J Mol Evol (1986) 23:119–126

Journal of Molecular Evolution

Self-Organization in Prebiological Systems: Simulations of a Model for the Origin of Genetic Information

D.S. Rokhsar,* P.W. Anderson, and D.L. Stein

Princeton University, Princeton, New Jersey 08544, USA

Summary. Computer simulations of a "spin glass" model for the origin of biological information are discussed. Selection is found to occur among a wide diversity of possible species, and in addition competition, adaptation, and hysteresis are all exhibited.

Key words: Molecular evolution — Spin glass

Introduction

Any attempt to understand the origin of life on Earth must address the problem of the origin of features that are universally shared by all modern organisms, for example, the genetic code, protein synthesis, and primitive metabolic pathways such as glycolysis. There are several levels on which such questions may be asked. We would ultimately like to find a plausible series of chemical steps that would lead to, for example, a self-replicating genetic code. Such an ambitious program must address questions such as the original molecular representation of the code (nucleotides? amino acids? inorganic ions in clays?) and the origin of the modern, tightly closed DNA→RNA→enzyme→DNA loop. Eigen (1971) has constructed a model in which proteins and RNA coevolve in a "hypercycle" that requires at the start highly developed "quasispecies" of polypeptide and polynucleotide chains perhaps already containing a great deal of biological information.

Offprint requests to: D.L. Stein

* *Present address:* Cornell University, Ithaca, New York 14853, USA

In this article, we shall be interested in the general question of how such information may originate. We would like to understand in principle how, given a collection of related molecules (for example, guanine and cytosine) and random events, biologically useful information can spontaneously arise in the form of molecular sequences along macromolecular strings. We do not worry about the initial molecular representation of such strings [for further discussion on this point, see Anderson (1983) and Stein (1984)]. We ask how, given "letters of the alphabet" that are available in an open system subject to an external energy flux, might the letters organize into "words"? What are the general principles involved in a transition from little to much biological information?

One of us (Anderson 1983) has constructed a simple mathematical model that attempts to describe the essential features of such a transition, features that may be called "universal" in the sense that a large number of specific biochemical models leading to information growth would exhibit them, independently of chemical details.

To understand the meaning of this model, we must define what we mean by a transition to biological information. Start with a soup of monomers (call them A and B) that interact at random and gradually form long strings of As and Bs. To determine whether the evolving soup of polymer strings contains information, we use the simplest measure of information content due to Shannon. There are 2^N possible strings of length N; if all are equally probable then the information content of the soup is

$$I = \log_2(W_0/W_1) \qquad (1)$$

where W_0 is the number of possible outcomes (2^N

in our case) and W_1 is the number of realized outcomes. Hence, the key to high information content is the presence of both diversity (or no predetermined outcomes), to maximize W_0, and selection (only a few of many possible outcomes realized), to minimize W_1 (see Anderson 1983). We found earlier (Anderson and Stein 1983) that it is not overly difficult to find a model that exhibits either diversity or selection; the problem is finding one that does both. Such a search may lead to the general principle that lies behind the "universality class" of transitions to biological information.

An early model of ours used "temperature cycling" to build up long strings from A and B monomers (for details, see Anderson and Stein 1983; also Blum 1962; Usher 1977; Kuhn and Kuhn 1978). In that model, longer strings served as templates for building copies of themselves, using base-pairing recognition and strong temperature gradients as a driving force. However, we know that nature must differentiate among strings, selecting some and discarding others. We cannot specify ahead of time how this might occur, but we do have evidence that a large variety of strongly competing factors influence the birth, growth, and death rates of strings; in a general mathematical model we can account for these factors in a random, statistical way [for a list of some of the chemical and physical factors that led us to this approach, see Anderson (1983)]. In essence, the complex chemistry of any macromolecular species capable of storing information will, we believe, necessarily give rise to a string survival probability that can be mathematically modeled as long-range, random interactions between monomers on the chain. This in turn causes many peaks and valleys in the string survival probability in the space of all polymers, giving an appearance reminiscent of the energy surface of a statistical-mechanical system, namely the spin glass.

A spin glass is a system of spins S_i localized at sites i that interact via quenched (i.e., frozen), random interactions; the Hamiltonian is

$$H = -\sum_{i,j} J_{ij} S_i S_j \qquad (2)$$

where the J_{ij} are random variables that can take on both positive and negative values. A key property of a spin glass is *frustration,* illustrated in Fig. 1. If the product of the J_{ij}'s around any closed loop is negative, not all "bonds" (i.e., J_{ij}'s) can be satisfied. Spin glasses are highly frustrated, and as a result possess a large number of inequivalent spin configurations that are metastable (i.e., lie at local minima of the free energy). The number of such minima diverges as $e^{\alpha N}$, $\alpha \approx 0.2$, for a system of N spins (we assume each S_i can take on only the values ± 1).

As a result, a spin glass has both diversity (many possible ground states) and stability (barriers preventing escape from a single state). Our problem can then be mapped onto the following spin-glass problem (Anderson 1983): Treat an "A" monomer as an "up" spin ($S_i = +1$) and a "B" monomer as a "down" spin ($S_i = -1$), and write the probability function for individual-chain (non)survival as

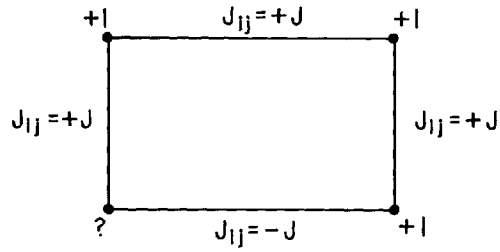

Fig. 1. Frustration in a spin glass on a two-dimensional square lattice. The spins sit on lattice points at the corners of the square shown. The lines connecting the spins (bonds) represent J_{ij}'s. Three are ferromagnetic (+J) and one antiferromagnetic (−J) in the picture shown. If we let the spin at the upper left point up ($S_i = +1$), then so will the next two spins (going around the square clockwise). However, we cannot specify whether the spin at lower left is up or down, since it is receiving conflicting signals from its two nearest neighbors. The square is "frustrated"—it has no unique ground state

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

for a given polymer sequence α of length N. We stress that the random variables J_{ij} do not correspond to specific chemical parameters of a certain chemical species; rather, they reflect what we believe to be a universal requirement of any chemical species capable of storing information—namely that its chemical complexity give rise to frustrated interactions among the monomers.

The "death function" of Eq. (3) may vary in value from $-\infty$ to $+\infty$ as $N \to \infty$; we now convert this to a rate, namely the probability per unit cycle that a given chain belonging to monomer sequence $S^{(\alpha)}$ will survive or die. This rate, $d_N{}^{(\alpha)}$, will vary from zero to one. Any smooth, monotonic mapping from the line ($D_N{}^{(\alpha)}$) to the unit interval ($d_N{}^{(\alpha)}$) should preserve the essential qualities of the death function and leave our model within the same universality class. We therefore choose

$$d_N{}^{(\alpha)} = \frac{\exp[D_N{}^{(\alpha)} + \mu(N)]}{1 + \exp[D_N{}^{(\alpha)} + \mu(N)]} \qquad (4)$$

where we introduce the function $\mu(N)$, which determines the overall death rates of *all* strings of length N in the soup.

The present paper reports the results of computer simulations of the above model.

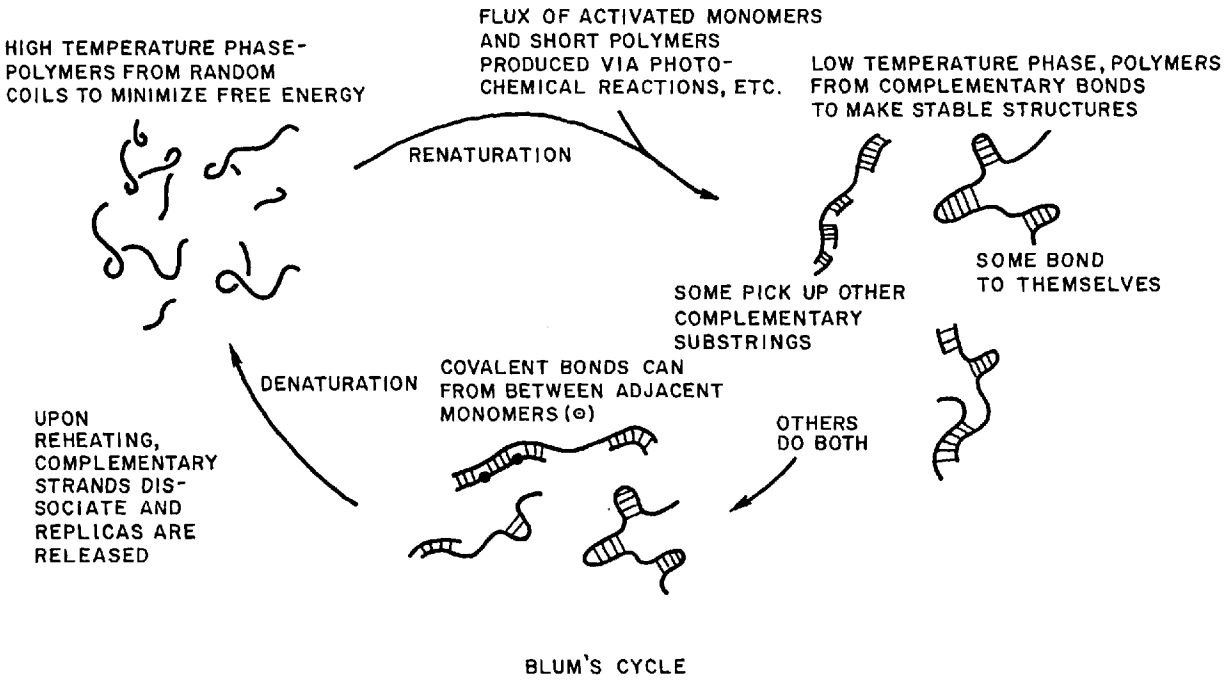

Fig. 2. Temperature cycling for building up longer polymer chains, after Blum (1962)

Method of Simulation

We begin the simulation with a small number of dimers and trimers from which we choose one type acting as a template for repliciation. This "soup" of small polymers contains equal numbers of each type of dimer and trimer to preclude any bias toward particular strings, which might later appear to have been "selected." At the beginning of each cycle (or "generation"), a flux of both A and B monomers is added to provide raw materials for growth and replication. The polymer chosen as template is compared with the remaining polymers one by one to determine if they are sufficiently complementary (i.e., for each A or B on the template string, a corresponding B or A should appear on the complement string) to form a double-stranded complex. With a probability dependent on the degree of complementarity, the two polymers stick together. Another possible complement is then selected from the soup and tested for complementarity with the part of the template that is not already matched with the first complement. This procedure continues until a certain number of consecutive unsuccessful attempts have been made to find a complement, or until the template has been completely matched. The template plus complements complex is then set aside, and a new template is chosen at random from the remaining single-stranded polymers. This process is repeated until all the strings have been used or until a few templates in a row have not found any complements. The process corresponds to the renaturation shown in Fig. 2.

To determine whether or not a string is sufficiently complementary to the template to form a double helix, the two strings are placed parallel to each other, shifted by a random amount. Each weak, complementary bond is assigned a probability p(A–A), p(A–B), or p(B–B) that describes the accuracy of recognition. For the probability that the two strings will match, we use the product of the probabilities for each weak bond. An error corresponds to a non-zero value of p(A–A) and/or p(B–B). When the complementary strand hangs over the end of the template, i.e., when some monomers in the complement are not matched with monomers in the template, a factor of p(A–end) or p(B–end) is included for each unmatched monomer. Physically, this accounts crudely for the different behavior of the ends of polynucleotides, which can be bound to ions, proteins, etc. that can interfere with the binding of complement to template. Practically, p(x–end) controls the lengthening rate of polymers; if p(x–end) is zero, the complementary copy cannot be longer than the template, and no growth can occur. To decide whether or not a complement will bond to a template, a computer-generated "random" number chosen from a uniform distribution between 0 and 1 is compared with the total probability of bonding; if the random number is less than this probability, the polymers form a double-stranded complex.

After the "renaturation" part of the cycle each complex is examined for juxtaposed complements; these can then form strong covalent bonds to combine two short strings into a longer one. This process is governed by the probabilities p(A=A), p(A=B),

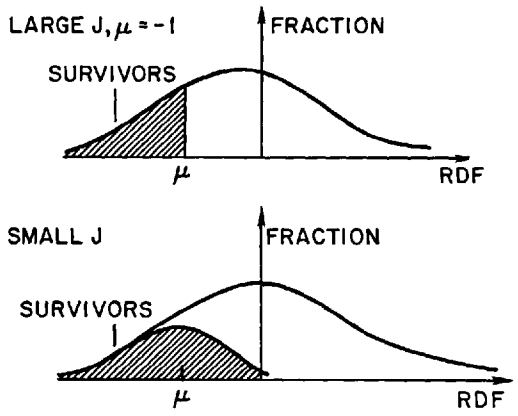

Fig. 3. The sharpness of the cutoff between surviving species and nonsurviving ones depends on the magnitude of J, which in some sense acts as an inverse temperature. The graph shown is the product of the density of states with the survival probability

and p(B=B). For example, in the complex

```
complement  A=A=A=B=B=A  B=A=B=A    A=B=A
            | | | | | | | | | |    | | |
template    B=B=B=A=B=A=B=A=B=A=B=A=B=A=B
```

which contains an "error" at the third site of the template, a covalent bond can form between the sixth and seventh monomers on the complement side. The complement of length three cannot be joined to the others, since it is not adjacent to another complement. Note that for small "double-bonding" probabilities, more subreplicas of the template are produced, while for large probabilities, a more complete complementary copy can be made. After the covalent bonds have formed, the weak bonds are broken (the "denaturation" part of the cycle) and we are left with a new soup of single-stranded polymers.

The application of the death function $D_N(S)$ occurs immediately after the denaturation step; $D_N(S)$ and the corresponding death probability are computed for each polymer as it is peeled off its template–complement complex. Again, a random number between 0 and 1 is generated to determine the fate of the polymer. The "death" of an RNA strand can be envisioned in many ways—decay into monomers; breakage; reproductive "death," in which the polymer is chemically altered to prohibit future pairing; or simply the removal of the polymer from the soup, perhaps by adsorption onto a rock; etc. To prevent a proliferation of polymers that would rapidly exceed the computer memory, we choose the latter, and discard all "dead" strings. Once all paired complexes have been denatured and checked for survival, the new soup of single strands is used to repeat the cycle.

The death probability is given by the Fermi–Dirac-like law of Eq. (4) but with $D_N^{(\alpha)}$ normalized to a Gaussian distribution of unit standard deviation:

$$D_N^{(\alpha)} \to \frac{D_N^{(\alpha)}}{J\sqrt{N(N-1)/2}}$$

The reason for this is that the distribution of possible values of $D_N^{(\alpha)}$ for strings of length N is a Gaussian of width $J\sqrt{N(N-1)/2}$. The normalization employed here removes the length dependence of the death rate.

The death process is controlled by the two parameters J and μ. The "chemical potential," μ, prescribes the value of the reduced death function (RDF) at which a polymer has a 50% chance of survival. The coupling constant, J, behaves like an inverse temperature, and describes the severity of the death function. For large J (zero temperature) the cutoff at μ is sharp, and all polymers S such that $d(S) < \mu$ survive, whereas all others die. Reducing the coupling constant results in a more "forgiving" environment where a low value of the death function becomes less important for survival (Fig. 3).

In all results reported below, we have maintained the complete dynamical symmetry between the two types of monomers by choosing p(A–A) = p(B–B) and p(A=A) = p(B=B) [$D_N(\vec{S}) = D_N(-\vec{S})$ by construction (Anderson 1983), so death is symmetric]. If one begins with an initial soup that also preserves this symmetry, one expects that two complementary strings S and −S will be produced at an identical rate, so their concentrations should be equal. For computer simulations with complementary kinetics, i.e., with p(A–A) small and p(A–B) equal to unity, this is indeed the case, with $\vec{S}$ and $-\vec{S}$ appearing equally to well within $\pm\sqrt{N}$, as expected for counting errors. Adding a death function with various μ and J does not alter this equivalence.

To reproduce itself, a string $\vec{S}$ must first make a complementary copy −S to be used as a template for production of more $\vec{S}$. Under optimal conditions of perfect recognition [p(A–A) = 0, p(A–B) = 1] and perfect polymerization [p(A=A) = p(A=B) = 1], this process can take two generations. Allowing "mutations" or errors will slow this process down, as will the possibility of producing S by linking fragments made by other polymers. An excess of $\vec{S}$ over $-\vec{S}$ will result in the production of more $-\vec{S}$, tending to restore the equality of the concentrations of the two polymers. There is a built-in resistance to breaking the symmetry between $\vec{S}$ and $-\vec{S}$.

Having demonstrated that our computer simulation with complementary pairing (like that observed in nature) does preserve the equality of $\vec{S}$ and $-\vec{S}$, even for error rates [p(A–A) = p(B=B)] on the order of 10–15%, it is convenient to eliminate the intermediate step of producing a complementary copy to reproduce $\vec{S}$. This is useful for two reasons: First, the growth rate is approximately doubled, and second, the data analysis is simplified by eliminating

the need for $\vec{S}$ and $-\vec{S}$ to be considered part of the same "species," as defined below. Note that the restoring force of complementary kinetics is now absent, since $\vec{S}$ and $-\vec{S}$ are no longer dependent on each other for reproduction. They behave as independent polymers, competing for the same raw materials, and need not be present in the same quantities. As will be seen below, their concentrations are in fact rarely comparable: One of the two dominates if either of them exists. For the intermediate case of random reproduction (no recognition; any string can be a template for any other) the concentrations of $\vec{S}$ and $-\vec{S}$ are completely random and fluctuate in time.

In what follows, we shall always be considering self-replicating kinetics unless otherwise specified, concentrating on the competition between polymers that are not mirror images of each other. Our concern is the possibility of selection of prebiological information, so we choose the simplest member of the "universality class" of these models to study. Note that comparisons with complementary kinetics data have shown that self-replicating runs show the same properties as complementary runs with the replacement of a pair $(\vec{S}, -\vec{S})$ by $(\vec{S})$ or $(-\vec{S})$. Furthermore, the symmetry of $\vec{S}$ and $-\vec{S}$, which led us to our choice of $D_N(\vec{S})$, is still preserved.

Even in the absence of a death function, one cannot say a priori what results the above model will yield. A template will often replicate only fragments of itself; these fragments can in turn reproduce wholly or in part, can be joined together to recreate the original template, or might join in a different order with other polymers to form a string completely different from the original. There is a great deal of competition for small fragments, which have not yet "specialized" and can be used to replicate many different templates; longer strings are more restricted in their choice of a template to bind to, but can also be used as templates themselves.

Species and the Death Function

The addition of a death function to our procedure provides the model with growth, differentiation, and stability. The cases we investigated most thoroughly correspond to $\mu \approx -0.7$ to -10.0 and various J, so that about 10–15% of all polymers survive (see shaded area in Fig. 3). Each run is dominated by a few species after approximately 100 generations. These dominant sequences differ from run to run under identical initial conditions, exhibiting the diversity needed for life. Furthermore, the sequences were stable and persisted (see below) for the length of most runs (~500–600 generations). If we observe the relative concentration (mole fraction) of any individual sequence, however, in almost all cases it peaks and then decays, with the time of occurrence of the peak dependent on the length of the sequence. As time progresses, the average sequence length of the population grows (see Fig. 4) and sequences are superseded by longer versions of themselves. It is therefore natural to consider polymers with similar sequences and subsequences as part of the same "species," as defined below.

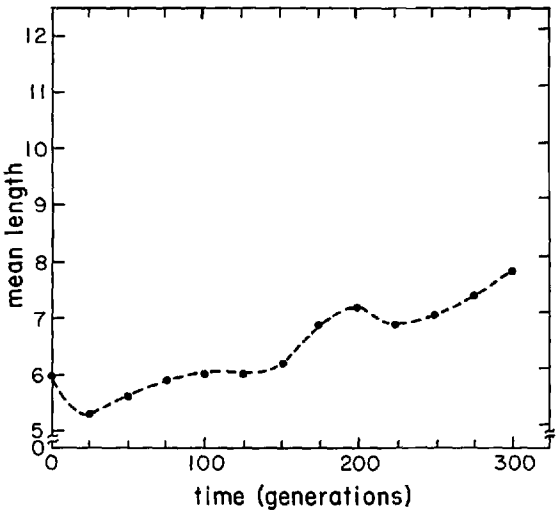

Fig. 4. As time increases, the mean length of surviving polymer species also increases

A species is constructed by choosing a generation and considering the sequences of polymers of average length. The soup is then searched for all smaller strings ("ancestors") and large strings ("descendants") whose sequences match that of the respective average-length string wherever comparison is possible. Small (length $\lesssim 5$) polymers are not included in the species, since these short polymers are ancestors to many species and can be considered "undifferentiated" materials. In the presence of mutations, ancestors and descendants that differ only by one monomer from the prototype are included. In practice, there is very little overlap between species, because the death function prohibits many sequences, thereby separating species in information space (which is quite large). The biological picture of a species is all those polymers that contribute to the propagation of the prototype. Each "valley" of the death function corresponds to one species.

The diversity required of prebiotic systems is manifested in the selection of a few species that dominate each run; which species is chosen varies from run to run. When two sequences of the same length and death function are placed in direct competition by "seeding" the initial soup, one species rapidly dominates the other (Fig. 5). In these direct-competition simulations we have not yet observed the development of a "symbiotic" relationship be-

tween the polymers, i.e., one in which they coexist, although this happens often when the initial conditions are not "seeded." Under seeded conditions, the survival of a sequence depends strongly on the structure of that particular valley of d(S) that lies below $\sim\mu$. If the valley is broad, the seeded polymer will immediately produce many ancestors, giving it an advantage over the narrower-valley sequence, which can produce only a few types of ancestors. Without seeding, longer ancestors are produced slowly, via the undifferentiated small polymers that are ancestors to all species. It is plausible that the large concentration of small polymers is more favorable to symbiosis.

Lengthening of a polymer sequence corresponds to the increasing "specificity" of an organism—while there are more ancestors, it is more difficult to put them together in the correct way. It is possible that, on lengthening of a given sequence, the new polymer produced will have "crawled out" of the minimum; i.e., d(S) may then be greater than $\sim\mu$. The tendency toward increasing length [which can, of course, be tempered by lowering p(A=A), etc., and p(x–end)] pressures species into being able to increase in complexity and "develop" to keep up with the others. Theoretically, one valley could branch into two species, each better adapted to some niche in the environment [i.e., some valley of d(S)]. Presumably this would require runs considerably longer than those we have made, which are limited to mean lengths of 15–20 monomers. Mutation rates of $\lesssim 15\%$ do not destroy the effects of competition and the subsequent selection of species.

Evolution and Adaptation

The previous section showed that competition for nutrients can lead to the coexistence of a few species. The representative sequence of each species steadily lengthens in time as the species searches for the optimal sequence for the given environmental conditions (i.e., given J_{ij}). This increase in complexity to make the best use of the environment is one prerequisite for evolution. A changing environment requires a population that can adapt to new conditions. This variation of the environment was simulated by allowing the σ_{ij} to vary in time (recall that $J_{ij} = J\sigma_{ij}$). Two random sets $\sigma_{ij}^{(1)}$ and $\sigma_{ij}^{(2)}$ were chosen, and $\sigma_{ij}(t) + [1 - a(t)]\sigma_{ij}^{(2)}$, where

$$a(t) = \begin{cases} 1 & t < t_c \\ 1 - (t - t_c)/t_v & t_c < t < t_c + t_v \\ 0 & t_c + t_v < t \end{cases}$$

Also, periodic a(t)s have been used, in which $\sigma_{ij}(t)$ oscillates between $\sigma_{ij}^{(1)}$ and $\sigma_{ij}^{(2)}$ every few hundred generations.

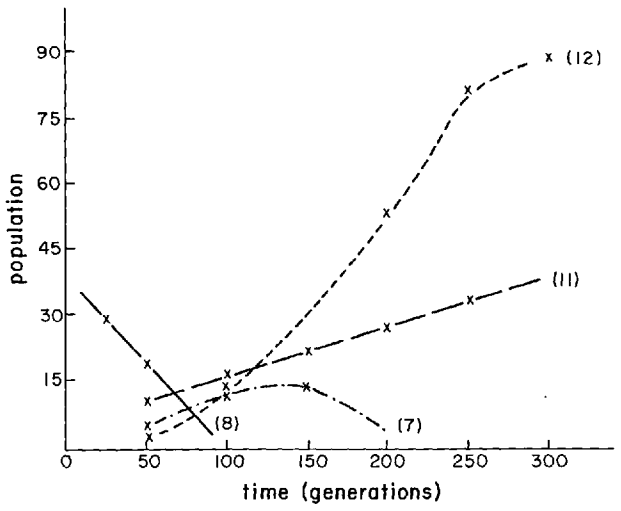

Fig. 5. A sample run showing the fates of different selected related species. The numbers in parentheses represent the lengths of the strands chosen

One result of a time-varying σ_{ij} is shown in Fig. 6a with a mutation rate (see below). As usual, initial growth under $\sigma^{(1)}$ leads to the selection of a few species (A–C). As the environment changes to $\sigma^{(2)}$, one of the species cannot survive (A), since this minimum of $d^{(1)}(S)$ does not correspond to a minimum of $d^{(2)}(S)$. Another species (B), present in trace amounts before the change, does coincide with a minimum of $d^{(2)}$, and dominates after $t_c + t_v$. The most interesting example of true adaptation, however, is species C, which suffers initially, but then mutates to align itself with a more favorable valley of $d^{(2)}(S)$. This alteration of C can be observed by studying the individual sequences in detail. The mutation allows the species to "adapt." Note that "species" here means the same sequence to within one error.

Periodic environmental changes also lead to interesting phenomena (Fig. 6b). Beginning at generation 200, $\sigma(t)$ alternated between $\sigma^{(1)}$ and $\sigma^{(2)}$, changing every 150 generations. Species 2 thrives on $\sigma^{(1)}$ but dies under $\sigma^{(2)}$. Polymer 1 grows modestly on $\sigma^{(1)}$ and takes over on $\sigma^{(2)}$. When $\sigma^{(1)}$ conditions are restored, species 2, which we know can dominate under these conditions, is not restored to its previous concentration. Species 1, having been in the very favorable environment $\sigma^{(2)}$, can now grow in $\sigma^{(1)}$, in which it was not overly successful in the previous cycle. Being suited to growth in $\sigma^{(1)}$ and $\sigma^{(2)}$, it continues to grow in the periodic conditions. The sharpness of the transition from $\sigma^{(1)}$ to $\sigma^{(2)}$, modeled by t_v, does not qualitatively change this behavior, although a sharp transition (small t_v) does cause a more severe depletion of the soup, which "prunes" the family trees of the species and lets them grow anew. A small t_v allows all species to start off relatively even, while a large t_v favors those species that

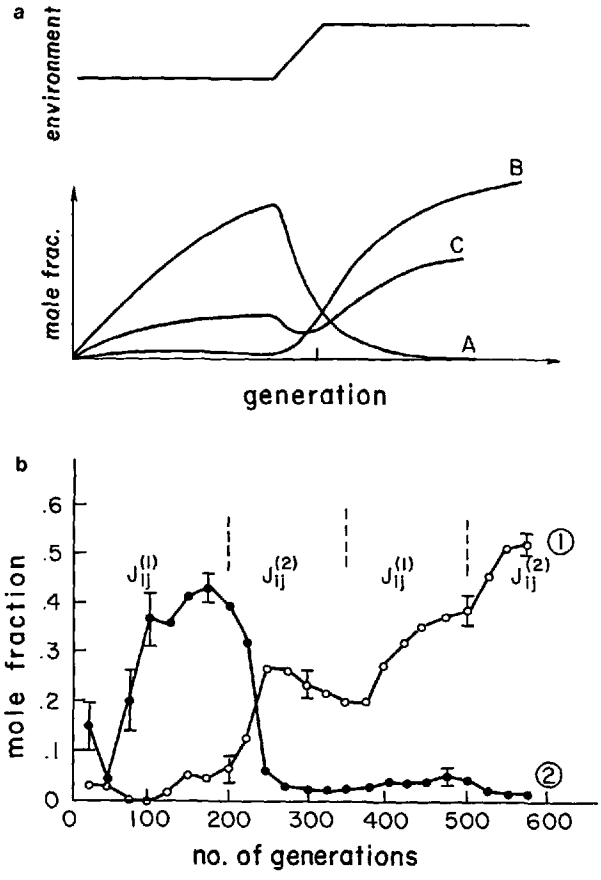

Fig. 6a, b. **a** Adaptation. The environment (i.e., the J_{ij}'s) was changed during the course of 100 generations. In this time, species A, which had been doing well, died out; species B, which had been doing poorly, began to grow; and species C began to die out, but then "mutated" (i.e., a replication error was made), and the new species, which had found a new minimum of the death function, began to grow. **b** Changing environment. The dashed vertical lines at 200, 350, and 500 generations represent environmental changes; i.e., different J_{ij}'s. Polymer 1 responded to the first change and thrived. Polymer 2 could not adapt, and died off. For this run $|J| = 2.0$, $m = -0.7$, error rate = 0.025. Error bars include noise

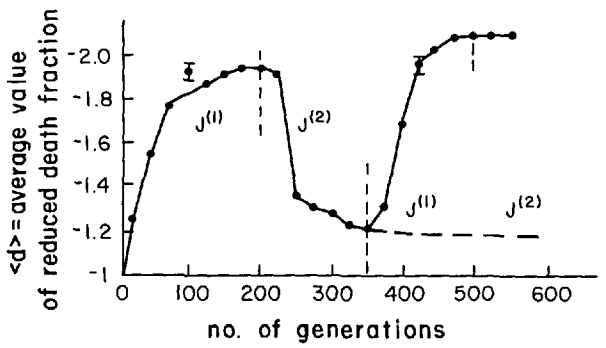

Fig. 7. Hysteresis. As in Fig. 6b, the environment suddenly changes at 200, 350, and 500 generations. Here we plot the death function average $\langle d \rangle$ with $D_N = J_{ij}^{(1)} S_i S_j$. That the polymer "remembers" that it used to have $J_{ij} = J_{ij}^{(1)}$ is shown by the fact that $\langle d \rangle$ does not fall to zero on resumption of $J_{ij}^{(1)}$

can be assembled from the family trees of other species that are not able to survive in $\sigma^{(2)}$. The growth of these species on $\sigma^{(1)}$ provides raw materials to fuel new species when the change comes.

Not only do certain species adapt to the new conditions, but in some sense the entire population "remembers" that it came from $\sigma^{(1)}$ conditions, and can therefore adapt faster when the $\sigma^{(1)}$ environment returns. This is demonstrated in Fig. 7. The ordinate is the average RDF $d^{(1)}(S)$ for the soup obtained using $\sigma_{ij}^{(1)}$. In the $\sigma^{(2)}$ time period, one would expect the totally independent death function $d^{(2)}(S)$ to govern the population, so that $d^{(1)} \to 0$. This is not the case; the population remembers its past environment, and adapts to it even better when it returns (compare the situation at 550 generations with that at 250). Changing initial conditions have in this case "shaken up" the soup, so that it lands in better minima than had been previously selected.

This "hysteresis," or dependence on past conditions, reminds one of the human appendix—selection does not always eliminate vestigial features. It is interesting to consider these data from the viewpoint of Hopfield's memory model (Hopfield 1982). The population has been "taught" to survive in $\sigma^{(1)}$, and when it learns to survive in $\sigma^{(2)}$ it does not forget what it had learned before. Unlike in Hopfield's model, the "genetic" memory is not stored in the J_{ij}.

Conclusion

We have demonstrated that the model described above, a self-replicating collection of polymers whose value or viability in a given environment is described by a chaotic, frustrated death function, can exhibit the required properties of growth, differentiation (diversity), and adaptation. These properties are relatively independent of the parameters of the model—a mutation rate $\lesssim 15\%$, a survival rate of ~10–20% (chosen by μ,J)—over a wide range of lengthening conditions. These bounds represent only the regions investigated, not actual limits. Furthermore, the absence of value does not lead to all of these properties. The magnitude of $\langle d(s) \rangle$ is directly related to value (in some sense it is the average value of the system). Like an order parameter, it describes the organization of the system but not which species are chosen; these are the arbitrary "phases" of the order parameter. External fluxes of large polymers that seed the soup act as the generalized forces in this problem, since they can influence the phase.

Present results indicate that the above model is a reasonable scenario for the emergence of valuable information in a prebiological context. In the model

we have fixed the definition of value by choosing the J_{ij}'s in advance, or allowing them a simple time variation, in order to examine more carefully the effects of a clearly defined value and how this value can cause a collection of polymers to order themselves. As mentioned above, value is determined only with respect to a "trigger system"—in our case the nucleic acid soup itself, which can be "triggered" to select a polymer species. A more complex and perhaps more realistic model would allow the soup itself to define value, and later would account for the influence of polypeptides, which is not required for our model.

Authors' Note

After the above was completed, Dr. Lloyd Demetrius called our attention to the papers on de novo replication of RNA by Biebricher et al. (1981a, b). In these papers, the essential characteristics of our model seem to be embodied in an actual biochemical experiment, although the mechanism is different. We consider it remarkable verification of our results that in essence the nature of their resulting RNA polymers is essentially identical with the nature of ours. Specifically, the outcomes of repeated experiments are similar but not identical to each other and they show the same tendency for a few quasispecies to develop. Thus, in essence their experiments are a confirmation of our model's biological relevance and at the same time a beautiful model of the beginnings of evolution.

Acknowledgment. The work of P.W. Anderson and D.L. Stein was supported in part by NSF grant DMR 8020263.

References

Anderson PW (1983) Suggested model for prebiotic evolution: the use of chaos. Proc Natl Acad Sci USA 80:3386–3390

Anderson PW, Stein DL (1983) Broken symmetries, dissipative systems, emergent properties and life. In: Anderson PW (ed) Basic notions of condensed matter physics. Benjamin/Cummings, Menlo Park, California pp 262–275

Biebricher CK, Eigen M, Luce R (1981a) Product analysis of RNA generated de novo by QB-replicase. J Mol Biol 143: 369–390

Biebricher CK, Eigen M, Luce R (1981b) Kinetic analysis of template-instructed and de novo RNA synthesis by QB-replicase. J Mol Biol 143:391–410

Blum HF (1962) On the origin and evolution of living machines. Am Sci 49:474–501

Eigen M (1971) Self-organization of matter and the evolution of biological macromolecules. Naturwissenschaften 58:465–523

Hopfield JJ (1982) Proc Natl Acad Sci USA 79:2554–2558

Kuhn H, Kuhn C (1978) Evolution of a genetic code simulated with the computer. Orig Life 9:137–150

Stein DL (1984) A model for the origin of biological information. Int J Quant Chem: Quant Biol Symp 11:73–86

Stein DL, Anderson PW (1983) A model for the origin of biological catalysis. Proc Natl Acad Sci USA 81:1751–1753

Usher DA (1977) Early chemical evolution of nucleic acids: a theoretical model. Science 196:311–313

Received April 23, 1985/Revised and accepted January 31, 1986