Pure & Appl. Chem., Vol. 65, No. 6, pp. 1143–1151, 1993.
Printed in Great Britain.

From amino acids to prebiotic active peptides: A chemical reconstitution

André Brack

Centre de Biophysique Moléculaire, C.N.R.S. 1A, Avenue de la Recherche Scientifique - 45071 - Orléans Cedex 2 - France
Fax : 38 63 15 17

Abstract. Primitive life can be schematically described as organic molecules processed by liquid water. Amino acids were most likely available on the primitive Earth, home-made from the primitive atmosphere or in hydrothermal vents. Import of extraterrestrial amino acids may represent an alternative supply. Experimental evidence is given for selective condensation of amino acids driven by the hydrolytic power of liquid water. When hydrophobic and hydrophilic amino acids are forced to coexist within the same chain in liquid water, the duality generates interesting topologies such as stereoselective and thermostable β-pleated sheet structures. Polycationic alternating peptides containing lysyl and hydrophobic residues strongly accelerate the hydrolysis of oligoribonucleotides. The β-sheet geometry plays an determinant role in the observed chemical activity, as in contemporary enzymes. β-sheet forming peptide sequences may represent a support for primordial information amplification.

According to the primordial soup hypothesis proposed by Oparin in 1924, small organic molecules were formed in a reducing atmosphere dominated by methane. When reaching terrestrial liquid water, the organic molecules were processed and generated the constituants of the first living systems and their nutrients. Since the historical experiment of Stanley Miller (ref. 1), nearly all of the biogenic elementary building blocks have been successfully synthesized from methane (CH_4) or its derivatives, formaldehyde (HCHO) and - in the presence of nitrogen - hydrogen cyanide (HCN). By analogy with contemporary living systems, it was generally believed that primitive life emerged as a cell, thus requiring at least three families of organic compounds : boundary molecules able to isolate the system from the environment, informative molecules allowing the storage and the transfer of the information needed for replication and chemically active molecules providing the basic chemical work of the system. It was also believed that chemists would be skilful enough to make small-scale versions of these molecules in order to construct in the laboratory an artificial primitive cell.

The present paper reviews some small-scale versions of proteins which have been synthesized in the laboratory.

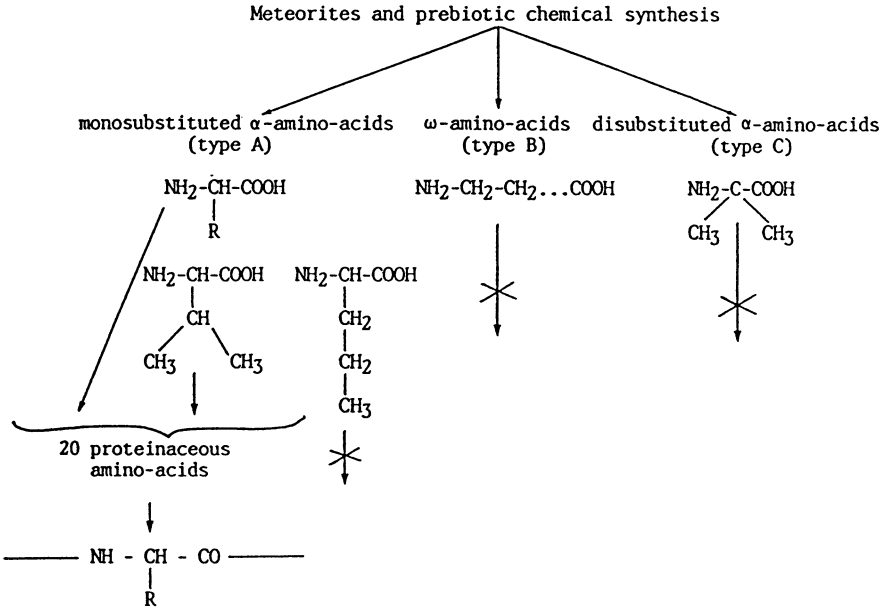
Most of the chemical reactions in a contemporary living cell are achieved by proteinaceous enzymes. Proteins are built up with twenty different amino acids. Each amino acid, with the exception of

glycine, exists under two enantiomeric forms L and D but proteins use only L ones. Proteins adopt asymmetrical rigid geometries, α-helices and β-sheets, which play a key role in the catalytic function. Proteins, even the smallest ones, are too sophisticated to be considered to be the products of an organic chemistry working at random, without any chemical selection. The chemist has therefore to understand how primitive proteins were selected and how they began to exhibit chemical activity.

1. SOURCES OF AMINO ACIDS

Amino acids were most likely available on the primitive Earth. They have been synthesized in the laboratory under very simple conditions using different sources of energy such as electric discharges (ref. 2), U.V. (ref. 3), shock-waves (ref. 4), X-rays (ref. 5) or heat (ref. 6). Amino acids have also been found in carbonaceous chondrites. For instance, eight proteinaceous amino acids have been detected in the Murchison meteorite, i.e. glycine, alanine, proline, leucine, isoleucine, valine, aspartic acid and glutamic acid (ref. 7). However, the mixture of amino acids is complex : it contains different types of compounds exemplified in scheme 1.

Scheme 1. Present-day proteins use only one family of amino acids (type A). Within the selected family the side-chain R must also fulfil certain conditions



In addition, it is generally admitted that the first amino acids available on the primitive Earth were racemic mixtures since their synthesis were run in a symmetrical environment. However, in view of the importance of optical purity in present life, it is difficult to believe that, at the beginning, a completely racemic life using simultaneously biomolecules of both configurations in the same protocell arose. Recently, Engel et al. (ref. 8) reported an excess of L-type alanine in the Murchison meteorite (D/L of 0.85). If the data of Engel are valid, a meteoritic preference for L-amino acids would push the problem of the origin of biological chirality out into the cosmos. Among the possible extraterrestrial sources of circularly polarized light, sunlight is generally discounted as being probably too weak and not of a consistent handedness for

sufficiently long periods. According to Bonner and Rubenstein (ref. 9), synchrotron radiation from the neutron star remnants of supernova events is a better candidate. Interaction of neutron star circularly polarized light with interstellar grains in dense clouds could produce chiral molecules in the organic mantles. The delivery of the chiral molecules to the Earth may have been achieved via comets or/and asteroids.

How could selected peptides, i.e. short condensates of homochiral proteinaceous α-amino acids, emerge from the mixture of organic compounds in aqueous solution ?

2. AQUEOUS CONDENSATION OF AMINO ACIDS

Peptides are formed when water molecules are removed from amino and carboxylic acid functions. Peptide chemistry offers a whole range of activating agents to condense amino acids in organic solvents. In water, the number of condensing agents is restricted, especially when looking for prebiotically plausible compounds.

Carbodiimides, R-N=C=N-R, are commonly used in organic media. They can be used in water providing a careful choice of the substituent R and of experimental conditions. Under these conditions, Cavadore and Previero (ref. 10) obtained long peptides in water up to the 30-mer. The simplest carbodiimide, H-N=C=N-H, can be considered as a tautomeric form of cyanamide NH_2-CN which can be obtained by UV irradiation of an aqueous solution of ammonium cyanide, NH_4CN. The formation of cyanamide requires the presence of iron which helps the absorption of UV energy. In fact, cyanamide is not stable and forms a dimer, dicyandiamide or cyanoguanidine NH_2-C(=NH)-NH-CN which is as reactive as carbodiimides. Peptides were obtained with cyanamide and cyanoguanidine. However, the reactions were very slow and did not proceed beyond the tetrapeptide (ref. 11-14). With diaminomaleonitrile, NC-$C(NH_2)$=$C(NH_2)$-CN) the formation of diglycine in 3.1 % yield has been observed (ref. 15).

Clays can also be used to condense amino acids in water. In an homogeneous aqueous solution, mixed anhydride alanyladenylate condensed partially up to heptaalanine but deactivation via hydrolysis remained the main pathway. In the presence of clays, such as montmorillonite, hydrolysis was suppressed and discrete polyalanines were obtained (ref. 16). However, montmorillonite-mediated polymerization of alanyladenylate did not lead to high oligopeptides in quantitative yield in other laboratories (ref. 17, 18). Lahav et al. (ref. 19) subjected mixtures of glycine and Na-kaolinite or Na-bentonite to wet-dry and temperature fluctuations (25-94 C) and observed the formation of oligopeptides up to five glycine residues in length. Only trace amounts of diglycine formed without clays. White and Erickson (ref. 20) studied the effects of the dipeptide histidyl-histidine in the polymerization of glycine during fluctuating moisture and temperature cycles on kaolinite. A turnover of 52 was observed, i.e. each molecule of dipeptide helps the polymerization of 52 molecules of glycine.

Surfactant aggregates have a spectacular effect on the polymerization of aminoacyladenylates as shown by Armstrong et al. (ref. 21). Amino esters oriented in monolayers (ref. 22, 23) or in micelles (ref. 24, 25) polymerize easily.

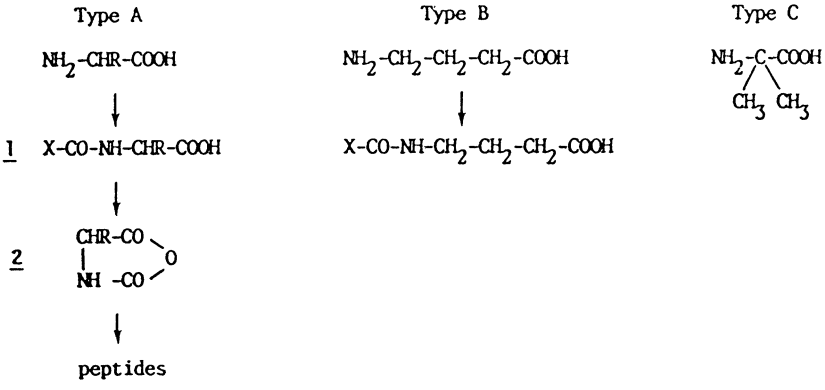
Thermal condensation of amino acids has been largely described by Fox and co-workers (ref. 26). They have shown that dry mixtures of amino acids polymerize when heated at 130 C to give "proteinoids".

In the presence of polyphosphates, the temperature can be decreased to 60 C. High molecular weights were obtained when an excess of acidic or basic amino acids are present. When heated in aqueous solutions at 130-180 C, the proteinoids aggregate spontaneously in microspheres of 1-2 μ, presenting an interface resembling the lipid bilayers of living cells. Under appropriate conditions, these microspheres increase slowly in size from dissolved proteinoids and are sometimes able to bud and to divide like bacteria. These microspheres catalyze the decomposition of glucose and work as esterases and peroxydases. The main advantage of proteinoids is their polymeric character and their organization into particles, but they represent a dramatic increase in complexity. When heating a mixture of selected L-amino acids, one gets a polycondensate which is only 50 % peptidic, the peptidic fraction is racemized, the peptide linkages are ambiguous since they include the α, β and γ functions of the dicarboxylic amino-acids and the sequences are multiple although not completely random.

3. SELECTIVE AQUEOUS CONDENSATION OF AMINO ACIDS

N-carboxyanhydrides (compound 2 in scheme 2) are good candidates for the selective polymerization of proteinaceous amino acids in water. Their formation requires an N-derivatization (1) - which is known to be difficult with disubstituted amino-acids due to steric hindrance - followed by a ring closure involving the α-carboxylic function. Ring-closure has proven to be possible with 5-membered rings (α-amino acids) and with 6-membered rings to a lesser extent (β-alanine). It should not be effective with larger rings (γ-amino-acids, γ-carboxylic group of glutamic acid, dipeptides, etc).

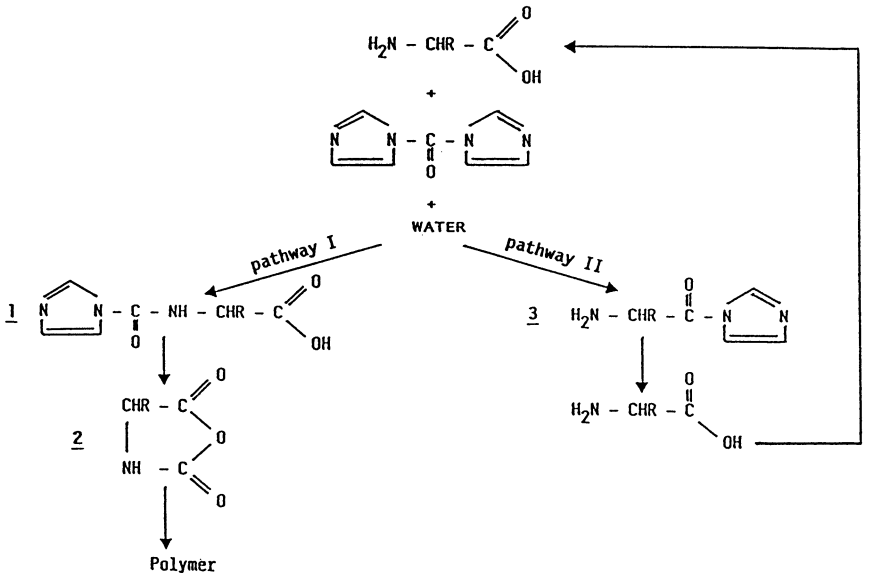
Scheme 2. Selective condensation of amino acids via N-carboxyanhydrides



Direct activation with N,N-carbonyldiimidazole (CDI) proceeds in water through the formation of the N-imidazoylcarbonyl amino acid (ref. 27) which ring-closes to give N-carboxyanhydrides (ref. 28). With glutamic acid, the condensation involves only the α-carboxylic function and not the side-chain. A mixture of amino acids close to that found in the Murchison meteorite was dissolved in water and treated with CDI. The condensate was found to be enriched in amino acids of type A (ref. 28).

Activation of amino acids by CDI in aqueous solution illustrates a chemical peculiarity of liquid water. In organic solvents, CDI is known to activate N-protected amino acids R-COOH to give the corresponding imidazolide R-CO-N⟨ ⟩N (ref. 29, 30). Added to free amino acids in organic solvents, CDI

Scheme 3. Liquid water drives the polymerization of amino acids via N-carboxyanhydrides



led to compound 3 in scheme 3 (ref. 31, 32). In the presence of water, aminoacylimidazolides are quite stable between pH 3 and pH 9, with nevertheless some hydrolysis (pathway II in scheme 3). They polymerize to afford oligopeptides but also substantial amounts of diketopiperazine, a cyclic dipeptide which is a dead end (ref. 32). The formation of imidazolides directly in water seems unlikely since only 6 % of the expected acetylimidazolide could be found in situ by IR spectroscopy when CDI was added to acetic acid in D_2O solution. Imidazolide absorption band could not be detected when CDI was added to free amino acids (ref. 28). Under those conditions, most organic chemists would consider as hopeless to use CDI for the polymerization of amino acids in water. Ehler and Orgel (ref. 27) nevertheless did the experiment and obtained oligopeptides via the intermediary formation of N-imidazoylcarbonyl amino acid 1. By continuous extraction of the reaction mixture with chloroform, we were able to isolate a compound presenting the IR absorption bands characteristic of N-carboxyanhydrides 2.

4. SELECTIVE ACCUMULATION OF HOMOCHIRAL ALTERNATING POLYPEPTIDES

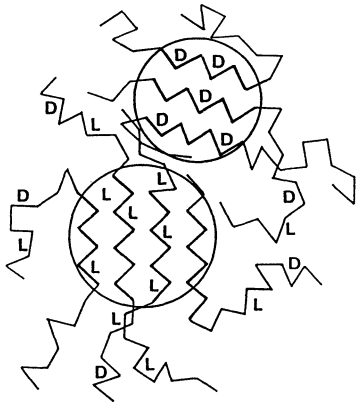
Abiotic amino acids, produced from primitive terrestrial atmospheres or in space, can be divided into two families according to the chemical composition of the side-chains : hydrocarbons which try to escape the water molecules and ionizable groups which have an affinity for water through hydrogen bonding. When these two species are forced to coexist within the same molecules, the duality generates interesting topologies. Over short distances, hydrophobic and ionizable groups generate chain confomations which depend strongly on the sequence.

Strictly alternating homochiral poly (L-valyl-L-lysyl), hereafter referred to as poly(Val-Lys), is soluble in water. At neutral pH, the lysyl side-chain amines are ionized as NH_3^+ groups. Due to charge repulsion, the chain cannot adopt a regular conformation. Addition of salt to this solution, for instance 0.1 M NaCl, produces a screening of the charges and allows the polypeptide to adopt a β-sheet structure (ref. 33). Because of the alternating sequence, all hydrophobic residues are confined to one side of each stand. The chains aggregate into asymmetrical bilayers with a hydrophobic interior and a hydrophilic exterior because of hydrophobic side-chain clustering.

Due to bilayer formation, strictly alternating hydrophobic-hydrophilic sequences are thermostable. Non-alternating sequences form α-helices which are thermolabile. Heating samples in which α- and β-structures coexist increases the amount of β-structure with a loss of α-helix. Alternating sequences are also more resistant to chemical degradation than α-helical sequences (ref. 34). To get a β-sheet and therefore a high resistance to degradation, the hydrophobic amino acids must display their property to a marked degree. For instance, poly (αAbu-Lys) associating L-α amino butyric acid with L-lysine, does not form β-sheets and was found to be 15 times more sensitive to mild hydrolysis than poly (Leu-Lys) (ref. 28).

Aggregation of alternating sequences to form β-sheets is possible only with homochiral (all-L or all-D) polypeptides. For instance, racemic alternating poly(D,L-Leu-D, L-Lys) is largely unable ot adopt the β-structure and remains mostly unstructured (ref. 35). When increasing amounts of L-residues are introduced into the racemic alternating polypeptide, the proportion of β-sheets increases and there is a good relationship between the percentage of β-form and the amount of L-residues in the polymer. The molecules can be described as a mixture of β-sheets and disordered segments. Those segments containing six or more homochiral residues aggregate into stable nuclei of optically pure β-sheets surrounded by the more fragile heterochiral disordered segments (ref. 36, 37).

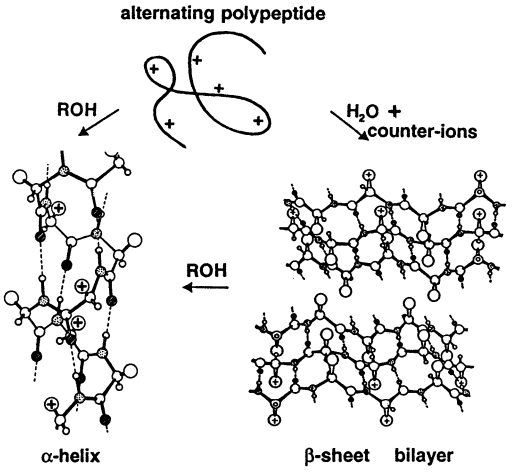
Scheme of a poly D,L-peptide chain with β-sheet nuclei circled



The samples were subjected to mild hydrolysis. The kinetic measurements showed two rate constants in agreement with the existence of two conformational species. After partial hydrolysis, the remaining polymeric fraction was enriched in the dominant enantiomer (ref. 38). The enantiomeric excess (absolute value of % L-%D) had increased from 54 to 68 %.

When the properties of water are modified by addition of increasing amounts of alcohol, the repulsion between charges and the stress on the hydrophobic groups are progressively released. Under these conditions, an α-helix is obtained for poly(Leu-Lys) instead of the β-bilayer described above. Even more interesting, when the alcohol is added to an aqueous solution of alternating poly(Leu-Lys) already transconformed into the β-structure by addition of salt, a β-to-α transition is obtained, probably by relaxation of the water constraint on the hydrophobic groups (ref. 39) (scheme 4).

Scheme 4. Liquid water drives the conformation of alternating polypeptides



5. CHEMICAL ACTIVITY OF SIMPLE PEPTIDES

We have shown in our laboratory that polypeptides containing basic amino acids strongly accelerate the hydrolysis of oligoribonucleotides at pH 7.5. Alternating poly(Leu-Lys) has been found to be the most active (ref. 40). For instance, the rate of hydrolysis of the oligonucleotide $(Ap)_9A$ is increased by a factor 185. Analysis of the hydrolysis products indicated that the polypeptide accelerates the classical base-induced hydrolysis of polyribonucleotides.

The basic polypeptides interact with oligoribonucleotides to form a complex. When increasing the ionic strength in the reaction mixture, the formation of the complex becomes less likely. When the salt molarity varies from 0 to 2, the hydrolytic activity decreases regularly and becomes negligible when the complexes do not form.

The β-sheet conformation is induced when the polypeptides are complexed to oligoribonucleotides as shown by infrared spectroscopy. The structure plays an important role in the hydrolysis since poly(Pro-Lys-Leu-Lys-Leu), which cannot adopt the β-sheet conformation because of the prolyl residues, is practically inactive as a hydrolytic catalyst. Racemic, alternating poly(D,L Leu - D,L Lys) was also found to have very little activity. As already mentioned, the simultaneous presence of both L- and D- residues impedes the formation of the β-structure. A set of alternating poly(leucyl-lysyl) samples ranging from the racemic to the homochiral all-L polymer has been synthesized. Their conformations can be described as a mixture of random coil and β-sheet structures, the amount of β-sheet increasing with the proportion of

L-residues in the polymer. The hydrolytic activity follows linearly the proportion of β-sheets, indicating that the β-sheets are involved in the hydrolysis (ref. 41).

The alternating polypeptides tested in the hydrolysis experiments are polydisperse with regard to chain-length and average about 50 residues per chain. Shorter well-defined monodisperse peptides were prepared in order to evaluate the critical chain-length required for the hydrolytic activity. Oligopeptides, acetyl-$(Leu\text{-}Lys)_n$-ethylamide, were prepared with n = 1,3 and 5. The hydrolytic activity of the di-, hexa- and decapeptide represents respectively 4, 13 and 92 % of that of $(Leu\text{-}Lys)_{26}$ (ref. 41).

CONCLUSION

Nucleotide chemists have failed to demonstrate that accumulation of substantial quantities of relatively pure oligonucleotides on the primitive Earth was a plausible chemical event. Therefore, the RNA world appears as an episode in biology but not at the very beginning of life.

The chemist has now to invent a chiral self-replicating system able to amplify simple information, with a small error rate to allow selection, without the participation of RNA. Was the first amplifying machinery based on thioester chemistry, on organic molecules adsorbed on pyrite or on organic molecules associated with clays ?

Amino acids were most likely available on the primitive Earth, either home-made or imported from space. Starting from a mixture of amino acids in liquid water, peptides can be selected on the basis of chemical selection and chemical resistance toward degradation through simple pathways which have been demonstrated in the laboratory. Some simple well-defined peptides have also been shown to develop active sites with efficient chemical activity. Accumulation of chemically active peptides on the primitive Earth appears therefore plausible via thermostable and stereoselective β-sheets made of alternating polypeptides. Large avenues are thus offered to peptides and should lead to new results.

REFERENCES

1. S.L. Miller, *Science* **117**, 528-529 (1953).
2. G. Schlesinger and S.L. Miller, *J. Mol. Evol.* **19,** 376-382 (1983).
3. C. Sagan and B.N. Khare, *Science* **173**, 417-420 (1971).
4. I. Barak and A. Bar-Nun, *Origins of Life* **6**, 483-506 (1975).
5. K. Dose and S. Risi, *Z. Naturforschung* **236**, 571-587 (1968).
6. J.G. Lawless and C.D. Boynton, *Nature* **243**, 405-407 (1973).
7. J.R. Cronin, S. Pizzarello and D.P. Cruikshank *in Meteorites and the Early Solar System* pp 819-857, University of Arizona Press (1988).
8. M.H. Engel, S.A. Macko and J.A. Silfer, *Nature* **348**, 47-49 (1990).
9. W.A. Bonner and E. Rubenstein, *BioSystems* **20**, 99-111 (1987).
10. J.C. Cavadore and A. Previero, *Bull. Soc. Chim. Biol.* **51**, 1245-1253 (1969).
11. M. Halmann, *Arch. Biochem. Biophys.* **128**, 808-810 (1968).
12. D.W. Nooner, E. Sherwood, M.A. More and J. Oro, **10**, 211-220 (1977).

13. J.R. Hawker and J. Oro, **17**, 285-294 (1981).
14. G. Steinman, R.M. Lemmon and M. Calvin, *Science* **147**, 1574 (1965).
15. S. Chang, J. Flores and C. Ponnamperuma, *Proc. Nat. Acad. Sci.* **64**, 1011-1015 (1969).
16. M. Paecht-Horowitz and F.R. Eirich, *Origins of Life* **18**, 359-387 (1988).
17. J.T. Warden, J.J. Mc Cullough, R.M. Lemmon and M. Calvin, J. *Mol. Evol.* **4**, 189-194 (1974).
18. A. Brack, *Clay Minerals* **11**, 117-120 (1976).
19. N. Lahav, D.H. White and S. Chang, *Science*, **201**, 67-69 (1978).
20. D.H. White and J.D. Erickson, *J. Mol. Evol.* **16**, 279-290 (1980).
21. D.W. Armstrong, R. Seguin, C.J. Mc Neal, R.D. Mac Farlane and J.H. Fendler, *J. Amer. Chem. Soc.* **100**, 4605-4606 (1978).
22. K. Fukuda, Y. Shibasaki and H. Nakahara, *J. Macromol. Sci. Chem.* **A 15**, 999-1014 (1981).
23. T. Folda, L. Gros and H. Ringsdorf, *Makromol. Chem., Rapid Commun.* **3**, 167-174 (1982).
24. N. Kunieda, M. Watanabe, K. Okamoto and M. Kinoshita, *Macromol. Chem.* **182**, 211-214 (1981).
25. K. Hanabusa, K. Kato, H. Shirai and N. Hojo, *J. Polymer Sci.* Part C **24**, 311-317 (1986).
26. S.W. Fox and K.D. Dose, *in Molecular evolution and the origin of life*, Marcel Dekker, New York (1977).
27. K.W. Ehler and L.E. Orgel, *Biochim. Biophys. Acta* **434**, 233-243 (1976).
28. A. Brack, *Origins of Life* **17**, 367-379 (1987).
29. G.W. Anderson and R. Paul, *J. Amer. Chem. Soc.* **80**, 4423 (1958).
30. R. Paul and G.W. Anderson, *J. Amer. Chem. Soc.* **82**, 4596-4600 (1960).
31. B.P. Gottikh, A.A. Krayevsky, N.B. Tarussova, P.P. Purygin and T.L. Tsilevich, *Tetrahedron* **26**, 4419-4433 (1970).
32. A.L. Weber and J.C. Lacey Jr., *Biochim. Biophys. Acta* **349**, 226-234 (1974).
33. A. Brack and L.E. Orgel, *Nature* **256**, 383-387 (1975).
34. A. Brack and G. Spach, *in Origins of Life*, Y. Wolman (ed.), D. Reidel Pub. Cy 487-493 (1981).
35. A. Brack and G. Spach, *J. Mol. Evol.* **13,** 35-46 (1979).
36. G. Spach and A. Brack, *J. Mol. Evol.* **13,** 47-56 (1979).
37. A. Brack and G. Spach, *Origins of Life* **11**, 135-142 (1981).
38. A. Brack and G. Spach, *J. Mol. Evol.* **15**, 231-238 (1980).
39. A. Brack and G. Spach, *J. Amer. Chem. Soc.* **103**, 6319-6323 (1981).
40. B. Barbier and A. Brack, *J. Amer. Chem. Soc.* **114**, 3511-3515 (1992).
41. A. Brack and B. Barbier, *Origins of Life* **20**, 139-144 (1990).