

CRITERIA FOR DISTINGUISHING BIOGENIC AND ABIOGENIC AMINO ACIDS – PRELIMINARY CONSIDERATIONS*

KEITH A. KVENVOLDEN

Planetary Biology Division, NASA-Ames Research Center, Moffett Field, Calif. 94035, U.S.A

(Received 3 March, 1972)

Abstract. Criteria to determine the mode of origin of amino acids can be established by consideration of their structure, enantiomeric distribution, composition, and relative abundance. A population of dominantly protein amino acids with one enantiomeric configuration most likely had a biological origin. Biological amino acids do racemize, however, so the absence of optical activity would not rule out the possibility that the amino acids in a racemic mixture were originally synthesized biologically. For racemic amino acids, therefore, structure, composition and relative abundance become important in ascertaining the origin of these compounds. Abiotically synthesized amino acids have a population composed of both protein and nonprotein structures present as racemic mixtures. N-alkyl compounds are common and α - β - and γ -amino carboxylic acids are expected. Also most amino acid isomers with four or less carbon atoms should be present in an abiotically derived mixture.

For studies of the beginnings of life on Earth and for the design of experiments to search for extraterrestrial life, the ability to distinguish between biogenic and abiogenic molecules is becoming increasingly important. In recent investigations of the geochemistry and cosmochemistry of amino acids, data have been obtained from four sources: (1) organisms of the biosphere, (2) modern and ancient sediments of the lithosphere, (3) carbonaceous chondrites possibly from the Asteroids or from comets, and (4) laboratory experiments simulating prebiotic chemical processes. From these data it is now possible to develop criteria for characterizing amino acid molecules with regard to their biological or non-biological origins. It must be recognized, however, that criteria assigned at this time represent working hypotheses which will be tested and possibly modified as new information about the occurrence and fate of amino acids is gathered.

Amino acids are composed of some of the most abundant reactive elements in the universe. Carbon, hydrogen, oxygen, nitrogen, and sometimes sulfur can theoretically be combined to produce an almost infinite variety of amino acid molecules. In the four major sources of amino acids, however, distinctive populations of molecules are present. Information regarding their origin can be found in four aspects of amino acid molecules: (1) molecular structure of individual amino acids; (2) the distribution of amino acid enantiomers; (3) the composition of a population of amino acids; and (4) the relative abundance of individual amino acids within a population. Similar considerations were originally developed by Eglinton (1969) particularly for lipids. It appears that amino acids carry within their structures and distributional patterns clues to their history, and these clues make possible the establishment of criteria for determining their origin.

* Lunar Science Institute Contribution.

1. Living Systems

Amino acids are common constituents of living organisms. About 26 different amino acid structures are present in proteins which are the 'substance of life' and serve enzymatic and structural functions (Morrison and Boyd, 1969). All of these amino acids are α -amino carboxylic acids and all except glycine contain at least one asymmetric center. Asymmetric amino acids in proteins are always of the L-configuration. A number of other amino acids occur in organisms, however. In fact more than 170 natural acids are known (Meister, 1965), and of these about 140 are nonprotein amino acids which are found free or as components of relatively small molecules such as peptides. Although they are many in number, nonprotein amino acids in organisms are present in much lower concentrations than the more common protein amino acids. Little is known about the function of nonprotein amino acids, but some appear to serve as intermediates and end products of metabolic cycles.

Although there is no conclusive evidence for the presence of D-amino acids in proteins of organisms, D-amino acids do occur both free and in peptide linkages in a variety of microorganisms. Meister (1965) lists 20 different D-amino acids which have been found in bacteria, antibiotics, insects and worms. D-alanine and D-glutamic acid are particularly important constituents of the cell walls of lower protists (Stanier *et al.*, 1963). Like nonprotein amino acids, the D-amino acids are generally present in organisms in much lower abundances than are protein amino acids.

The distribution of amino acids in organisms is variable. Because of the dominance of protein amino acids, these are usually the only ones reported when consideration is given to the bulk composition of organisms. Table I illustrates the relative abundance of amino acids in three organisms – a bacterium, a green alga and a mollusk. Only one nonprotein amino acid is listed, diaminopimelic acid, and it occurs in the bacterium as a constituent of the cell wall.

When the amino acids in organisms are considered relative to the four aspects of amino acids that were previously defined, the following generalizations can be made: (1) the most common and abundant amino acids are α -amino carboxylic acids, (2) L-amino acids dominate but D-amino acids are present in low concentrations, (3) about 17 to 20 protein amino acids are the most abundant while nonprotein amino acids constitute a minor part of most total amino acid populations of organisms, and (4) the relative abundance of individual amino acids in organisms is variable although aspartic acid, glutamic acid, glycine, and alanine are among the most common.

2. Geological Systems

Most of the amino acids which enter geological systems come from organisms in the biosphere. Those amino acids which survive the transition from the biosphere to the lithosphere become incorporated in modern sedimentary deposits where they undergo a number of diagenetic changes. Proteins begin to hydrolyze and the process of racemization is initiated. Some new amino acids are formed and the concentration of all

TABLE I
Amino acid content of organisms^a

	<i>Escherichia coli</i> ^b (bacterium)	<i>Chlorella vulgaris</i> ^c (green alga)	<i>Mercenaria</i> ^d (mollusk)
Alanine	164	124	31
Arginine	68	255	19
Aspartic acid	127	104	194
Cysteine	22	—	—
Cystine	—	3	—
Diaminopimelic acid	6	—	—
Glutamic acid	134	126	38
Glycine	100	100	100
Histidine	12	53	3
Isoleucine	59	57	12
Leucine	102	98	20
Lysine	90	169	31
Methionine	44	23	1
Phenylalanine	42	45	17
Proline	59	94	45
Serine	78	53	79
Threonine	60	47	23
Tryptophan	15	34	—
Tyrosine	27	45	27
Valine	71	89	14

— Indicates 'not reported'.

^a Abundances reported relative to glycine = 100.

^b From Table V (Luria, 1960, p. 20).

^c From Table IV (Fowden, 1962, p. 202).

^d From Table 26 (Hare and Mitterer, 1967).

amino acids is reduced with time through degradation. The occurrence of amino acids in fossils, soils, recent and ancient sediments has been reviewed by Hare (1969). These compounds have been reported to be present in many kinds of geological samples ranging in age from the present to as old as 3.4×10^9 yr, but it is generally believed that the amino acids in early Precambrian sediments are not as old as the rocks themselves and may represent modern contamination (Kvenvolden, 1972).

Because most amino acids in sediments are derived originally from organisms, the kinds and distributions of these compounds should represent contributions from all the organisms of that environment. Indeed, a study of the occurrence of amino acids in geological systems shows that the compounds most often detected are the common protein acids. Nonprotein amino acids have not generally been reported to occur in very modern sediments and fossils, but with increasing geologic age a few nonprotein amino acids are formed probably from protein amino acids. For example alloisoleucine, ornithine, citrulline, α -aminobutyric acid, γ -aminobutyric acid, and β -alanine are found in ancient geological samples. It is quite likely that the sources of these com-

TABLE II
Amino acids in fossils and sediments^a

	<i>Mercenaria</i> ^b (modern)	<i>Mercenaria</i> ^b (Middle Miocene) (15–20 × 10 ⁶ yr old)	Saanich Inlet ^c Sediment (< 100 yr old)	Saanich Inlet ^c Sediment (9000 yr old)
Alanine	446	79	232	46
β-alanine	—	—	trace	4
Alloisoleucine	—	11	4	2
α-aminobutyric acid	—	7	—	—
γ-aminobutyric acid	—	10	trace	11
Arginine	274	—	1	1
Aspartic acid	2808	17	490	3
Cysteine	—	—	9	1
Glutamic acid	555	60	387	11
Glycine	1457	54	472	46
Histidine	40	—	57	9
Isoleucine	178	8	184	47
Leucine	288	18	275	66
Lysine	456	15	77	42
Methionine	16	1	7	5
Ornithine	—	6	trace	8
Phenylalanine	251	12	114	33
Proline	651	68	249	27
Serine	1152	6	265	3
Threonine	338	1	112	3
Tyrosine	396	5	50	6
Valine	199	30	228	49

— Indicates 'not found'.

^a Abundances reported in nanomoles amino acid/gram of sample.

^b From Table 26 (Hare and Mitterer, 1967).

^c From Table 1 (Kvenvolden *et al.*, 1970) and from Brown *et al.* (1970).

pounds were respectively isoleucine, arginine, arginine, glutamic acid, glutamic acid, and aspartic acid (Hare and Mitterer, 1967; Hare, 1969).

Table II shows two examples of the kinds and distributions of amino acids found in modern and ancient *Mercenaria* and in sediments spanning the last 9000 yr. This table calls attention to the fact that amino acid concentrations tend to decrease with time and that the relative abundances of individual amino acids change with time.

Amino acids with asymmetric carbon atoms are dominantly of the L-configuration in modern geological materials. This fact indicates that biological precursors have provided these amino acids, but these compounds begin to undergo interconversions and show a progressive increase in the degree of racemization with increasing age of the fossil or sediment. Hare and Abelson (1968) showed that the percentage of D-amino acids increased from zero in Recent *Mercenaria* to about 50 in *Mercenaria* of Miocene age. This observation means that L-amino acids were likely present in *Mercenaria* shells living during Miocene time and that during the last 10 to 20 × 10⁶ yr

these amino acids have racemized so that at present the fossil shells contain equal concentration of the D- and L-enantiomers of individual amino acids. Partial racemization of amino acids with time has also been noted in marine sediments (Kvenvolden *et al.*, 1970; Bada *et al.*, 1970; Wehmiller and Hare, 1971).

The following generalizations can be made with regard to amino acids in geological systems: (1) As in living systems the dominant compounds are those classified as protein amino acids and have the structure of α -amino carboxylic acids, (2) In very modern geological materials L-amino acids are most common, but with time these amino acids interconvert so that amino acids older than about 15×10^6 yr should be present as racemic mixtures, (3) Although only protein amino acids have been reported in modern materials, specific nonprotein amino acids such as β -alanine, α - and γ -aminobutyric acids, alloisoleucine and ornithine have been found in ancient geological specimens, (4) the relative abundances of individual amino acids appear to change and decrease with time, and glycine and alanine are common in populations of ancient amino acids.

3. Carbonaceous Chondrites

Amino acids have been found in a number of carbonaceous chondrites, but before 1970 a general consensus prevailed that meteorite amino acids likely represented terrestrial, biological contamination. The kinds of amino acids detected and their distributional patterns were very similar to those that occur in fingerprints (for a review see Hayes, 1967). Sixteen common protein amino acids were noted and glycine and serine were the most abundant species observed. Later, Oró *et al.* (1971b) showed that the amino acids in three carbonaceous chondrites were dominantly of the L-configuration. This finding strongly supported the idea of terrestrial contamination.

The fall of the Murchison meteorite in September of 1969 in Australia provided fresh type II chondrite material for organic chemical studies. Analyses of this meteorite revealed a population of amino acids quite different from that seen earlier in other meteorites. The Murchison was shown to contain at least 18 amino acids, but of these only 6 were protein amino acids and the remaining 12 were nonprotein amino acids (Kvenvolden *et al.*, 1970; *et al.*, 1971). Recently, the same suite of amino acids has been found in a freshly broken sample of Murray meteorite (Lawless *et al.*, 1971), and an independent confirmation of ten of the amino acids in both the Murchison and Murray was obtained by Cronin and Moore (1971). Table III lists the amino acids recently found in these meteorites. The six protein amino acids are given in order of decreasing abundances as observed in the Murchison meteorite. Besides the six protein amino acids, eight of the nonprotein amino acids are α -amino carboxylic acids while three have the amino group in the β -position and one has that group in the γ -position. The most abundant nonprotein amino acids are α -aminoisobutyric acid, isovaline, N-methylglycine and α -amino-*n*-butyric acid.

The enantiomeric distributions of nine of the asymmetric amino acids in the Murchison and Murray meteorites were found to be racemic (Kvenvolden *et al.*, 1970; Kvenvolden *et al.*, 1971; Lawless *et al.*, 1971). These amino acids are designated

TABLE III
Amino acids in meteorites and abiotic syntheses

Murchison and Murray meteorites	Electric discharge ^a	Fischer-Tropsch ^b	Other abiotic experiments ^c
Protein Amino Acids			
Glycine	+	+	+
Alanine*	+	+	+
Glutamic Acid*	+	+	+
Valine*	—	—	+
Proline*	—	—	+
Aspartic Acid*	+	+	+
Nonprotein Amino Acids			
Isovaline**	—	—	—
α -aminoisobutyric acid	+	—	—
α -amino- <i>n</i> -butyric acid*	+	—	+
N-methylglycine	+	—	+
β -aminoisobutyric acid*	—	—	—
β -amino- <i>n</i> -butyric acid**	—	—	—
Norvaline*	—	—	—
β -alanine	+	—	+
γ -aminobutyric acid	—	—	—
N-methylalanine**	+	—	—
N-ethylglycine	—	—	—
Pipecolic acid*	—	—	—

+ Positive identification.

— Not found including also not looked for.

^a Data from Miller and Urey (1959).

^b Data from Table 4 of Hayatsu *et al.* (1971).

^c Data summarized from Tables II and III of Lemmon (1970).

* Asymmetric amino acids present as racemic mixtures.

** Ratio of enantiomers not yet determined.

(*) in Table III. Oró *et al.* (1971a) provided confirmation of the racemic nature of alanine and proline in the Murchison meteorite.

The distinctive properties of the amino acids from the Murchison and Murray meteorites are the following: (1) Structurally fourteen of the compounds are α -amino carboxylic acids, and four have the amino group located either β or γ to the carboxyl group; N-alkyl (methyl and ethyl) amino acids are present, (2) Individual amino acids are racemic, (3) the number of nonprotein amino acids exceeds the protein amino acids and with the exception of two compounds, all isomers of amino acids with two through four carbon atoms are present, and (4) glycine and α -aminoisobutyric acid are dominant members of the amino acid populations.

4. Abiotic Syntheses

Laboratory experiments designed to simulate the primitive, prebiotic environment of the Earth have produced a variety of amino acids. The early work of Miller (1953) utilized an electric discharge in a mixture of methane, ammonia, and water. This

classic experiment has been followed by innumerable variations the results of which have been clearly summarized by Lemmon (1970). Table III shows which amino acids that have been prepared in abiotic synthesis experiments are the same as those recently identified in meteorites. All but one of the amino acids synthesized are α -amino carboxylic acids. Two N-methyl amino acids have been made. Four out of six protein amino acids and five out of twelve nonprotein amino acids of the meteorites have been positively synthesized abiotically in the laboratory. Glycine and alanine are usually the most abundant amino acids produced in these experiments, and, of course, both enantiomers of asymmetric amino acid are created in equal quantities. The similarities between amino acids synthesized in these experiments and those found in meteorites suggests that the amino acids in meteorites were formed naturally by abiotic, chemical processes.

5. Criteria: Biogenic vs. Abiogenic

From the evidence that has been presented it is now possible to establish criteria which will help distinguish biologically produced amino acids from those that have been synthesized without the aid of living systems. Other investigators have attempted to set up criteria for distinguishing the mode of origin of lipids particularly hydrocarbons (McCarthy and Calvin, 1967; Eglinton, 1969; Pering, 1971). Although evaluation of the origin of organic materials requires that all classes of compounds be examined, only amino acids will be considered here. It is likely that initial wet chemical experiments carried out on planetary surfaces, such as the surface of Mars, will be directed toward a very limited number of classes of compounds. If only amino acids are sought then criteria of origin based only on these compounds will have to suffice.

Hare (1969) established criteria for the mode of origin of amino acids based only on the isomeric composition of isoleucine. His criteria are (1) Only L-isoleucine = recent biological origin, (2) L-isoleucine and D-alloisoleucine = biological origin (probably older than Pliocene if isomers are in equilibrium); and (3) D, L-isoleucine and D, L-alloisoleucine = abiotic origin. These criteria are useful but are restrictive in that they require that isoleucine and its isomers be present. A more general set of criteria can be established for all amino acids and these are listed in Table IV. Finding amino acids with only one configuration would certainly indicate the operation of a biological process; at least up till now there is no evidence that abiotic syntheses can produce only one enantiomeric configuration. But finding a racemic mixture of protein amino acids would not necessarily rule out a biological origin for these compounds, because racemization may have altered the original asymmetric distribution. If a population of amino acids contains both protein and nonprotein amino acids other than β -alanine, α - and γ -aminobutyric acids, and alloisoleucine, and if these amino acids are racemic, an abiotic process is strongly indicated. With the criteria established here it should be possible to determine with some certainty the mode of origin of amino acid populations found on Earth, in meteorites and on planets. These criteria, however, are subject to change as additional information concerning the biochemistry, geochemistry and cosmochemistry of amino acids is gathered.

TABLE IV
Criteria for the mode of origin of amino acids

	Biogenic	Abiogenic
Structure	α -amino carboxylic acids dominant with occasional β - and γ -amino carboxylic acids in ancient geological deposits of biologic origin.	α -amino carboxylic acids common along with β and γ -amino carboxylic acids. N-alkyl (methyl and ethyl) amino acids present.
Enantiomers	L-amino acids are most abundant with small contributions of D-amino acids. For ancient amino acids the abundance of D-amino acids increases. Amino acids older than about 15×10^6 million yr should occur as racemic mixtures.	Racemic mixtures of amino acid enantiomers.
Composition	Protein amino acids dominate. For ancient amino acids specific nonprotein amino acids such as β -alanine, α - and γ -aminobutyric acids and alloisoleucine would be expected.	Both protein and nonprotein amino acids. Most amino acid isomers with two through four carbon atoms present.
Distribution	Variable with glutamic acid and aspartic acid common in young amino acid populations and glycine and alanine common in ancient amino acids.	Variable with possibly glycine and α -aminoisobutyric acid dominant.

Note added in proof: Since this paper was prepared, Miller has synthesized with the electric discharge all of the amino acids which have been found in the Murchison meteorite (Ring *et al.*, 1972; Wolman *et al.*, 1972).

References

- Bada, J. L., Luyendyk, B. P., and Maynard, J. B.: 1970, *Science* **170**, 730.
 Brown, F. S., Baedeker, M. J., Nissenbaum, A., and Kaplan, I. R.: 1970, *Geol. Soc. Am., Abst. with Programs* **2**, 503.
 Cronin, J. R. and Moore, C. B.: 1971, *Science* **172**, 1327.
 Eglinton, G.: 1969, in G. Eglinton and M. T. J. Murphy (eds.), *Organic Geochemistry, Methods and Results*, Springer-Verlag, New York, p. 20.
 Fowden, L.: 1962, in R. A. Levin (ed.), *Physiology and Biochemistry of Algae*, Academic Press, New York, p. 189.
 Hare, P. E.: 1969, in G. Eglinton and M. T. J. Murphy (eds.), *Organic Geochemistry, Methods and Results*, Springer-Verlag, New York, p. 438.
 Hare, P. E. and Abelson, P. H.: 1968, *Carnegie Inst. Washington Yearbook* **66**, 526.
 Hare, P. E. and Mitterer, R. M.: 1967, *Carnegie Inst. Wash Yearbook* **65**, 362.
 Hayatsu, R., Studier, M. H., and Anders, Ed.: 1971, *Geochim. Cosmochim. Acta* **35**, 939.
 Hayes, J. M.: 1967, *Geochim. Cosmochim. Acta* **31** 1395.
 Kvenvolden, K. A.: 1972, 'Organic Geochemistry of Early Precambrian Sediments', Proc. 24th Intl. Geological Congress, Montreal Section 1, 31.

- Kvenvolden, K. A., Peterson, E., and Brown, F. S.: 1970, *Science* **169**, 1079.
- Kvenvolden, K. A., Lawless, J. G., Pering, K., Peterson, E., Flores, J., Ponnampereuma, C., Kaplan, I. R., and Moore, C.: 1970, *Nature* **228**, 923.
- Kvenvolden, K. A., Lawless, J. G., and Ponnampereuma, C.: 1971, *Proc. Natl. Acad. Sci.* **68**, 486.
- Lawless, J. G., Kvenvolden, K. A., Peterson, E., Ponnampereuma, C., and Moore, C.: 1971, *Science* **173**, 626.
- Lemmon, R. M.: 1970, *Chem. Rev.* **70**, 95.
- Luria, S. E.: 1960, in I. C. Gunsalus and R. Y. Stanier (eds.), *The Bacteria – A Treatise on Structure and Function*, Vol. 1, Academic Press, New York, p. 1.
- McCarthy, E. D. and Calvin, M.: 1967, *Nature* **216**, 642.
- Meister, A.: 1965, *Biochemistry of the Amino Acids* (2nd ed.), Vol. 1, Academic Press, New York, p. 57.
- Miller, S. L.: 1953, *Science* **117**, 528.
- Miller, S. L. and Urey, H. C.: 1959, *Science* **130**, 245.
- Morrison, R. T. and Boyd, R. N.: 1969, *Organic Chemistry*, Allyn and Bacon, Inc., p. 1098.
- Oró, J., Gibert, J., Lichtenstein, H., Wikstrom, S., and Flory, D. A.: 1971a, *Nature* **230**, 105.
- Oró, J., Nakaparksin, S., Lichtenstein, H., and Gil-Av, E.: 1971b, *Nature* **230**, 107.
- Pering, K. L.: 1971, NASA TMX-62, 020, 118 p.
- Ring, D., Wolman, Y., Friedmann, N., and Miller, S. L.: 1972, *Proc. Natl. Acad. Sci.* **69**, 765.
- Stanier, R. Y., Doudoroff, M., and Adelberg, E. A.: 1963, *The Microbial World*, 2nd ed., Prentice-Hall, p. 142.
- Wehmiller, J. and Hare, P. E.: 1971, *Science* **173**, 907.
- Wolman, Y., Haverland, W. J., and Miller, S. L.: 1972, *Proc. Natl. Acad. Sci.* **69**, 809.