Journal of Biological Physics 31: 599–606, 2005.
DOI: 10.1007/s10867-005-6063-7

Homochirality in Bio-Organic Systems and Glyceraldehyde in the Formose Reaction

S. TOXVAERD
Department of Chemistry, H. C. Ørsted Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark; e-mail: tox@st.ki.ku.dk

Abstract. The article explores the possibility that the ordering of bio-organic molecules into a homochiral assembly at the origin of life was performed not in aqueous solutions of amino acids or related materials but in racemic glyceraldehyde in the "formose" reaction at high concentration and temperature. Based on physical chemical evidence and computer simulations of condensed fluids, it is argued that the isomerization kinetics of glyceraldehyde is responszible of the symmetry break and the ordering of molecules into homochiral domains.

Key words: homochirality, formose reaction, origin of life

1. Introduction

This article explores the possibility that the ordering of bio-organic molecules into a homochiral assembly at the origin of life was performed not in solutions of amino acids or related material; but in racemic glyceraldehyde (GLA), $CH_2OHCHOHCHO$, in the "formose" reaction at high concentration and temperature. The formose reaction is the spontaneous condensation of formaldehyde into sugar which was discovered in 1861 [1]; the aldol-like mechanism is described in [2].

Formaldehyde, CH_2O, is the simplest carbohydrate and can be synthesized from carbon dioxide. It is produced in a reducing atmosphere containing CO_2 and methane [3]; but what is more interesting it is also synthesized from volcano volatiles by reduction with either methane or sulfide minerals [4]. The condensation into sugar is catalyzed by not only amino acids [5], but also by naturally occurring aluminosilicates at hydrothermal springs and the "black smokers" [6]. So the formose reaction is well known and is believed to be the source of sugars and related bio-organic molecules at the origin of life [6, 7]. Based on physical-chemical evidence and computer simulations of condensed fluids, it is here argued that this reaction is also responsible for the symmetry break and the ordering of molecules into homochiral domains.

2. Chiral Discrimination

The interactions between chiral molecules will in general be stereospecific: Interactions between similar ($++$ or $--$) and dissimilar ($+\,-$) molecules are different. This distinction is called "chiral discrimination". Whereas the effect of chiral discrimination in the crystalline phase has been known since Pasteur's famous experiment [8], it is generally believed that the effect is small in the fluid phase [9] as it is the case for diluted aqueous solutions. The first two achiral reactants in the formose reaction, formaldehyde and glycolaldehyde, are miscible with water, and the other chiral carbohydrates are in general very soluble due to hydrophilic alcohol groups, but an exception is GLA which is only moderately soluble in water (3 g in 100 ml $\approx$ 0.3 M at 20 °C). In diluted aqueous solutions, the effect of chiral discrimination is small, and chiral sugar molecules will end in a racemic mixture due to the active keto-enol isomerization kinetics. Also, the isomerization kinetics for chiral amino acids will give a racemic aqueous solution [10] of D- and L-amino acids, although this isomerization kinetics is much slower than, for example the conversion between D- and L-GLA [11]. (This fact is a main objection against many models for symmetry break and creation of homochirality. For example, if L-amino acids arrived on Earth via a meteorite, as was indeed recently the case [12], synthezised at solid surfaces, or obtained by spontaneous crystallization [9], the racemization starts when the L-amino acids are dissolved in water).

The chiral discrimination in the crystalline phase favours, in general, the racemic crystal [13]. The reaction Gibbs free energy for the reaction

$$\text{D-crystal} + \text{L-crystal} = \text{racemic crystal} \tag{1}$$

is negative for most chiral systems and proportional to the difference, $T_R - T_E$, between the melting point temperature, T_R, of the racemic crystal and the corresponding melting point temperature, T_E, of the enantiomers [13]. For most of the crystals investigated in [13], racemic crystals have higher melting point temperatures than their corresponding enantiomer crystals. The reaction Gibbs free energy for (1) is only positive when the crystal of an enantiomer has a melting point temperature which is more than $\approx$25 degrees higher than the melting point temperature of the racemic crystal. Glyceraldehyde is not only unique with respect to its solubility in water. The stability of the enantiomer crystals is unique as well. The melting point temperatures are: T_R = 130–132 °C [14, 15] and T_E = 194–198 °C [16], which gives a temperature difference of $\approx$65 °C, strongly favouring the enantiomer molecular arrangement.

The stability of D-GLA(s) and L-GLA(s) is in fact unique, e.g. a related amino acid, serine has the opposite property. Its two melting temperatures are: T_R = 231–233 °C and T_E = 228 °C, respectively, in favor of the racemic arrangement. But L-Serine is especially interesting because it exhibits chiral discrimination in diluted aqueous solutions and forms clusters with D-GLA [17], and it exhibits

corresponding chiral discrimination with other L-amino acids [18] whereby it establishes the stereospecific link between D-sugars and L-amino acids in aqueous solutions.

When GLA is dissolved in water it forms oligomers [19, 20], with dimers and five-membered rings predominant. Its solubility in a typical organic solvent, 1-octanol, is low, and equal to the solubility in water [21]. The stability of D-GLA(s) and L-GLA(s) indicates that it energetically prefers homochiral domains. The isomerization kinetics is active in aqueous solutions and the fluid phase, but not in crystals. The hypothesis is that the kinetics, which ensures a racemic solution of GLA in a diluted aqueous solution, will on the contrary order a racemic mixture at high concentration in the fluid state into homochiral domains by a "domain catalyzed isomerization kinetics" [22, 23] due to GLA's strong chiral discrimination. But the domain-catalyzed isomerization kinetics also breaks the racemic symmetry and favors **one** homochiral domain.

3. Molecular Dynamics Simulations of Domain Catalyzed Isomerization Kinetics

The systems of GLA molecules are set up in a qualitative way by taking a system of D- and L-particles with different attraction strength. A pair of enantiomers, D,D or L,L, will attract each other more strongly than a pair of D-and L-GLA, and it is this energy gain which causes the physical-chemical behavior described in the introduction and the high melting point temperature of the enantiomer crystals. In a simple qualitative model, which does not account explicitly for interactions such as hydrogen bonds, the chiral discrimination can be obtained by varying the strength of attractions between pairs of D,D (=L,L) and D,L [22].

The isomerization kinetics is active in the fluid state and in an aqueous solution especially at high temperatures. The kinetics for the conversion

$$\mathrm{D} + \mathrm{D} \underset{E_{\mathrm{DL}}}{\overset{E_{\mathrm{DD}}}{\rightleftharpoons}} \mathrm{D} + \mathrm{L} \underset{E_{\mathrm{LL}}}{\overset{E_{\mathrm{DL}}}{\rightleftharpoons}} \mathrm{L} + \mathrm{L} \tag{2}$$

is implemented e.g. as a collision-activated kinetics [22] where the activation energy, E_{DL}, for DL-collisions, which may convert a D-molecule into a L-molecule or *vice versa*, is less than the corresponding activation energy $E_{\mathrm{DD}} = E_{\mathrm{LL}}$, allowing a conversion of one of the molecules in the collisions. This inequality

$$E_{\mathrm{DL}} < E_{\mathrm{DD}} = E_{\mathrm{LL}} \tag{3}$$

accounts for the chiral discrimination with a lower potential energy of pairs and domains of enantiomers. The system is bistable due to the isomerization kinetics, and the kinetics results in an eventual dominance of one of the species at late times.

The time evolution of systems of many thousand D- and L-particles, started as racemic fluid- and solute mixtures, are obtained by Molecular Dynamics (MD) [22, 23] in $\approx$10 ns. Let particle i at time t in the system of N particles be a D-particle.

If a high energy collision with e.g. an L-particle, j, activates particle i to a (pair-) potential energy, $u_{i,j}(t) \geq E_{\mathrm{DL}}$, the conversion is performed on the basis of the energy difference between the two states, ΔE. The total intermolecular potential energy of the i-particle with all the other N-1 D- and L-particles at time t, $u_{i,\mathrm{D}}(\mathbf{r}^{\mathbf{N}}(t))$, is calculated as well as the corresponding energy, $u_{i,\mathrm{L}}(\mathbf{r}^{\mathbf{N}}(t))$, if the activated particle i is converted to a L-GLA. The difference

$$\Delta E = u_{i,\mathrm{L}}(\mathbf{r}^{\mathbf{N}}(t)) - u_{i,\mathrm{D}}(\mathbf{r}^{\mathbf{N}}(t)) \tag{4}$$

is used to perform the kinetics. If ΔE is small, as is the case for a diluted racemic solution of D- and L-particles, the mixture is observed to remain racemic. Correspondingly, a diluted solution of either D- or L-particles ends in a racemic mixture, as do diluted solutions of sugars and amino acids in nature. But in a concentrated solution or pure fluid phase of the particles a strong chiral discrimination results in the fact that there is a significant energy difference ΔE. When a molecule is energetically excited, e.g. at the collision to the (intramolecular) saddle point energy between its two conformational states, the chiral discrimination given by ΔE will tend to bring the excited molecule to the state with the smallest energy. By choosing the conversion with a Boltzmann probability

$$p = \frac{e^{-\Delta E/kT}}{1 + e^{-\Delta E/kT}} \tag{5}$$

one describes the domain effect on the particle as a thermodynamic force on the chiral molecule [22].

The gain in potential energy by simple diffusion-driven segregation into homochiral enriched D- and L-domains is enhanced by the active isomerization kinetics; but the kinetics also breaks the symmetry in the racemic mixture. This intermolecular domain-catalyzed kinetics can be implemented in the MD in different ways [22] which, however, all lead to a symmetry break and a total dominance of one of the species for sufficient energy difference between the activation energies (3).

Figure 1 shows an example of symmetry breaks, given by the excess fraction between D- and L-particles. It is not possible to predict when the symmetry break appears, nor which fraction will eventually dominate. But it is usually the largest of the established homochiral subdomains in the fluid or the biggest homochiral droplet in the solution that finally dominates. The systems end in the equilibrium states with (for e.g. a D-particle dominance)

$$\frac{[\mathrm{D}]}{[\mathrm{L}]} = e^{-\frac{(E_{\mathrm{DL}} - E_{\mathrm{DD}})}{kT}} \tag{6}$$

provided that the pre-exponential collision parameter is the same for the two kinds of conversions, as it must be in the present simple model.

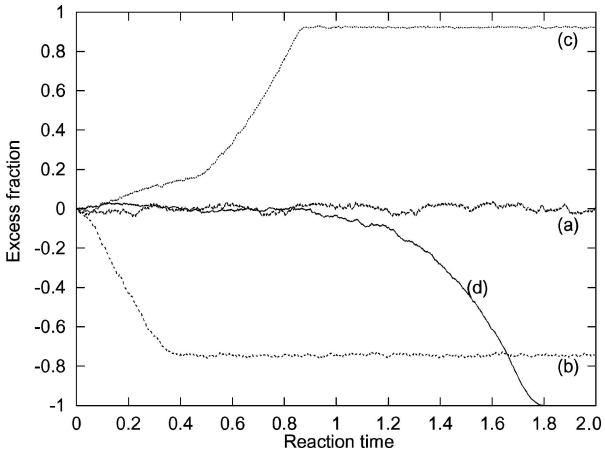

Figure 1. Excess fraction, $(N_D - N_L)/(N_D + N_L)$, as a function of reaction time in units of *ns*. The reactions are all for an activation energy $E_{\mathrm{DL}} = 3kT$ and different values of $E_{\mathrm{DD}} = E_{\mathrm{LL}}$. (a): $E_{\mathrm{DD}} = E_{\mathrm{DL}} + kT$; (b): $E_{\mathrm{DL}} + 2kT$; (c): $E_{\mathrm{DL}} + 3kT$; and (*d*): $E_{\mathrm{DD}} = E_{\mathrm{LL}} = \infty$. The racemic mixture symmetry is broken for $E_{\mathrm{DD}} - E_{\mathrm{DL}} > kT$.

When GLA was synthezised by the spontaneous formose reaction it must have been in the presence of other organic molecules. GLA is amphiphilic in the sense that it is moderately solvable in water and in an organic solvent. Such amphiphiles separate together with hydrophobic compounds in water in various ways, including emulsion-like separation and micelle formation [21]. A commonly used bio-organic solvent is 1-octanol. The distribution coefficient for GLA in a water/1-octanol system at 25 °C has been measured for different concentrations of GLA (1, 2, 4 and 8 mg/ml) and is $\approx$1 (mole fraction) [21]. A distribution coefficient equal to 1 is obtained in the qualitative MD-model when the GLA has the same attractive energy per GLA-particle in a diluted solution in "water" (W) and in an organic solute (O). The organic molecules in a ternary mixture of the diluted solute O- and GLA-particles segregate by cooling the mixture down. The GLA-particles segregate together with the O-particles with an emulsion-like droplet formation for a reduced attractive energy of the GLA-particles to W- and O-particles between 40 to 75% respectively. The GLA-particles segregate with the O-particles in droplets; but with the amphiphilic GLA also spread in the droplet interface in a surfactant-like manner. Figure 2 shows a distribution of the GLA particles after droplet segregation, in a large droplet containing an equal number of O-particles (not shown).

The possibility of an emulsion-like separation with a subdomains containing only the amphiphiles is important for the stereospecific ordering described in [22, 23]. If a multicomponent mixture of organic components including GLA is cooled down and performs this kind of phase segregation, the GLA can undergo an ordering into one of its enantiomer configurations by a domain-catalyzed isomerization

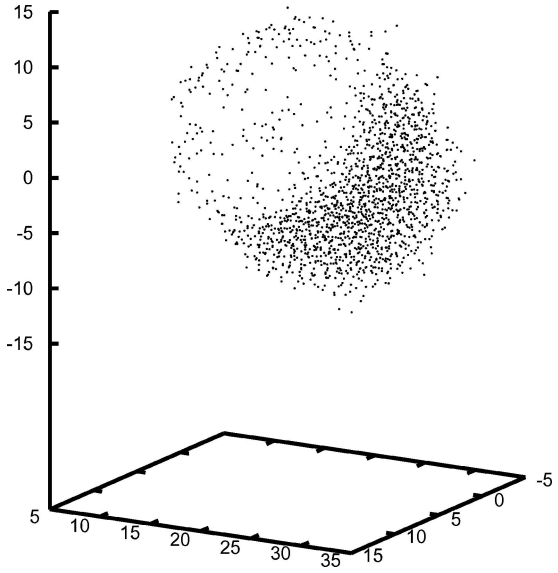

Figure 2. Distribution of GLA-particles in a droplet of GLA- and O-particles (the O-particles in the interior of the droplet and the solvent W-particles are not shown). The GLA-particles perform their own subphase in the droplet and with some GLA-particles in the remaining solvent-droplet interface.

kinetics of the molecules in the GLA-subdomain. At the same time the GLA-molecules are present in the interface between the hydrophobic components and water, where they can act as the seeds for synthesis of glycerides and phosphoroglycerides. Once they react, e.g. with carboxylic acids by creation of esters (at carbon atoms No. 1 and No. 2), the resulting molecules behave as stereospecific surfactants (the keto-enol isomerization kinetics is then turned off) and will remain in the interface and ensure the stability of the micelle structure.

4. Discussion

The creation and the amplification of enantiomeric excess from a racemic mixture are only the first steps on a long pathway toward a living organism. The hypothesis that spontaneous homochirality originates from the formose reaction, obtained by an active isomerization kinetics in the fluid phase at high pressure, concentrations and temperature, deviates on several points from other hypothesis about the origin of homochirality in bio-organic matter. It is obtained by a spontaneous reaction between flexible molecules, synthesized in large quantities at the hydrothermal springs and not a rare event between stable molecules which appear in (very) low concentrations, as it must have been the case for amino acids. If GLA was present in high concentrations in coherent fluid domains the keto-enol kinetics could have ensured homochirality. That it gave D-GLA dominance instead of

L-GLA dominance can be characterizes by a morphological principle, which one can describe as "self-stabilizing coincidence". But once the D-GLA dominance was established the effect of chiral discrimination with the amino acids ensured a bio-organic environment of D-GLA and L-amino acids [17, 18].

Another crucial task in morphogenesis is the formation of a boundary that separates the aqueous medium containing optically pure material from the external achiral environment. All known organisms are cellular and each cell is separated from the aqueous environment by a thin bilayer composed of stereospecific membrane molecules. Glyceraldehyde is an interesting molecule in this context. It is the simplest carbohydrate with a chiral center and chiral glyceroderivatives in the membranes are homochiral and can be synthesized from GLA in a reducing environment, as has been the case in the early oceans [24]. The membranes of archaeobacterias are ether-like glyceroderivatives from carboxylic alcohols whereas procaryote and eucaryote membranes consist of esters with carboxylic acids [25]. Membranes of homochiral glyceroderivatives are common for all living organisms including the most primitive forms [25, 26].

D-glyceraldehyde-3-phosphate is the central molecule in glycolysis which is also active in an anaerobic environment, and glycolysis requires only sugar for spontaneous reaction, all of which could be obtained from the formose reaction, in the deep oceans. Existing organisms could carry the chemical records of chemical evolution at the origin of life. The homochirality of the "anchor" molecules in the membranes and the central role of D-GLA in the basal anaerobic metabolism could be the track laid by the morphogenesis of the origin of homochirality in living organisms.

There is of course a very long route of self-assambly from simple emulsion-like droplets with a stereospecific ordering of their surfaces of D-GLA to even the simplest procaryote or archaea with double-layered membranes, a very complex cytokinesis (cell division) and a metabolism. The present hypothesis offers a possible explanation of the very first steps on this route.

References

1. Butlerow, A.: Formation Synthetique d'une Substance Sucrée, *Compt. Rend. Acad. Set.* **53** (1861), 145–147.
2. Breslow, R.: On the Mechanism of the Formose Reaction, *Thetrahedron Lett.* **21** (1959), 22–26.
3. Schlesinger, G. and Miller, S.L.: Prebiotic Synthesis in Atmospheres Containing CH_4, CO, and CO_2, *J. Mol. Evol.* **19** (1983), 383–390.
4. Vladimirov, M.G., Ryzhkov, Y.F., Alekseev, V.A., Bogdanovskaya, V.A., Otroshchenko, V.A. and Kritsky, M.S.: Elektrochemical Reduction of Carbon Dioxide on Pyrite as a Pathway for Abiogenic Formation of Organic Molecules, *Origin Life Evol. Biosphere* **34** (2004), 347–360.
5. Weber, A.L.: The Sugar Model: Catalysis by Amines and Amino Acid Products, *Origin Life Evol. Biosphere* **31** (2001), 71–86.
6. Gabel, N.W. and Ponnamperuma, C.: Model for Origin of Monosaccharides, *Nature* **216** (1967), 453–455.

7. Washington, J.: The Possible Role of Volcanic Aquifers in Prebiologic Genesis of Organic Compounds and RNA, *Origin Life Evol. Biosphere* **30** (2000), 53–79.
8. Pasteur, M.L.: Recherches sur les Proprietes Specifiques des deux Acides qui composent Acide Racémique, *C. R. Acad. Sci.* **26** (1848), 535; *Ann. Chim. Phy.* **28** (1850), 56–99.
9. Avalos, M., Babiano, R., Cintas, P., Jimenez, J.L. and Palacios, J.C.: Symmetry Break by Spontaneous Crystalization – Is It The Most Plausible Source of Terrestrial Handedness We Have Long Been Looking For? – A Reappraisal, *Origin Life Evol. Biosphere* **34** (2004), 391–405.
10. Bada, J.L.: Kinetics of Racemization of Amino Acids as a Function of pH, *J. Amer. Chem. Soc.* **94** (1972), 1371–1373.
11. Fedoroňko, M. and Königstein, J.: Kinetics of Mutual Isomerization of Trioses and their Dehydration to Methylglyoxal, *Coll. of Czechoslov. Chem. Commun.* **34** (1969), 3881–3894.
12. Pizzarello, S.: Chemical Evolution and Meteorites: An Update, *Origin Life Evol. Biosphere* **34** (2004), 25–34.
13. Leclercq, M., Collet, A. et Jacques, J.: Mésure de la Stabilité des Racémiques Vrais, *Tetrahedron* **32** (1976), 821–828.
14. Gresham, W.F. and Grigsby, W.E.: An Improved Synthesis of DL-Glyceraldehyde, *J. Org. Chem.* **14** (1949), 1103–1107.
15. Vik, J.-E.: Base-Catalyzed Retro Condensation Reactions of Some ot-Hydroxymethyl Carbonyl and Nitro Compounds, *Acta Chem. Scand.* B **28** (1974), 509–516.
16. Morgenlie, S.: Oxidation of Carbohydrate Derivatives with Silver Carbonate on Celite, *Acta Chem. Scand.* **27** (1973), 1557–1564.
17. Takats, Z., Nanita, C. and Cooks, R.G.: Serine Octamer Reactions: Indicators of Prebiotic Relevance, *Angew. Chem. Int. Ed.* **42** (2003), 3521–3523.
18. Kock, K.J., Gozzo, F.C., Nanita, S.C., Takats, Z., Eberlin, M.N. and Cooks, R.G.: Chiral Transmission Between Amino Acids: Chirally Sellective Amino Acid Substitution in the Serine Octamer as a Possible Step in Homochirogenesis, *Angew. Chem. Int. Ed.* **41** (2002), 1721–1724.
19. Bystrický, S., Sticzay, T., Polÿaková, M. and Fedoroňko, M.: Study of D-glyceraldehyde by Circular Dicroism and Ultraviolet Spectroscopy, *Coll. Checoslovak. Chem. Comm.* **46** 240–245.
20. Yaylayan, V.A., Harty-Majors, S. and Ismail, A.A.: Investigation of DL-glyceraldehyde-dihydroxyacetone Interconversion by FTIR Spectroscopy, *Carbohydrate Res.* **318** (1999), 20–25.
21. Toxvaerd, S.: Droplet Formation in a Ternary-Fluid Mixture: Spontaneous Emulsion and Micelle Formation, *J. Phys. Chem.* **108** (2004), 8641–8645.
22. Toxvaerd, S.: Molecular Dynamics Simultions of Isomerization Kinetics in Condensed Fluids, *Phys. Rev. Lett.* **85** (2000), 4747–4750.
23. Toxvaerd, S.: Domain Catalyzed Chemical Reactions: A Molecular Dynamics Simulation of Isomerization Kinetics, *J. Chem. Phys.* **120** (2004), 6094–6099.
24. Holland, H.D.: *The Chemical Evolution of the Atmosphere and Oceans,* Princeton University Press, Princeton 1984.
25. Itoh, Y.H., Sugai, A., Uda, I. and Itoh, T.: The Evolution of Lipids, *Adv. Space Res.* **28** (2001), 719–724.
26. Peretó, J., López-García, P. and Moreira, D.: Ancestral Lipid Biosynthesis and Early Membrane Evolution, *TRENDS in Biol. Sci.* **29** (2004), 469–477.