J Mol Evol (1982) 19:47–61

Journal of Molecular Evolution

Stages of Emerging Life – Five Principles of Early Organization*

M. Eigen[1] and P. Schuster[2]

1 Max Planck Institut für Biophysikalische Chemie D–34 Göttingen, Federal Republic of Germany
2 Institut für Theoretische Chemie und Strahlenchemie der Universität Wien, A 1090 Wien, Austria

Summary. Five principles underly the evolution of the genetic language: formation of stereoregular heteropolymers, selection through self-replication, evolution of quasispecies towards optimal structures, regulated cooperation between competitors through catalytic hypercycles and evaluation of translation products through compartmentalization. These principles are formulated and illustrated by means of experimental results

Key words: Prebiotic evolution – Polynucleotide replication – Quasispecies – Hypercycles – Compartmentalization

Introduction

New fundamental concepts emerge in the transition from non-living to living matter. Biological information and the machinery to process it originate in this early period of evolution. Ensembles of macromolecules organize themselves and develop the genetic language. In the recent past experiments have been made which examine the forces that have driven this early development and that are still in operation at present under modified and more subtle conditions. The results of these experiments put contraints upon theories of the origin of life and bring them down from the realm of pure speculation to the more solid grounds of experimental tests.

The enormous variety of studies in prebiotic chemistry is reviewed in other contributions to this journal, particularly in the special issues dedicated to the memory of H.C. Urey (Oró and Miller 1981, 1982). Through these systematic investigations the identity of the most abundant molecules on the very early Earth became clear. Most of the natural amino acids were among them. The building blocks of polynucleotides – the four bases, ribose and phosphate – were available too under prebiotic conditions. An enormous amount of low molecular weight organic material and energy-rich compounds was synthesized in those early days. Material was provided from steadily refilling pools for the formation of polymers, among them polypeptides and polynucleotides. The appearance of these molecules brought into operation new principles of organization.

1. Formation of Heteropolymers Introduces an Otherwise Unknown Richness into the Structures and Properties of Molecules

The two most important classes of biopolymers, the polypeptides and the polynucleotides, share a common building principle: a regular, repetitive backbone with variable side chains. The structural concept of such heteropolymers allow a "hyper-astronomically" large number of different molecules to be made even for moderate polymer lengths. (This number is 20^{ν} for natural polypeptides and 4^{ν} for polynucleotides, when ν is the number of segments. For a polynucleotide with $\nu = 300$, a fragment which is just large enough to encode for one of the smallest proteins, 4.10^{180} different sequences are possible, while there are e.g. "only" 10^{11} stars in our galaxy and the total observed matter in the universe corresponds to something like 10^{80} hydrogen masses. In ordinary chemistry we deal with much smaller numbers of different molecules: "Beilstein" lists about 10^7 organic molecules and Avogadro's number is only 6.10^{23}.

In order to visualize this special aspect of biopolymers thermodynamically we assume a set of polymers with the same chain length ν and the same average com-

* Dedicated to the memory of Harold C. Urey (1893–1981)

Offprint requests to: P. Schuster

0022–2844/82/0019/0047/$03.00

position. The stabilities of the sequences will differ little, and to this degree only because they have different secondary and tertiary structures. We are thus dealing with a huge number of almost equally stable hypothetical molecules. Every sequence corresponds to a local potential energy minimum, or to a set of minima for different conformations. Minima belonging to different sequences are separated by rather high energy barriers since at least two chemical bonds have to be broken in order to convert one particular sequence into another. At room temperature no interconversions between sequences occur. The material available in a single experiment (or on the entire Earth or even in the universe) would not allow one to synthesize even a single molecule of every possible sequence. Thus, only a tiny fraction of what is theoretically possible is really accessible.

Experimental evidence for polymerization under prebiotic conditions has been provided by several authors. Sidney Fox and coworkers were able to produce polymers by cooking aqueous solutions of amino acids and heating to dryness. These so-called "protenoids" have a limited repertoire of primitive cytalytic properties (Fox 1973). Leslie Orgel and his group have presented a variety of experimental results that show how spontaneous formation of polynucleotides might have occurred under prebiotic conditions. These experiments showed in particular that some divalent metal ions, in particular Zn^{2+}, have strong catalytic effects. Polycondensation of activated ribonucleosides in presence of Zn^{2+} or Pb^{2+} ions gives a much higher yield of regular linkages (3′–5′) than can be obtained in the absence of these ions. In a recent series of experiments they were able to study the influence of a polymer template on polymer synthesis (Lohrmann et al. 1980). They polymerized activated guanylic acid in the presence of a poly (C) template. Again Zn^{2+} ions play an important role. In their presence oligo (G) and poly (G) with chain lengths up to $\nu = 50$ and high predominance of 3′–5′ linkages (85%) wee obtained. These studies on template-induced polymerization of activated nucleotides are a strong indication that polynucleotide replication occurred even in enzymefree systems.

2. Autocatalysis or Self-Replication Under Conditions Far from Equilibrium Introduces Selection into Molecular Ensembles

Polynucleotides are the only known biomolecules which can act as templates for the replication. They can be regarded as "constitutional" autocatalysts since this unique property is a result of their molecular structure. The complementarity of base pairs allows the synthesis of a unique negative copy of each polynucleotide. Replication of this negative then yields a second copy of the original strand. A plus-minus ensemble thus behaves very much like a simple autocatalytic system. The kinetics

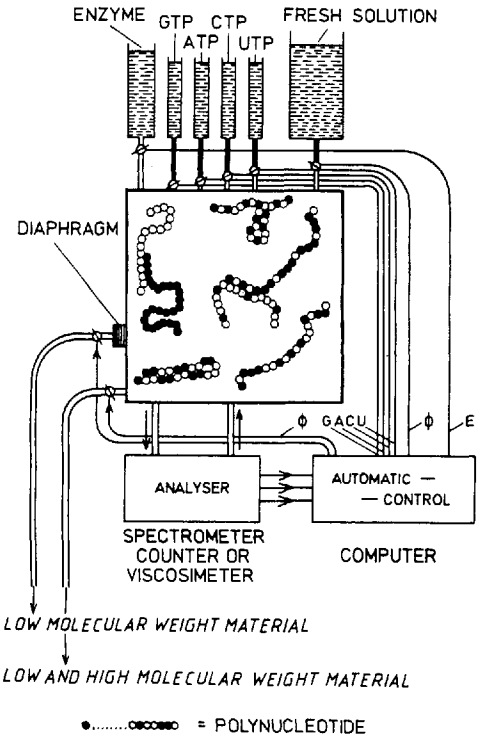


Fig. 1. The evolution reactor. This kind of flow reactor consists of a reaction vessel which allows for temperature and pressure control. Its walls are impermeable to polynucleotides. Energy rich material is poured from the environment into the reactor. The degradation products are removed steadily. Material transport is adjusted in such a way that the concentration of monomers is constant in the reactor. A dilution flux ϕ is installed in order to remove the excess of polynucleotides produced by replication. Thus the sum of the concentrations

$$[I_1] + [I_2] + \ldots + [I_n] = \sum_i c_i = c$$

may be controlled by the flux ϕ. Under "constant organization" ϕ is adjusted such that $c = c_o$ is constant. The regulation of ϕ requires internal control, which may by achieved by analysis of the solution and data processing by a computer as indicated above

of plus-minus replication has been analysed in detail (Eigen 1971; Eigen and Schuster 1979, p. 13). A straightforward result was obtained: after an initial phase of internal equilibration, the plus-minus ensemble grows as a single autocatalytic entity.

Competition between different self-replicating units, i.e. different sequences of RNA or DNA can be studied most easily under idealized conditions that could be realized in an "evolution reactor" (Fig. 1). An evolution reactor is an open system with adjustable flows of low molecular weight material. The composition of the solution in the reactor can be controlled by means of a general dilution flux ϕ. Suppose n different replicating molecules, $I_1, I_2, \ldots, I_n$ are present in the reactor at concentrations $[I_1] = c_1$, $[I_2] = c_2$, ..., $[I_n] = c_n$. The total concentration is then $c = \Sigma c_i$. The dynamics of processes in the reactor are described by the differential equations

$$\frac{dc_1}{dt} = \dot{c}_1 = (f_1 - d_1 - \frac{1}{c}\phi)c_1 = (E_1 - \frac{1}{c}\phi)c_1$$

$$\frac{dc_2}{dt} = \dot{c}_2 = (f_2 - d_2 - \frac{1}{c}\phi)c_2 = (E_2 - \frac{1}{c}\phi)c_2 \qquad (1)$$

$$\vdots$$

$$\frac{dc_n}{dt} = \dot{c}_n = (f_n - d_n - \frac{1}{c}\phi)c_n = (E_n - \frac{1}{c}\phi)c_n$$

Here f_i denotes the number of molecules I_i synthesized per time unit at unit concentration ($c_i = 1$), d_i similarly denotes the rate of decomposition, usually by hydrolysis. The quantity $E_i = f_i - d_i$ is called the excess production. The rates f_i, d_i and E_i will be constant when the reactor is run under constant internal conditions, i.e. constant temperature, constant pressure and constant concentrations of low molecular weight materials.

Differential equation (1) is non-autonomous since the total concentration $c(t)$ is a function of time that depends on the flux $\phi(t)$ (Eigen and Schuster 1979, p. 31, 32 and 35):

$$c(t) = c(0) + \int_0^t \{\sum_i E_i c_i - \phi(\tau)\}\, d\tau \qquad (2)$$

With a simple transformation we can eliminate the time dependent total concentration $c(t)$: we introduce relative concentrations $x_i = c_i/c$, ($\Sigma x_i = 1$) and find

$$\dot{x}_i = (E_i - \bar{E})x_i;\ i = 1, 2, ..., n \qquad (3)$$

$\bar{E}$ is the mean excess production $\Sigma E_i x_i$. The solution of equation (3) can be visualized easily (Fig. 2). The relative concentrations of those polynucleotides which fulfil $E_i > \bar{E}$ at a certain instant $t = t_1$ will increase whereas those with $E_i < \bar{E}$ will decrease. $\bar{E}$ increases as a consequence of these concentration changes. At $t > t_1 + \Delta t$ more molecules fall below the critical limit $\bar{E}$. Since the mean excess production $\bar{E}(t)$ is a monotonously increasing function, more and more molecules are taken out of the competition. Ultimately only the polynucleotide with the largest excess production E_m remains. Then $\bar{E}$ also reaches its maximum value $\lim_{t\to\infty} \bar{E} = E_m$, and "selection of the fittest" takes place. Selection brings order into the enormous diversity of the initial formation of heteropolymers.

We did not apply a physically sensible constraint to the time program of the dilution flux. Equation (3) and the results derived from it hold for any $\phi(t)$. Selection between selfreplicating polynucleotides takes place under stationary conditions as well as in growing or shrinking populations. The only assumption which has to be made does not concern $\phi(t)$; it is simple $\bar{E} > 0$, which means that the formation rate has to be larger than the mean decay rate (Eigen and Schuster 1979, pp. 35, 36).

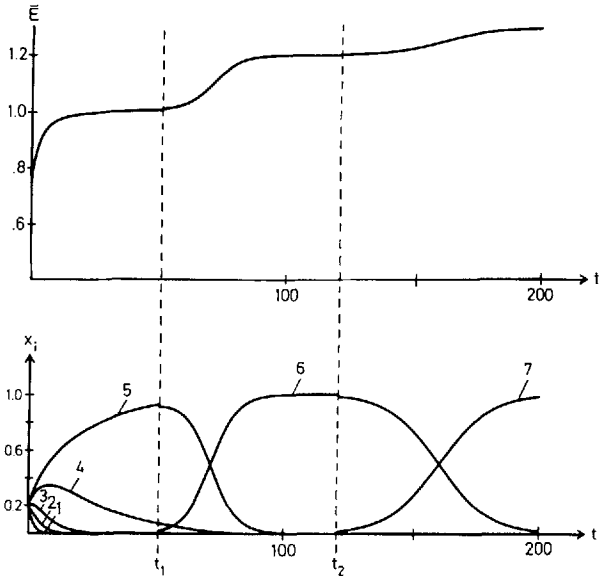


Fig. 2. Selection and optimization in an ensemble of polynucleotides. We present solution curves of the differential equation (3) $\dot{x}_i = x_i(E_i - \bar{E})$, $i = 1, 2, ..., 7$ with $E_1 = 0.3$, $E_2 = 0.65$, $E_3 = 0.82$, $E_4 = 0.96$, $E_5 = 1.01$, $E_6 = 1.2$ and $E_7 = 1.3$ in arbitrary time units and $\bar{E} = \Sigma E_i x_i / \Sigma x_i$. The initial conditions are: $x_1(0) = x_2(0) = ... = x_5(0) = 0.2$ and $x_6(0) = x_7(0) = 0$. The mean excess productivity $\bar{E}$ starts from an initial value of $\bar{E} = 0.748$ and increases steadily as the population becomes homogeneous: $\bar{E} \to 1.01$ and $x_1, x_2, x_3, x_4 \to 0$ as $x_5 \to 1$. One easily recognizes the cause for the increase of $\bar{E}$: The less efficiently growing polynucleotides are eliminated. They disappear in the same sequence as their excess productivity E_i increases, namely 1, 2, ... 5. At $t = t_1 = 50$ we observe a fluctuation $x_6 = 0 \to x_6 = \delta$. The appearance of I_6 as a favourable mutant leads to a further increase in $\bar{E}$, which approaches the value $\bar{E} = 1.2$ as I_5 disappears and the population becomes homogeneous again: $x_1 = ... = x_5 = 0$, $x_6 = 1$. The same story happens when a more efficient mutant appears at $t = t_2 = 120$. I_6 is replaced by I_7 and $\bar{E}$ approaches the temporary optimum $\bar{E} = 1.3$. This example illustrates the nature of the selection process and the role of the mean excess productivity $\bar{E}$ as the quantity which is optimized

Experimental studies on selection between replicating polynucleotides were performed first by Sol Spiegelman more than ten years ago (Spiegelman 1971). He replicated RNA of the bacteriophage Qβ in serial transfer experiments (Fig. 3). The stock solution contained the specific enzyme Qβ replicase as well as the four nucleoside triphosphates, GTP, ATP, CTP and UTP. Starting out from infectious RNA strands, sequences were obtained through serial transfer which replicated faster and had lower molecular weights than the original molecules. The overall rate of synthesis, comparable to $\bar{E}$, increased steadily during the sequence of transfers. The more rapidly replicating polynucleotides were unable to infect bacteria, since the constraints applied in serial transfer experiments select exclusively for higher replication rates. The

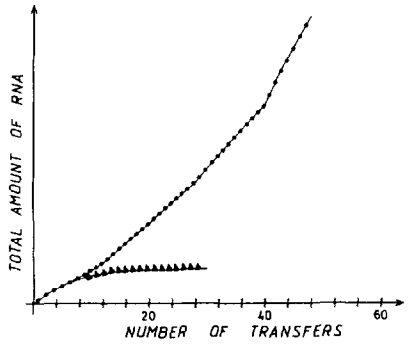


Fig. 3. Serial tranfer experiments on RNA replication by Qβ-replicase according to Spiegelman (1971) RNA replication in the test tube is initiated by addition of Qβ RNA. Two measurements are recorded: the number of infectious particles (▲) and the amount of nucleoside triphosphate incorporated into RNA, a measure of total polynucleotide synthesis (●). The rate of RNA synthesis increases spontaneously in jumps. The percentage of infectious polynucleotides, however, becomes smaller and smaller. New RNA molecules were formed which grow faster but are no longer infectious. Since serial transfer selects for excess productivity only, infectiousness is readily lost. Several spontaneous, stepwise increases in the rate of RNA synthesis are observed until an optimal value is reached.

mechanism through which the more rapidly replicating variants were formed leads us to the next principle.

3. Replication Errors Lead to Mutant Distributions Called Quasispecies. The Accuracy of the Replication Process Sets a Limit to the Amount of Genetic Information that Can Be Transmitted

No physical process can proceed with ultimate accuracy. Hence, we have to account for replication errors in polynucleotide reproduction. The incorporation of mutation into the deterministic kinetic equations of selfreplication is straightforward (Eigen 1971; Eigen and Schuster 1979, pp. 8–15). We account for error copies by means of a quality factor Q, which is used here in a manner slightly different from our earlier papers. It can be expressed as a matrix

$$Q = \{Q_{ij}\};\ i,j = 1,...,n \tag{4}$$

which characterizes the synthesis of a distribution of polynucleotides from each template. Q_{ii} is the fraction of copies of template I_i which are free of errors and Q_{ji} the fraction of molecules I_j obtained as mutants of I_i. A conservation law of probabilities holds: $\sum_j Q_{ji} = 1$, as every copy has to be either correct or erroneous. The rate constants for error-free replication and mutant formation are obtained by multiplication with the total synthetic efficiency of the template (for template I_i we use f_i as before). Additionally, we account for degradation (d_i) in the diagonal terms

$$\omega_{ii} = f_i Q_{ii} - d_i \quad \text{and} \quad \omega_{ji} = f_i Q_{ji} \quad \text{for } i \neq j$$

The kinetic equations for polynucleotide replication under the condition of an evolution reactor are then of the form

$$\dot{x}_i = (\omega_{ii} - \bar{E})x_i + \sum_{j \neq i} \omega_{ij} x_j;\ i = 1, ..., n \tag{5}$$

The mean excess production $\bar{E}$ is as before $\sum_i (f_i - d_i)x_i$. Note that equation (5) converges to the simple selection equation (3) if Q approaches the unit matrix or, in other words, in case the replication process becomes absolutely correct[1].

Equation (5) can be converted into a linear differential equation by means of suitable integral transformaton (Thompson and McBride 1974; Jones et al. 1976). In case we are interested in relative concentrations as we are here, common factors cancel and the system is completely determined by the differential equation

$$\dot{z}_i = \sum_{j=1}^{n} \omega_{ij} z_j;\ i = 1, ..., n \tag{5a}$$

with $x_i(t) = z_i(t) \exp\{-\int_0^t E(\tau)\,d\tau\}$. This equation can be studied by standard techniques of linear algebra. Ultimately we are dealing with an n-dimensional eigenvalue problem. Apart from kineticly degenerate cases there always exists a largest eigenvalue λ_{max} which is related to the largest selective value ω_{mm} (ω_{ii}; i = 1, ..., n). As in the error-free case an optimization process $\bar{E}(t) \rightarrow \lambda_{max}$ takes place. The eigenvector corresponding to λ_{max} represents the "long-lived" mode of the differential equation (5). After long enough time all the other modes have died out and the mixture of polynucleotides or mutant distribution, described by this eigenvector is all that remains. We have called it a "quasispecies" in order to indicate the analogy to the notion of a species in biology.

In order to learn more about the nature of the "quasispecies" distribution we distinguish two cases: (1) high and (2) low accuracy of the replication process. At sufficiently high accuracy the quasispecies is the stable long-term distribution of polynucleotide sequences. This distribution is dominated by the most efficiently replicating molecule, that is the master sequence I_m, with $\omega_{mm} = \max(\omega_{ii})$, which is accompanied by its most frequent and most efficient mutants.

Approximate formulas for λ_{max} and the relative frequencies of master sequence and mutants in the steady state have been obtained by perturbation theory (Eigen and Schuster 1979, p. 13; Thompson and McBride 1974;

1 Part of this section is more technical than the previous ones. The reader not interested in these details may continue with the paragraphs following equation (9)

Jones et al. 1976). These approximations describe ensembles of polynucleotides above a certain chain length ($\nu > 20$) quite well (Swetina and Schuster 1982). The most important result for our present concern is the existence of a sharply defined threshold for the accuracy of replication

$$Q_{mm} > Q_{min} = \sigma_m^{-1} \tag{6a}$$

The superiority σ_m of the master sequence contains the relevant kinetic parameters

$$\sigma_m = \frac{f_m}{d_m + \bar{E}_{-m}} \tag{6b}$$

wherein

$$\bar{E}_{-m} = \sum_{\substack{i=1 \\ i \neq m}}^{n} E_i x_i \Big/ \sum_{\substack{i=1 \\ i \neq m}}^{n} x_i$$

is the mean excess production in absence of the master sequence.

Now let us study the quality factor of the replication process more closely. Template induced replication proceeds digit per digit[2]. Therefore we can assign a single digit accuracy to every propagation step of the growing chain. For a polymer with ν segments the quality factor Q can be written as a product of ν individual factors. For the correct replication of a given polynucleotide say I_k we obtain:

$$Q_{kk} = q_1^{(k)} \cdot q_2^{(k)} \cdot \ldots \cdot q_\nu^{(k)} = \bar{q}_k^{\nu} \tag{7}$$

Herein q_1 is the accuracy for the incorporation of the first base, q_2 that for the incorporation of the second base etc. In general the single digit accuracy depends on the nature of the base to be incorporated, on the neighboring base pairs, on the mechanism of replication and on environmental factors. Nevertheless, we can define a mean single digit accuracy $\bar{q}_k$ which accounts for all these influences implicitly and which is characteristic for a given sequence I_k. For long enough natural polynucleotides with similar base compositions these mean single digit accuracies will mainly depend on the mechanism of replication since specific neighbour effects cancel out in long sequences. For a given mechanism of replication the limiting case of a minimum accuracy is equivalent to a maximum chain length ν_{max}:

$$Q > Q_{min} = \bar{q}^{\nu_{max}} = \sigma^{-1}$$

or

$$\nu_{max} = -\frac{\ln \sigma}{\ln \bar{q}} \cong \frac{\ln \sigma}{1 - \bar{q}} \tag{8}$$

Equation (8) has been used to explain the lengths of genomes in primitive organisms like RNA viruses and bacteria (Eigen and Schuster 1979, pp. 17–24). In the case of RNA viruses, moreover, the existence of nucleotide sequence heterogeneities as predicted by the concept of a quasispecies has been proved through cloning experiments (Domingo et al. 1978, 1980; Ortin et al. 1980; Fields and Winter 1981).

In order to apply equation (8) to the development of early replicating systems under prebiotic conditions we may use the results obtained in Leslie Orgel's laboratory (Lohrmann et al. 1980). He and his coworkers determined the accuracy of enzyme-free oligo (G) synthesis on a poly (C) template. They used a 50 : 50 mixture of activated A and G and found an incorporation of A:G in the ratio 1 : 200, which is equivalent to a single digit quality factor for G of $q_\nu \cong 0.995$ under these conditions. What they have determined, of course, is not the mean single digit accuracy ($\bar{q}$) of enzyme-free replication which we would like to know, but an upper limit for it: the interaction of the GC base pair is stronger than that of the AU pair and hence we have to expect a lower quality factor in the latter case. The maximum chain length for replication in enzyme-free systems thus can be estimated to lie around $\nu_{max} \cong 100$ bases or somewhat lower. The molecules which can be replicated with sufficient fidelity are of the size of the tRNA's.

Equation (7) is also useful for estimating the frequencies of mutants. For this we have to assign single digit mutation frequencies to the various base exchange processes. A systematic way to do this is shown in Table 1. Again we have to assume that neighbouring base pairs influence the muation frequencies, which then are specific for the replacement of the nucleotide at position "i". Following Table 1 we denote the three possibilities by α_i, β_i, and γ_i. In absence of insertions and deletions we have the conservation relation

$$q_i + \alpha_i + \beta_i + \gamma_i = 1$$

which expresses the fact that one of the four bases must occupy position "i" in the replica. For the correct copy and the three possible mutations at position "i", say $I_k \to I_l, I_k \to I_m, I_k \to I_p$, we thus have the following distributions of quality factors and mutation frequencies

$$I_k \to I_k : Q_{kk} = q_1 q_2 \ldots q_i \ldots q_\nu$$
$$I_k \to I_l : Q_{lk} = q_1 q_2 \ldots \alpha_i \ldots q_\nu$$
$$I_k \to I_m : Q_{mk} = q_1 q_2 \ldots \beta_i \ldots q_\nu \quad \text{and}$$
$$I_k \to I_p : Q_{pk} = q_1 q_2 \ldots \gamma_i \ldots q_\nu$$

[2] For the sake of simplicity we dispense with a detailed discussion of other sources of replication errors besides point mutations. In principle deletions or insertions could be incorporated into the model as well

Table 1. Systematics of point mutations

Original base	Base in the mutant: Purine-Purine Pyrimidine-Pyrimidine (Transition)[a]	Frequency	Intrapair exchange (Transversion)[a]	Frequency	Intrapair exchange (Transversion)[a]	Frequency
G	A	α(G)	C	β(G)	U	γ(G)
A	G	α(A)	U	β(A)	C	γ(A)
C	U	α(C)	G	β(C)	A	γ(C)
U	C	α(U)	A	β(U)	G	γ(U)

a Notation of Volkenstein (1979)

Now let us consider a realistic situation as observed with RNA replication in bacteriophages. In this particular example purine-purine or pyrimidine-pyrimidine exchanges are the most probable mutations by far. We may account for this situation by the approximative assumption

$$1 - q \cong \alpha > > \beta, \gamma$$

In this case it is possible to calculate the frequency factor Q_{jk} for a certain mutation of this kind ($I_k \rightarrow I_j$) from the Hamming distance D_{jk} of these two sequences. (The Hamming distance is the smallest number of base exchanges with which it is possible to convert a given sequence into another.) Again we use mean single digit accuracies ($\bar{q}_k$) and mutation frequencies ($\bar{\alpha}_k \cong 1 - \bar{q}_k$)

$$I_k \rightarrow I_j : D_{jk} = d \rightarrow Q_{jk}\, \bar{q}_k^{\nu-d}\ \ \bar{\alpha}_k^d \cong \bar{q}_k^{\nu-d}\ (1 - \bar{q}_k)^d$$

Remembering that $\bar{q}$ is close to one and hence $1 - \bar{q}$ is very small we realize a coarse graduation of mutation frequencies: one error mutations are formed with higher frequencies than two error mutations, three error mutations have still lower frequencies, etc.

The concentration of any mutant in the quasispecies is determined by a superposition of the corresponding mutation frequency and the relevant rate constants. Perturbation theory leads to an approximate formula for the ratios of concentrations within the quasispecies

$$\bar{x}_i = \bar{x}_m \cdot \frac{Q_{im} \cdot f_m}{\omega_{mm} - \omega_{ii}}\ ;\ i = 1, .., n;\ i \neq m \qquad (9)$$

where the stationary concentrations of the master sequence and some mutant are denoted by $\bar{x}_m$ and $\bar{x}_i$ respectively. By "m" we denote tha master sequence as above. Before we turn to the low accuracy case we consider the fraction of a master sequence within this quasispecies as a function of the accuracy of replication (Fig. 4). We realize a steady decrease of this fraction ($\bar{x}_m$) until it becomes exceedingly small at the critical value Q_{min}.

At low replication accuracy the relative frequencies of the individual mutants are no longer determined by the kinetic constants. The selective values ω_{ii} then have no influence on the long term polynucleotide distributions. These are determined instead exclusively by the

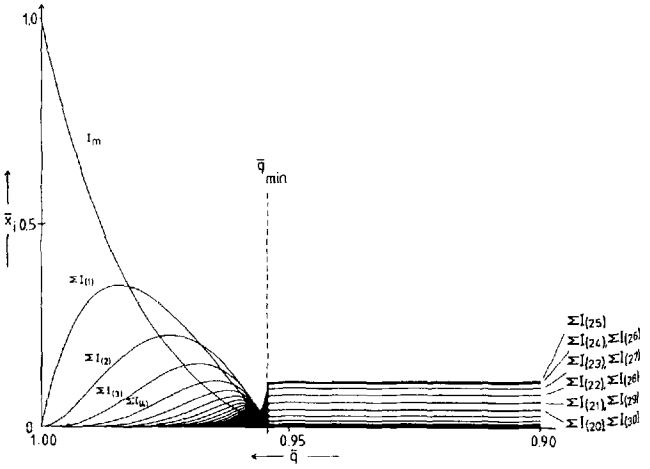


Fig. 4. Stationary mutant distribution for the prevailing conditions (quasispecies) in an replication ensemble of polynucleotides. The curves represent the relative concentration of the master sequence (I_m), the sum of the relative concentrations of all one error mutants ($\Sigma I_{(1)}$), of all two error mutants ($\Sigma I_{(2)}$) etc. as functions of the mean single digit replication accuracy ($\bar{q}$). In this example the polynucleotides length is $\nu = 50$. The formation rate constant for the master sequence is chosen to be $f_m = 10$, all other sequences $f_1 = f_2 = ... = 1$ in arbitrary time and concentration units. All decomposition rate constants including that of the master sequence are set equal and hence do not enter into the differential equation (5). From these parameter values we calculate $\sigma = 10$ and a critical single digit accuracy of $\bar{q}_{min} = (Q_{min})^{1/\nu} = 0.945$ according to equation (8). With decreasing replication accuracy a pronounced decrease in the relative concentration of the master sequence is observed. One-error mutants, then two-error mutants dominate the polynucleotide distribution. The critical value $\bar{q}_{min}$ calculated by perturbation theory (8) is an excellent approximation to the exact solution. Below the critical accuracy the concentrations are exclusively determined by the statistical weight of the corresponding sequences. Hence, the sum of the concentrations of 25-error mutant ($\Sigma I_{(25)}$) is largest, followed by 24- and 26 error mutants ($\Sigma I_{(24)}$, $\Sigma I_{(26)}$) etc. (For details see Swetina and Schuster 1982)

statistics of mutations. In case there are no statistical degeneracies all sequences are present at equal concentrations (see e.g. Fig. 4). In such a system replicating at a low level of accuracy there is no inheritance. Following a consecutive series of replica from generation to generation there is no conservation of the polynucleotide sequence. Such a series of polynucleotides thus approaches a random walk in an abstract space of all possible sequences.

So far we have studied our problem by means of ordinary deterministic kinetic equations. This approach assumes that all molecules under consideration are present in large numbers of copies. What "large" means depends on the particular reaction mechanisms studied. In most cases a few thousand molecules are sufficient to guarantee for all practical purposes, coincidence of the deterministic solution and the corresponding stochastic process. Facing the reality of polynucleotide replication we have to account for the "hyperastronomically" high numbers of different sequences. These numbers, as we pointed out earlier, exceed by far the numbers of molecules available in any realistic population. It is impossible to interpret the kinetic equations naively to include all possible sequences, because the dimension n is so large that most variables take exceedingly small values. (A meaningful concentration could really not be less than one molecule per earth's volume of water.) The differential equations thus have to be restricted at least to those sequences which are actually realized in the system.

Let us visualize the evolution of a replicating system by means of an abstract but nevertheless illustrative model. We construct a point space by assigning a point to each sequence. The points are arranged such that sequences which differ by single base exchange, i.e. pairs of sequences with Hamming distance D = 1, are first (closest) neighbours. The Hamming distance, thus, defines a kind of metric on this point space. Due to the hyerastronomically large number of possible sequences this set of points is giant. A quasispecies occupies a small part of it only which can be understood as a connected subset or subspace[3]. We distinguish three different scenarios:

(1) The replication process is very accurate, so that only a few mutants are present. The quasispecies then corresponds to a stationary or practically stationary subset in a low dimensional subspace of the giant space of possible sequences. The appearance of mutants outside this occupied subspace is a rare event and consequently the system is described well by deterministic differential equations.

(2) The replication process is of low accuracy. Then the evolution of the system corresponds to a random walk through the giant space. No stationary distribution of polynucleotides can be reached since frequently occurring replication errors create a high percentage of new sequences all the time.

(3) The replication process is just sufficiently accurate to sustain a master sequence at low stationary concentration. Replication in this case operates close to the minimum accuracy (Q_{min}) and we encounter a great variety of mutants. Eventually a sequence turns up that replicates more efficiently than the master and thus will replace it. The ensemble of polynucleotides as a whole follows a trajectory through the giant point space along which the efficiency of replication is optimized. This process resembles the schematic diagram shown in Fig. 2.

Case (3) is most relevant for evolution since it allows the polynucleotide structures to optimize with respect to properties relevant for replication speed and accuracy. At first glance it is not at all obvious that such a "hill climbing process" really leads to any respectable peak, and we may well ask why optimization does not soon stop at some relatively "low hill". If by a similar procedure we wanted to reach any of the major peaks on the Earth we soon would have to carry out practically forbidden large jumps between lesser peaks.

The optimization procedure based on RNA- or DNA-replication, however, is not well represented by a hill climbing process in an ordinary mountainous landscape. Consider a sequence having ν positions. Any of them can undergo three possible changes at any instance, e.g. from A to G, U or C (see Table 1). Our metapher, "hill climbing in some mountain-range" considers the optimization of height within a landscape erected upon a two-dimensional grid. We have two degrees of freedom to chose a step of given length. In polynucleotide replication we have 3^ν degrees of freedom for single base exchange in a given polynucleotide of length ν (in case the three exchanges are not equivalent the number of degrees of freedom may reduce to ν with some peculiar constraints). The topologies of two- and high-dimensional space – ν is the range of a few thousand for viral RNA already – are quite different. One of the most dramatic consequences of this is the reduction of (average) distances. With ν positions there are alltogether 4^ν possible sequences. If these 4^ν sequences are mapped onto a linear space, i.e. arranged in a row, the average distance of two sequences is very large. In an ν-dimensional space with unit edge length, the largest possible distance – e.g. between some arbitrary initial and the optimal final sequence – never exceeds ν. Hence the landscape in ν-dimensional space is much more bizarre most points being saddle points that allow change to continue uphill in at least one direction. Among the accessible points one might thus find – for virtually all initial positions – always some respectable nearly peak. Moreover, if the

[3] "Connected" means in this context that the subset does not contain isolated members, i.e. points which have no other member of the subset in a neighbourhood of Hamming distance D = 1

dimension ν is not too large and it is possible to populate the mutant spectrum with k-error mutants, k being larger than 1, one might always jump from any local peak to a quite different environment. How far one can get in this way, whether one reaches the absolute maximum, depends largely on the distribution of peaks. Theoretical studies of this question with various topologies of protein structures are underway. Experiments carried out with "de novo" synthesized RNA chains using virus replicases show that RNA populations really adapt quickly to optimal performance of replication in a changing environment (e.g. increasing salt concentrations or inhibitors of replication such as ethidium bromide, Sumper and Luce 1975; Biebricher et al. 1981).

4. Cooperation Between Otherwise Competitive Self-Replicating Elements is Introduced by Higher Order Catalytic Action in the Form of Positive Feedback Loops Called Hyercycles

Self-replication inevitably leads to competition and selection of the most efficiently replicating species. The combination of selection and mutation allows structures and properties to optimize in the Darwinian sense. The accuracy of the replication process, however, sets a limit to the size of the optimizable structures. The lengths of the genetic message that can be transmitted from generation to generation is limited by the properties of the replication machinery. In evolution such a replicating system thus runs into a dead end when it approaches the critical polynucleotide length λ_{max}. In order to build a more accurately replicating machinery more genetic information is needed, but more information can be processed only when a more accurately operating machinery is available. How can this machinery – i.e. better replicase function – be made? There is a straightforward answer to this problem: several information carriers must be present at the same time and in controlled relative amounts. In order to be sufficiently general, control has to be exerted internally by a suitable reaction mechanism. Let us consider the mathematical problem first. If we apply mass action kinetics the analysis is simplest: first order template-induced replication leads to selection and, hence, we have to consider second-order terms in order to search for cooperation between self-replicating elements. The differential equation to be integrated is thus of the form

$$\dot{c}_i = (k_i + \sum_j k_{ij} c_j - \frac{1}{c}\phi)\, c_i; \quad i = 1,2, ..., n \qquad (10)$$

For sake of brevity we dispense with terms which arise from mutations. The rate constant k_i, in general, will be equal to the excess production of template I_i, i.e. the difference in rate constants between first order formation and degradation: $k_i = E_i = f_i - d_i$. The total concentration c is defined as in equation (1) to be $c = \sum_i c_i$. Note that we use concentrations instead or relative concentrations (x_i) here, because the dynamics of the system depends on the total concentration c. In order to visualize this dependence we adjust the flux ϕ in such a way that it leads to the condition of constant organization, $c = c_o$ = constant, and study equation (10) as a function of c_o. At low total concentration ($\lim c_o \to 0$) the first order terms will dominate and we are back to equation (1), which leads to selection. In the high concentration limit the first order terms can be neglected. The dynamics of the system then may be studied appropriately by means of the simpler equation

$$\dot{c}_i = c_i (\sum_j k_{ij} c_j - \frac{1}{c}\phi); \quad i = 1, ..., n \qquad (11a)$$

As we have shown (Eigen and Schuster 1979; pp. 31, 32) the topology of equation (11a) is independent of c, since the kinetics of the system is determined by a homogeneous function of the concentrations ($\sum_i k_{ij} c_i c_j$ is homogeneous of degree 2) and therefore we may formulate the problem in relative concentrations without loosing generality

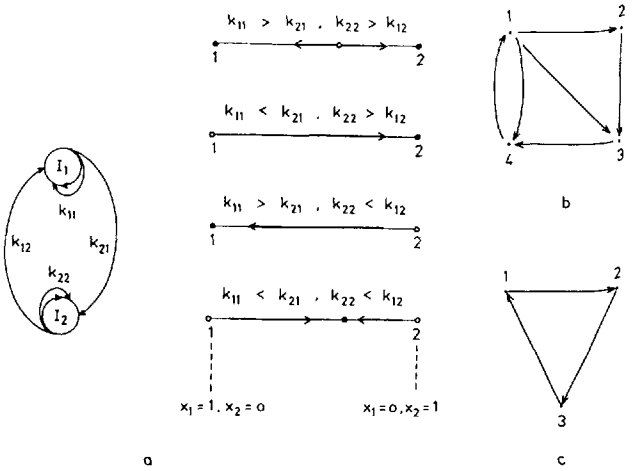


Fig. 5 a–c. Catalytic action and cooperation. a Different qualitative behaviour of the differential equation (11b) with n = 2. We distinguish four cases: (1) $k_{11} > k_{21}, k_{22} > k_{12}$ leads to selection of either I_1 or I_2 depending on the initial concentrations; (2) $k_{11} < k_{21}, k_{22} > k_{12}$ leads to selection of I_2; (3) $k_{11} > k_{21}$, $k_{22} < k_{12}$ to selection of I_1; and (4) $k_{11} < k_{21}, k_{22} < k_{12}$, finally, is the only case where both I_1 and I_2 are present at the stable stationary state. The relations of rate constants for cooperation ($k_{11} < k_{21}, k_{22} < k_{12}$) can be subsumed under the general formulation that mutual enhancement has to exceed self-enhancement; b A graph with a Hamiltonian arc: $1 \to 2 \to 3 \to 4 \to 1$; c The graph corresponding to a hypercycle (n = 3). Note that it consists only of a single Hamiltonian arc

$$\dot{x}_i = x_i\left(\sum_j k_{ij}x_j - \phi\right);\ i = 1, \ldots, n \qquad (11b)$$

Equation (11b) has been studied in great detail (Hofbauer et al. 1980). For present purposes we require two general results. Cooperation between self-replicating elements is observed in systems in which the elements are connected by a positive feedback loop of catalytic actions (Fig. 5). In graph theory such a closed loop is called a Hamiltonian arc. In low dimensional systems (n = 2, 3 and 4) the existence of such an Hamiltonian arc is a necessary condition for cooperation. Independently of the dimension n another general theorem of cooperation could be proved (Schuster et al. 1979; Hofbauer et al. 1981): the elements of a system show cooperative behavior if the network of catalytic action consists exclusively of such a positive feedback loop. In context of the rate constants in equation (11b) this means that only particular off-diagonal elements of the matrix $K = \{k_{ij}\}$ may by different from zero.

$$K = \begin{bmatrix} 0 & 0 & \ldots\ldots & 0 & k_{1n} \\ k_{21} & 0 & \ldots\ldots & 0 & 0 \\ 0 & k_{32} & \ldots\ldots & 0 & 0 \\ & & & & \\ 0 & 0 & \ldots\ldots & k_{n,n-1} & 0 \end{bmatrix}$$

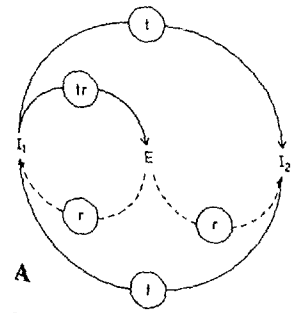


Fig. 6. RNA-phage infection of a bacterial host, as a simple hypercyclic process. Using the translation machinery (tr) of the host cell the infectious plus strand (I_1) first instructs the synthesis of a protein subunit (E) which associates with three host proteins to form a phage-specific RNA-replicase. This replicase complex (r) exclusively recognizes both phage-RNA strands, plus and minus, and replicates them. The result is a burst of phage-RNA production which follows a hyperbolic growth law. We show solution curves obtained by numerical integration of the differential equations which describe the most simple reaction mechanisms of this kind:

$$I_1 \xrightarrow{k} E + I_1 \qquad [I_1] = x_1$$

$$I_1 + E \xrightarrow{f_2} I_2 + I_1 + E \qquad [I_2] = x_2$$

$$I_2 + E \xrightarrow{f_1} I_1 + I_2 + E \qquad [E] = e$$

The rate constants and initial concentrations were: k = 0.01, $f_1 = 0.02$, $f_2 = 0.04$ and $x_1(0) = 0.01$, $x_2(0) = 0.01$, $e(0) = 1$ in arbitrary time and rate constant units. Despite the fast growth of x_1 and x_2 their ratio soon approaches the stationary value $(f_1/f_2)^{1/2}$

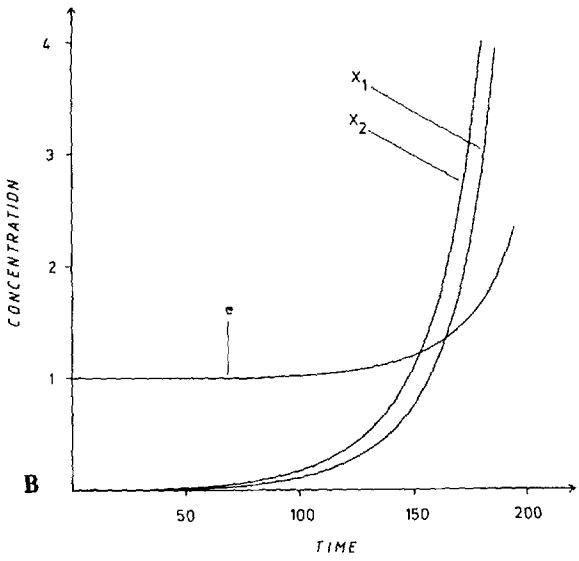
Equation (11b) than simplifies further ($k_{i,i-1} = \gamma_i$ all indices are understood modulo n: n + 1, 0 = n):

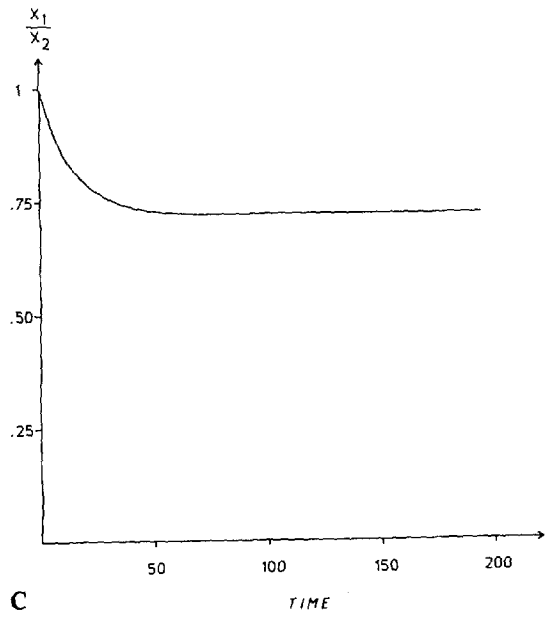
$$\dot{x}_i = x_i\left(\gamma_i x_{i-1} - \phi\right);\ i = 1, \ldots, n \qquad (11c)$$

A system which is described by an equation of type (11c) is called an elementary hypercycle (Eigen and Schuster 1979; pp 41 ff.).

For a physical interpretation of the theorem of cooperation we consider the simple two-dimensional case n = 2 (Fig. 5). Co-operation requires that mutual catalytic enhancement exceeds self-enhancement. In terms of rate constants this requirement implies $k_{12} > k_{22}$ and $k_{21} > k_{11}$. Higher dimensional systems (n > 2) have to meet additional requirements: mutual enhancement has to be such that it sustains a closed catalytic feedback loop.

Do we find hypercyclic organization in nature? In order to start from simple systems let us consider the repli-





cation of a simple RNA bacteriophage like Qβ in a host cell (Fig. 6). Although this system represents a rather trivial example of a hypercycle – the plus-minus ensemble is one replicating entity and thus, the cycle contains only a singe element – it is illustrative with respect to the nature of catalytic action. The plus and minus strands both act as templates in replication. Catalytic action upon the replication process is the origin of the second factor (x_j) in equation (11b). RNA molecules lack the functional groups needed to serve as catalysts, so it is not hard to understand that almost all catalysis in biochemistry is done by proteins. The role of RNA in this particular catalysis (aside from the template function) is an indirect one, to be seen in the action of the plus strand as the messenger for translation into the specific replicase. More precisely, the plus strand codes for one of the four subunits of Qβ-replicase. Positive catalytic feedback of RNA on RNA replication thus occurs via translation. Polynucleotide translation is a very complicated, many step process. Under suitable conditions, however, it can be described well by a simple "over-all" kinetics which is closely related to the dynamics of differential equation (11b) (see also Eigen and Schuster 1979, pp. 50 ff.).

Considering the growing plus-minus ensemble in Fig. 6 we observe a faster than exponential increase in the concentration of RNA, which we denote as hyperbolic growth (Eigen and Schuster 1979, pp. 29, 30). Moreover, we recognize internal regulation of the ratio of concentrations of plus and minus strand: after a short initial period this ratio approaches the square root of the corresponding ratio of rate constants (Eigen 1971).

Some higher RNA viruses, like influenza, have a genome split into independent RNA molecules (Palese 1977). RNA synthesis in such viruses has to be organized in such a way, that the individual molecules are present in equal molar amounts for packing into the new virus particles. Although the mechanism of this regulation is not known, it seems to be a good candidate for some kind of hypercycle.

Other examples of hypercyclic organization are found with higher organisms when different species cooperate in symbiosis. In these systems again mutual enhancement has to exceed self-enhancement in order to stabilize the symbiontic organization.

Among other examples from the biology of higher organisms we mention one from sociobiology that has been analyzed in detail by different mathematical models. Often we observe a kind of self-organization of behaviour in animal societies which can be analyzed formally by a differential equation of type (11b). The mathematical basis for co-existence of two or more strategies of behaviour is precisely the same as in hypercycles: mutual support has to exceed self-enhancement (Schuster et al. 1981; Hofbauer et al. 1982).

Let us return to the question of how hyercycles may have been realized under prebiotic conditions. Replicating polynucleotides organize themselves into quasispecies under conditions such that the error limit for replication is rather low. According to the experiments by Leslie Orgel and coworkers (Lohrmann et al. 1980) it lies around $\nu_{max} \cong 100$ or somewhat lower in A,U-containing molecules. The RNA molecules available for further evolution were thus of the size of present day tRNAs. Polynucleotides are not efficient catalysts, so in order to improve catalytic efficiency instructed protein synthesis is required. Randomly polymerized polypeptides with the desired catalytic properties help very little. They can exert their catalytic function only temporarily, because they are hydrolysed after some time and as a consequence of the "hyperastronomically" large number of different sequences the same molecule will never be formed again. How could translation or a primitive version of it then arise? A detailed mechanism of this process is not known yet. Several models have been suggested (see e.g. Crick et al 1976; Eigen and Schuster 1979, pp. 62–67) and these have to be exposed to experimental tests now. We do not intend to go into too much detail here and to repeat the arguments that have been discussed extensively elsewhere (Eigen and Schuster 1979; pp. 62–67; Schuster 1981; Eigen and Winkler-Oswatitsch 1981) but we would like to stress some features of a "minimum requirement system" for primitive translation. Such a minimal system definitely requires more than a single RNA sequence. With reduced demands on structural and functional perfection we can assert multiple functions for primordial polynucleotides: a straightforward suggestion in this direction is that the primordial adaptors, the tRNA precursors, were also the first genes, and that eventually the plus strands had one, the minus strands the other function (Eigen and Winkler-Oswatitsch 1981). Nevertheless, one would need at least as many RNA molecules as there are amino acids to be incorporated into the primitive code. Four different amino acids might well have been sufficient to build up the first instructed polypeptides. Accordingly, we would require at least four polynucleotides to be present at the same time. Polynucleotides can be available for longer periods in case they are members of the same quasispecies. Four polynucleotide sequences might have originated from a single quasispecies distribution in which the relative frequences of the individual molecules were determined exclusively by their capability to replicate and by the frequencies of mutation. After an initial phase, during which the simultaneous presence within the same quasispecies was sufficient, the primitive translation system needed more reliable control on the relative concentrations of its constituents. This can be seen from a straightforward consideration: Let us assume that the system succeeded in building a primitive replicase. Then, catalyzed replication soon will outweight the uncatalyzed process and the rate constants of replication are no longer the same as for enzyme-free replication. Another quasispecies forms. The new and the old mutant distri-

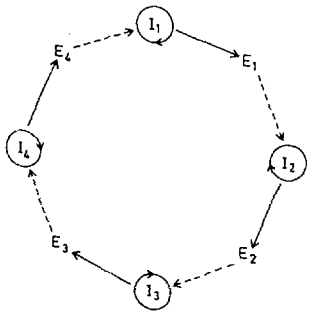


Fig. 7. Hypercycle with translation. The polynucleotides I_i are carriers of two kinds of instruction, one for their own replication, due to their capability to act as template, and the other for their translation into the proteins E_i, which are the effective catalysts. The proteins act as specific replicases or specific subunits in a multi-subunit enzyme

butions may differ substantially. Eventually, the former master sequence is replaced by another one which is better adapted to the enzyme-catalyzed reaction. Very likely the system loses control of the relative concentrations of the four essential polynucleotides. This deficiency in regulation is a substantial handicap for further evolution of the primitive translation system. Hypercyclic organization avoids this difficulty: once a catalytic feedback loop has been formed the mechanism of catalyzed replication controls internally the relative concentrations of all the constituents. One of the major problems concerns the origin of hypercyclic coupling. The conceptually simplest suggestion, but perhaps not that which is easiest to realize physically, introduces specificity into RNA replication: the translation products of the primordial genes act as specific catalysts in the replication of other RNA molecules (Fig. 7). Deriving all of the RNA molecules from a single quasispecies can be of advantage in this context: the individual RNA-molecules and their translation products differ only slightly in the beginning. Thus, they may have similar properties and closely related structures.

How can hypercycles evolve after they are formed and have passed through an internal stabilization period? A system with hypercyclic organization approaches either an internal equilibrium or a state with oscillating concentrations. Thereafter we can consider the whole system as an integrated functional entity. In competition with the external world such a hypercycle behaves like a single individuum growing with a non-linear rate. In the simplest case we deal with quadratic growth rates. Competition between hypercycles then is determined by the differential equation

$$\dot{C}_i = k_i C_i^2 - C_i \cdot \frac{\phi}{C} \qquad (12)$$

Here C_i is the total concentration of hypercycle H_i, i.e. the sum of the concentrations of all polynucleotides belonging to H_i, and $C = \Sigma C_i$. Equation (12) is of some interest since it describes an example of restricted optimization. In an initial phase during which several hypercycles are present in comparable concentrations, selection of the system with the largest value of k_i takes place. After on hypercycle has won this initial competition the chance of replacing it is practically zero for any new although more efficient hypercycle. Even large fluctuations in concentrations are not enough to defeat the established system. There is no optimization of properties through adaption by successive replacement of less efficient competitors between hypercycles like that we observed with competitive polynucleotide sequences.

One way how hypercycles can nevertheless evolve is shown schematically in Fig. 8. Suitable mutations can be incorporated into the system when they fit properly into the network of catalytic reactions. Stepwise enlargement of the system through the incorporation of mu-

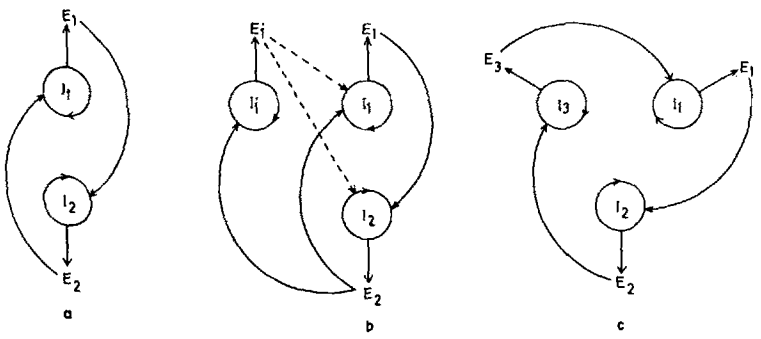


Fig. 8 a–c. Evolution of hypercycles. A mutation – incorporation mechanism allows the stepwise extension of hypercycles. **a** A two membered hypercycle with translation; **b** A mutant of I_1 here denoted by I_1' appears. It has a certain catalytic effect on the replication of I_1 and I_2 by means of its translation product. **c** Extension of the hypercycles by one member occurs in the case the mutant fulfils the following two conditions: (1) I_1', the later I_3, is recognized better by E_2 than I_1 is and (2) the translation product of I_1' (= I_3), the protein E_1' (= E_3) is a better catalyst for the replication of I_1 than E_2 is. In other words, catalytic coupling has to fulfil the criterion of a three membered hypercycle: mutual enhancement along the cycle $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ has to prevail over self-enhancement

tants is thus a highly conservative process: what has been established once is kept in the future. We may speak of "once for ever" decisions.

How can one test models of the origin of tranlation? Primarily, further experimental research on polynucleotide-polynucleotide and polynucleotide-polypeptide interactions will help to bring order into the various suggestions by making some more, others less plausible. The study of specific aggregates has to be carried on. How can we hope to speculate constructively on the origin of translation when our knowledge of the ribosome, its structure and function as well as the details of subunit interactions within it, is still far from being complete? In paleontology, research on the course of biological evolution is based on the fossil record. Prebiotic chemistry is guided by the results of simulation experiments in which reactions are studied under physico-chemical conditions which mimic the most likely environment on the primordial Earth. Research on the origin of translation is in a much worse situation: there are no fossils available for this very early period of emerging life and the molecules involved in this development are too complicated to allow straight forward simulation. So one has to look for the basic features of possible reaction mechanisms first. How can one check these predictions? Did this important stage of the origin of life pass by without leaving any remnants or traces for us? Attempts have been made to search for "intellectual fossils", remnants of early evolution which are hidden in present day structures and functions of biological macromolecules. Particularly wellsuited candidates for such investigations are the sequences of polynucleotides in regions where the restrictive demands on their functional properties are low or where redundancies in the genetic code allow some freedom in the choice of bases. Some of the recent sequence studies (Eigen and Winkler-Oswatitsch 1981; Cedergren et al. 1981; Sheperd 1981) are conclusive with respect to regularities in the present polynucleotides. Studies were performed on tRNA's and on DNA from various sources. These regularities suggest that primordial polynucleotides preferred a regular pattern of repeated triplets of the type -RNY- where R stands for a purine base, Y for a pyrimidine and N for any of the four bases. This pattern agrees well with the postulates of repetitive patterns made in models of primitive translation (Crick et al. 1976; Eigen and Schuster 1979, pp. 62–67). Stability considerations suggest a high GC content of the early information carriers. Combining the various arguments we end up with a suggestion for the first four codons in the primordial translation system: these four triplets and the corresponding amino acids are GGC = glycine, GCC = alanine, GAC = aspartic acid and GUC = valine. It is worth noting that these four amino acids are the most abundant found under simulated prebiotic conditions.

Further search for intellectual fossils and conclusive experimental studies on the dynamics of polynucleotide-polypeptide systems will bring more light into this exceedingly important and fascinating period of early evolution.

5. Formation of Compartments and Individualization Allows Efficient Evaluation of the Relevant Functional Properties of Translation Products

The fifth and last principle of emerging life we discuss here is ubiquitous and easy to observe in living matter: life is everywhere cellular. The most obvious question that comes to our mind then is: why and when did the cellular structure arise in the course of early evolution? Let us concentrate on the "why" first and visualize a system which consists of polynucleotides and polypeptides and which is capable of replication and translation. Is such a system also capable of precellular evolution in the Darwinian sense? Selection occurs at the level of polynucleotides, as has been verified in great detail experimentally. What about their translation products? In order to visualize this problem, consider two idealized classes of mutations which we denote as "phenotypic" and "genotypic" (Fig. 9). The major effect of a phenotypic mutation is a change in the tertiary structure of the molecule. Accordingly, the polynucleotide becomes a better or a worse target for replication (as well as it gains or loses in stability towards hydrolysis). Selection immediately sets in and the phenotypic properties are readily improved step by step until an optimal rate of replication is achieved. The properties of the translation product may be changed very little in case of a typical phenotypic mutation. In the language of molecular genetics an extracistronic mutation would be an example of a phenotypic mutation. Genotypic mutations behave oppositely in this respect: they lead to a substantial change in the translation product without having much of an effect on the tertiary structure of the polynucleotide. Let us assume that the modified translation product is more efficient than its precursor. In a homogeneous medium this advantageous polypeptide does not specifically favour the production of the mutant gene. It favours the synthesis of different polynucleotides or leads, eventually, to a general enhancement of polynucleotide and polypeptide synthesis. The result of a favourable genotypic mutation thus will be a local increase in the rate of synthesis and also a local increase of concentrations provided diffusion is sufficiently slow. At the steady state we shall find a concentration gradient and all molecules from the more productive volume element will spread slowly together with the favourable mutant gene. This scenario suggests a possible solution to the problem: in spatial isolation the favoured volume element would have the most pronounced advantage.

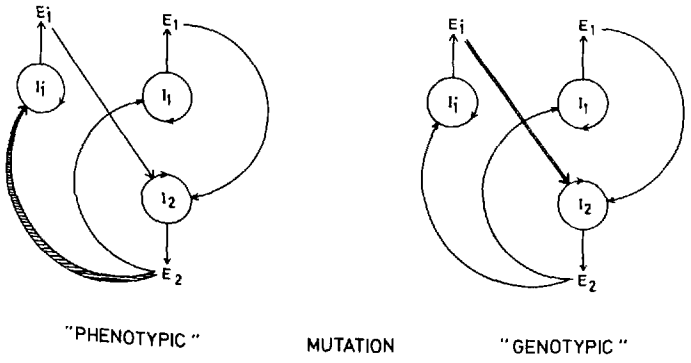


Fig. 9. Two idealized classes of mutations in a primitive replication-translation system. The "phenotypic" mutations lead to mutants (I'_1) which are better targets for the specific replicase, whereas the properties of their translation products (E'_1) are about the same as in the wild type (E_1). The "genotypic" mutant, in contrast, is characterized by a better translation product but more or less unchanged recognition by the replicase

Phenotypic and genotypic mutations are abstractions which we have used here for the purpose of illustration. In practice, every change of the sequence of a primordial gene alters both the tertiary structure of the polynucleotide and the properties of its translation product. In most cases one effect will dominate and a classification is meaningful. In principle, the whole argumentation can be given also with four classes:

(1) phenotypically favourable – genotypically favourable,
(2) phenotypically favourable – genotypically deleterious,
(3) phenotypically deleterious – genotypically favourable and
(4) phenotypically deleterious – genotypically deleterious.

The picture we obtain in this case is much more detailed but the general conclusions are the same.

Slow diffusion limits spatial spreading and has a weak isolating effect. Nevertheless, it cannot really support selection of favourable genetic mutation effectively. The real solution to the problem, spatial isolation and individualization, is the fifth principle of self-organization. In primitive systems this is achieved by the formation of compartments. A compartment can be defined as a volume which is separated from the surrounding medium by structurally stable, spatial barriers. These barriers consist e.g. of ampholytic organic material with lipid like properties and thus resemble present day cell membranes. Barriers built from other material are also thinkable. The compartment includes a set of macromolecules. Some of these molecules are polynucleotides undergoing replication. In the compartment the fate of the mutant gene is essentially coupled to the efficiency of its translation product, because the entity on which selection acts is the whole compartment. Replication of RNA in compartments is shown schematically in Fig. 10. The evolving system is back to Darwinian evolution. Full optimization of all properties which are relevant for compartment replication is possible now.

Isolation of genes in compartments solves some problems that cannot be solved in homogeneous solution in a simple way. However, it also creates problems that do not exist in homogeneous solution. Apart from the capability to select for advantageous genotypic mutations, the major positive effect of compartments is to maintain high local concentrations of macromolecules. The membranes surrounding the compartment act as diffusion barriers for all high molecular weight materials. Other advantages that may result from the presence of membranes are catalytic action of surfaces, activation of monomers through processes coupled to membrane transport, etc. The problems created by compartment formation can be summarized in four groups:

(1) Material transport has to be organized in some way. The growing system requires an input of the activated building monomers. Hydrolysis products have to be re-

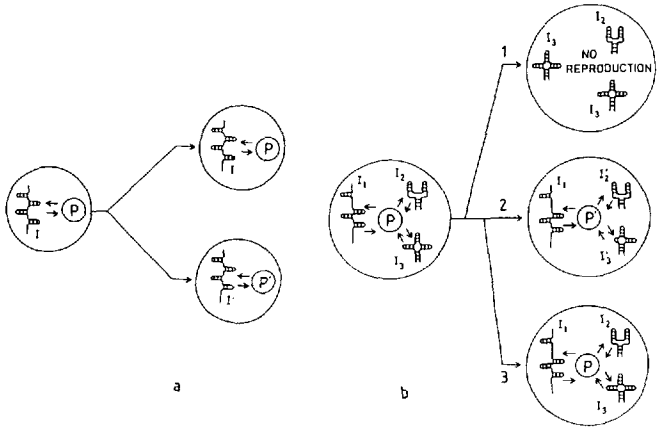


Fig. 10 a and b. Replication in compartments. Compartmentation puts selection pressure on genotypes. Suppose information carrier I encodes machinery (P) for translation and replication a After mutation of I to I' and translation of I' to P', the system may be recompartmentalized into I/P and I'/P' daughter compartments. Such compartmentation allows selection of the gene encoding the more efficient machinery. For selection to occur, the superiority of the machinery must compensate for errors made in each round of replication. Compartmentation by itself, however, does not suffice for selection of superior information divided among several information carriers. b Daughter compartments with incomplete sets of components are not viable (1). Compartments with sets that are complete but include mutations (2) allow evolutionary change. Even if the original set is the optimal one (3), the rate of proliferation required to maintain stability is much higher than it would be in a homogeneous solution without compartmentation. The net effect, in other words, is a severe sharpening of the error-threshold condition.

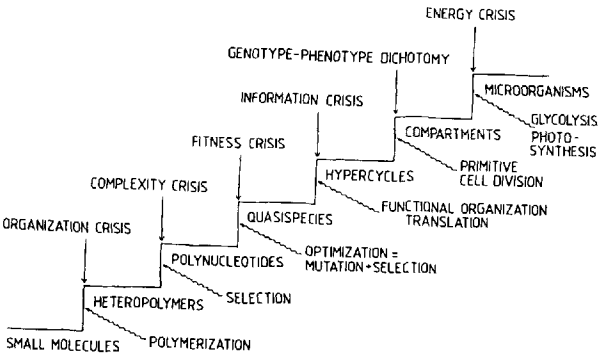


Fig. 11. Six critical steps during early evolution. On its way from small molecules to microorganisms the evolving system had to pass at least the six "crises" indicated by arrows. A "technical" innovation or new principle helped to overcome the obstacle the system had run into. The new concept is indicated by wavy arrows

moved or reactivated to enter into polymers synthesis anew. This problem requires e.g. membranes which sustain the diffusion of low molecular weight compounds by means of gates or pores which are small enough to hinder diffusion of macromolecules.

(2) Formation and division of compartments lead to packing problems. A complete set of molecules is necessary to make a compartment viable.

(3) The synthesis of macromolecules and the formation or division of compartments have to be synchronized in order to guarantee stationary concentrations of polymers inside the compartment.

(4) The replicating set of macromolecules in the compartment is subject to an error propagation problem similar to that we have discussed before.

Compartments which are able to master problems (1) to (4) would appear to require a rather elaborate molecular organization. At present it is difficult to suggest plausible models for this phase in prebiotic evolution. Based on the arguments presented above we suggest that compartments were incorporated into the evolving system as soon as necessary, but as late as possible (The interested reader is referred to a recent summary: Eigen et al. 1981).

The critical stages of early evolution are summarized in Fig. 11. We use the expression "crisis" to indicate that a new principle of organization had to come into operation in order to overcome an obstacle into which the previous development had run. The various crises represent a logical sequence of steps in prebiotic evolution from a mixture of small molecules to the first prototypes of procaryotic cells. Each of these steps requires the achievements of the previous ones, although two or even more might have been taken at nearly the same time during the historical course of prebiotic evolution. It is necessary to stress one point in order to avoid misunderstanding. Some of the critical steps incorporate new materials into the evolving system. For example, polypeptides came under the control of the replicating system after the information crisis had been overcome, membranes after the problem of genotype-phenotype dichotomy had been solved, and the utilization of sunlight after the primordial energy crisis had been surmounted through the development of the photosynthetic machinery. Emphasis is laid on the formulation "came under control of the replicating system". We indicate thereby that polypeptides, lipids and membranes, energy rich compounds must have existed first as necessary environmental factors and were used as resources before the replicating polynucleotides "learned" to get control of these essential constituents of present day life. This control consisted of the instruction of polypeptide synthesis through translation, of membrane synthesis and cell division, and of the devlopment of diverse primitive metabolic pathways in order to utilize environmental resources.

Acknowledgements. The authors are indebted to Professor W.C. Gardiner for many suggestions and careful reading of the manuscript. The work in Wien has been supported financially by the Fonds zur Förderung der Wissenschaftlichen Forschung (Project no 3502 and 4506).

References

Biebricher CK, Eigen M, Luce R (1981) J Mol Biol 148:369–390 and 391–410

Cedergren RJ, Sankoff D, LaRue B, Grosjean H (1981) CRC-Crit Rev Biochem 11:35–104

Crick FHC, Brenner S, Klug A, Pieczenik G (1976) Orig Life 7: 389–397

Domingo E, Davilla M, Ortin J (1980) Gene 11:333–346

Domingo E, Sabo D, Taniguchi T, Weissmann C (1978) Cell 13: 735–744

Eigen M (1971) Naturwissenschaften 58:465–526

Eigen M, Schuster P (1979) The hypercycle. Springer Verlag, Berlin Heidelberg New York

Eigen M, Winkler-Oswatitsch R (1981) Naturwissenschaften 68: 217–228 and 282–292

Eigen M, Gardiner W, Schuster P, Winkler-Oswatitsch R (1981) Sci Am 244(4):88–118

Fields S, Winter G (1981) Gene 15:207–214

Fox SW (1973) Naturwissenschaften 60:359–368

Hofbauer J, Schuster P, Sigmund K (1981) J Math Biol 11:115–168

Hofbauer J, Schuster P, Sigmund K (1982) Biol Cybern 34: 51–57

Hofbauer J, Schuster P, Sigmund K, Wolff R (1980) SIAM Appl Math C 38:285–304

Jones BL, Enns RH, Ragnekar SS (1976) Bull Math Biol 38: 15–28

Lohrmann R, Bridson PK, Orgel LE (1980) Science 208:1464–1465

Oró J, Miller SL (1981) J Mol Evol 17:263–306

Oró J, Miller SL (1982) J Mol Evol 18:149–220

Ortin J, Najera R, Lopez C, Davilla M, Domingo E (1980) Gene 11:319–331

Palese P (1977) Cell 10:1–10

Schuster P (1981) Prebiotic Evolution. In: Gutfreund H (ed) Biochemical Evolution, Cambridge Univ Press, Cambridge, pp 15–87
Schuster P, Sigmund K, Wolff R (1979) J Diff Equations 32: 357–368
Schuster P, Sigmund K, Hofbauer J, Wolff R (1981) Biol Cybern 40:1–25
Sheperd JCW (1981) Proc Natl Acad Sci USA 78:1596–1600
Sheperd JCW (1981) J Mol Evol 17:94–102
Spiegelman S (1971) Quart Rev Biophys 4:213–253
Sumper M, Luce R (1975) Proc Natl Acad Sci USA 72:162–166
Swetina J, Schuster P (1982) Biophys Chem, in press
Thompson CJ, McBride JL (1974) Math Bioscience 21:127–142
Volkenstein MV (1979) J Theor Biol 80:155–169

Received May 19, 1982/ Revised July 29, 1982