Hypotheses on the role of the free enthalpy in origin of life and biological evolution

Gérard Berger

Abstract

A model is proposed to account for the origin of life and evolution of living organisms, based on the necessary decrease of the free enthalpy of the biosphere. This would be possible by the role of the enzymes which activate the biological reactions. We hypothesize that their active sites have been produced not by chance and natural selection but by condensation of aminoacids in the presence of substrates. The information would have been transferred to nucleic acids by the reverse protein synthesis mechanism. A lot of genes of enzymes would have been distributed between different species of monocellular organisms.

Some of them would have aggregated into colonies, including eventually species of complementary metabolisms. As different kinds of reactions might have occurred at the same time, a more rapid decrease of the free enthalpy received from outside would have resulted, promoting the phenomenon of aggregation.

As the places and the numbers of the different cells of the living organisms are maintained by reproduction, a strict mechanism must exist. We hypothesize that short repeated sequences present in the genome would have played this role.

After reunions into colonies in the early times of evolution, only punctual mutations and modifications of repartition of genomes, but no new activities, would have occurred.

The different theories of origin of life and evolution are discussed, with respect to our model.

Keywords

Free enthalpy, Gibbs free energy, cell colonies, repeated sequences, genome.

Introduction

The origin of life and the causes of the evolution of the living organisms have been the subjects of countless debates. As nucleic acids can be used as templates for their own reproduction, they have been considered by most biologists as the first molecules of the living world. Different theories have been developed to account for the evolution of the organisms. Apart from the theory of the creation of fixed species, the theory of natural selection of favorable mutations is, nowadays, the mainstream one. The theory of the role of selection by chance of particular characters, when the number of individuals is low, has been developed. The theory of the action of the environment has been in vogue at a given time, has declined, but is now rising again. A possible thermodynamic perspective of biologic evolution has been also considered, and is preferentially exposed in this article: from birth of life to the present organisms, the decrease of the free enthalpy of the environment, at constant pressure and temperature, would be the principal motor of evolution: $\Delta G<0$ with $G=H-TS$, where H is the enthalpy, S the entropy and T the absolute temperature .The modifications which fulfill this condition would be fixed in the populations.

Hypotheses on the origin of life

The theory of the origin of life accepted by the majority of the scientific community is based on the role of RNA molecules, which can take multiple conformations and have some particular enzymatic activities. In vitro selection experiments have been performed to test models in the evolution of the RNA world (Eigen and Schuster 1977; Hirao and Ellington 1995; Schmidt 1999; Schuster 1993).

However these activities are restricted, up to now, to RNA as substrate, because of the limited number of chemical groups of its four molecules.

This darwinian evolution of molecules must have created, by random trials, the whole set of the enzymes of the living world, in a short interval of time: 100 million years, perhaps 25 (Oberbeck 1989) after the last impacts of giant meteorites, knowing that for a favorable property, numerous other solutions have been remained unsuccessful. Moreover, the intermediary steps, which are supposed to have occurred, would not have been functional and would not have been selected.

By another way, several authors (Dose 1983; Fox 1980; Matsuno 1983 and others), have hypothesized that proteinoids presenting different metabolic activities have been formed by thermic condensation of aminoacids during the prebiotic period.

We have proposed that the sequences of their active sites would have been determined not by chance, but by the presence of high chemical potential substances, according to the affinities between the aminoacids and the chemical groups of the substrates (Berger 2003), Figure 1. Specific substrate binding polymers have been synthesized on this principle, using chiral compounds (Akkelah 1981; Akelah and Sherrington 1981).

Earth’s surface receives heat from the inner core and light from Sun. A part of this energy is sent back to space, the rest is transformed into thermodynamic and chemical energies.

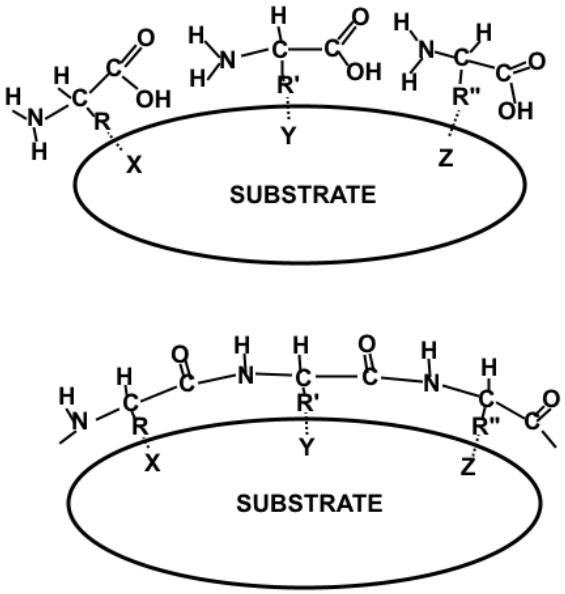

Figure 1 : Formation of the binding sites of a proteinoid.
The different aminoacids (side chains R, R', R") are held in position by specific interactions with different groups (X, Y, Z). They are then polymerized and form the specific binding site of the proteinoid.

During the abiotic period, the chemical energy was distributed between different compounds, among which some have high chemical potentials. All systems tend to evolve from high to low chemical potential. The transfer is favored by catalysts which lower the free energies of activation of the intermediary complexes (Eyring 1935).

As the proteinoids were permanently destroyed by water and diluted by divisions of the protocells, a mechanism of synthesis must have occurred, different from the molding on substrates, certainly too rare, and from direct replication, not possible. We have previously proposed that this mechanism could have been similar to the reverse of the present synthesis of the proteins (Berger 2003). It could have occurred as follows (Figure 2).

-unwinding of the proteinoid on the protoribosome,
-formation of a ternary complex between the C terminal aminoacid, the anticodon loop of the corresponding proto tRNA, and their specific protoaminoacyl tRNA synthetase,
-rupture of the peptidic bond and transfer of the aminoacyl group to the 3' extremity of the proto tRNA. This exchange would be simple if, as Hopfield (1978) hypothesized, the anticodon triplet was close to the 3' extremity in the primitive tRNA,
-same operation with the second aminoacid,
-two proto tRNAs would then be side by side with their anticodon loops in a similar position,

-the accessible parts of these two loops may be copied according to base pairing and the copies linked together by a proteinoid possessing a RNA synthetase activity. Why should only three nucleotides of each loop be copied? It corresponds perhaps to the number of nucleotides implied in the complex with the aminoacid. Sometimes, this number is reduced to two, the nature of the 5' extremity being of no importance (wooble) (Crick 1966). Perhaps steric hindrance or the necessity to have linear fragments may limit the length of the anticodon.

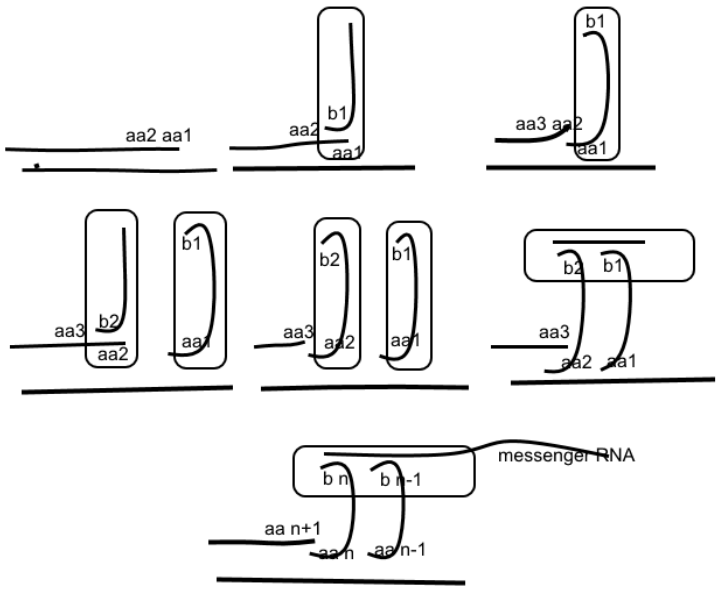

Figure 2 : Proposition of schema of the transfer of the information from proteinoids to nucleic acid (see text).

Enzymes catalyze the forwards and backwards reactions equally. They do not alter the equilibrium itself but the speed at which it is reached. The level of the equilibrium can be modified by the concentrations of the reactants. For example, carbonic anhydrase catalyzes the reaction in either direction, depending on the concentration of CO_2 (high in tissues and low in lungs). At the expense of the degradation of one molecule of proteinoid, obtained by molding on substrate, the information regarding its aminoacid sequence would be transmitted to a ribonucleic acid, then, through a proteinoid possessing a RNA dependent DNA synthetase activity, to a DNA molecule, able to share it between the two daughters of a dividing protocell.

We have shown that the specificity of the affinity of a nucleic codon for an aminoacid could be enhanced by a ternary association with an aminoacyl-t RNA synthetase (Berger 2003):

Let us consider the different aminoacids a_i, a_j ,…a_n and the corresponding tRNAs b_i, b_j, …b_n.

We have the equilibria : $a_i + b_i \leftrightarrow a_i b_i$ with $K_{ii}=\frac{(a_i b_i)}{(a_i)(b_i)}$

The selectivity being partly defective, we have also:

$$a_i + b_j \leftrightarrow a_i\, b_j \text{ with } K_{ij}=\frac{(a_i b_j)}{(a_i)(b_j)}$$

If we admit, for simplification that the concentrations of the different tRNAs are equal, and are large with respect to those of the aminoacids, it comes:

$$\frac{(a_i b_i)}{(a_i b_j)}=\frac{K_{ii}}{K_{ij}}$$

If the specificity of interaction is low, there is an important proportion of hybrid complexes: if K_{ii} =10 K_{ij} the percentage of error is 10%.

We suppose that an abiotic proteinoid synthesis could occur in the medium. According to the hypothesis exposed above, proteinoid adapted to the different present complexes may be formed: P_{ii} and P_{ij}, respectively specific to $a_i b_i$ and $a_i\, b_j$. We have the equilibria:

$$a_i b_i + P_{ii} \leftrightarrow a_i b_i\, P_{ii} \qquad \text{with } K=\frac{(a_i b_i P_{ii})}{(a_i b_i)(P_{ii})}$$

and

$$a_i b_j + P_{ij} \leftrightarrow a_i b_j\, P_{ij} \qquad \text{with } K'=\frac{(a_i b_j P_{ij})}{(a_i b_j)(P_{ij})}$$

It is logical to assume that the bindings between proteinoids and complexes are of the same type and that $K=K'$, so

$$\frac{(a_i b_i P_{ii})}{(a_i b_j P_{ij})}=\frac{(a_i b_i)}{(a_i b_j)}.\frac{P_{ii}}{P_{ij}}$$

Another logical hypothesis is that the synthesis rate of these proteinoids is proportional to the concentration of their substrates.

$$\frac{d(P_{ii}\,total)}{dt} = k(a_i b_i)$$

At the beginning of the synthesis $(a_i b_i)= (a_i b_i)_0$, initial concentration of the complex $a_i b_i$ and $P_{ii}\ total=k(a_i b_i)_0 t$.

If only a small fraction of P_{ii} is bound to the complex, $a_i b_i\, P_{ii} \approx P_{ii}\ total,$ and it comes: $\frac{(a_i b_i P_{ii})}{(a_i b_j P_{ij})}=\frac{(a_i b_i)_0}{(a_i b_j)_0}.\frac{k(a_i b_i)_0 t}{k(a_i b_j)_0 t}=\left(\frac{(a_i b_i)_0}{(a_i b_j)_0}\right)^2=\left(\frac{K_{ii}}{K_{ij}}\right)^2$

It can be seen that, under these conditions, the selectivity of association between an aminoacid a_i and a tRNA b_i is improved by the association of a proteinoid P_{ii} (proto aminoacyltRNA synthetase) to the complex $a_i b_i$: if K_{ii} =10 K_{ij}, the percentage of errror is fallen to 1%.

Remark: Some of the simplifications of calculation used above are justified, others do not correspond to the actual facts, but have been assumed to compare the specifities in the extreme cases.

Furthermore, enzymes can couple two or more reactions, so a thermodynamically favorable reaction can be used to drive a thermodynamically

unfavorable one. The hydrolysis of ATP is often used to drive other chemical reactions.

Later on, the permanent presence of nucleic acids in the cells would have favored the present mechanism: from a single molecule of DNA, several messenger RNAs would be transcribed, each one giving rise to multiple copies of proteins.

The initial environmental conditions (pH, ionic strength, substrates) would have determined the conformation of the aminoacids as well as their associations with the nucleic acids. Would they have been different, the genetic code would not have been the one we know.

The genes of the enzymes would have been distributed randomly among the different protocells and those involved in complementary reactions would have get together.

Life would have fully appeared when the associations of genes in a same cell have enabled its autonomy and its reproduction using the energy of the high chemical potential substrates from the environment. The different functions such as digestive, muscular, nervous, already exist in the primitive monocellular organisms : Paramecium has muscular fibrils, some protozoa are sensitive to light, others are able to transform light into chemical energy. It is admitted that mitochondria and chloroplasts have resulted from the engulfment of bacteria and algae by primitive organisms. In each monocellular organism, apart from the house-keeping genes which provide the enzymes for the cell vital reactions, there are numerous genes which are expressed by the transcription factors, when they are useful to interact with the environment.

We assume that later on, during evolution, no new activity has been created, only modifications and repartitions of the genes between organisms would have occurred, nevertheless giving rise to a large biodiversity.

Association into colonies

The next step of the evolution seems to have been the formation of homogenous colonies. The repetitive division of the same cell and of its daughters at the same place can give rise to a colony. Each cell may have received metabolites from a neighbouring cell of the same type, among which other genes were expressed, which would increase the metabolism of the whole. The biological mass and the number of cells would have increased, at the expense of the materials and the environment free enthalpy. The reunion of monocellular organisms into colonies would be then favored during evolution.

There are some to day's traces of this primitive stage :

-Bacteria or yeast form homogenous colonies on agar plates, but these associations are loose and are destroyed by washing. Other organisms, such as Volvox, aggregate in more stable colonies, able to carry out sexual and asexual reproduction.The cellular adhesion is performed by thin strands of cytoplasm (protoplasmates), or by an extracellular matrix made of gelatinous glycoproteins.

-The first divisions of the fertilizided eggs of sea urchin give rise to a morula, a spherical aggregate of almost identical cells. Although the Haeckel's law (ontogeny

recapitulates phylogeny) is in part discredited, the morula could be compared to an ancient cell colony.

Different kinds of colonies of complementary metabolisms could have merged, each one taking advantage of the expressed genes by the other one. Sponges and Cnidarians are constituted of a small number of types of cells. Each type corresponds to the expression of a lot of genes, carrying out a particular function; these cells tend to regroup in rough shapes of organs. In Physalia physalis, the division of work is extended to high level: there are cells that function as floaters, others work as fishing filaments (dactylozooids), others which ensure the motion, the digestion (gastrozooids) or the reproduction (gonozooids).

All the cells of a primitive colony would have had the same total genome, the different genes being expressed in turn by the transcription factors, in the specialized cells, when they are needed. In addition, we hypothesize that a short sequence of nucleotides, specific of the type of the colony would be added at the beginning of the genome, when a new cell joins the colony. When several colonies of different types merge, in order to complete their metabolisms, we hypothesize that their genomes link to each other, with their specific short sequences placed at the beginning of each ones, and that their expression is triggered by proteins, that we call connection factors, only present in certain cells, and which give rise to new types of tissues. (Figure 3).

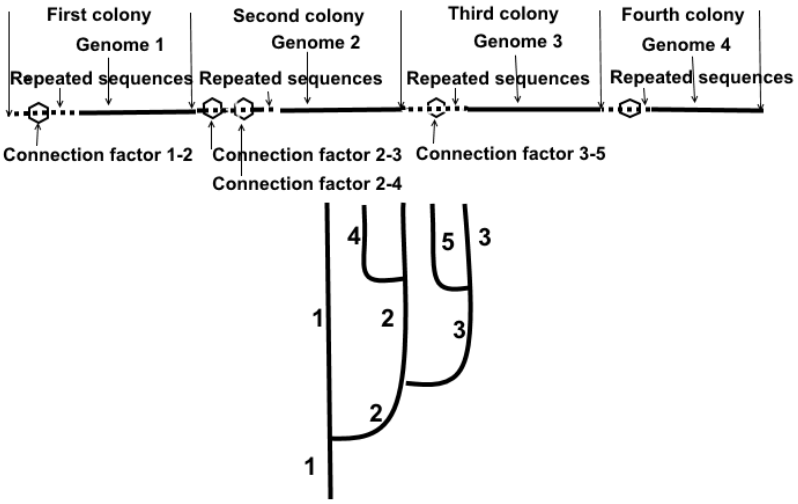

Figure 3 : Hypothetical genome of multicellular organisms.

The genomes of the different colonies would be linked together, in the order of their arrival. In front of each one, respective repeated sequences would determine the number maximum of cells and would be expressed one after the other, in the 5'-3' direction. When a connection factor, present in specific cells, is reached, the next type of cells would be expressed and would give rise to a new lineage. The cells of the precedent type would continue dividing until the end of their repeated sequences.

Present nucleic repeated sequences, the role of which are more or less unknown (Schlmenti and Duncan 1984; Shapiro and Sternberg 2005), are perhaps the remnants of these associations. Microsatellites are situated near or in the centromeres and are constituted of 2 to 10 nucleotides, repeated 10 to100 times, varying with the species. There are also minisatellites of 15 to 20 nucleotides, repeated up to 150 times. The telomeres are short repetitive nucleotide sequences (300 bps to many kilo bps) at each end of a chromatid. Their presence protects the genes located in front of them on the chromosome from being truncated instead. Over time, due to each cell division, the telomere ends become shorter (Olovnikov 1973). When they are too short, the cell detects this uncapping and begins apoptose. The consequence is that the somatic cells of eukaryotes can divide only a limited number of times (Hayflick 1965).

Repeated sequences and morphology of the organisms

The shapes and the positions of the organs are the result of different mechanisms and are determined by the genome. During ontogeny, each short repeated sequence would be expressed in turn, in the 5'-3' direction, which would give rise to the duplication of the corresponding cells, then to the expression of their genes, according to the transcription factors.

This hypothesis is to be compared with the fact that, in a normal differentated cell culture, the number of divisions cannot exceed a limit (50 for the human fibroblasts), related to the length of the telomeres of the chromosomes, which are shortened at each duplication. It has been already proposed that the telomeres could have a role in aging (Aubert and Lansdorp 2008).

In the same way, the fertilized eggs of sea urchins, ascidians and amphibians, clived in certain plans of symmetry, give rise to normal, but smaller adults.

In the case of sexual reproduction, each allele is specific of one parent and the morphologies of the descendants are intermediate. If, during meiosis, an interruption of the repeated sequences of the genome of a kind of cells occurs at the n'^{th} one of the father allele and at the n''^{th} one of the mother allele, the recombinant fragments would have $n+n'-n''$ and $n-n'+n''$ repeated sequences, giving rise to $2^{(n+n'-n'')-1}$ or $2^{(n-n'+n'')-1}$ cells. The children have intermediate morphologies but may look more like one of the parents. How could be explained otherwise the forms of the organs and particularly the specific forms of the fingerprints or of the specificity of a human facial traits?

However, contact inhibition mechanism may stop the division of normal cells before the last repeated sequence has been reached.

The expression of the genomes of the different colonies would be carried out by the connection factors, in the lineages descended from cells where these factors are present.

Obviously, the modifications of the numbers of repeated sequences and of connection factors would be hereditary only when they occur in the genomes of the germinal cells. They would be fixed in the populations when the modifications of the morphology are beneficial to the organisms (beak or wings shape in birds, number of

descendants, etc), or when there is a genetic drift, due to a bottleneck in the populations. If they give rise to a decrease of the enthalpy of the genome, we hypothesize below that they are also favored. It must be noted that the possible role of the repeated sequences in evolution has been already considered (Dover 1982).

Punctual mutations and biological clock

Punctual mutations are attributed to the DNA polymerase errors during duplication or to ionizing radiation or mutagenic chemicals. When they occur in the coding sequences they may give rise to modifications of the sequence of the proteins.

The comparison of the sequences of the homologous proteins clearly shows that the number of differences depends only on the nature of the proteins and of the time spent since the divergence of the species from a common ancestor: proteins from human, snake, carp or worm are equally distant from yeast corresponding proteins, at the 10% level. The rate at which appear the differences is practically constant with time and is as low as the proteins are submitted to strong structural or functional constraints: it is the molecular clock.

Our model is compatible with this theory. The creation of a new species would be due to the association of a new colony to an existing organism and the addition of their genomes. The gene of a protein of the first species would be also present in the total genome of the second species and would receive quite the same mutations, since submitted to the same constraints.

When a mutation gives rise to a non functional protein, it is removed, because the organism dies or does not reproduce. It is generally agreed that favorable punctual mutations are extremely rare, but when they occur in the germinal cells, they are fixed in the populations. According to Kimura (1986), the observed punctual mutations would be neutral mutations, fixed by genetic drift. However, the neutral character of a mutation is not a sufficient condition to be fixed, since all the synonymous substitutions are not observed.

The punctual mutations could have an effect on association constants of complexes or on the maximum rates of enzymes but, to our knowledge, they almost never lead to new activities, which would have been created only during the prebiotic period.

In the theory which is proposed below, in order to be fixed, the mutations must correspond to a decrease of the free enthalpy of the genome, and not to be deleterious.

Evolution of the genome

Large range of DNA mutation rate values have been reported, depending on species and genes. They have been measured by the number of mutations in the whole genome between an organism and its descendant or indirectly by evolutionary scenarios and expressed per generation or per year.

As the homologous proteins diverge regularly with time, regardless of the generation time of the species, or of their conditions of life (temperature), but according to the constraints of structure and function of the protein, it could be supposed that the DNA mutation rate is not the only determining factor.

In order to account for this discrepancy, we propose the hypotheses that the genome of the eukaryotes, in a closed system, separated from the remainder of the cell, would tend towards a state of minimum free enthalpy, by the formation of nucleosomes, then of chromosomes, and that the mutations allowing this formation would be fixed in the population. Histone-DNA interaction free energy in nucleosomes is well defined. It has been measured by dialysis-based approach (Thastrom et al 1999, Thastrom et al 2004).

The packaging of the genome would also facilitate the expression and the replication of the genome by avoiding the entangling of the chains and would be also favored for this reason.

The eukaryote nucleosome core consists of 147 base pairs of DNA wrapped in 1.67 left-handed superhelical turns around an histone octamer consisting of two copies each of four different histones. These complexes are stabilized by hydrogen bonds between the OH and NH_2 groups of the histones and the DNA phosphates. The DNA of the prokaryotes associates also with basic proteins, analogous to histones.

The rate of formation of complexes, according to the transition state theory of Eyring (1935), is :

$$v = \frac{\gamma kT}{h} e^{-\frac{\Delta G_a}{RT}},$$

where ΔG_a is the free enthalpy of activation of a certain number of base pairs, in order to ensure the winding in left-handed helix, necessary to the formation of the complex.

γ, the probability that the transition state actually leads to the complex and not returns to the initial separated compounds.

k, the Boltzmann constant.

h, the Planck constant.

R, the gas constant.

T, the absolute temperature.

In order that a new complex could be formed (γ =1), it is necessary that its stability would be higher than that of the separated compounds (ΔG<0), otherwise the reverse reaction of dissociation would occur (Figure 4).

The term ΔG_a depends mainly on the nucleotidic sequences, which vary from a gene, or from a part of a gene, to the other one. It would be the same for two homologous genes, since these sequences have the same origin, accounting for the equality of the mutation rates of homologous proteins.

The term ΔG, decrease of the free enthalpy of the genome, corresponds to the formation of a new nucleosome, leading to a new local structure.

The rate of formation of the complexes would vary with the absolute temperature, but only between 273 and 313°K for the majority of the organisms.

It is conceivable that the rate of punctual mutations of DNA, which are fixed in the population and observed in the evolution of the proteins, would be far lower than when initially produced by radiations or by polymerase mistakes, since those which do not lead to the formation of nucleosomes would not be favored.

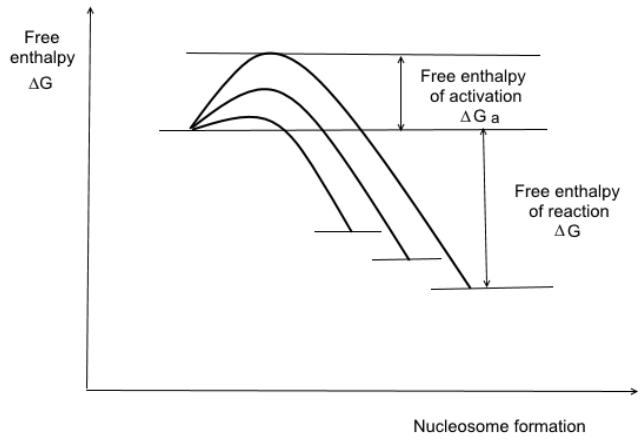

Figure 4: Free enthalpy of genomes.

In order to pass from an energy level to a lower energy one, the genome must acquire a free enthalpy of activation. Once this step is passed, several levels of lower energies are allowed.

The fixation of the mutations depends also on the functional and structural constraints: efficiency of the protein, when the mutation occurs in a coding region, or at a place where the nucleosome acts on its expression.

When a gene is duplicated, the mutation rate of the duplicate is identical to that of its model for a time, then it differs progressively. This is consistent with the present theory: the nucleotide sequences being the same in the two genes, the rates of formation of the new complexes are identical. But the proteins expressed by the new gene would not be submitted to the same constraints and would differ later on from the former ones.

Phylogeny and ontogeny

Do these hypotheses account for phylogeny and ontogeny of living organisms?

Phylogeny

We have hypothesized that during the abiotic period, a lot of enzymatic sites were induced in proteinoids, by the presence of substrates, and that this information was transferred to nucleic acids through an unique genetic code by a mechanism reverse to the protein synthesis. During the primitive biological period, these genes would be spread among a multitude of monocellular organisms, able, owing to their house-keeping genes, to use the energy of the high chemical potential substrates from their environment, to reproduce and to carry out particular reactions.

We have also assumed that the next step in biological evolution has been the aggregation of monocellular organisms of the same species, after several duplications. The genome of all the cells from a colony would be identical, it would begin by a number of small repetitive sequences which would determine the number of the cells in the colony.

Two colonies of complementary functions would have had tendancy to enter into partnership. We hypothesize that their genomes would have linked together by the means of the connection factors, with the repetitive sequences in front of each genome, in the lineages descended from the cells in which these connection factors were present. Their complementary metabolisms would have involved an increase of the biological mass and a more rapid decrease of the environment free enthalpy.

These reunions of different colonies would have occurred in the early part of evolution. Afterwards, less significant modifications would have arisen, as the following ones:

- punctual mutations, favored by the decrease of the free enthalpy of the genome and leading to modification of the expression of genes by creation or displacement of nucleosomes.

-other punctual mutations, leading to synonymous mutations, without action on phenotype, could also be fixed by simple genetic drift (Kimura 1986).

- under the pressure of the environment and the needs of the organisms (feeding, breathing, swimming, walking, etc), the number of the repeated sequences, the numbers and places of the connection factors in specific somatic cells may have been modified. If, at least in the early times, the Weissmann barrier between soma and germen was not absolute, a transfer of these modifications might have been made possible.

-the attachment of a new colony to an existing organism generates an increase of its complexity. This is partly in agreement with Lamarck's theory, according which evolution of organisms is driven from simplicity to complexity.

- each reunion has been certainly a rare event, and a long space of time may have separated two of them. When it would occur, it would allow rapidly multiple new possibilities of forms to develop, which would account for the punctuated equilibria theory of Gould (1977) (Figure 4).

-Large chromosomal rearrangements can also occur, such as deletions, duplications, inversions and translocations, caused by breakage at two different locations and the joining of the broken ends, giving rise sometimes to new species.

-Convergence: there are numerous examples of similarity of shape between organisms without link or relationship: the wing of the birds and the bat's wing membrane, the dorsal fin of shark, cetaceans and ichthyosaur, the parallel evolution

of marsupials and placentals, etc. These convergences can be explained, according to our model, by the common presence, in these organisms, of the genomes of cells from the primitive colonies which had the property to develop the same role. Under the same conditions and constant pressions, the number of the repeated sequences of particular somatic cells could be modified in order to express the needed characteritics. In the rare cases where the modifications would occur in the germinative cells, they could be hereditary.

-Mimicry: surprising mimicries exist between very distantly related organisms, such as insects and plants. As common primitive cell colonies are unlikely, a possible explanation could be that the mimics have acquired a group of genes from their models, allowing them to reproduce some of their characteristics (color, shape, odor, behavior). Viruses common to the two species could have been the vectors of this horizontal gene transfer. Caterpillars could have been infected by high concentrations of viruses and could have received from them genes of the plants they ate, infected by the same viruses. This phenomenon has been shown particularly in prokariotes and unicellular eukaryotes, and there is some evidence that terminal protein-containing genomes could be a vehicle of inter-kingdom genetic information carrier all throughout evolution (Redrejo-Rodrigez et al 2012). The result is obvious when the modification brings an advantage to the mimic and is fixed in the population, for instance protection or better reproduction. It is possible that numerous features of organisms are due to this mechanism, with no visible effect but with real profit, or are neutral and have been fixed by random drift.

Ontogeny

The development of the fertilized egg always begins by several divisions, leading to identical totipotent cells. According to our theory, these undifferentiated cells would correspond to those of the first primitive colony. The genes are expressed under the action of the transcription factors and each of them is read in the 5'-3' direction, and in the order depending on that of the junctions of the colonies.

When the n^{th} repeated sequence of a particular type of cells is expressed, an aggregated mass of 2^{n-1} cells would be formed. When the connection factor of the second colony is reached, the cell containing it gives rise to a lineage of $2^{n'-1}$ cells of the second type (Figure 5), while the cells of the first type continue dividing, up to the expression of the n^{th} last repeated sequence. Two cell lineages would be developed according to the number of their repeated sequences. There are bindings between cells of the same type and between cells of different types by cell adhesion molecules (CAM). The shapes of the organs would then be determined by the different *n* values, by the numbers and places of the connection factors and by the cell adhesion molecules. The linear information of the DNA genome would be transformed into three-dimensional information. There could be also mutual influence between cells of different types, by diffusion of compounds capable of activating new genes (induction).

It is conceivable that during its development, an embryo could pass through the same stages as those of the embryos of their ancestors: according to the Haeckel's law, "ontogeny recapitulates phylogeny". However this theory has been in part discredited: there are some resemblances but not identities. This law could be valid

only if the programs of development of the organisms would have not varied since their apparition, which is wrong in the majority of cases (mutations, heterochrony, differences in environment).

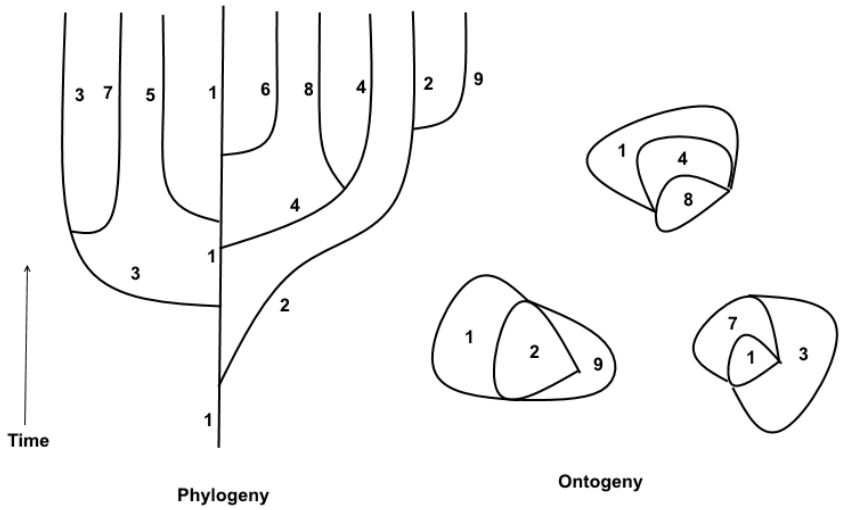

Figure 5: Hypothetical phylogeny and ontogeny of multicellular organisms. The development of the multicellular organisms would be determined by the order in which the colonies have been aggregated.

Conclusion : the different motors of evolution

Different theories have been developped in order to account for the biological evolution. Each of them contains certainly a part of truth, depending on the considered organisms and phenomena. Their relative credibility and importance is discussed below, with regard to the assumptions we have made.

The natural selection, Darwinism and Neodarwinism

Natural selection has a direction guiding evolution towards heritable adaptation to the current environment. It favors the spread of alleles whose effects on phenotype increase survival and reproduction of their carriers, and conversely decrease the frequency of alleles that cause unfavorable effects and ignores those that are neutral.

This theory, exposed by Darwin, in his book “On the origin of species by means of natural selection” and simultaneously by Wallace, then developed later on in neodarwinism or modern synthesis of darwinian evolution through natural selection

with mendelian genetics, by Mayr, Dobzanski, Huxley, Haldane, Simpson, Morgan, Wilson, Dawkins and others, is until now the mainstream theory of evolution.

The logic of their demonstrations makes credible their theory of the microevolution (the variation of the beaks of the finches of Darwin, as an example), and explains the consensus in its favor. However, the small variations of the morphology of a species, due to the random mutations, must accumulate in the same direction or for the same purpose, in order to be beneficial and fixed in the population. Some dissidents, in particular Grassé (1973) and Denton (1988), have not accepted it for sure, since the effect of the mutations is only revealed at the end, and the intermediary forms of fossils are lacking.

Lamarckism and Neolamarckism

According to the Lamarck's theory, an organism could pass on to its offspring characteristics that it acquires during its lifetime. This idea has been also developped by Erasme Darwin, Wallace, Geoffroy Saint Hilaire, Cabanis, Cope and others.

Experiments about this theory are more or less open to criticism, very little are convincing, due to difficulties of interpretation. However, some environmentally induced features seem to be transmitted to descendants by epigenetic means (Mendizabal et al 2014). It has been shown that the mechanism underlying epigenic inheritance can also lead to saltational changes that reorganize the epigenome (Jablonka and Lamb 2008).

Insofar this phenomenon would be real and important for evolution, it could be explained by our model of development of the organisms. When the environment needs the expression of a gene in particular cells, and when this gene is present in the genome of a particular colony, the number of the repeated sequences and of the connection factors may be modified in the convenient somatic cells. However the corresponding development in the germinative cells is needed. The Weissmann barrier may have not been absolute, at least in the early times.

The genetic random drift

The regularity of the molecular clock has been interpreted by Kimura (1986) as the result of the fixation of the neutral mutations in the population. These mutations would be fixed after a number of generations equal to four times the number of the effective population size, and the time between two successive mutations would be the inverse of the mutation rate.

Genetic random drift has no direction and is guided only by the mathematics of change. It acts upon the genotyping frequencies within a population, without regard to their phenotypic effects.

The hypothesis that the decline in genetic variation in small populations following the founder effect would be important for new species to develop has been tested repeatedly through experimental research, but without clear results.

In large populations, the importance of genetic drift seems to be negligible and unable to induce the formation of new species.

The decrease of the free enthalpy (Gibbs free energy)

Numerous authors have expressed the idea that the thermodynamic laws govern the biological world, as they do for the chemical environment: Clausius,

Schrödinger, Boltzmann, Lehninger, Oparin, Prigogine, Lotka, Krebs, Avery and others.

As a consequence of the second law of thermodynamics, the metabolism and the reproduction of the living organisms participate to the decrease of the free enthalpy of the environment.

We have hypothesized that, at the beginning, the proteinoids would have allowed this evolution, by the role of catalysts of their active sites, by lowering the activation enthalpy of the intermediary complexes of the biochemical reactions. By these means, the energy of high chemical potential compounds formed by the Sun radiations could be more rapidly released.

By an inverse protein synthesis mechanism, the informations contained in the proteinoids would have been transferred to the ribonucleic and deoxyribonucleic acids, allowing an easy and abundant reproduction of the enzymes.

This abiotic evolution would have been favored, since it gave rise to reactions in agreement with the decrease of the environment free enthalpy.

During the biologic period, the reunions of complementary genes in the same cell, or the aggregation of cells into colonies and the fusion of colonies of complementary metabolisms would have worked towards the same purpose. The repeated sequences would have been a consequence of the reunion into colonies and would have determined the shapes of the organisms and of their organs.

Inside the nucleus, the chromosomes would have evolved towards a state of minimum free enthalpy. In order to decrease their energy level, a higher energy than the enthalpy of activation required for their production must be supplied to the nucleosomes. This energy would be obtained by one (or several) point mutation(s), giving rise to a new sequence, which would lead finally to a decrease of the energy of the genome and which would then fix it in the population.

However the second law of thermodynamics is defined for thermodynamic equilibrium systems and the living organisms operate far from this state. This is also the case for much of the physical processes and nevertheless, this law has been useful for the understanding of the phenomena.

The law of decrease of free enthalpy have determined the direction of the reactions giving rise to proteinoids with different activities, to nucleic acids, to the genetic code and to different enzymes spread in mono and pluricellular organisms, which have increased in mass and number.

In its turn, the increase of the living mass have led to a more rapid decrease of the free enthalpy of high chemical potential compounds from the environment, so that the two phenomena are tightly linked. A virtuous circle has been produced, or rather a vicious circle, on account of our increasing consumption of the energy of a planet with limited resources.

Acknowledgements

I gratefully acknowledge Dr Christian Prenant for his helpful suggestions and reading of the manuscript.

References

Akelah A.(1981). Heterogeneous organic synthesis using functionalised polymers. Synthesis. 6: 413-438.

Akelah A. and D. Sherrington. (1981). Application of functionalized polymers in organic synthesis. Chem. Rev. 81: 557-587.

Aubert G. and P. Lansdorp. (2008). Telomeres and aging. Physiological Reviews. 88: 557-579.

Crick F. (1966). Codon-anticodon pairing: the wooble hypothesis. J.Mol.Biol. 19: 548-555.

Berger G. (2003). Deterministic Hypotheses on the origin of life and of its reproduction. Med. Hypotheses. 61: 586-592.

Denton M. (1988). in Evolution, une théorie en crise. Londreys, Paris.

Dose K. (1983). Prebiotic polypeptides and the origin of biological information. Naturwissenschaften. 70: 378-382.

Dover G. (1982). Molecular drive: A cohesive mode of species evolution. Nature. 299: 111-117.

Eigen M. and Schuster P.(1977)) The hypercycle. A principle of natural self organization. Part; Emergence of the hypercycle. Naturwissenschaften. 64:541-565.

Eyring H. (1935). The activated complex in chemical reactions J. Chem. Phys. 3: 107-115.

Fox S. (1980). Metabolic microspheres. Origins and evolution. Naturwissenschaften. 67: 378-383.

Gould S.and N. Eldredge. (1977). Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology. 3:115-151.

Grassé P. (1973). in L'évolution du vivant. Albin Michel ed. Paris.

Hayflick L. (1965). The limited in vitro lifetime of human diploid cell strains. Exp Cell Res.37: 614-636.

Hirao I. and Ellington A.(1995) Recreating the RNA world. Curr. Biol.5: 1017-1022.

Hopfield J. (1978). Origin of the genetic code: a testable hypothesis based on tRNA structure, sequence and kinetic proofreading. Proc. Natl. Acad. Sci. USA.75: 4334-4338.

Jablonka E.and M. Lamb. (2008). Soft inheritance: Chalenging the modern synthesis. Genetics and Molecular Biology.31:389-395.

Kimura M. (1986). DNA and the neutral theory. Philos Trans R Soc London (Biol). 312: 343-354.

Matsuno K. (1983). Evolutionary changes in the information content of polypeptides. J Theor Biol.105: 185-199.

Mendizabal I., T. Keller, J. Zeng and S. Yi. (2014). Epigenetics and Evolution. Integr Comp Biol. 54: 31-42.

Oberbeck V. and Fogelman G.(1989). Estimates of the maximum time required to originate life. Origins Life.19: 549-560.

Olovnikov A. (1973). A theory of marginotomy. The incomplete copying of template margin in enzymatic synthesis of polynucleotides and biological significance of the phenomenon. J Theor Biol. 41: 181-190.

Redrejo-Rodriguez M., D. Munoz-Espin, I. Holguera, M. Mencia and M. Salas. (2012). Functional eukaryotic nuclear localization signals are widespread in terminal proteins of bacteriophages. Proc Natl Acad Sci USA.109: 18482-18487.

Schlmenti J. and C. Duncan. (1984). Ruminant globin gene structures suggest an evolutionary role for Alu-type repeats. Nucleic acids res. 12: 1641-1655.

Schmidt F.(1999) Ribozymes - why so many, why so few? Mol. Cells. 9: 459-463.

Schuster P.(1993). RNA based evolutionary optimization. Origins Life Evol. Biosphere. 23: 373-391.

Shapiro J. and R. von Sternberg. (2005). Why repetitive DNA is essential to genome function. Biol Rev Camb Philos Soc. 80:227-250.

Thaström A., Lowary P., Widlund H., Cao H. Kubista M. and Widom J.(1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. J. Mol.Biol. 288(2): 213-229.

Thaström A., Lowary P., and Widom M.(2004). Measurementof histone-DNA interaction free energy in nucleosomes.Methods. 33:33-44.
