Amyloid world—the impossible requirement for identical prebiotic peptides in high concentration

Royal Truman, Chris Basel, and Boris Schmidtgall

High concentrations of *identical or very similar* peptides are required by origin of life models like the amyloid world hypothesis. The most likely peptide candidates would be as short as possible and use the most abundant amino acids (AAs). Assuming that two glycine and two alanine AA could have been used, the smallest suitable peptides would likely be octapeptides. However, octapeptides would have had a steady-state concentration ≈ 10^{-44} M in water at ca. 25°C or would have been entombed in solid salt beds. An analysis based on the expected distribution of prebiotic AAs and contaminants implies that a proportion <10^{-12} of the already extremely dilute octapeptide-like sequences would have been available to produce the same amyloid. Furthermore, these would have had to incorporate only L-enantiomers and avoid side-chain reactions. In contrast, *identical* octapeptides at a concentration of 10^{-4} M–10^{-2} M would have been necessary to form amyloids. This concentration could not have been achieved prebiotically.

Functional proteins are indispensable for all organisms and play a prominent role in many origin of life (OoL) theories. Proteins contain 50 or more amino acids (AAs) and have a distinct complex three-dimensional structure. When the term 'peptide' is used here, chains ≤50 AAs are meant; therefore, all the statements referring to peptides apply to proteins also. Evolutionists have struggled to find natural processes capable of producing the relevant sequences without a cellular genetic system. The problem is exacerbated by the fact that peptides and proteins must satisfy ten fundamental constraints *concurrently* to have any relevance for OoL purposes, as summarized in table 1.[1]

Requirement no. 1 is a serious problem for OoL researchers. Therefore, speculations have revolved around the notion that much shorter sequences might have been used for life to have initiated. Examples include early prions,[2] the Protein World Hypothesis,[3,4] and especially the amyloid world hypothesis.[5,6]

Forming amyloid aggregates is analogous to crystallization, depending on precise synergistic interactions, as observed by Route *et al*.[7] However, a vast number of *identical* peptides must aggregate correctly to form the necessary 3-dimensional structures under plausible prebiotic conditions.

What are amyloids?

According to Greenwald and Riek, amyloids are "an unbranched protein fiber whose repeating substructure consists of β strands that run perpendicular to the fiber axis, forming a cross-β sheet of indefinite length".[8] The stacked structures result in complex filaments with diameters typically of 6–12 nm.[8,9]

The focus of this paper is on the impossibility of obtaining high concentrations of *identical* peptides under plausible prebiotic conditions, as used in OoL experiments, to form a particular kind of amyloid.

Decades of research have provided a deep understanding of which peptide sequences could produce β-sheets, the usual minimal secondary structure necessary to form amyloids.[10–14] Some general principles include:[15,16]

1. A pattern of hydrophobic-hydrophilic AAs.
2. Small hydrophobic AAs which reduce steric interferences are favoured in the interior of the β-sheet. This includes alanine, valine, isoleucine, and glycine. Beta-branched amino acids, such as isoleucine and valine, however, would cause steric clashes in adjacent β-strands.
3. Hydrophilic and charged AAs, such as lysine, arginine, glutamic acid, and aspartic acid, are usually found on the surface of β-sheets. These can engage in electrostatic interactions and hydrogen bonding with water.
4. Proline usually introduces a kink in the polypeptide chain, disrupting H-bonding.
5. Aromatic AAs, such as phenylalanine, tyrosine, and tryptophan, can sometimes help by forming hydrophobic interactions.
6. Amino acids with a carbonyl group in their side chain, such as asparagine and glutamine, can be useful at the N-terminus of β-strands to form H-bonds with the neighbouring strand.

Table 1. Constraints all prebiotically relevant peptides must fulfil

No.	Requirement large peptides/proteins must fulfil for life to be possible
1	Many amino acids must be linked together, about 300 on average for proteins.
2	Only the L-amino acid enantiomers must be included.
3	Only the linear polymers must form, i.e., the side chains of amino acids must not react.
4	Precise sequences of amino acid residues must form to perform useful functions.
5	Other molecules, including non-biological amino acids, must be avoided in the peptides.
6	The long chains must adopt a suitable 3-dimensional structure.
7	A vast number of peptide copies must be produced continually for millions of years.
8	The correct proportion of peptides having a specific sequence must be co-located.
9	The entire system or organism must self-replicate, including all necessary peptide copies.
10	The polymers and 3-dimensional structure must be formed under relevant conditions.

Table 2. Propensity of customized heptapeptides to form extended β-sheets at 22°C, from Rufo *et al.*[22] β-sheets were surmised based on circular dichroism (CD) spectra reported in their 'supplementary information' figure S2. Esterase catalysis was tested through hydrolysis of *p*-nitrophenylacetate to form *p*-nitrophenol. Numbering of the sequences as used by Rufo *et al.* shown in red are residues present in trace amounts prebiotically.

No.	Peptide sequence	Esterase catalysis with 1 nM Zn^{2+}
β-sheets not found or not reported		
8	A H A H A R A	Very little / no activity
11a	V H V H V Q V	Catalytically active
β-sheets only formed with high [Zn^{2+}]		
1	L H L H L D L	Very little / no activity
2	L H L H L E L	No activity
3	L H L H L Q L	Catalytically active
10	V H V H V R V	Catalytically active
6	L H L H L K L	Moderately active
7	L H L H L R L	No activity without adding Zn^{2+}
Evidence for β-sheets without Zn^{2+}		
4	L H L H L Y L	Moderate activity
5	L H L H L H L	Very little / no activity
9	I H I H I R I	Catalytically active
11	I H I H I Q I	Catalytically active
12	I A I H I R I	Very little / no activity
13	I H I A I R I	Very little / no activity

7. Amino acids with a hydroxyl group or amine group in their side chain, such as serine, threonine, and lysine, can be useful at the C-terminus of β-strands to form H-bonds with the neighbouring strand.

An important principle is that longer peptides could benefit from many inter-strand attractive interactions to form β-sheets, whereas *short peptides would need to be optimized* by stronger interactions, using appropriate AAs.

Although β-sheets are required by most amyloid fibrils, *not all β-sheets can form amyloids*; side chain interactions play a major role.[8,17,18]

We will examine next the peptide sequences used by OoL researchers in key experiments to produce amyloid-like properties. Remarkably only one of the experiments used AAs exclusively from the so-called *reduced set* of 10 AAs discussed below, meaning the AAs expected to have been available prebiotically and minimally necessary for life.

Tjernberg *et al.* (2002)—peptide sequences resembling KFFE and KVVE

Tjernberg *et al.* examined a series of short model peptides to ascertain the minimum length to form amyloid-like fibrils, focusing on AA combinations known to have high propensities to form β-strands and residues with complementary charges.[17] The peptides underwent a 10-day incubation at 37°C, either at 200 or 300 μM, and at pH 5.0 or 7.0. Two peptides succeeded in producing fibrils, KFFE and KVVE (K = lysine; F = phenylalanine; E = glutamate; V = valine). Two other tetrapeptides, KFFK and EFFE, which did not form fibrils individually, did so when mixed in equimolar amounts. Peptides shorter than four AAs failed to form amyloid fibrils.

Although the authors did not claim prebiotic relevance, two of the key AAs they used, K and F, *are not members of the prebiotically relevant reduced set*.

Rufo *et al.* (2014)—peptide sequences resembling LHLHLRL

Heptapeptides with alternating apolar and polar residues often form extended β-sheets when the hydrophobic alternating periodicity is retained.[15,19–21]

Rufo *et al.* began with a carefully optimized peptide LHLHLRL (L = Leucine; H = Histidine) and then modified some residues, in particular replacing Leu with Ile and Val since these have strong β-sheet-forming propensity.[22] The results are summarized in table 2.

Even though the sequences had been designed to optimize forming β-sheets, more than half did not display amyloid-like properties or did so only in the presence of unrealistically high concentrations of Zn^{2+}. The concentration of ≈ 1 mM Zn^{2+} cofactor used to facilitate hydrolysis and formation of fibril was about a million times higher than present in oceans.[23]

Maury, the leading advocate of the amyloid world hypothesis, claimed, in a 2018 review article, that these amyloids had been synthesized under allegedly plausible prebiotic conditions, even though all of the peptides used AAs *not expected to have been present prebiotically*, or in vanishingly low concentrations at best.[24]

Rout *et al.*—peptide sequences FEFEFEFE with RFRFRFR

Rout *et al.* experimented in 2018 with carefully designed pairs of peptides able to form amyloids at pH 7.4 and 37°C when in an approximately 1:1 molar proportion.[7] The two peptides had alternating hydrophilic and hydrophobic side chains which were complementary, to produce intermolecular β-strands.

One peptide called the 'template' had the sequence Ac-FEFEFEFE-NH_2, whereas the various 'substrate' peptides had the sequence RFRFRFR-NH_2, FRFRFR-NH_2, or RFRFR-NH_2. The '–Ac' at the beginning of the 'template' indicates that the N-terminus had been acetylated; i.e., the NH_2 converted to $NHCOCH_3$.[25] The '–NH2' at the end of all the peptides indicates that their C-terminus had been amidated; i.e., COOH converted into $CONH_2$. The reason for these *prebiotically irrelevant modifications* was not explained.

Three amino acids were used to construct the peptides in this study[7]: F = phenylalanine; E = aspartic acid R = arginine, of which F and R would *not have been prebiotically relevant*, i.e., not part of the reduced set of AAs. Their chemical structures are shown in figure 1.

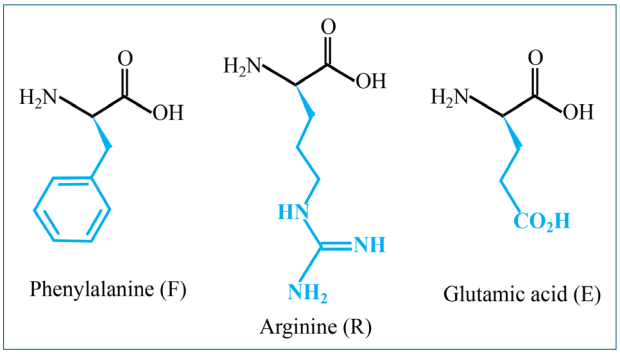

Figure 1. Chemical structure of phenylalanine (F), arginine (R), and glutamic acid (E) used in peptides studied by Rout *et al.*[7] Sidechain groups are shown in blue.

As pointed out by Rout *et al.*, amyloids formed in these experiments due to the complementarity of the side chains, and because the template Ac-FEFEFEFE-NH_2 had a net negative charge at neutral pH, whereas substrates like (RFRFRFR-NH_2) had a net positive charge.[7]

The mixed peptides required *several special processes to form amyloid-like substances*, which were not prebiotically relevant, namely:

- Sonication to improve solubility before mixing, since the individual peptides formed flocculent aggregates.
- Ideal stoichiometries in implausibly high, pure concentrations: 130 μM template and 100 μM substrate, i.e., ≈ 1×10^{-4} M.
- Agitation in an Eppendorf thermomixer at *800 rpm* at 37°C for ~18 hours.

Maury *et al.* (2012): peptide sequence EGGSVVAAD-amide

Amyloid expert Maury designed in 2012 the most plausible amyloid-forming peptide found in the literature from a prebiotic point of view, namely EGGSVVAAD-amide (as the acetate salt, where E = Glutamic acid, G = Glycine, S = Serine, V = Valine, A = Alanine and D = Aspartic acid).[5] These six AAs are members of the reduced set of AAs. Unfortunately, the peptide itself was not studied but its sodium acetate salt. Maury does not explain why this particular choice of salt was used. Counter-ions influence multiple peptide properties, including self-assembly tendencies, and stability.[26] Different counter-ions (e.g., Cl^-, HCO_3^-, and SO_4^{2-}) have very different effects on peptide secondary structure.[27] If the peptide had formed in an oceanic environment, the sodium chloride would have predominated.

A mixture of fibres of different lengths and shapes formed, including fibrils with amyloid character, after 14 days at room

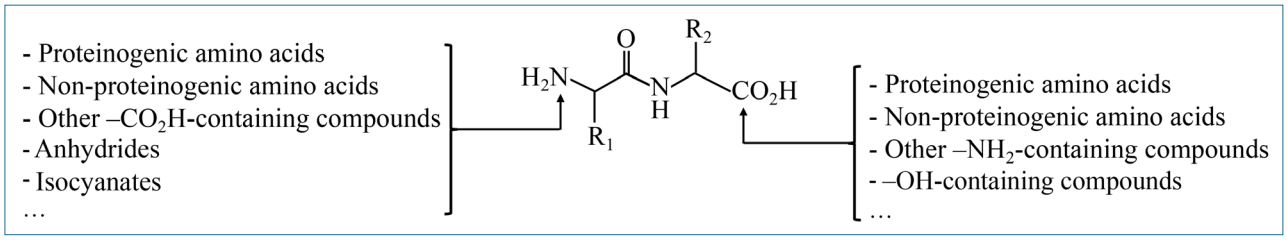

Figure 2. Peptide elongation could occur through reactions at each end, using a wide variety of chemicals. Reactions at the side chains have been ignored, as well as intramolecular reactions.

temperature using an extremely high concentration of the pure peptide.[5,24] This led Maury to conclude:[5]

> "The demonstration of spontaneous polymorphic amyloid formation of a prebiotically plausible short peptide provides experimental support for the amyloid world theory of the origin of life on the primordial Earth."

The *pure* 5 mg/ml EGGSVVAAD-amide acetate samples correspond to prebiotically absurd high concentrations ≈ 5.7×10^{-3} M and the 10 mg/ml samples to 1.13×10^{-2} M.[28]

The amorphous polymorphic fibrillar amyloid network seen with an electron microscope showed no symmetry nor regularity, nothing providing a consistent infrastructure for reproducible and repetitive chemical or metabolic processes to occur. This was in no manner analogous to the highly consistent structures of folded proteins or DNA coils.

Expected peptides under prebiotic conditions

The condensation of amino acids in water is endergonic, i.e., $\Delta G > 0$, and thus unfavourable. This is why the experimental section of OoL experiments inevitably mentions that the AAs had been chemically altered (i.e., 'activated') to place them in a reactive, high-energy state. Otherwise, the steady-state abiotic concentration would have been exceedingly low. Nevertheless, assuming that peptide polymers had formed, what might the sequences have resembled?

A peptide could react at either end with an AA, but, as shown in figure 2, non-proteinogenic AAs and other kinds of compounds could interfere with this reaction. The mixture of peptide-like substances formed prebiotically would have depended on the proportion of each chemical present.

Scores of compounds have been identified in meteorites and Miller-type spark discharge experiments which would have competed with proteinogenic AAs. In five papers, Truman and Schmidtgall analyzed the relative abundance of proteinogenic and non-proteinogenic AAs as well as other chemicals able to react with AAs.[29–33] Based on several studies, in particular Kobayashi *et al.* (2023),[34] evolutionary models indicated that about 1,000 times more AAs would have resulted from precursors formed in the atmosphere from energy sources such as lightning, solar energetic particles (SEPs), and galactic cosmic rays (GCRs) than the combined contribution from UV radiation, hydrothermal vents, and extra-terrestrial input.[30,31,35]

Kobayashi *et al.* experimented with mixtures of CO_2, N_2, H_2O, and CH_4, emphasizing proportions believed to reflect the prebiotic atmosphere, leading to several conclusions:[30]

- The proton irradiation and spark discharge experiments *produced only Gly* (with possibly trace amounts of Ala) from among the proteinogenic AAs under the most realistic atmospheric mixtures, i.e., the fraction CH_4 <5%. At unrealistically high levels of CH_4, very small amounts of Ala were found, but in the various compositions, Gly + Ala always represented about 99–100% of the proteinogenic yield.
- Under all conditions and gas mixtures, 5–10 times more carboxylic acid was produced than proteinogenic AAs.
- Proton irradiation with realistic gas mixtures also produced a trace amount of serine, but more than 10 times more non-proteinogenic AAs such as α-amino butyric acid, β-alanine, and γ-amino butyric acid were produced.
- Kobayashi *et al.* did not report the concentration of other substances, such as amines, alcohols, aldehydes, and ketones.

Other researchers have also reported that Gly + Ala represented about 99.4%–99.6% of the moles of proteinogenic AAs formed using a wide range of $CO_2/N_2/H_2$ mixtures.[36] A proportion of about 99.5% Gly + Ala was also typically found in earlier Miller-type experiments based on reducing gases like CH_4, H_2, and NH_3 which are no longer considered prebiotically relevant.[36,37] A minority of researchers have argued, however, that such mixtures might have existed temporarily, caused by massive meteors containing a large amount of reducing iron or reducing volatile gases, as discussed extensively by Truman and Schmidtgall.[30]

In almost all of the experiments to simulate prebiotic conditions, only a small minority of the proteinogenic AA was obtained.[30,32] One consistent observation has been that when conditions are used to optimize the formation of AA, such as including unrealistically high amounts of

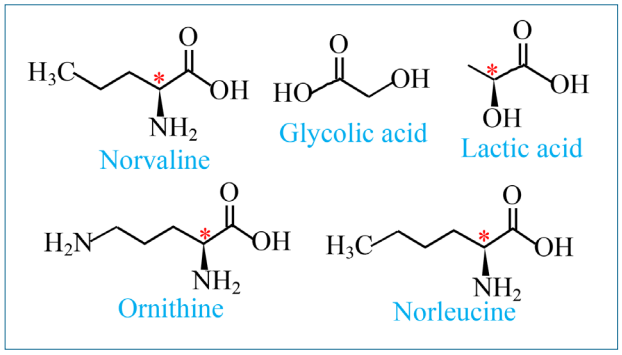

Figure 3. Some non-proteinogenic amino acids and hydroxy acids which condense readily with amino acids, believed to have been present in high concentrations on prebiotic earth. Chiral carbons are shown with a red asterisk. Redrawn by R. Truman from Fried *et al.*, ref. 36.

H_2S, or prebiotically implausible anti-oxidants, then an over-proportional amount of non-proteinogenic AAs are also formed.

Depending on the goals of the projects, researchers usually identified only the AA formed, or other specific chemicals. By comparing all the studies, it has become apparent that many amines, carboxylic acids, α-hydroxyl carboxylic acids, aldehydes, alcohols, etc. would also have formed or been delivered extraterrestrially in high proportions. These would have interfered with forming proper peptides. To illustrate, five examples of compounds often reported in high yields include norvaline, glycolic acid, lactic acid, ornithine, and norleucine, as shown in figure 3.[38]

α-hydroxy acids like glycolic acid and lactic acid are analogues of amino acids, where the –OH group has replaced an $–NH_2$ group. α hydroxy acids could also incorporate into peptide-like chains leading to esters—C(O)OR—instead of amides—C(O)NHR—as formed in peptides. These depsipeptides[39] have been shown to form during wet–dry cycle prebiotic simulations using combinations of α-AAs and α-hydroxy acids.[36,40,41]

As mentioned above, Truman and Schmidtgall analyzed various potential abiotic sources of AAs and other chemicals, noting that most would have arisen from the effects of lightning, solar energetic particles, and galactic cosmic rays. Since AAs and many other chemicals would have readily mixed in oceanic water irrespective of their origin, the overall proportion of proteinogenic AAs would have reflected the relative proportion which derived from all the potential sources.

A reduced set of amino acids

Evolutionary biochemists recognize the statistical impossibility of obtaining multiple copies of precise protein sequences abiotically when each position could include any of 20 proteinogenic AAs. In addition, only the L-enantiomers would be acceptable. This dilemma led to evolutionary proposals, like the amyloid world hypothesis, early prions,[42] and Ikehara's Protein World Hypothesis.[43,44]

The assumption is usually made that primitive life may have begun with a *reduced set* of AAs, typically defined as Alanine (A), Aspartic acid (D), Glutamic acid (E), Glycine (G), Isoleucine (I), Leucine (L), Proline (P), Serine (S), Threonine (T), and Valine (V).[45,46,47] Some claim cysteine might have been available also, but it is unstable, oxidizing easily.[36] This simplification was accepted in this paper since the other proteinogenic AAs obtained abiotically would have been in irrelevantly low concentrations. One consequence is that only one of all the peptide sequences mentioned in the studies above could have conceivably been produced in a relevant concentration prebiotically. It is impossible to obtain high concentrations of identical peptides if an AA found in only a trace concentration must have been incorporated at the same position.

Although almost none of the peptides published would have formed prebiotically, the experiments illustrated the difficulty of forming amyloids from short peptides, even when expertly designed. Often special temperatures, pH, incubation times, rapid agitation, sonication, and so on were necessary, and even then, amyloids were frequently not obtained.

The optimal prebiotic scenario

Many kinds of amyloids would have been necessary in a hypothetical amyloid world. Whether this is realistic can be evaluated by considering the peptides that could most easily have led to amyloids. These are:

1. peptides that are as short as possible
2. peptides that contain as many of the highest abundance AAs as possible.

The peptides should be as short as possible

Only a portion of proteins are involved in forming amyloids, and the minimum length has been the subject of much research. The only high-resolution description of a complete biological amyloid fibril seems to be based on a protein domain 72 AAs long.[8] This is a large region, but the core region of fibrils is generally believed to consist of a parallel in-register arrangement of β-strands ≈20 AAs or somewhat longer.[48]

Many specific regions within proteins have been shown to initiate amyloid aggregation. As Tjernberg *et al.* wrote:[17]

> "For example, the peptides NFGAIL from islet amyloid polypeptide, HQKLVFFAED from the amyloid β-peptide (Aβ), and VQIVYK from tau form

fibrils and are crucial for fibril formation of the full-length peptides."

Based on experiments conducted by Tjernberg *et al.*, tetrapeptides are the smallest peptides able to form amyloids, but their few successful examples required one or more AAs which were not part of the reduced set of AAs.[17]

Kinetic modelling by Truman, Tan, and McCombs, reported in 2024, implied that the prebiotic equilibrium concentration of octaglycine in water at around 25°C would have been $<10^{-44}$ M, assuming a high equilibrium availability of 10^{-4} M glycine.[49] The steady-state concentration of larger peptides would have decreased by a factor of about 5×10^5 per AA in the peptide.

Consistent with the kinetic analysis, peptide bonds are estimated to have a half-life in water at 25°C of between only 350 and 600 years *per bond*.[50] One can assume that the rate constants of condensation and hydrolysis of the interfering chemicals would have been comparable to that of AAs. That leads to $<10^{-44}$ M of octapeptide-like polymers, which corresponds to *<10 copies throughout all oceanic water*.[51]

Therefore, for OoL models, the peptides used to form amyloids should be as short as possible or their concentration would have been essentially zero. The experiments of Tjernberg *et al.* indicated that 4 AAs would be too short using the reduced set of amino acids assumed to have been available in relevant concentrations.

Maury, a medical expert on amyloids and leading proponent of the amyloid world hypothesis, devoted considerable effort to finding the shortest suitable peptide based on the reduced set of AAs. The nonapeptide found, EGGSVVAAD-amide, consisted of 9 AAs.[5] At equilibrium in water at 25°C, a 10^{-4} M mixture of AAs would have equilibrated with a $\approx 10^{-50}$ M concentration of nonapeptides.[47] It will be assumed here that a similar but shorter octapeptide sequence might also work, also based on only AAs from the reduced set of AAs. The concentration would have been $\approx 10^{-44}$ M, in stark contrast with the concentration of $\approx 10^{-2}$ M at room temperature Maury had to use.[5]

The peptide should contain as many of the highest-abundance AAs as possible

Clearly, the peptides formed in the highest proportion would consist of high-probability AAs: glycine with some alanine, but the peptides must be able to produce amyloids.

To establish the most optimistic, but still reasonable, scenario possible for OoL purposes, it will be assumed that some 8-AA octapeptides based on the reduced set of AAs might form amyloids. It will be assumed that, analogous to Maury's sequence, it would contain 2 glycine and 2 alanine AAs. No restrictions will be placed on the sequences

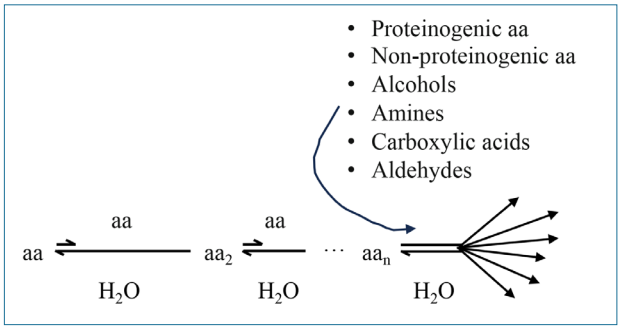

Figure 4. Many kinds of chemicals can react with peptides.

themselves, except that they must form amyloids. Suitable sequences might resemble, for example, (1) or (2):

GA*nnnn*AG (1)

AAG*nnn*G*n* (2)

where *n* represents any AAs except G or A. In theory, (1) or (2) might be candidates to produce a distinct amyloid if present in high enough concentration. Of course, *each candidate would decrease the concentration of the others by consuming available AAs*. Only rarely would very different short sequences collaborate to form extended β-sheets able to produce the same amyloid.

Probabilistic approximations

Many kinds of chemicals can react with AAs, as shown in figure 4. Multi-functional chemicals that contain two or more of any combinations of –OH, $–NH_2$, or –COOH functional groups could be inserted into a growing peptide (e.g., α-hydroxyl carboxylic acids, α-hydroxyl amines, or dicarboxylic acids). These contaminants will be referred to as 'peptide-like polymers'. Mono-functional contaminants would 'cap' one end of the peptide, permitting further condensation reactions at only the other end.

Different octapeptides could lead to an amyloid 'family', but multiple very similar or identical copies would have been required. For example, peptides with sequences like (1) above might produce one kind of amyloid, whereas peptides with sequences like (2) might produce a different kind. The proportion of acceptable peptides able to form a particular amyloid within the space of peptide-like polymers will depend on the probability p(AA) of introducing an acceptable AA at each position.

Truman and Schmidtgall observed that whenever OoL experiments were designed to maximize the yield of AAs, the proportion of undesirable interfering chemicals increased.[29–33] Realistic proportions will be proposed here to develop the

statistically most favourable evolutionary scenario that is reasonable. The assumptions are:

Gly = 80.0%; Ala = 10.0%; multi-functional contaminants = 8%; and other proteinogenic AAs = 2%.

Most of the assumed 2% other proteinogenic AAs would have been dominated by one or two AAs, with all the rest present in very low proportions. For example, according to Kobayashi *et al.*, the major source of AAs would have resulted from the action of solar energetic particles, and galactic cosmic rays, simulated using proton irradiation.[34] Experiments with CO_2, N_2, CH_4, and H_2O mixtures using CH_4 proportions between 0–20% led to ratios of Ser:Gly = 0 to 0.012; and Asp:Gly = 0 to 1.2×10^{-3}. *No other proteinogenic AAs were found*, and multifunctional contaminants were not looked for. For these CH_4 proportions, multi-functional contaminants α-amino butyric acid + β-alanine + γ-aminobutyric represented up to about 6% of the proteinogenic AAs when using proton irradiation, and up to about 24% for spark discharge experiments. In other words, most proteinogenic AAs were not obtained at all, but the environment would have been flooded with contaminants.

The favourable scenario will assume that the octapeptide-like polymer could use two glycine and two alanine AAs; two of the proteinogenic AAs having next higher frequency (≈1% each), and two more AAs having ≈0.1% frequency each. These assumptions lead to the following probabilities, whereby only proteinogenic AAs and multi-functional contaminants have been taken into account. The effect of single-functional alcohols, amines, carboxylic acids, etc. will be considered later.

- p(Gly) = 0.80. Most of the latest experiments using slightly reducing or neutral gas mixtures implied the proportion of Gly could be >0.99, which would drastically lower all the other probabilities.
- p(Ala) = 0.1 Alanine or alternative AAs also able to function at that position. Acceptable alternative AAs would have been present in relative proportion <<[Ala] and not have affected the value of 0.1 much.
- $p(n) = 10^{-2}$ for the two higher probability proteinogenic AAs after Gly and Ala.
- $p(n) = 10^{-3}$ for lower probability proteinogenic AAs.

Discussion

Suppose a 7-AA peptide had been present, and either end could incorporate another AA. A resulting sequence might resemble, for example, the octapeptide GGAA*nnnn*. The assumed probabilities led to a proportion of potential candidates / all octapeptides:

$$p(\mathrm{GGAA}n_h n_h n_l n_l) = (0.80)^2(0.1)^2(10^{-2})^2(10^{-3})^2 < 10^{-12} \qquad (3)$$

where n_h represents the two proteinogenic AAs in highest concentration besides Gly and Ala and n_l the lowest concentration proteinogenic AAs.

However, the proportion of octapeptide-like polymers able to form one specific kind of amyloid would have been far less than shown in eqn (3), for several reasons:

- A wide variety of monofunctional alcohols, amines, and carboxylic acids could also have added to heptapeptide-like compounds, increasing the proportion of unacceptable end-capped octapeptides.
- The AAs would have been racemic. The probability of all the AAs (except for Gly) being L-enantiomers would have decreased the proportion by a factor of $(0.5)^n$. For $n = 6$ AAs, $(0.5)^6 = 0.016$.
- Side chain reactions were not taken into account. These would have interfered with formation of amyloids and decreased the concentration of linear peptides.
- $p(n_h)$ and $p(n_l)$ were assigned identical probabilities, but assuming non-identical concentrations would have lowered the overall probability. For example, $p(n_h)p(n_h) = (10^{-2})^2 = 10^{-4}$, whereas *assigning the same total amount* of 0.02 (i.e., 0.01 + 0.01) but distributed unequally between the two AAs produces a lower probability, e.g., $p(n_h)p(n_h) = (0.018)(0.002) \leq 4 \times 10^{-5}$. The same applies to the $p(n_l)$ pair. To illustrate, in Maury's sequence, EGGSVVAADE (glutamic acid) and D (aspartic acid) would not have been present in identical proportions.

Therefore, considering factors such as these indicates that eqn. (3) is too high. We conclude conservatively that:

$$\text{Proportion suitable octapeptides / all octapeptide-like} < 10^{-14}. \qquad (4)$$

Any suitable sequence would have had to be present in a high concentration of (almost) identical copies. Larger candidate peptides or those using AAs not found in the reduced set would have had a much lower probability than shown in equation (4).

High concentrations of similar peptides impossible

The nonapeptide Maury designed, EGGSVVAAD-amide, did not form an amyloid effectively, requiring very high concentrations of 10^{-3}–10^{-2} M.[24] Forming an amyloid from a shorter peptide would have been even more difficult, so a minimum concentration of ≈ 10^{-3} M pure octapeptide would have been necessary. Equation (4) states that only a fraction of about 10^{-14} of all octapeptide-like sequences would have been able to form one kind of amyloid. Recall from above that there would not have been even 10 copies of octapeptide-like sequences in a prebiotic ocean.

In principle, wet–dry cycles could have driven the condensation reaction shown in figure 4 forward, such as by water evaporation / rain cycles in a salty pond, but this has been shown incapable of providing a feasible solution.[52] Wet–dry cycles decompose AAs and peptides after a few cycles, steadily decreasing the yield of peptides. Analysis of highly optimized experiments, recalibrated to prebiotically realistic concentrations of AAs, contaminants, and pH, indicated that peptides of length 8 or longer would have been present, at best, in trace concentrations. The very dilute larger peptides would have been embedded in solid NaCl, along with many other chemicals where amyloids could not have formed, or remained dissolved if enough water was present, where they would have hydrolyzed.[53]

Smaller peptides would have dominated the mixtures

A serious problem would have been the inevitable concomitant presence of smaller peptides, as shown in figure 4.[50] In water at around 25°C at equilibrium, the proportion of $[\text{peptide}]_n : [\text{peptide}]_{n-1} \approx 2 \times 10^{-6}$,[47] e.g.:

$$\text{GGAA}nnnn : \text{GAA}nnnn : \text{AA}nnnn \ldots = 2 \times 10^{-6} : 2 \times 10^{-6} : 2 \times 10^{-6} \ldots \quad (5)$$

Using wet–dry cycles, the requisite $[\text{peptide}]_8$ of 10^{-3}–10^{-2}M to form amyloids would have been embedded in $[\text{peptide}]_2$–$[\text{peptide}]_7$ in several orders of magnitude total high concentration. The slurry would have either solidified or contained far too little peptide_8 if remaining in solution.

Inevitable contamination by additional chemicals

OoL chemical experiments use pure chemicals to avoid the realities described, whereas almost 100% of the chemicals present in solution prebiotically would not have been related to forming octapeptide-like polymers.

Removing water concentrates the residues, but amyloids only form in water, not as solids. However, reintroduced water to dissolve peptides would dilute the peptides and allow their hydrolysis.

Forming an amyloid world ecosystem

Finally, a hypothetical amyloid world would have required many different kinds of amyloids, each *providing useful collaborative functions*. Fried *et al.* recognized this, pointing out, in a 2022 mini-review article, that[36]

> "While there are examples of catalytic amyloids, evidence of their catalytic utility being directed toward the synthesis of other prebiotically relevant molecules is lacking and would represent an important discovery."

The statistical analysis has shown that a single amyloid would not have formed prebiotically, so clearly an amyloid world never existed.

References

1. Truman, R. and McCombs, C., Insignificant concentrations of peptides form in water: part 1—using hot temperatures or high pH, *J. Creation* **38**(1):126–135, 2024.
2. Jheeta, S., Chatzitheodoridis, E., Devine, K., and Block, J., The way forward for the origin of life: prions and prion-like molecules first hypothesis, *Life (Basel)* **11**(9):872, 2021.
3. Ikehara, K., [GADV]-Protein World Hypothesis on the Origin of Life, *Orig. Life Evol. Biosph.* **44**:299–302, 2014.
4. Ikehara, K., Evolutionary steps in the emergence of life deduced from the bottom-up approach and GADV hypothesis (top-down approach), *Life (Basel)* **6**(1):6.
5. Maury, C.P.J., Liljeström, M., and Zhao, F., Was the first molecular replicator on the primitive Earth an informational amyloid? EGGSVVAAD, a prebiotically plausible peptide, spontaneously forms amyloid assemblies, *J. Biol. Res.***18**:332–335, 2012.
6. Maury, C.P.J., Self-propagating, β-sheet polypeptide structures as prebiotic informational molecular entities: the amyloid world, *Orig. Life Evol. Biosph.* **39**:141–150, 2009.
7. Rout, S.K., Friedmann, M.P., Riek, R., and Greenwald, J., A prebiotic template-directed peptide synthesis based on amyloids, *Nature Commun.* **9**:234–242, 2018.
8. Greenwald, J. and Riek, R., Biology of amyloid: structure, function, and regulation, *Structure* **18**(10):1244–60, 2010.
9. Taylor, A.I.P. and Staniforth, R.A., General principles underpinning amyloid structure, *Front. Neurosc.* **16**:878869, 2022.
10. Subramani, A. and Floudas, C.A., β-sheet topology prediction with high precision and recall for β and mixed α/β proteins, PLoS ONE **7**(3):e32461, 2012.
11. Tang, L., Zhao, Z., Zhang, L., Zhang, T., and Gao, S., Predicting the outer/inner betastrands in protein beta sheets based on the random forest algorithm; in: Huang, DS., Han, K., Gromiha, M. (Eds.), *Intelligent Computing in Bioinformatics*, ICIC 2014. Lecture Notes in Computer Science, Springer, Cham., vol. 8590, 2014.
12. Mao, W., Wang, T., Zhang, W. *et al.*, Identification of residue pairing in interacting β-strands from a predicted residue contact map, *BMC Bioinformatics* **19**:146, 2018.
13. Guilloux, A, Caudron, B., and Jestin, J.L., A method to predict edge strands in beta-sheets from protein sequences, *Comp. Struct. Biotechnol. J.* **19**(7):e201305001, 2013.
14. Cheng, J. and Baldi, P., Three-stage prediction of protein β-sheets by neural networks, alignments and graph algorithms, *Bioinformatics* **21**:i75–i84 Issue suppl_1, 2005.
15. Schneider, J.P., *et al.*, Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide, *J. Am. Chem. Soc.* **124**:15030–15037, 2002.
16. Zhang, S., Marini, D.M., Hwang, W., and Santoso, S., Design of nanostructured biological materials through self-assembly of peptides and proteins, *Curr. Opin. Chem. Biol.* **6**:865–871, 2002.
17. Tjernberg, L., Hosia, W., Bark, N., Thyberg, J., and Johansson, J., Charge attraction and β propensity are necessary for amyloid fibril formation from tetrapeptides, *J. Biol. Chem.* **277**:43243–43246, 2002.
18. Zanuy, D. and Nussinov, R., The Sequence dependence of fiber organization. A comparative molecular dynamics study of the islet amyloid polypeptide segments 22–27 and 22–29, *J. Mol. Biol.* **329**:565–584, 2003.
19. DeGrado, W.F. and Lear, J.D., Induction of peptide conformation at apolar water interfaces. 1. A study with model peptides of defined hydrophobic periodicity, *J. Am. Chem. Soc.* **107**:7684–7689, 1985.
20. West, M.W., *et al.*, De novo amyloid proteins from designed combinatorial libraries, PNAS **96**:11211–11216, 1999.
21. DeGrado, W.F., Wasserman, Z.R., and Lear, J.D., Protein Design, a minimalist approach, *Science* **243**:622–628, 1989.
22. Rufo, C., Moroz, Y., Moroz, O., *et al.*, Short peptides self-assemble to produce catalytic amyloids, *Nature Chem.* **6**:303–309, 2014.
23. Truman, R. and Basel, C., Prebiotically plausible peptides don't self-assemble to produce useful catalytic amyloids, *J. Creation* **37**(3):135–140, 2023.

24. Maury, C.P.J., Amyloid and the origin of life: self-replicating catalytic amyloids as prebiotic informational and protometabolic entities, *Cell. Mol. Life Sci.* **75**:1499–1507, 2018.
25. Amidation and acetylation, lifetein.com, accessed 16 Feb 2024.
26. Sikora, K., Jaśkiewicz, M., Neubauer, D., Migoń, D., and Kamysz, W., The role of counter-ions in peptides—an overview, *Pharmaceuticals (Basel)* **13**(12):442, 2020.
27. Johansson, J., Gudmundsson, G.H., Rottenberg, M.E., Berndt, K.D., and Agerberth, B., conformation-dependent antibacterial activity of the naturally occurring human peptide LL-37, *J. Biol. Chem.* **273**:3718–3724, 1998.
28. Using a molecular weight of 884.86 g/mol for EGGSVVAAD-amide (as the acetate salt).
29. Truman, R., Schmidtgall, B., and Basel, C., Relative proportion of prebiotic amino acids: part 1—experiments using reduced gas mixtures, *J. Creation* **38**(3):87–98, 2024.
30. Truman, R., Schmidtgall, B., and Basel, C., Relative proportion of prebiotic amino acids: part 2—models to form highly reduced atmospheres, *J. Creation* **38**(3):99–104, 2024.
31. Truman, R., Schmidtgall, B., and Basel, C., Relative proportion of prebiotic amino acids: part 3—experiments using oxidized gas mixtures, *J. Creation* **38**(3):105–111, 2024.
32. Truman, R., Schmidtgall, B., and Basel, C., Relative proportion of prebiotic amino acids: part 4—the case for the primitive atmosphere having been weakly reducing, *J. Creation* (in press).
33. Truman, R., Schmidtgall, B., and Basel, C., Relative proportion of prebiotic amino acids: part 5—amino acids from non-atmospheric reaction, *J. Creation* (in press).
34. Kobayashi, K., Ise, J.-i., Aoki, R., Kinoshita, M., Naito, K., Udo, T., Kunwar, B., Takahashi, J.-i., Shibata, H., Mita, H., *et al.*, Formation of amino acids and carboxylic acids in weakly reducing planetary atmospheres by solar energetic particles from the young sun, *Life (Basel)* **13**:1103, 2023.
35. Truman, R. and Basel, C., Contamination of meteorite samples with terrestrial amino acids, *J. Creation* **38**(2):7–11, 2024.
36. Johnson, A.P., Cleaves, H.J., Dworkin, J.P., Glavin, D.P., Lazcano, A, and Bada, J.L., The Miller volcanic spark discharge experiment, *Science* **322**(5900):404, 2008.
37. Miller, S.L. and Urey, H.C., Organic compound synthesis on the primitive earth, *Science* **130**:245–251, 1959.
38. Fried, S.D., Fujishima, K., Makarov, M., Cherepashuk, I., and Hlouchova, K., Peptides before and during the nucleotide world: an origins story emphasizing cooperation between proteins and nucleic acids, *J. R. Soc. Interface* **19**:20210641, 2022.
39 Depsipeptides, goldbook.iupac.org, accessed 16 Feb 2024.
40. Forsythe, J.G., Yu, S.S., Mamajanov, I., Grover, M.A., Krishnamurthy, R., Fernández, F.M., and Hud, N.V., Ester-mediated amide bond formation driven by wet-dry cycles: a possible path to polypeptides on the prebiotic Earth, *Angew. Chem.* **127**:9871–9875, 2015.
41. Chandru, K., Guttenberg, N., Giri, C., Hongo, Y., Butch, C., Mamajanov, I., and Cleaves, H.J., Simple prebiotic synthesis of high diversity dynamic combinatorial polyester libraries, *Commun. Chem.* **1**:30, 2018.
42. Jheeta, S., Chatzitheodoridis, E., Devine, K., and Block, J., The way forward for the origin of life: prions and prion-like molecules first hypothesis, *Life (Basel)* **11**(9):872, 2021.
43. Ikehara, K., [GADV]-protein world hypothesis on the origin of life, *Orig. Life Evol. Biosph.* **44**:299–302, 2014.
44. Ikehara, K., Evolutionary steps in the emergence of life deduced from the bottom-up approach and GADV hypothesis (top-down approach), *Life (Basel)* **6**(1):6, 2016.
45. Parker, E.T., Cleaves, H.J., Dworkin, J.P., Glavin, D.P., Callahan, M., Aubrey, A., Lazcano, A., and Bada, J.L., Primordial synthesis of amines and amino acids in a 1958 Miller H2S-rich spark discharge experiment, PNAS **108**:5526–5531, 2011.
46. Burton, A.S., Stern, J.C., Elsila, J.E., Glavin, D.P., and Dworkin, J.P., Understanding prebiotic chemistry through the analysis of extraterrestrial amino acids and nucleobases in meteorites, *Chem. Soc. Rev.* **41**:5459–5472, 2012.
47. Higgs, P.G. and Pudritz, R.E., A thermodynamic basis for prebiotic amino acid synthesis and the nature of the first genetic code, *Astrobiology* **9**(5):483–490, 2009.
48. Margittai, M. and Langen, R., Fibrils with parallel in-register structure constitute a major class of amyloid fibrils: molecular insights from electron paramagnetic resonance spectroscopy, *Q. Rev. Biophys.* **41**(3–4):265–297, 2008.
49. Truman, R., Tan, C., and McCombs, C., Insignificant concentrations of peptides form in water: part 2—using moderate temperatures, *J. Creation* **38**(1):136–149, 2024.
50. Radzicka, A. and Wolfenden, R., Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases, *J. Amer. Chem. Soc.* **118**(26):6105–6109, 1996.
51. 1.31 × 10^{21} L terrestrial water)(10^{-44} moles/L)(6.02 × 10^{23} molecules/mole) ≈ 8 octapeptides would have been present.
52. Truman, R. and Basel, C., The limited value of wet–dry cycles to produce large peptides prebiotically, *J. Creation* **38**(3):124–128, 2024.
53. At concentrations of ≈6 M NaCl precipitates at ambient temperatures. The maximum solubility = 358 gm NaCl / L at 20°C, and 364 gm NaCl / L at 40°C. Solubility table of compounds in water at temperature, sigmaaldrich.com, accessed 16 Feb 2024.

Royal Truman *has bachelor's degrees in chemistry and computer science, an M.B.A., a Ph.D. in organic chemistry, and a two-year post-graduate* Fortbildung *in bioinformatics. After working 40 years for the world's largest chemical firm he retired, and is currently a visiting scientist at a German university.*

Chris Basel *has a bachelor's degree in chemistry from William Jewell College, a master's degree in analytical chemistry from the University of Kansas, and a certificate in ADMET Process (pharmacology) from the University of California-San Diego. He worked as a research manager in the pharmaceutical industry for 30 years and now serves as an associate professor of chemistry and department chair at a Christian university in Missouri.*

Boris Schmidtgall *obtained his M.S. (German Diploma) and Ph.D. in chemistry from the University of Göttingen, followed by postdoctoral work in biochemistry in Strasbourg, France. Since 2018 he has been employed by the German creation research group Wort und Wissen. He has published extensively in secular international chemistry journals and also on origin of life research and biomolecular chemistry from a creation science perspective.*