Available online at www.sciencedirect.com

SCIENCE DIRECT®

*Mendeleev Commun.*, 2007, **17**, 7–9

Mendeleev Communications

# Kinetic models of the prebiological evolution of macromolecules. Thermocycle as the motive force of the process

**Sergei D. Varfolomeev**[a,b]

[a] *N. M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, 119991 Moscow, Russian Federation. Fax: +7 495 939 5417; e-mail: sdvarf@sky.chph.ras.ru*

[b] *Department of Chemistry, M. V. Lomonosov Moscow State University, 119992 Moscow, Russian Federation*

DOI: 10.1016/j.mencom.2007.01.003

**The kinetic simulation of processes involved in the synthesis of macromolecules in thermocyclic mode reveals that the possibility of efficient selection of macromolecules exists and absolutely improbable events can be turned to absolutely certain ones.**

The appearance of self-propagating polymeric molecules with a huge narrowing of the diversity of macromolecular structures is the key aspect in the origination of life.[1] The spontaneous formation of phosphodiester bonds (nucleic acids) and peptide bonds (proteins) can give a virtually unlimited diversity of molecules with different actual structures. The proteins and nucleic acids that currently exist on the Earth constitute an insignificant part of all molecules that could be formed in primary processes and subsequently underwent selection to form real biological structures. For example, the human genome consists of $3\times10^4$ genes. However, $20^{100} = 10^{130}$ different proteins containing about 100 amino acid residues can be constructed from 20 amino acids. The situation with the number of active sites of enzymes is even more paradoxical. The entire today's biological world is 'served' by a limited number of structures (100–150) that have formed as enzyme catalytic sites.[2–5]

Thus, to explain the phenomenon of the origination of life, one has to explain, from the standpoint of natural sciences, the mechanism of molecular convergence of macromolecule structures, their selection based on certain principles, and the emergence of a mechanism of their proliferation (self-replication).

In his book devoted to the evolution of macromolecules at a prebiological level, Eigen[6] gave a kinetic description of selection processes of macromolecules using the 'competitive advantage' concept. However, beyond the scope of these studies, a question remains open regarding the mechanism of processes of macromolecule self-replication, the motive force that ensures 'ordering' and 'competitive advantage' and, ultimately, molecular convergence of the diversity of polymers to a limited number of types of biomacromolecules.

We believe that this mechanism can be identified, namely, it is related to the utilisation of solar energy and rotation of the Earth around its axis. As a result of rotation, any point of Earth's surface undergoes cyclic variations in temperature. On the primordial Earth, when the atmosphere was rather rarified, the amplitude of these variations was rather high. Depending on specific conditions (radiation, heat exchange and thermal conductivity), the variations can have amplitudes ranging from +200 to –100 °C. Various chemical reactions and phase transitions occur under these conditions.

Let us consider how thermocycling conditions can affect the synthesis and evolution of macromolecular structures. It is assumed that the system contains a sufficient amount of monomers.[1]

In the chemical sense, contemporary biomacromolecules are the products of polycondensation, a reaction that occurs with formation of water and gives the amide bond (proteins and peptides) or phosphodiester bond (nucleic acids). A monomeric molecule should be trifunctional and should contain a reaction centre of polymerization and a set of functional substituents that make it possible to chemically discriminate one monomeric molecule from another. Potential functional substituents can react with each other in a supramolecular way and with some selectivity determined by the free energies of complex formation due to hydrogen bonds, ionic and hydrophobic interactions.

Let us consider the processes that occur in the monomer–polymer system if the temperature varies in a cyclic manner.

*1. Temperature increase above 100 °C, phase transition of water to the gas state, transition of monomers to the solid phase.* The following processes occur under these conditions: polycondensation, formation of chemical bonds, formation of polymers with irregular structure and transition of water to the gas phase. Reactions of this type are base reactions for the synthesis of polyamides and polyesters in the present-day chemistry of macromolecular compounds.[7] The reactions occur spontaneously; the reaction rates and depths (degree of polymerization) depend on the temperature.

*2. Temperature decrease below 100 °C, water condensation, formation of a mixed solution of a polymer and monomers.* Sorption of monomers on the polymer can occur under these conditions. If there is a certain degree of complementarity that is provided by supramolecular interaction, the selection of monomers on the polymer occurs.

*3. Temperature increase above 100 °C, phase transition of water, transition of the polymer and monomers to the solid phase.* Polycondensation processes occur under these conditions; interactions between monomers sorbed on the polymer are advantageous to some extent. This process is known as matrix polymerization in the chemistry of macromolecular compounds. The rate and selectivity of polymerization depend on the monomer-to-polymer affinity, the conditions of mutual orientation of monomer functional groups, and the environment temperature.

*4. Temperature decrease below 100 °C, water condensation and formation of a liquid phase of solution containing double polymer chains parts of which are complementary.* If the temperature decreases and the system is considerably diluted with condensed water, the chains are separated to give a polymer–monomer system.

Subsequent processes occur in a cycle with self-propagation of partially complementary polymeric chains. Note that each of the four process stages occurs spontaneously under thermodynamically favourable conditions.

Since the process is cyclic, it is most convenient to relate time ($t$) to the number of cycles

$$t = n\tau_c, \tag{1}$$

where $n$ is the number of cycles and $\tau_c$ is the duration of one cycle. In our particular case, $\tau_c$ equals the Earth's daily rotation cycle. To provide a kinetic description based on the model described above, it is convenient to introduce kinetic parameter $a$ that characterises the behaviour of the system within one cycle. If the number of macromolecules in the system is doubled during a cycle, the number of molecules will increase according to the law

$$N = N_0 2^n, \tag{2}$$

where $N_0$ is the initial concentration of macromolecules.

In the general case,

$$N = N_0 a^n, \tag{3}$$

where $a$ is a kinetic parameter, $0 < a \leq 2$, which can be called the replication factor.

The power function can be expressed as

$$a^n = e^{kt},\ k = \ln a/\tau_c. \tag{4}$$

This makes it possible to analyse the system behaviour within classical concepts of chemical kinetics using a differential equation with continuous functions.[8]

Let us consider several kinetic models of how events can develop.

*The evolution-segregation model.* Let us imagine that, as a result of primary processes in the system at high temperatures, polycondensation of monomers occurred to give a primary set $N_1, \ldots N_i \ldots N_p \ldots N_n$ of macromolecules with a broad distribution by molecular mass, composition and sequence of functional substituents. As a result of thermal cycling in accordance with the mechanism described above, each molecule has certain capability for proliferation:

$$dN_i/dt = k_i N_i \text{ for all } 1 \leq i \leq n. \tag{5}$$

The parameter $k_i$ includes constants $a_i$ [equation (4)] and it is an efficient characteristic that reflects the balance of increase in the amount and destruction of a certain macromolecular structure

$$k_i = k_s^i - k_h^i, \tag{6}$$

where $k_s^i$ is the rate constant of the synthetic process and $k_h^i$ is the rate constant of the destruction process, primarily hydrolysis, which is a process opposite to polycondensation. At $k_i > 0$ ($a_i > 1$), an exponential increase in the number of this type of macromolecules will occur, whereas at $k_i < 0$ ($a_i < 1$), an exponential decrease in their number or concentration will be observed.

$$N_i = N_{i,0}\exp(k_i t), \tag{7}$$

where $N_{i,0}$ is the number or concentration of the $i$th type of macromolecules at the initial moment in time. The independent development of each type of macromolecules will occur in the system.

Let us assume that, due to some structural features in the broad distribution of parameters $k_i$ in the system, a polymer is synthesised, whose parameter $k_p$ is somewhat higher (maybe, quite insignificantly) than the parameters $k_i$, $i \neq p$ for the other macromolecules. Let us name this macromolecule $N_p$. This means that either polymer $N_p$ is more stable or its synthesis is somewhat faster

$$N_p = N_{p,0}\exp(k_p t). \tag{8}$$

The initial fraction of $N_{p,0}$ can be very small,

$$N_{p,0} =<< \sum_{i \neq p} N_{i,0}.$$

The competitive advantage $k_p > k_i$ can be due to either higher synthesis rates ($k_s^p > k_s^i$) or higher stability, immunity to hydrolysis, *i.e.* $k_h^p < k_h^i$. It can be stated from general considerations that the more efficiently the monomers are organised on the polymer matrix, *i.e.*, the higher the free energy of supramolecular interaction of a monomeric unit with a complementary area of the polymer, the more efficiently matrix synthesis occurs, that is, the higher its rate.

For a polymer of type $N_p$, the fraction of a given type of macromolecules will be

$$\eta_p = N_p/(N_p + \sum_{i \neq p} N_i) = 1/\{1 + \sum_{i \neq p}(N_{i,0}/N_{p,0})\exp[(k_i - k_p)t]\}. \tag{9}$$

At $k_i < k_p$ and $t \to \infty$, $\eta_p \to 1$, *i.e.*, after a time, the entire population of polymers will be represented by a single structure $N_p$. The kinetic behaviour of the system is determined by the parameters $N_{i,0}/N_{p,0}$ and $(k_i - k_p)$.

The process involves an induction period; transition to a monostructure can occur starting from a single successful molecule. Structures that are synthesised faster or have higher stability displace all others in the course of time.

*Autocatalytic process – convergence explosion.* The situation changes considerably if the structure $N_p$ acquires catalytic properties that increase its proliferation rate. This means that a structure appears that accelerates one of the stages of $N_p$ formation.

It is important here that a nonlinear term appears in the differential equations that describe the process dynamics

$$dN_p/dt = k_p' N_p^2, \tag{10}$$

where $k_p'$ is the second-order rate constant. In this case,

$$N_p(t) = N_{p,0}/(1 - k_p' N_{p,0} t). \tag{11}$$

Accordingly, the fraction of productive molecules will change with time according to the law

$$\eta_p = N_{p,0}/[N_{p,0} + (1 - k_p' N_{p,0} t)\sum_{i \neq p} N_{i,0}\exp(k_i t)]. \tag{12}$$

The function has a jump at $t_{cr} = 1/k_p' N_{p,0}$.

Evolution of the system accelerates abruptly when the value of $t_{cr}$ is approached. The fraction of 'positive' molecules rapidly reaches unity. In this case, we deal with an explosion of a kind, where the system is quickly converted to a unified structure.

The molecular mechanisms that result in the influence of $N_p$-type molecules on their own synthesis can vary greatly. They can include catalytic hydrolysis of 'inactive' polymers, creation of systems to protect their own structures from nonspecific hydrolysis, participation in the synthesis of monomers within the scope of a compartment, and, finally, direct participation in acceleration of matrix polymerization.

To compare the equations obtained with real times of prebiological evolution, it is now convenient to obtain time estimates based on the number of thermocycles $n$ taking into account equation (4).

In this case, equation (9) can be transformed to

$$\eta_p = (N_{p,0}/N_{n,0})/[N_{p,0}/N_{n,0} + (a_n/a_p)^n], \tag{13}$$

where $N_{n,0}$ is the entire set of non-productive molecules with smaller values of $a_n$.

For a quantitative evaluation it is convenient to introduce the parameter $n_{k0}$, such that $\eta_p = 1/2$. It follows from equation (13) that in the case of 'evolution segregation'

$$n_{cr} = \ln(N_{p,0}/N_{n,0})/\ln(a_n/a_p). \tag{14}$$

A quantitative analysis of the equations obtained allows us to make a number of principally important conclusions.

*Thermocycling is a method to change absolutely improbable events (probability of $10^{-50}$) to absolutely certain ones (prob-*

**Table 1** Time required for the evolution segregation of macromolecules with a change in $\eta_p$ from $10^{-50}$ to 1 at different ‘competitive advantage’ values $a_p/a_n$.

| $a_p/a_n$ | Number of cycles | Time, millions of years |
|---|---|---|
| 1.000000006 | $1.91\times10^{10}$ | 52.6 |
| 1.000000008 | $1.43\times10^{10}$ | 39.4 |
| 1.000000013 | $8.85\times10^{9}$ | 24.3 |
| 1.000000025 | $4.6\times10^{9}$ | 12.6 |

*ability of 1).* Estimates reveal the extraordinary ability of a system for molecular convergence even if the ‘competitive advantages’ are exceedingly small.

Let us assume that the initial fraction of ‘productive’ molecules is vanishingly small, $N_{p,\mathrm{o}}/N_{n,0} = 10^{-50}$. The times required for virtually all the population of macromolecules to be converted to unified structure $N_p$ depending on the parameter $a_p/a_n$ are listed in Table 1.

It is well known that the age of the Earth is estimated to be 4.56 milliard years and that the first microorganisms are observed in formations 3.5 milliard years old.[1] Thus, the duration of prebiological evolution is estimated to be within several hundred million years. One can see in Table 1 that the evolution of a system to single-structure polymers from one molecule in a population of $10^{50}$ molecules can take 10–50 million years.

A system evolves to a single structure (or a group of structures with similar values of ‘competitive advantage’) even at vanishingly small values of the parameter that characterises the ‘competitive advantage’. Table 1 indicates that the prebiological evolution process is successful even with very small values of ‘competitive advantage’. A six-milliardths fraction of the ‘replication factor’ is sufficient for complete conversion of a system to a single structure in 50 million years. Estimates of the parameter $1 < a_p < 2$ based on concepts about the nature of physico-chemical interaction of a monomer with a polymer molecule show that this parameter should be considerably higher. This gives us reasons to believe that the prebiological process of macromolecule selection can be much faster.

If macromolecules acquire catalytic functions, this results in ‘autocatalytic explosion’ and accelerates dramatically the conversion to single-structure systems. It follows from equation (11) that, at $t \to 1/k'_p N_{p,0}$, the variation in the fraction of ‘productive’ molecules accelerates abruptly and the function has a jump at $t_{\mathrm{cr}} = 1/k'_p N_{p,0}$.

On transition from continuous time to the number of cycles, we obtain

$$n_{\mathrm{cr}} = 1/k'_p N_{\mathrm{p},0}\tau_c.$$

It can be shown that, if evolution and autocatalytic processes occur at equal rates at initial moments in time,

$$k'_p N_{p,0} = \ln\,(a_p/a_n)/\tau_{\mathrm{c}}.$$

Let us consider the process regularities in case of hypothetical evolution of small RNA. Assuming the tentative mean length of molecules to be 100 bases, the initial number of possible synthesised variants of macromolecules will be

$$N_{n,0} = 4^{100} \cong 10^{60}.$$

Let us assume that a single molecule with ‘competitive advantage’ expressed as a ‘replication factor’ has appeared in this set of molecules, and that this molecule has catalytic activity that favours its replication. If we assume $a_p/a_n = 1.000001$, in accordance with ref. 14, the evolution–segregation process will be characterised by $n_{\mathrm{cr}} = 1.36\times10^8$ cycles or a time of 0.378 million years. The transition is very sharp and is completed within $1.36\times10^8$–$1.4\times10^8$ cycles.

For ‘autocatalytic explosion’, we have $n_{\mathrm{cr}} = 1/\ln\,(a_p/a_n)$, which gives $10^6$ cycles or $2.73\times10^3$ years.

Thus, the catalytic properties of a molecule accelerate the evolution process by a factor of ~140.

The above estimates are of interest for comparative purposes, but they demonstrate that, given quite an insignificant ‘competitive advantage’ ($a_p/a_n = 1.000001$), a single molecule synthesised accidentally (one out of $10^{60}$) can completely change the environment and become the only structure within a period of time that is small on the scale of the Earth evolution (~3000 years).

The transformation of prebiological macromolecules into biomacromolecules occurred through a number of stages, each of which could have an exceedingly small probability ($10^{-10}$ or lower) with an overall trend for increase in $a_p$ from 1 to 2. Prebiological evolution could include the following stages:

i. Parallel evolution of the worlds of polyesters and polyamides with the participation of PNA molecules as catalysts of certain stages.

ii. Appearance of peptides that served as catalysts in RNA and DNA replication (protection of ‘productive’ structures from hydrolysis, hydrolysis of ‘nonproductive’ structures, transportation of nucleotides to replication centres, acceleration of replication process).

iii. Appearance of the trinucleotide–amino acid match.

iv. Appearance of prebiological ribosomes and a coding mechanism for the synthesis of catalysts.

v. Appearance of catalysts with selection of a limited number of active centres.

vi. Abandonment of thermocycle – transition to the synthesis under isothermal conditions.

Each of the stages can have a very low probability and can occur in a random way at the level of separate molecules or molecular formations. In this process, each stage can develop according to one of the models presented above (evolution segregation or autocatalytic explosion). The increase in $a_p$ is a necessary condition of the process that ensures the survival and domination of productive structures. Eventually, during transition from prebiological to biological evolution, $a_p$ has increased from 1 to 2.

The following conclusion is obvious from our analysis: the emergence of life is a natural and deterministic development of macromolecular structures; a prerequisite of the process is the presence of trifunctional monomers forming irregular polymers, which evolve at the initial stage under thermocycling conditions. Life could not but appear under these conditions.

I am grateful to Academician E. M. Galimov for initiating my interest in the problem of origination of life and for valuable discussions. The study was supported by the Presidium of the Russian Academy of Sciences programme (25 p) ‘Problems of Biosphere Origination and Evolutions’.

## References

1 E. M. Galimov, *Fenomen zhizni* (*The Phenomenon of Life*), URSS, Moscow, 2001 (in Russian).
2 S. D. Varfolomeev and K. G. Gurevich, *Izv. Akad. Nauk, Ser. Khim.*, 2001, 1629 (*Russ. Chem. Bull., Int. Ed.*, 2001, **50**, 1709).
3 S. D. Varfolomeev, *Mendeleev Commun.*, 2004, 185.
4 S. D. Varfolomeev, I. V. Uporov and I. A. Gariev, *Usp. Khim.*, 2005, **74**, 67 (*Russ. Chem. Rev.*, 2005, **74**, 61).
5 S. D. Varfolomeev, *Khimicheskaya enzimologiya* (*Chemical Enzymology*), Akademiya, Moscow, 2005 (in Russian).
6 M. Eigen, *Self-Organization of Matter and the Evolution of Biological Macromolecules*, Springer Verlag, Berlin, 1971, p. 214.
7 Yu. D. Semchikov, *Vysokomolekulyarnye soedineniya* (*Macromolecular Compounds*), Akademiya, Moscow, 2003, p. 367 (in Russian).
8 S. D. Varfolomeev and K. G. Gurevich, *Biokinetika* (*Biokinetics*), Fair-Press, Moscow, 1999 (in Russian).

*Received: 8th August 2006; Com. 06/2769*