Salt-induced formation of mixed peptides under possible prebiotic conditions

Michael G. Schwendinger and **Bernd M. Rode***
Institute of Inorganic and Analytical Chemistry, University of Innsbruck, Innrain 52a, A-6020 Innsbruck (Austria)

(Received December 4, 1991; revised April 9, 1991)

Abstract

Investigations on the copper mediated peptide formation in aqueous solution in the presence of high amounts of NaCl have been extended to systems simultaneously containing glycine and alanine. Preferential formation of ala–gly compared to gly–ala is observed, total yields as high as 10.5% and 5.8% of initial glycine and alanine are found, and various tripeptides also form. Racemization occurs only to a limited extent.

1. Introduction

Numerous attempts have been made up to now to establish reasonable pathways for the formation of peptides under possible prebiotic conditions. Although little is known about the actual conditions that prevailed on the primitive earth, e.g. suggestions about the amount of oxygen in the prebiotic atmosphere range from 10^{-70} [1] to 10^{-1} [2] of the present atmospheric level (PAL), an aqueous environment has to be considered as the most probable scenario for peptide formation, e.g. salty water droplets dispersed in the atmosphere [3], the primordial ocean, or hydrothermal vents [4]. Due to unfavourable thermodynamics [5, 6], the ionic nature of the amino acids prevents the nucleophilic attack of the amino group to the carboxylic C atom. Even from glycine ethyl ester only very low yields (max. 4%) of diglycine ethyl ester can be obtained in aqueous [7] or non-aqueous [8] solution. To overcome this thermodynamic barrier, complete evaporation and subsequent reaction of the dry amino acid mixtures in melt [9], or the use of large amounts of condensation reagents like inorganic polyphosphates [10, 11] or cyanamides [12, 13] were proposed as possible prebiotic conditions. The recently discovered salt-induced peptide formation catalyzed by Cu^{2+} ions and NaCl [14, 15] gave good yields (more than 5% of diglycine) under very simple conditions at moderate temperatures in aqueous solution and thus makes the discussion about the abundance of specific condensation reagents on the primitive earth obsolete.

An important feature of a reaction model claiming to have prebiotic relevance is its applicability to amino acids other than the simplest one, glycine. Further the investigation of systems containing more than one amino acid should give valuable information about preferential formation of specific sequences which could have dominated in prebiotic peptides.

Therefore we have extended our investigations in a first step to alanine, which — due to its simple structure and stability — is formed in simulated prebiotic amino acid synthesis reactions with relatively good yields [16] and therefore can be assumed to have been second in availability among biologically relevant amino acids.

Optimization of various conditions for this new reaction have already been performed with the system containing glycine as the only amino acid [15]. Summarizing these results, a ratio of amino acid/copper concentration near 2, 5 M NaCl, high amino acid concentrations, unchanged pH, and moderate temperatures (c. 80 °C) were found as optimal conditions, and only these where applied in the present study.

2. Experimental

Gly, Ala, NaCl and $CuCl_2 \cdot 2H_2O$ were obtained from Fluka Co; the necessary reference peptides were purchased in analytical grade quality from Sigma Chemical Co. OPA/MPA reagent was obtained from

*Author to whom correspondence should be addressed.

0020-1693/91/$3.50

Hewlett Packard Co. All chemicals were used without further purification.

Solutions were prepared in distilled water and heated without stirring on a thermostatized sand bath in glass flasks with reflux cooler under argon atmosphere or air; micro experiments with dipeptides were performed in sealed 1 ml vials.

Analysis of the 50-fold diluted reaction mixture was carried out by HPLC on a HP 1090M LC system with diode array detector at 338 nm, after online derivatization of amino functions with *o*-phthalaldehyde/3-mercaptopropionic acid reagent (OPA/MPA). A Shannon Hypersil column (ODS, 5 μ, 200×2.1 mm) was used with the mobile phase composition (solvent A: 30 mM $NaOAc \cdot 3H_2O$, 0.2 mM EDTA, 0.2% THF; solvent B: 100 mM $NaOAc \cdot 3H_2O$, 0.1 mM EDTA, 80% MeCN) and gradient according to Schuster [17], at a flow of 0.45 ml/min.

Enrichment of tripeptides was achieved by the following procedure: after precipitation of Cu^{2+} as sulfide by H_2S and filtration, the solution was freed from inorganic components by a cation exchanger in its ammonium form and subsequent elution of the amino acids and peptides with 0.5 M ammonia. The excess of amino acids was removed by means of open column separation on silica gel 60 (Merck) with 0.1 M NH_4OAc/MeCN (1:3) as mobile phase. The last fraction, containing gly_2 as the main component, was concentrated by vacuum distillation, and finally the buffer salts were removed by the ion exchanger.

Products were identified by comparison of retention times of authentic samples. As the peptide formation has been unambiguously proven in the case of glycine, together with the fact that the chemical behavior of the formed compounds (sequence inversion, see Section 3) is consistent with that of authentic dipeptides, this method of identification seemed sufficient.

3. Results and discussion

Figure 1 demonstrates the formation of all possible dipeptides in a solution containing 0.49 M of gly, L-ala, Cu^{2+} and 5 M NaCl after 522 h, together with the reference peaks obtained with a standard mixture of amino acids and peptides. Perfect agreement of retention times proves the identity of products. Under the chosen chromatographic conditions the two diastereomers of ala_2 are also separated and therefore the extent of racemization in our system could be observed.

Figure 2(a) and (b) compares the increase of peptide yields with time under argon and air atmosphere under otherwise identical conditions, while

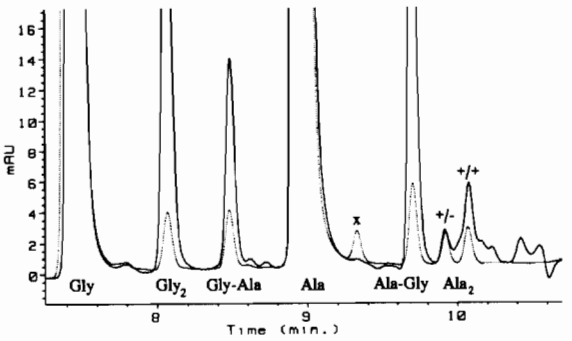

Fig. 1. Chromatogram obtained after 522 h at 75 °C from a solution 0.4 M in gly, L-ala, Cu^{2+} and 5 M in NaCl, under air (solid line) at a pH of 2.8, together with a reference curve (dashed line). The additional peak (X) belongs to glycine; its absence in the sample chromatogram is due to the lower [gly].

Fig. 2(c) shows the dependence of total amino acids incorporated into dipeptides on reaction time.

A clear order in concentrations results after a reaction time of 200 h.

$$[gly_2] > [ala\text{–}gly] \gg [gly\text{–}ala] > [ala_2]$$

Higher yields of peptides can be obtained under air. Reduction of Cu^{2+} by the amino acid continuously lowers the concentration of this catalyst. As in the case of pure glycine, under inert conditions the solutions change their color from dark green via light green to pale yellow, and peptide formation stops. From there, at temperatures above 80 °C, acidic hydrolysis takes place and gradually reduces the yield of peptides. Under air Cu^{2+} is regenerated, and peptide formation carries on. As long as the excess of amino acids is high enough to occupy all of the binding sites of copper, peptides are stable with respect to oxidation, but nevertheless the long term fate of peptides under air would also be their destruction due to oxidation. Repetitive addition of fresh seawater and subsequent evaporation could have served as a continuous supply of an excess of amino acids and, therefore, suppressor of this process.

The need for Cu^{2+} in the reaction demands at least partially oxidizing conditions on the primitive earth. According to Ochiai [1] a p_{O_2} higher than 10^{-35} atm. would have been sufficient to form Cu^{2+}.

A possible source for O_2 in the primitive atmosphere is electro- and photochemical cleavage of CO_2 and H_2O. The production rate of O_2 therefore strongly depends on the moisture content of the upper atmosphere. Taking into account that the surface temperature probably was higher than nowadays, the oxygen concentration could have reached values of 0.1–0.001 PAL [2]. Additionally geological records indicate at least local concentrations of

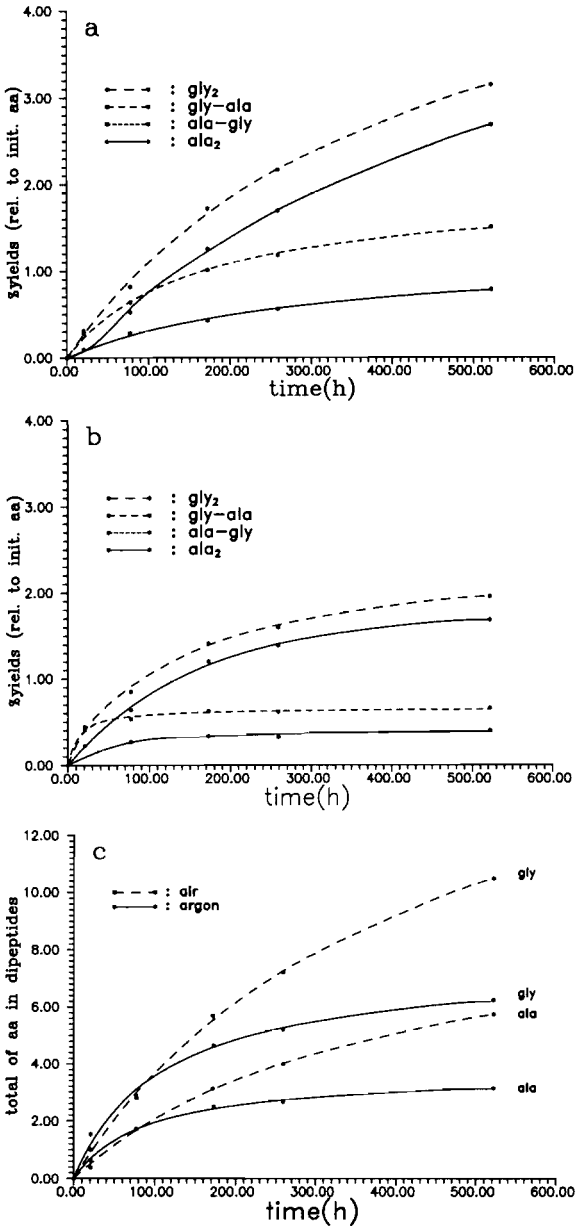

Fig. 2. Dependence of yields of dipeptides (in % of total of initial amino acids) on time, under air atmosphere (a), and under argon (b), at a pH of 2.8. (c) % of amino acids incorporated into dipeptides under air and argon, respectively.

oxygen [18]: the east siberian Aldan group, formed *c.* 4 billion years ago contains carbonates, sulfates and iron formations, which demand the presence of O_2. Therefore, together with the characteristic presence of 'green zones', in precambrian rock formations, mainly consisting of malachite and azurite, it seems to be reasonable to assume that both O_2 and Cu^{2+} were always present. So the experiments under air can be assumed to come closer to prebiotic conditions than those under pure argon atmosphere.

Except for early stages of the reaction, higher amounts of ala–gly than gly–ala are formed. This cannot be attributed to copper catalysis, but rather to acidic sequence inversion via the mixed cyclic anhydride of glycine plus alanine, methyl-piperazinedione. As illustrated in Fig. 3, ala–gly and gly–ala quickly approach an equilibrium ratio [gly–ala]/([gly–ala]+[ala–gly]) of 0.29, both in the presence and absence of copper, within a short time. The same ratio is obtained in our experiments, showing that the condensation of alanine and glycine would favor gly–ala, but the much faster ring-closing/ring-opening reaction inverts this distribution. This sequence inversion, already known for some time [19], is also possible with other dipeptides [20] and hence will also rule the formation of higher peptides under our conditions.

Table 1 summarizes some selected results obtained in peptide condensation experiments. The highest temperature gave highest yields within the shortest times, but as oxidation of amino acids becomes significantly faster, the latter process fully compensates this initial advantage. After 450 h only a small amount of amino acids and peptidic material is left.

The last column of Table 1 allows the estimation of stereoselectivity in our system. A threefold amount of optically pure ala_2 compared to the other diastereomers is formed, and longer reaction times reduce this excess only slowly.

Control experiments carried out without Cu^{2+} under otherwise identical conditions did not lead to measurable amounts of dipeptides.

As can be seen in Fig. 1, peaks of amino acids and dipeptides are the dominating ones in the chromatogram, nevertheless some minor peaks following gly–ala and in the region of ala_2 indicate the formation of additional peptides. In order to verify the presence of higher peptides, chromatography as described in Section 2 was applied. Figure 4, showing the chromatogram of the last fraction, clearly proves the formation of gly_3, $glyala_2$, and all tripeptides containing 2 glycines ($alagly_2$, gly–ala–gly, gly_2–ala). Their low yields (<0.1%) and the incomplete separation on the column did not allow a quantitative determination of their relative amounts.

The mechanism of chain elongation becomes evident by two simple experiments, in which ala–gly and gly–ala are brought to reaction with an equimolar amount of glycine for a short time, under our usual reaction conditions (Cu^{2+}/NaCl). As can be seen from Table 2, the only detectable tripeptide in the case gly/ala–gly is ala–gly–gly, while in the case gly/gly–ala the main product is gly–ala–gly. The minor formation of ala–gly–gly in the latter reaction can be attributed to the fast inversion of gly–ala (see

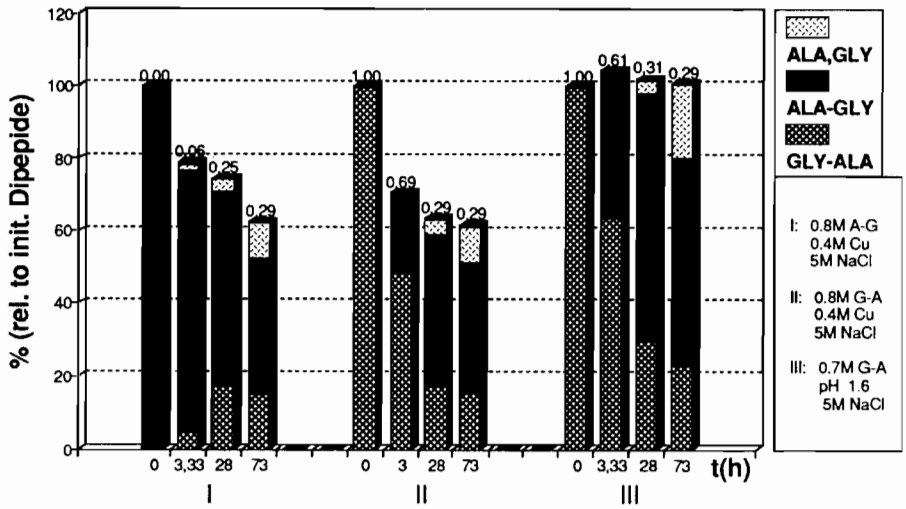

Fig. 3. Sequence inversion and hydrolysis of ala–gly and gly–ala as a function of time. The numbers above the bars are the molar fractions of gly–ala.

TABLE 1. Peptide formation in solutions containing gly, ala, Cu^{2+} and NaCl at selected reaction times[a]

Conditions	Time (h)	$pept_{gly}$[b]	$pept_{ala}$[b]	gly	gly_2	glyala	ala	alagly	ala_2	$x^{+/+c}$
gly, L-ala, Cu^{2+}: 0.49 M	173	4.7	2.5	76	2.8	0.63	90	1.2	0.7	0.77
NaCl: 5 M	259	5.2	2.7	73	3.2	0.62	85	1.4	0.7	0.74
78 °C, argon	522	6.3	3.1	73	3.7	0.66	85	1.7	0.8	0.69
gly, L-ala, Cu^{2+}: 0.49 M	172	5.7	3.1	58	3.4	1.0	71	1.3	0.9	0.79
NaCl: 5 M	259	7.2	4.0	42	4.3	1.2	59	1.7	1.1	0.79
75 °C, air	522	10.5	5.8	20	6.3	1.5	34	2.7	1.6	0.79
gly, DL-ala, Cu^{2+}: 0.50 M	143	6.1	3.5	40	3.6	0.87	54	1.6	1.0	0.5
NaCl: 4.5 M	305	5.7	3.6	8.9	2.9	0.92	23	1.9	0.8	0.5
85 °C, air	449	0.3	0.7	1.2	0.05	0.06	7.7	0.2	0.4	0.5

[a]All yields are given in % of initial amino acids. [b]% of glycine and alanine, respectively, incorporated into dipeptides. [c]$x^{+/+}$: $[ala^{+/+}]/([ala^{+/+}]+[ala^{+/-}])$.

Fig. 3). It follows that chain elongation mainly proceeds through nucleophilic attack of the amino group of the amino acid to the free carboxylic group of the peptide.

As hydrolysis of Cu^{2+} and complexation automatically decreases the pH in our system to the rather low value of 2.8, the unprotonated $-NH_2$ function can only be delivered by complexation with Cu^{2+}, while the availability of the non-coordinated protonated carboxylic group of the peptides is not questionable at this pH.

This result is consistent with the previously reported data [21], that complexes of chlorocuprates with diglycine are weaker than the corresponding glycine complexes, and hence the binding sites of Cu^{2+} are preferably occupied by the amino acid.

It should be mentioned here, that a series of experiments with other metal ions did not lead to enhanced peptide formation [14, 15]. The unique role of Cu^{2+} therefore seems to be based on its specific binding ability of amino acids and peptides, and possibly also the value of its Cl^- complexation constant.

4. Conclusions

The dipeptides gly_2, gly–ala, ala–gly and ala_2 are readily formed in good yields in aqueous solution in the presence of Cu^{2+} and high concentrations of NaCl. Under optimal conditions more than 10% of glycine and 5% of alanine are incorporated into dipeptides; formation of higher peptides also occurs.

Preferential formation of a large excess of ala–gly over gly–ala can be observed; it is caused by acidic sequence inversion.

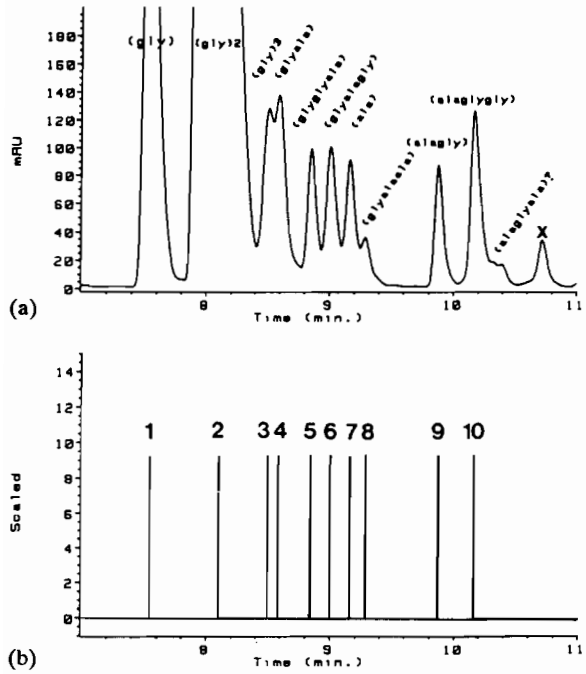

Fig. 4. Chromatogram of last silica gel fraction, proving formation of tripeptides, together with a schematic representation of the reference peaks. 1: gly, 2: gly_2, 3: gly_3, 4: gly–ala, 5: gly_2–aly, 6: gly–ala–gly, 7: ala, 8: gly–ala_2, 9: ala–gly, 10: ala–gly_2; X: unknown peak.

TABLE 2. Formation of tripeptides in a solution containing 0.4 M peptide, 0.4 M glycine, 0.4 M Cu^{2+} and 5 M NaCl after 11 h at 80 °C under argon

Peptide	gly–gly–ala	gly–ala–gly	ala–gly–gly
gly–ala	0.07	0.8	0.2
ala–gly			0.6

Yields are given in % of the initial peptides.

Racemization occurs only to a limited extent; the formed peptides therefore maintain widely the stereospecific structure of the original amino acids.

The salt-induced peptide condensation is in principal independent of the atmosphere, but secondary reactions (oxidations by Cu^{2+}) let traces of O_2 appear favourable for the reaction.

Acknowledgement

This work was supported by the Austrian 'Fonds zur Förderung der wissenschaftlichen Forschung' (Project P7393-CHE), who also supplied the HPLC equipment. This generous support is gratefully acknowledged.

References

1 E. Ochiai, *Origins Life, 9* (1978) 81.
2 J. Carver, *Nature* (London), *292* (1981) 136.
3 C. R. Woese, *J. Mol. Evol., 13* (1979) 95.
4 J. Baross, *Origins Life, 15* (1985) 327.
5 T. Oie, G. H. Loew, S. K. Burt, J. S. Binkley and R. D. MacElroy, *J. Am. Chem. Soc., 104* (1982) 6169.
6 T. Oie, G. H. Loew, S. K. Burt and R. D. MacElroy, *J. Am. Chem. Soc., 105* (1983) 2221.
7 A. L. Weber and L. E. Orgel, *J. Mol. Evol., 11* (1978) 189.
8 D. A. Buckingham, D. M. Foster and A. M. Sargeson, *J. Am. Chem. Soc., 92* (1970) 5701.
9 S. W. Fox and K. Dose, *Molecular Evolution and the Origin of Life,* Marcel Dekker, New York, 1977.
10 J. Rabinowitz, *Helv. Chim. Acta, 53* (1970) 1350.
11 H. Sawai and L. E. Orgel, *J. Mol. Evol., 6* (1975) 185.
12 J. Hulshof and C. Ponnamperuma, *Origins Life, 7* (1976) 197.
13 G. Steinman and M. N. Cole, *Proc. Natl. Am. Soc., 58* (1967) 735.
14 M. G. Schwendinger and B. M. Rode, *Anal. Sci., 5* (1989) 411.
15 B. M. Rode and M. G. Schwendinger, *Origins Life, 20* (1990) 401.
16 G. Schlesinger and S. Miller, *J. Mol. Evol., 19* (1983) 376.
17 R. Schuster, *J. Chromatogr., 431* (1988) 271.
18 H. D. Pflug, *Naturwissenschaften, 71* (1984) 63.
19 D. A. Long and T. G. Truscott, *Trans. Faraday Soc., 59* (1963) 1833.
20 A. A. Brewerton, D. A. Long and T. G. Truscott, *Trans. Faraday Soc., 66* (1970) 2297.
21 R. Tauler and B. M. Rode, *Inorg. Chim. Acta, 173* (1990) 93.