Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 1751–1753, March 1984
Evolution

A model for the origin of biological catalysis

(evolution/genetic code/computer modeling/origin of life/prebiotic)

D. L. STEIN AND P. W. ANDERSON

ABSTRACT **We propose a mathematical model for the next stage in the origin of life after that treated in our earlier work. At this stage we introduce the possibility of the modification of the environment by the information-containing entities and feedback between the environment and the population of macromolecules and hence provide a model for the development of the Eigen hypercycle.**

In earlier work (refs. 1 and 2; unpublished results), a mathematical model of the origin of biological information was proposed and studied using computer simulations. We confined ourselves to the simplest situation, which we anticipated belonged to the "universality class" of all such models: given an initial random "soup" of two types of monomer, A and B, in the presence of a large energy flux, how can the system give rise to information-carrying macromolecules (i.e., long strings of As and Bs with information content)? Information content here can be specified in the simple, standard way (see, for example, ref. 3): assuming *a priori* that either A or B is equally probable at any site on a polymer, then if W_0 is the total number of *possible* strings of length N and W_1 is the number of different kinds of realized *outcomes*, then the information content of the soup of strings of length N is

$$I = \log_2\left(\frac{W_0}{W_1}\right). \qquad [1]$$

Information content of a soup of strings of length N is then maximized if the following two conditions are met (2):

(*i*) **Diversity.** If only a handful of outcomes is likely, the information content of the system is of 0(1) rather than 0(N). That is, we must not be able to predict at initial times what the resulting strings will look like. This is one reason why transitions like the convective instability in fluids are probably poor analogies for the origin of life (4–6). Kolmogorov and Chaitin (see, for example, ref. 7) have given a more sophisticated treatment of problems of complexity and information content. They point out that a good measure of the true "randomness" or "information contained" in a number (or a string of symbols such as a polymer) is the $\log_2$ of the length of the program necessary to specify it. Thus, a string like $(\text{A-A-A} \ldots)_N$ contains no information to order N because it can be specified by a 2- or 3-bit program while a truly random number of length N contains N bits. [C. Bennett's (personal communication) version of "complexity" of an object is also important in biological questions: the time required to construct it from its minimal program. Both of these measures are large for truly living systems.]

(*ii*) **Selection and Stability.** If $0(2^N)$ outcomes are possible and $0(2^N)$ outcomes actually appear, again $I \sim 0(1)$. We require a macroscopic occupation of only a few possible "states" (i.e., polymers) at the end and further require that these final "states" remain macroscopically occupied over many generations: the living system must reproduce itself.

If both diversity and stability are present, it is then reasonable to conclude that we have "broken symmetry in information space." These two requirements have a natural analogy in the spin glass (2), where a system of localized spins with random interactions below a temperature T_f freezes into a random configuration that is highly metastable but is only one of a number of possible final configurations that diverges with the number of spins in the system. Just as the free energy surface in configuration space of a spin glass therefore has many peaks and valleys, it is expected that the complex requirements of RNA chemistry and interactions with the external world should result in a survival probability function that likewise has many peaks and valleys in the space of all polymers. One may therefore map the original problem onto an Ising spin-glass chain (A $\rightarrow +1$, B $\rightarrow -1$) with long-range random interactions (2).

A crucial point is that the method of string replication (effected by "temperature cycling"; ref. 1) ensures in the absence of copying errors exact symmetry between A and B at every site in detail. Therefore, a given string S and its inverse $-S$ (e.g., A-B-B-A-A-B and B-A-A-B-B-A) must have the same survival probability. The simplest possible "death function" $D_N(S)$ that determines the survival probability of a given string S of length N is (2)

$$D_N(S) = \sum_{i>j=1}^{N} J_{ij} S_i S_j, \qquad [2]$$

where the J_{ij}s are fixed and random, taking on the values ± 1 with equal probability. The higher the value of $D_N(S)$, the less likely the survival of string S. Note that our requirement of inversion symmetry precludes a term $\Sigma_i h_i S_i$ in $D_N(S)$, so that the model resembles a spin glass as opposed to a random-field magnet.

Computer runs (unpublished results) of the model showed that it indeed satisfied the requirements of diversity and selection; in every run (lasting about 800 cycles or generations), several hundred species (families of related polymers) were created out of which a few (5–10 or so) comprised of order 40–60% of the soup, depending on the run. We therefore see selection and stability. With the *same* parameters (e.g., J_{ij}) in different runs, different polymers were selected; hence, diversity. Other interesting properties exhibited by the model, such as adaptation, hysteresis, and memory effects, will be discussed elsewhere. We therefore concluded that the proposed model does indeed contain some of the essential properties needed in a model of the transition to biological information. [Incidentally, the results strongly resemble experiments of Biebricher *et al.* (8, 9) on "*de novo*" replication and may explain those.]

We would now like to carry the model a step further. So far we have succeeded in generating RNA-like (and presumably by extension protein, clay, and such) "quasi species" in the sense of Eigen (7) on the one hand and on the other in producing a possible general framework for the much later evolution of species. We would also like to ask whether similar considerations, or an extension of the model at hand, can be used to understand (in the same general way) how a genetic code could come about, in the sense of a macromolecular "blueprint" [DNA mostly today but early on probably RNA (10)] assembling other macromolecules (in our case, enzymes) that then serve to catalyze the formation of a new blueprint. We now wish to understand two problems: (*i*) How may one initiate such a closed, self-perpetuating loop?

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

and (*ii*) Why should such a code be universal? Eigen (8) has dealt with question *ii* in his theory of catalytic hypercycles, but as far as we can determine, no general model yet exists that sheds light on question *i*.

The primary modification to the previous model that we shall make is to note that the J_{ij}s are no longer fixed and external but should now be partially determined by the sequences of monomers (i.e., the S_is) on the strings themselves. (The associative memory model of J. J. Hopfield follows a similar procedure in this respect.) This requirement models a feedback loop mechanism that is meant to simulate, in the simplest possible manner, the feedback loop coupling the genetic code to protein synthesis.

Taking a cue from biology, we note that we need to go beyond specific single-site recognition and feedback. It is the spatial folding of a protein (polymer) that is most important in determining its catalytic abilities, not its sequence per se (although, of course, the latter helps to determine the former). In determining how the S_is determine the J_{ij}s, we therefore need to consider the sequences of monomers over each entire polymer and over all polymers present in the soup.

It would also be surprising if a successful feedback loop were to be established in every run. Because of the delicate interplay needed among a number of factors, it is reasonable to expect that a given population in some soup has a high probability of dying out before a successful loop can be established. On the early earth, one can imagine many attempts at life in many different regions before success occurred. We want a model in which success is not preordained in any given try, though the probability of *a* success over many tries approaches one.

We see every reason to believe that RNA molecules themselves may possess catalytic activity as a consequence of their rather random ternary structure. It may be that the first hypercycles did not involve proteins per se and that polypeptides were brought in only later as "passengers" of some sort on the RNA catalysts. The mechanism described below is neutral on this point.

With these considerations in mind, we propose the following model. (*i*) We start out as before with a death function

$$D_N^{(1)}(S) = \sum_{i>j=1}^{N} J_{ij}^{(1)} S_i S_j, \qquad [3]$$

where the $J_{ij}^{(1)}$s are fixed. This determines which strands are likely to survive because of their internal chemistry and interactions with the environment, as before (ref. 2; unpublished results).

(*ii*) Some sequences will be better able to act as templates for the synthesis of proteins or catalytic RNA than others. To keep things simple, we assume there are two amino acids (C and D) and describe the configuration of the proteins by a spin vector $\mathbf{M}_i$, with C = +1 and D = −1. Because of such complexities in RNA chemistry as chains folding back on themselves, self-entanglement, and so on (2), not all chains can act as templates for protein synthesis. Once again, we can only take these factors into account in a random, statistical fashion. We introduce a second set of couplings, $J_{ij}^{(2)}$, independent of the first, such that the number

$$P_N(\mathbf{S}, \mathbf{M}) = \sum_{i=1}^{M}\sum_{j=1}^{N} J_{ij}^{(2)} S_i M_j \qquad [4]$$

determines the probability of a given chain giving rise to a polypeptide by template synthesis.

The rule for generating catalytic sequences is similar to that for determining surviving species in step *i*. The probability per unit cycle of a given *RNA* chain $S^{(\alpha)}$ giving rise to a catalytic chain $M^{(\alpha)}$ (or of acting as a catalyst itself) is given by

$$d^{(\alpha)} = \frac{e^{[P_N(S^{(\alpha)}) + \mu(N)]}}{1 + e^{[P_N(S^{(\alpha)}) + \mu(N)]}} \qquad [5]$$

where we have introduced a "chemical potential" $\mu(N)$ determining the overall likelihood of a chain of length N giving rise to catalysis (this will be different from a similar chemical potential used in step *i* and in fact we will adjust μ to make catalysts relatively rare. In each cycle, $d^{(\alpha)}$ determines the probability of a chain $M^{(\alpha)}$ arising. After each cycle, we will then find some collection of $M^{(\alpha)}$s.

(*iii*) We now consider the assembly of enzymes formed in a given cycle acting back to catalyze formation of new nucleic acids. Once again, there should be randomicity in catalysts acting back on templates, and so we introduce a third set of fixed random couplings $J_{ij}^{(3)}$. Let α denote an entire protein. (S_i denotes the nucleotides A or B.) We then propose a second death function

$$D_N^{(2)} = \sum_{i,j}\sum_{\alpha} J_{ij}^{(3)} M_i^{(\alpha)} M_j^{(\alpha)} S_i S_j \qquad [6]$$

for nucleic acids. Note that $J_{ij}^{(3)}$ implies that an enzyme synthesized by a given RNA strand may catalyze *all* strands in different ways. The sum over α implies that all enzymes created are involved in catalysis, to the benefit of some nucleic acids and the detriment of others. There is a high degree of frustration (in the spin-glass sense) present in the feedback. Moreover, the effective couplings in $D_N^{(2)}$

$$J_{ij}^{(\text{eff})} = \sum_{\alpha} J_{ij}^{(3)} M_i^{(\alpha)} M_j^{(\alpha)} \qquad [7]$$

are not constant but evolve in time according to changes in the population of the nucleotide soup.

The algorithm we propose is similar to that to be used elsewhere with the addition of a slow generation of "enzymes" according to the probability (3) and the application of $D_N^{(2)}$ to the soup at the end of every cycle (so that $D_N^{(1)}$ is applied first). $D_N^{(1)}$ governs which species are suitable for replication at all, as discussed in ref. 2. $D_N^{(2)}$ then further selects among surviving strands, corresponding to the fact that some species are catalyzed more efficiently by the protein soup than others.

There is still some choice to be made about the order of events in this picture. One possibility is to use the algorithm described here from the beginning. A second is to use our unpublished algorithm for a hundred generations or so until some RNA quasi species are built up and then apply the new algorithm. A third possibility is to have a noninteracting coevolution of nucleic acids and proteins using our unpublished scheme for both (with different J_{ij}s for each, of course) and then turn on interactions in the manner prescribed here. Since so little is known about the order of appearance of nucleic acids and enzymes, all three possibilities should be examined on the computer.

The model presented here has complicated feedback properties and surely contains cooperativity as well as competition. Further, if $D_N^{(1)}$ and $D_N^{(2)}$ do not overlap, all species will die out. In the model presented here, the minima of $D_N^{(2)}$ need to find and evolve toward the minima of $D_N^{(1)}$. Overlap can be difficult to find. If this evolution cannot take place sufficiently quickly, the entire population in the soup will die out. We therefore can adjust parameters so that only a small percentage (depending on the parameters used) of runs succeed in establishing stable species coupled in a closed loop.

We acknowledge valuable discussions with J. J. Hopfield and F. E. Yates. The ideas here expressed developed in discussions at the Aspen Center for Physics (Aspen, CO), summer 1983. The work at Princeton University was supported in part by National Science Foundation Grant DMR 8020263.

1. Anderson, P. W. & Stein, D. L. (1984) in *Self-Organizing Systems: The Emergence of Order*, ed. Yates, F. E. (Plenum, New York), in press.
2. Anderson, P. W. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3386–3390.
3. Brillouin, L. (1962) *Science and Information Theory* (Academic, New York).
4. Landauer, R. (1984) in *Self-Organizing Systems: The Emergence of Order*, ed. Yates, F. E. (Plenum, New York), in press.
5. Anderson, P. W. (1981) in *Order and Fluctuations in Equilibrium and Nonequilibrium Statistical Mechanics*, eds. Nicolis, G., Dewel, G. & Turner, J. W. (Wiley, New York), pp. 289–297.
6. Stein, D. L. (1980) *J. Chem. Phys.* **72,** 2869–2874.
7. Kolmogorov, A. N. & Chaitin, G. J. (1982) *Int. J. Theor. Phys.* **21,** 941–954.
8. Biebricher, C. K., Luce, R. & Eigen, M. (1981) *J. Mol. Biol.* **148,** 369–390.
9. Biebricher, C. K., Luce, R. & Eigen, M. (1981) *J. Mol. Biol.* **148,** 391–410.
10. Miller, S. & Orgel, L. E. (1974) *Origins of Life on Earth* (Prentice-Hall, Englewood Cliffs, NJ).