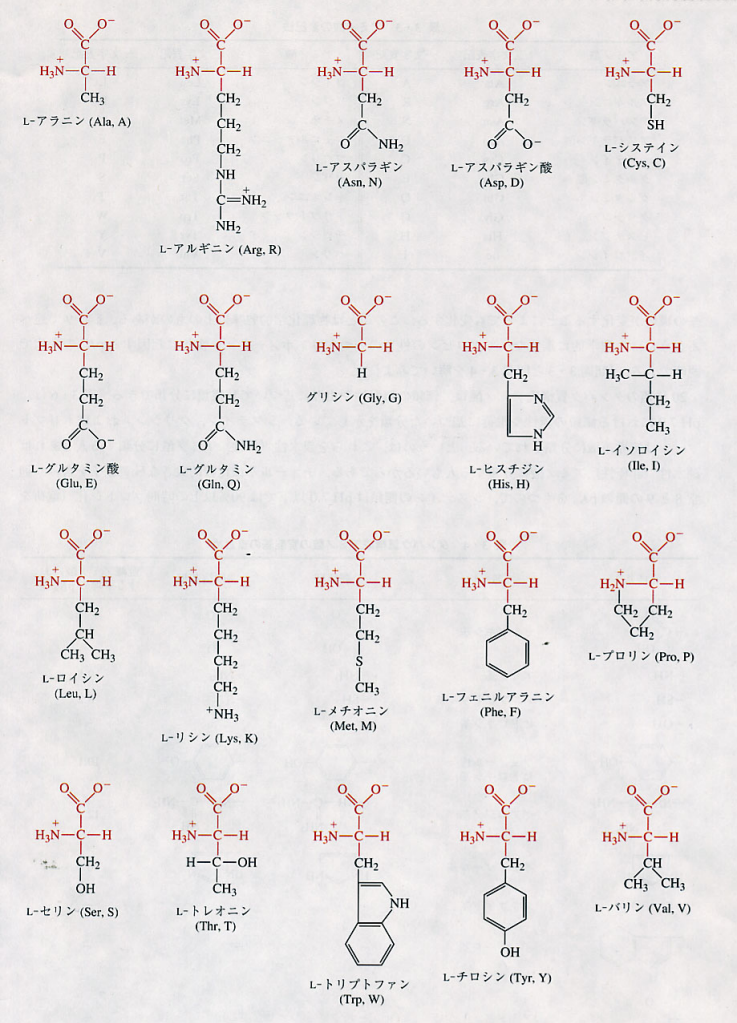
Proteins are synthesized with 20 α-amino acids

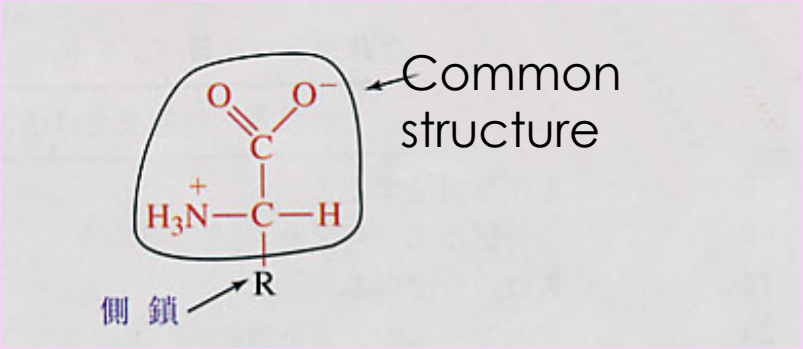
By post-translational modifications, proteins contains 30- different aminoacid residues, and these modified amino acids are extremely important for protein function (examples are phosphorylated aminoacids, hydroxyproline in collagen, etc)

Nucleic acids are synthesized with 4 nucleotides

Likewise, many different nucleotide residues are found in DNAs and RNAs

20 protein amino acids = Magic 20





Using Ray's results,[2] the proof also generalizes immediately to the case where particles are killed at a rate $V(x)$ throughout the region R as well as at the boundary. Only slightly more than continuity of $V(x)$ almost everywhere in R is required. Similar results are expected to hold for the elastic-barrier case.

[1] M. Kac, "On Some Connections between Probability Theory and Differential and Integral Equations," *Proc. Second Berkeley Symposium Math. Statistics and Probability*, pp. 189–215, 1951.
[2] D. Ray, "On Spectra of Second-Order Differential Operators," *Trans. Am. Math. Soc.* **77**, 299–321, 1954.
[3] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1 (New York: Interscience Publishers, Inc., 1953).

CODES WITHOUT COMMAS

BY F. H. C. CRICK, J. S. GRIFFITH, AND L. E. ORGEL

MEDICAL RESEARCH COUNCIL UNIT, CAVENDISH LABORATORY, AND DEPARTMENT OF THEORETICAL CHEMISTRY, CAMBRIDGE, ENGLAND

Communicated by G. Gamow, February 11, 1957

This paper deals with a mathematical problem which arose in connection with protein synthesis. We present the solution here because it gives the "magic number" 20, so that our answer may perhaps be of biological significance. To make this clear, we sketch in the biochemical background first.

It is assumed in one of the more popular theories of protein synthesis that amino acids are ordered on a nucleic acid strand (see, for example, Dounce[1]) and that the order of the amino acids is determined by the order of the nucleotides of the nucleic acid. There are some twenty naturally occurring amino acids commonly found in proteins, but (usually) only four different nucleotides. The problem of how a sequence of four things (nucleotides) can determine a sequence of twenty things (amino acids) is known as the "coding" problem.

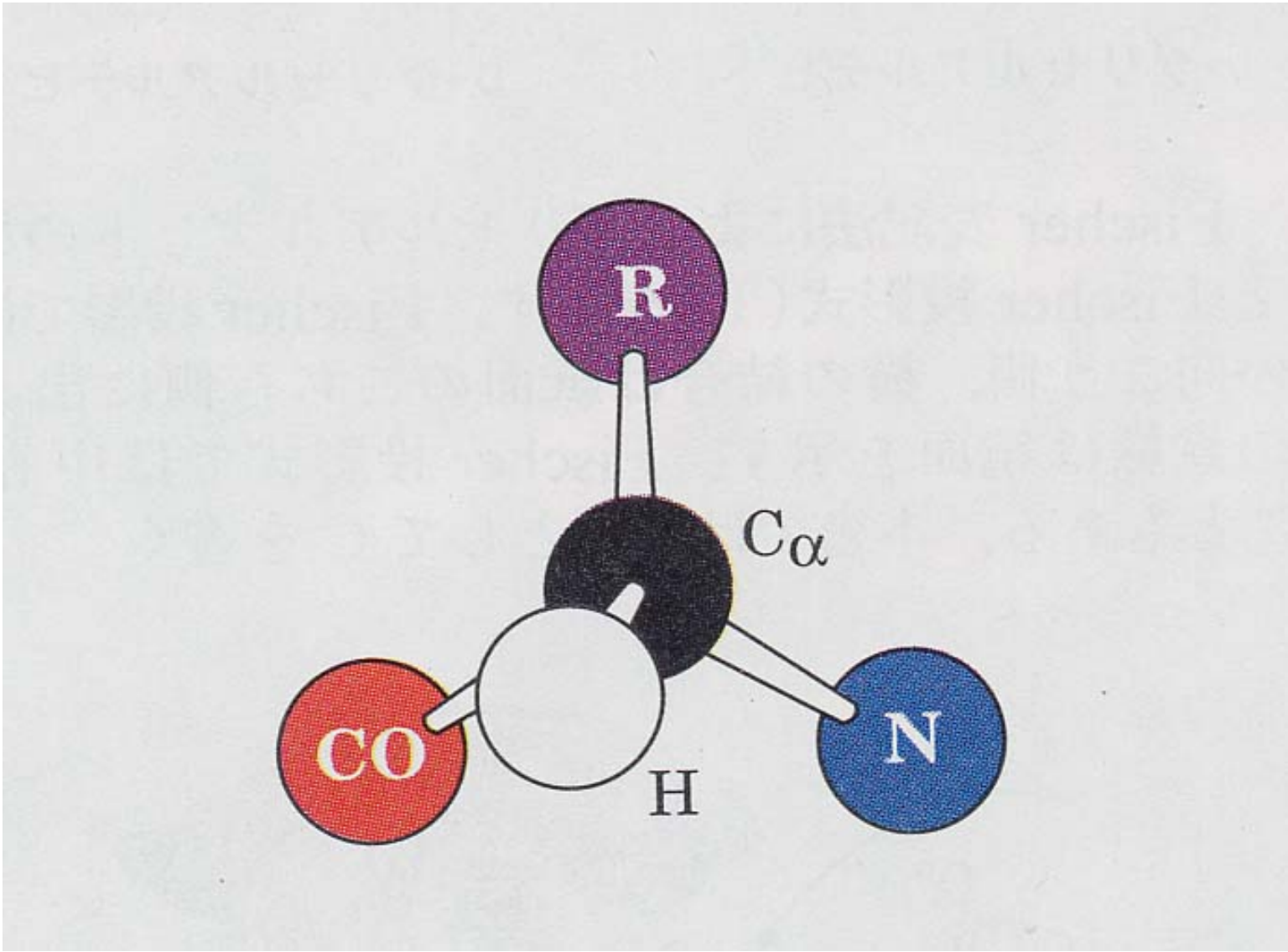
This problem is a formal one. In essence, it is not concerned with either the chemical steps or the details of the stereochemistry. It is not even essential to specify whether RNA or DNA is the nucleic acid being considered. Naturally, all these points are of the greatest interest, but they are only indirectly involved in the formal problem of coding.

The first definite proposal was made by Gamow.[2] His code, which was suggested by the structure of DNA, was of the "overlapping" type. The meaning of this is illustrated in Figure 1. Gamow's code was also "degenerate"—that is, several sets of three letters (picked in a special way) stood for a particular amino acid. However, all the 64 ($4 \times 4 \times 4$) possible sets of three letters stood for one amino acid or another, so that any sequence whatever of the four letters stood for a definite sequence of amino acids.

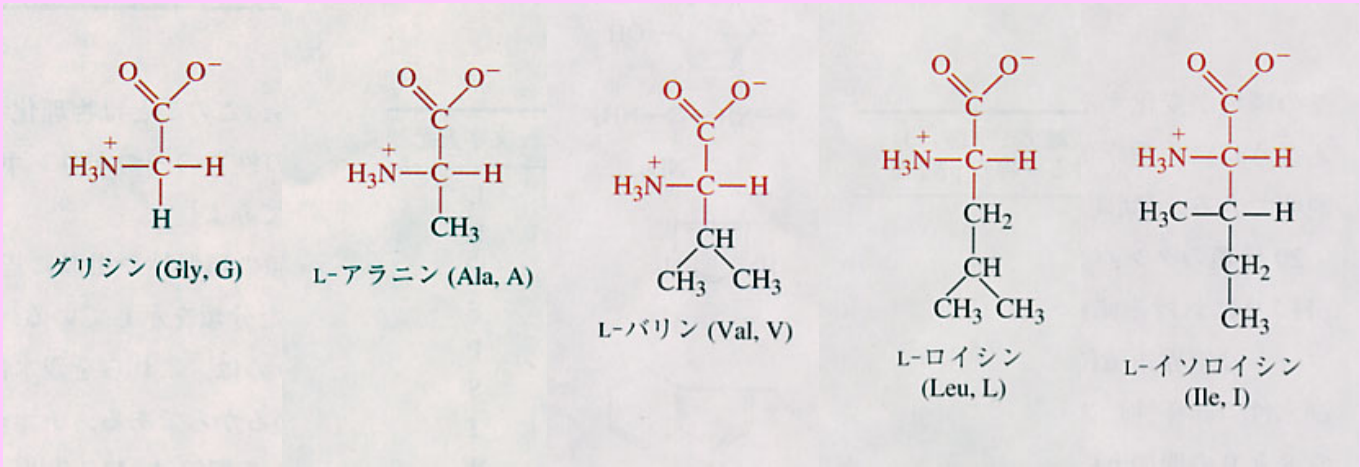
It is easy to see that codes of the overlapping type impose severe restrictions on the allowed amino acid sequences. Unfortunately, no such restrictions have been found, although considerable (unpublished) efforts have been made, by a number of workers, to find them. Part of this work has been reviewed by Gamow, Rich,

Magic20

L-amino acid



Very difficult to find out reason why they were selected



Very difficult to find out reason why they were selected

グルタミン酸　$HOOC\text{-}CH(NH_2)\text{-}CH_2\text{-}CH_2\text{-}COOH$

アスパラギン酸　$HOOC\text{-}CH(NH_2)\text{-}CH_2\text{-}COOH$

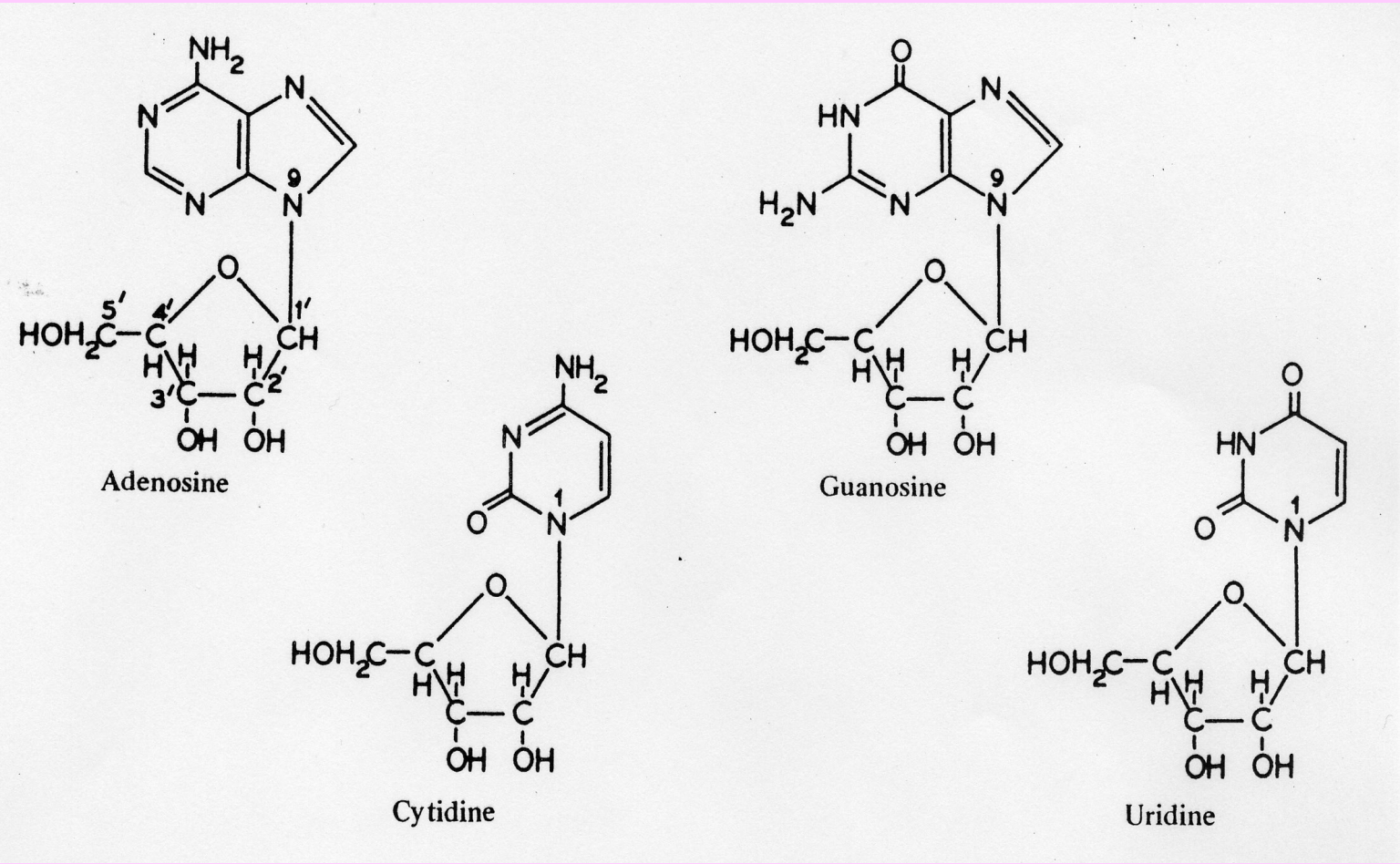
リシン　$HOOC\text{-}CH(NH_2)\text{-}CH_2\text{-}CH_2\text{-}CH_2\text{-}CH_2\text{-}NH_2$

オルニチン　$HOOC\text{-}CH(NH_2)\text{-}CH_2\text{-}CH_2\text{-}CH_2\text{-}NH_2$

Are 20 mino acids essential for a functional protein?

- No !
- Functional proteins have been chemically synthesized using 5-9 amino acids (JBC 280, 37742; Nature Struct. Biol. 4, 1039; etc)
- Cystein is missing in many proteins produced by aerobic, extreme thermophiles, since C-S bond is unstable at high temperature under aerobic conditions

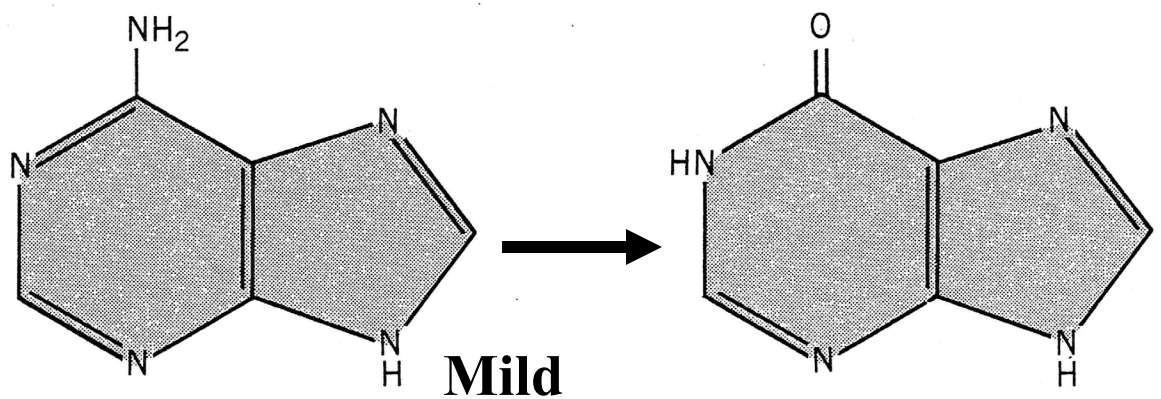
Nucleic acids are synthesized using 4 nucleotides



Why 4 ? Two is enough; Languages consisting 2 letters are perfect

- Computers and Morse code are 2-letter languages
- Why genetic system is 4 letter code?
- 2 letter system (A & T or G & C) is enough for DNA double stranded struclture

Most reasonable 2 letter code for genetic system would be A·I letter system



A

Mild oxidation, possible in the primitive ocean

I

Speculations 1

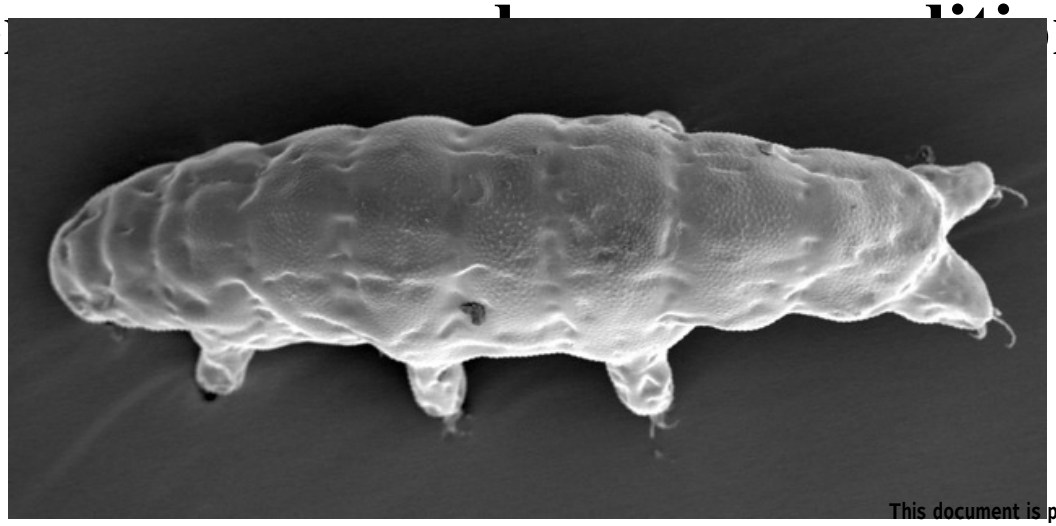
- H, O, C, N would be major elements for ETs
- P would be an important element for ETs
- H_2O would be solvent of Ets' cells
- For catalic functions, proteins=polymers of amino acids would be ideal, but Ets'protein amino acids could be more than 20 or less than 20. D-amino acids could be used by some of ETs
- For genetic information storage, nucleic acid double helical structure would be ideal, but ET's nucleic acids may not be synthesized using A, C, G, T(or U) nucleotides

The final question on biochemistry of extraterrestrial life is “life span”

- Biochemical basis of determination of life span (or generation time) is totally unknown
- Life span of organisms on our planet varies from a few hours to a few hundred years
- Some of unculturable organisms (especially those located in deep undergrand) may has abnormally long generation time

Life Span

- **It seems that some organisms can across the boarder between life and death**
- **Some microbes under natural conditions are dormant, and unculturable since we do not know haw to revive**
- **Most of bacteria and water bears and some others can survive under dried conditions. Water bears can revive after [illegible] for long time**



Speculations 2

- It might be very difficult to detect life on other planet if
- its biochemistry is entirelly different from ours
- or its life span (or generation time) is abnormally long
- or we donot know how to revive dormant cells

Thank you for your kind attention

The picture shows my institute, only 10 min drive form here