Richard J. Bagley, J. Doyne Farmer, and Walter Fontana
Complex Systems Group, Theoretical Division, and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545 and Santa Fe Institute, 1120 Canyon Road, Santa Fe, NM 87501

Evolution of a Metabolism

We demonstrate that when stochastic effects are taken into account, over long periods of time autocatalytic metabolisms evolve through a series of punctuated equilibria. We outline a rigorous theoretical treatment of the dynamics of autocatalytic metabolisms, and present some heuristic numerical simulations. We develop the analogy between the evolution of autocatalytic metabolisms and that of contemporary organisms, and argue that while the essential properties of variation and selection are satisfied, there are nonetheless some intriguing differences that merit further study.

1. INTRODUCTION

The word "evolution" is often narrowly construed to apply only to the process of variation and selection in biology. However, there is an older and broader usage of this word, originating with Spencer,[16] that views evolution as driving long-term organizational change in nature, with biological evolution as a special case. Spencer defines evolution as "a change from an incoherent homogeneity to a coherent heterogeneity." Analogies between evolutionary behavior in many areas, such as astronomy, geology, economics, and sociology ,have been developed at a qualitative

Artificial Life II, SFI Studies in the Sciences of Complexity, vol. X, edited by C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen, Addison-Wesley, 1991

level.[11,16] However, as yet no one has been able to formulate this analogy quantitatively, or to demonstrate that it has predictive value. The failure to articulate the broader notion of evolution as a quantitative scientific principle justifies the prevalence of the narrower view that evolution in biology is fundamentally different from evolution in other natural phenomena.

One form of evolution that probably played a dominant role prior to biological evolution is what Calvin has termed "chemical evolution."[3,4] In this paper we study chemical evolution in the context of an artificial yet in many respects realistic model for catalytic reactions in polymer chemistry.[1,7] These results build on those of a companion paper,[1] in which it was demonstrated that under appropriate conditions metabolisms spontaneously emerge from a chemical soup. This emergence takes place on short time scales, over which the behavior is well approximated by deterministic equations. This paper studies the time evolution of these metabolisms over longer time scales, where stochastic effects play a critical role. We show that the metabolisms make transitions through a series of different fixed points, exhibiting what appears to be an open-ended succession of "punctuated equilibria."

The form of chemical evolution that we study here should be contrasted with that based on self-replication through templating, studied extensively by Eigen and others.[5,6] Templating reactions form the basis for the reproduction of contemporary organisms; the replication dynamics of chemical systems is so closely analogous to that of biological populations that it is essentially biological evolution on a molecular scale. Self-replication pertains to the possible evolution of early life forms, rather than to alternative evolutionary processes that may have preceded contemporary life. The form of chemical evolution that we study here is much closer to that envisioned by Rössler.[13,14,15]

Autocatalytic metabolisms reproduce themselves autonomously, without templating reactions. The analogy to biological evolution is not as direct as it is for templating systems. However, we argue that it nonetheless involves a process of variation and selection and deserves the name "evolution," at least in the broad sense articulated by Spencer. Although there are differences between the evolution of autocatalytic metabolisms and the evolution of biological organisms, they are at the level that one would naturally expect for proto-life forms based on alternative principles from those of contemporary organisms. The evolution of autocatalytic metabolisms is similar enough to that of biological organisms that many aspects are immediately recognizable, yet at the same time that there are provocative differences. The chemistry in which they evolve is simple enough for quantitative study and simulation. By presenting an example of an alternative form of evolution, we hope to support the idea that evolution can indeed be regarded as a broader physical principle driving organizational change, and to illustrate some of its more general properties.

This paper describes work in progress. Our goals are to outline the scenario under which autocatalytic metabolisms evolve, to discuss the issues involved in simulating their behavior, to present some preliminary numerical results, and to make some remarks comparing their evolution to that of contemporary biological organisms.

This paper will draw heavily on the companion paper, "Spontaneous Emergence of a Metabolism," in this volume,[1] where we study the chemistry of catalyzed polymerization reactions. The catalysis of these reactions is assumed to be specific, i.e., a typical polymer catalyzes the formation of only a small subset of all possible reactions. Under appropriate circumstances, the system may contain an *autocatalytic set*, i.e., a set of polymers such that each polymer is produced by at least one catalytic reaction involving only other members of the autocatalytic set. The system is driven away from equilibrium by an influx of a few special polymers, called the *food set*. When parameters are in the appropriate regime, the autocatalytic set may boost its own concentrations many orders of magnitude above the background of spontaneous reactions. When this occurs, we call the result an *autocatalytic metabolism*.

Although we will attempt to summarize important aspects of the companion paper as we go along, we will assume that the reader can refer to it as necessary.

2. POSSIBLE METHODS OF SIMULATION

Since there are an infinite number of possible polymer species, and an infinite number of possible reactions, properly simulating the behavior of a polymer network is a formidable problem. This problem is particularly severe for catalytic reactions, over long time scales. We compare three approaches: continuous differential equations, stochastic molecular collisions, and deterministic metadynamics.

Continuous differential equations. As long as the concentration of each species is sufficiently large, the dynamics can be described by a system of deterministic differential equations with continuous concentration variables $x_i > 0$, where i labels the possible species. However, because there are an infinite number of possible equations, from a practical point of view such a simulation is obviously impossible. Furthermore, even if it were possible, this approach provides a poor approximation of reality, particularly over long time scales. This is because real reaction vessels are always finite, which induces a minimum concentration θ corresponding to a single molecule. Until at least one molecule of a given chemical species is present, it cannot cause any reactions. As a result reactions are initiated sequentially, as the system creates the necessary constituents. In contrast, for continuous differential equations, for any time $t > 0$ all reactions are switched on, and all species typically have nonzero concentrations. For autocatalytic reactions this is not merely an esoteric problem: It results in a qualitatively incorrect prediction of the true dynamics.

Stochastic molecular collisions assumes integer populations of each species and simulates reactions as discrete collisions by sampling at random. This method is faithful to reality (at least for the level of description we are interested in here). However, as discussed in the companion paper, if the difference between the largest and smallest concentrations is large, it can be time consuming in comparison with

differential equations. This problem taxes the resources of even the largest parallel computers. It is difficult to simulate a system of any size over long time scales.

Deterministic metadynamics uses a sequence of deterministic equations that change as the behavior of the system changes. This alternative attempts to make the best of both worlds.[2,7,8] The topological structure of the kinetics is represented by a graph, reflecting the dominant chemical species and chemical reactions at any given time. The graph changes to reflect either the creation of new species or the elimination of old species. This takes place relative to a concentration threshold θ, corresponding to the concentration when one molecule is present. The threshold thus specifies the size of the reaction vessel.

We wish to emphasize that while this procedure involves a sequence of changing graphs, *it is purely deterministic.*[1] Because of the finite threshold θ, the results are different from those that would be obtained with a fixed continuous system of equations. However, because the simulation is deterministic, over long time scales the results diverge from those that would be obtained with a stochastic simulator. In the companion paper we were interested in the problem of the *emergence* of autocatalytic metabolisms. This takes places on short time scales, over which the deterministic metadynamics procedure described above is reasonably accurate. One of our main purposes in this paper is to introduce a modification of the metadynamics procedure, called *stochastic metadynamics*, which combines the physical accuracy of a full stochastic molecular collision simulation with the speed of a metadynamics simulation, and allows us to follow the evolution of autocatalytic metabolisms for long periods of time.

For the equations studied here, for any fixed graph our simulations indicate that the dynamical equations always have a unique stable fixed point. We call this a *dynamical fixed point.* As described in detail in the companion paper, the existence of a unique fixed point makes it possible to speed up the metadynamics algorithm considerably by using an algebraic fixed point solver. For a given graph we find the corresponding dynamical fixed point; if there are new species over the threshold, we update the graph, and then find a new fixed point. We repeat this until there are no species that cross the threshold. Since no new species are created or destroyed, the graph remains fixed. We call the corresponding final state a *metadynamical fixed point.* Dynamical fixed points have no physical meaning, and are just computational conveniences. In contrast, metadynamical fixed points are physically meaningful. This is clarified in the next two sections.

[1] The kinetic equations should not be confused with the rules for the assignment of kinetic parameters. Though we may use a random assignment rule, once the assignment is completed the kinetic parameters are fixed and, for the simulations of the companion paper, the kinetic equations are completely deterministic.

3. AUTOCATALYTIC NETWORKS AS FLUCTUATION AMPLIFIERS

To understand why thresholds and spontaneous fluctuations have a large effect on autocatalysis, it is important to distinguish between internally and externally catalyzed reactions. An *internally catalyzed* reaction pathway is catalyzed by a species within the current autocatalytic metabolism. Let σ be a species that initially has a population of zero, and so is external to the metabolism. If σ is produced by a catalyzed reaction whose catalyst is already abundant, the concentration of σ will increase rapidly. In contrast, an *externally catalyzed* reaction pathway is catalyzed by a species outside the current autocatalytic metabolism. In order to be catalyzed, such reactions must wait for a spontaneous reaction to produce the catalyst. Since the spontaneous reactions are slow, when compared to the catalyzed reactions they can be regarded as discrete fluctuations. As pointed out by Rössler,[14] this introduces a delay in the activation of externally catalyzed reactions.

The simplest example is given in Figure 1(a). Let A and B be two polymers in the autocatalytic metabolism, and let σ be a third polymer that catalyzes its own formation.

$$A + B \xrightarrow{\sigma} \sigma, \tag{1}$$

Assume that σ is not produced by any other reactions involving only members of the autocatalytic metabolism, i.e., that it is an element of the *background* of polymers formed by spontaneous reactions. Assume that in the initial meta-dynamical fixed point its population is zero.[2] If a fluctuation produces σ, then, since it catalyzes its own formation, its concentration may increase by orders of magnitude.

Note that the boost in the concentration of σ may have other side effects; for example, σ may catalyze other reactions. This may cause the production of new polymer species, which in turn catalyze other reactions, etc. Since these events involve only internally catalyzed pathways, they take place on a rapid time scale, without delays. When σ is pumped up to high concentrations, the change in the dynamic equilibrium may also alter the concentrations of other species in the metabolism. The resulting competition might cause other polymer species to disappear from the metabolism. New species can be created, and old species can become extinct.

Purely deterministic metadynamics, as used in the companion paper, explores only the internally catalyzed pathways. Since the system cannot produce any of the external catalysts, it gets stuck at a metadynamical fixed point. For convenience, we will often call this a *pinned state*. Once in a pinned state, until a spontaneous fluctuation produces one of the external catalysts, the system cannot change. When a fluctuation occurs it may trigger another period of rapid change as the system

[2] When the simulation of the concentration of the background as an aggregate is included, assume that its concentration is less than the threshold, so that with the deterministic metadynamic rule of the companion paper,[1] it is not allowed to catalyze new reactions.

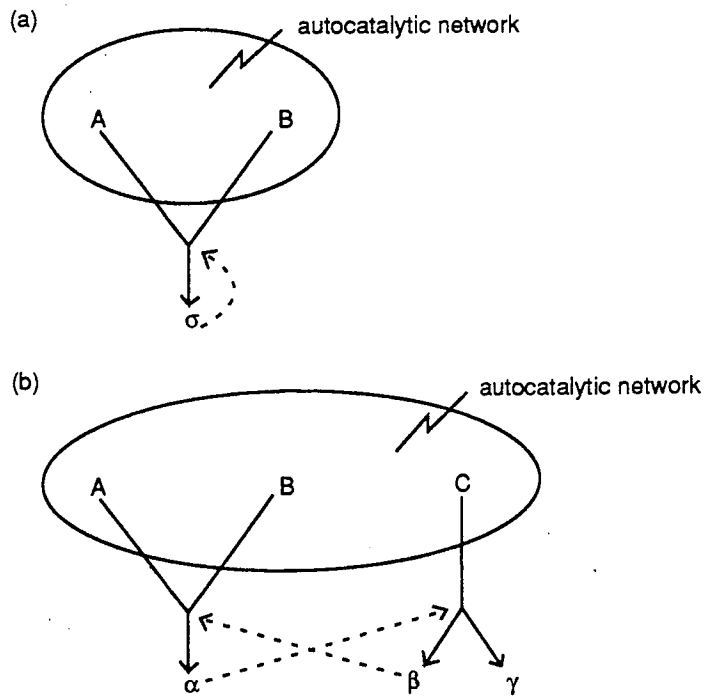

FIGURE 1 *Examples of autocatalytic mutations.* (a) A and B are members of an autocatalytic metabolism. $\sigma = A + B$ catalyzes its own formation. If the concentration of σ is initially zero, then except for the spontaneous reaction it will remain zero. If by chance the spontaneous reaction produces a molecule of σ, then the autocatalytic reaction may increase the concentration of σ by many orders of magnitude, creating a new pinned state which includes σ. (b) depicts a somewhat more complicated instance of the same phenomenon: A, B, and C are in the autocatalytic metabolism. Through spontaneous reactions A and B produce α, and C produces β and γ. Suppose α catalyzes the cleavage of C, and β catalyzes the condensation of A and B. If either α or β are produced by a spontaneous reaction, they may pump each other up and join the metabolism.

explores the newly activated internally catalyzed pathways, until the system settles into a new metadynamic fixed point. The end result is a change in the composition of the autocatalytic metabolism, and a transition to a new pinned state. We call a transition from one pinned state to another an *evolutionary modification* of the autocatalytic metabolism.

We will call an autocatalytic reaction such as the one shown in Figure 1(a), in which a species catalyzes its own formation, a *first-order autocatalytic loop*. There are autocatalytic loops of all orders. For example, Figure 1(b) depicts a second-order

autocatalytic loop, in which α catalyzes the formation of β and β catalyzes the formation of α. The number of possible graphical combinations grows exponentially with the order of the loop. There are thus an enormous number of possible autocatalytic loops, and an enormous number of possible evolutionary modifications of autocatalytic metabolisms.

If an evolutionary modification is *robust*, i.e., if the steady state concentrations of the new elements at the new pinned state are orders of magnitude above the threshold, then the modification is unlikely to reverse itself spontaneously, since the probability of a series of fluctuations that decrease the population by several orders of magnitude is virtually nil. Thus, even though all the reactions involved are reversible, from a stochastic point of view *robust evolutionary modifications are effectively irreversible.* Once an evolutionary modification is triggered, the system is unlikely to return to its previous pinned state.[3]

The evolution of the system through time is highly path dependent. At any given time there are many possible spontaneous fluctuations. Each fluctuation that actually occurs generates a series of irreversible changes, effecting both the probability that a given fluctuation will occur, and the probability it might initiate a modification. Externally catalyzed reactions are activated sequentially, in random order; the probability that a given reaction will be activated at a given time is altered by each preceding fluctuation. This is quite different from what one would observe with a purely deterministic model.

Thus we see how the long time dynamics of a chemical network involves a series of transitions between pinned states, which are similar to punctuated equilibria in evolutionary biology. For a network of any reasonable size, the number of possible pinned states is so large that the system may evolve for a very long time without ever repeating itself.

4. STOCHASTIC METADYNAMICS

In this section we present a stochastic extension of the metadynamical method that allows us to simulate the evolution of autocatalytic metabolisms reasonably quickly. The basic idea is to perturb the catalytic reaction graph by randomly adding new species that may cause evolutionary modifications. Doing this in a physically realistic manner, so that the perturbations occur with the proper relative probabilities and time scales, involves several complex issues. In this section we outline our approach to the problem, and describe the heuristic treatment that forms the basis for the simulations of the next section.

[3] Of course, if a fluctuation of species σ triggers an evolutionary modification, it is always possible that some later evolutionary modification might introduce competition and eliminate σ. However, when this occurs it is unlikely to cause a return to the original pinned state.

4.1 THE SPECIAL ROLE OF THE SHADOW

The shadow is the subset of species in the background that are produced by reactions involving only themselves and members of the catalyzed reaction network. (See Figure 4 of the companion paper.) The shadow plays a dominant role in initiating evolutionary modifications, for two reasons: First, because the autocatalytic metabolism is at high concentration, the shadow is typically maintained at higher concentrations than other parts of the background. Thus fluctuations creating elements of the shadow are more likely than others. Second, for a fluctuation of a *single* species to trigger an evolutionary modification, that species must be in the shadow. This is true almost by definition: if a background element is not in the shadow, then its production requires at least one other background element. Triggering catalytic production therefore requires at least two simultaneous fluctuations, which is unlikely. For both of these reasons, the majority of evolutionary modifications are initiated in the shadow.

4.2 ENUMERATION OF AUTOCATALYTIC SUBGRAPHS

Figure 1 shows two possible autocatalytic subgraphs that might trigger an evolutionary modification of an autocatalytic metabolism. There are an enormous number of other possible subgraphs. We will restrict attention to autocatalytic subgraphs that can be triggered by a single fluctuation in the shadow.

We begin by introducing a simplified graph description that describes the feedback structure of the subgraph. This simplified description ignores the identity of the "parents" in the autocatalytic metabolism, focusing attention on the catalytic relationships in the shadow. A situation in which σ catalyzes its own formation, as shown in Figure 1(a), is represented by a simple graph with a single vertex σ, and a single directed edge from σ to itself, as shown in Figure 2(a).

This reduced graph also lumps together any other simple autocatalytic reactions that produce σ. For example, there might be another set of polymers A' and B' in the autocatalytic set whose condensation is catalyzed by σ, or a polymer C' whose cleavage is catalyzed by σ. The reduced graph describes all of these reactions taken together.

Figure 2 enumerates the combinatorial possibilities for reduced graphs with up to three vertices. In addition to simple loops, there are many other possibilities. In many of these the autocatalytic feedback is maintained exclusively by a subset of the possible vertices; in this case we call the remaining vertices *parasites*. For example, Figure 2(c) has two vertices. The first vertex provides the autocatalytic feedback through a first-order loop. It also catalyzes the formation of the parasitic second vertex, which does nothing to support the first loop.

In principle it is possible to enumerate the autocatalytic modifications of arbitrary order. However, for simplicity in this paper we will restrict attention to modifications of order three or less, as shown in Figure 2.

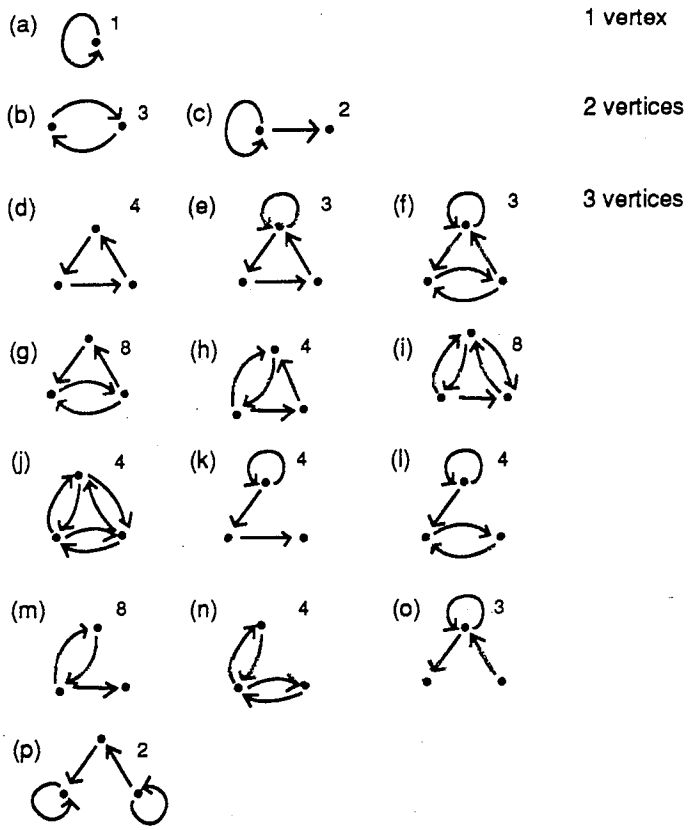

FIGURE 2 *Possible autocatalytic subgraphs in the shadow with three vertices or less.* These reduced graphs ignore the identity of the "parents" in the autocatalytic set. The vertices correspond to chemical species in the shadow, and the edges indicate their catalytic relationship to each other. (a) corresponds to a species that catalyzes its own formation, and is equivalent to the full graph of Figure 1(a); it also lumps together any other graphs of the same form with different parents but the same σ. (b) corresponds to a two-cycle, such as that of Figure 1(b). (c) also involves two vertices; however, it can be regarded as a one cycle with a "parasite." The remainder of the graphs enumerate the possibilities with three vertices. The counts with each graph give the number of realizations that contain the shown skeleton, and in which some of the vertices are additionally autocatalytic on their own.

4.3 ASSIGNING PROBABILITIES TO AUTOCATALYTIC MODIFICATIONS

All of the autocatalytic subgraphs listed in the previous section describe additions to the metabolism that can be triggered by a single spontaneous fluctuation. Once the fluctuation occurs it may activate the other elements of the subgraph. For example, in Figure 2(b), suppose a fluctuation produces α; unless α decays first, it produces β, which in turn may produce another α, etc. Whether or not these fluctuations grow so that they pump up the concentrations to a robust level, well above the threshold, is a complicated birth and death process with an uncertain outcome. After a long time there are two likely possibilities:

1. Creation events overcome decay events, and the concentrations of the subgraph are roughly at the level predicted by deterministic kinetics.
2. Decay events overcome creation events, and the concentrations are zero.

This neglects unlikely events, such as additional spontaneous fluctuations, and assumes that the kinetic parameters are such that the initial production rate exceeds the decay rate (otherwise the concentrations almost certainly go to zero).

The relative probability of these two outcomes can be approximated using a master equation. For example, consider the case of a simple first-order autocatalytic loop, as described in Figure 1(b). Neglecting saturation, the deterministic kinetics can be approximated by

$$\dot{\sigma} = (1 + \nu\sigma)(k_f AB - k_r H\sigma) - K\sigma, \tag{2}$$

where ν is the catalytic efficiency, k_f is the forward rate constant, k_r is the reverse rate constant, H is the concentration of water, and K is the global dissipation parameter, corresponding to the rate at which material diffuses out of the system. When σ is sufficiently small, we can neglect the quadratic term. If $\nu\theta \gg 1$ we can also neglect the spontaneous reaction, and this becomes

$$\dot{\sigma} = \nu k_f AB\sigma - (K + k_r H)\sigma. \tag{3}$$

When the concentrations are small the problem is more properly treated in terms of a master equation. Letting $P(n,t)$ be the probability that σ has population n at time t, the master equation corresponding to Eq. (3) is

$$\frac{\partial P(n,t)}{\partial t} = c\,(n-1)P(n-1,t) + d\,(n+1)P(n+1,t) - (c+d)nP(n,t), \tag{4}$$

where $c = \nu k_f AB$ is the "creation" rate, and $d = K + k_r H$ is the "death" rate.

Assume that at $t = 0$ the spontaneous reaction creates a single molecule. For the master equation this corresponds to the initial condition $P(1,0) = 1$. By solving the master equation (see e.g., Karlin[10]) and taking the limit as $t \to \infty$, we can compute the survival probability, $P_s = 1 - P(0,\infty)$. It is

$$P_s = \begin{cases} \frac{c}{(c+d)} & \text{if } c > d; \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

P_s is the probability a mutation will grow once it is triggered.

The asymptotic survival probability for an mth-order subgraph can also be derived using an m-dimensional master equation. The equations are linear, but they are complicated and we have not yet solved them. As an approximation, however, we first study the corresponding deterministic equations. In particular, it is clear that the survival probabilities will depend on the deterministic creation and death rates. If there is not a growing mode in the deterministic limit, the survival probability in the stochastic case will be zero.

Making the same approximation used above, for a simple mth-order cycle the linearized kinetic equations are

$$\begin{aligned} \dot{\sigma_1} &= c_1\sigma_m - d\sigma_1 \\ \dot{\sigma_2} &= c_2\sigma_1 - d\sigma_2 \\ &\vdots \\ \dot{\sigma_m} &= c_m\sigma_{m-1} - d\sigma_m, \end{aligned} \tag{6}$$

where $c_i = \nu k_f A_i B_i$ are the catalyzed production rates[4] of the m species external to the metabolism whose concentrations are given by σ_i. The kinetics is described by the eigenvalues (and eigenvectors) of the matrix

$$\begin{pmatrix} -d & 0 & 0 & \cdots & c_1 \\ c_2 & -d & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & c_m & -d \end{pmatrix} \tag{7}$$

The m eigenvalues turn out to be

$$\lambda_k = \Big(\prod_{j=1}^{m} c_j\Big)^{1/m} \exp(2\pi i \cdot k/m) - d, \quad k = 0, 1, \ldots, m-1. \tag{8}$$

Fluctuations will grow as long as the largest eigenvalue $\lambda_0 = (\prod_{j=1}^{m} c_j)^{1/m} - d$ is positive. This suggests that the probability for a fluctuation that generates a growing mth-order graph depends on the geometric mean of the catalytic production rates c_i. This has some interesting possible consequences: Since the mass of the system is constant, in order to increase the number of species, it must typically decrease the population of each species inside the autocatalytic metabolism. The growth rates c_i depend on the concentrations in the metabolisms; if these decrease, the analysis above indicates that the probability of survival also decreases. The system should thus evolve to a critical state in which the linear growth rate for a typical mutation is on average near zero.

[4] This form assumes that all the reactions are condensation reactions; for a cleavage reaction it is $c_i = \nu k_r HC_i$.

5. NUMERICAL EXPERIMENTS

In this section we present a few preliminary numerical experiments on the evolution of autocatalytic metabolisms. The purpose of the simulations presented at this stage is simply to demonstrate the basic principle behind the chemical evolution of autocatalytic mutations. The approach at this stage is strictly *ad hoc*: We ignore the important issue of the time between transitions from one pinned state to another, and make an arbitrary choice concerning the relative survival probabilities of autocatalytic subgraphs.[5] A simulation proceeds as follows:

1. Explore the internally catalyzed pathways by running the deterministic metadynamics algorithm until the system reaches a metadynamical fixed point.
2. Construct the allowed autocatalytic subgraphs of the shadow.
3. Assign a weight to each of the subgraphs.
4. Select a subgraph with probability given by its normalized weight.
5. Try to install the selected subgraph and check if its members acquire concentrations above threshold within the autocatalytic network.
6. If no, return to step 4.
7. If yes, install it in the network and return to step 1.

We call this procedure *stochastic metadynamics*.

A typical simulation is shown in Figure 3(a). We arbitrarily measure the time in "meta-steps," which correspond to the number of times the metadynamics algorithm finds new dynamical fixed points. Figure 3(a), for example, plots the number of polymer species above threshold, beginning with an initial state in which the only species with nonzero concentrations are those of the food set. The graph changes 12 times, corresponding to 12 different dynamical fixed points, before the system settles on a metadynamical fixed point with roughly 20 species. At this point we introduce a mutation, which triggers a series of new internally catalyzed pathways, until the system settles onto another metadynamical fixed point, roughly ten steps later. This process repeats itself until the number of species reaches 50 and the simulation terminates. During this time new species are created, and old species become extinct, as shown in Figure 3(b).

The "meta-step" time scale used in these simulations masks the physical correspondence to punctuated equilibria. In reality, after each mutation the system quickly reaches a new steady state, and then remains relatively unchanged for a much longer period of time, until the next mutation. To make a correspondence between the simulation and reality, at each time marked by a triangle in Figure 3, one should imagine an interval of indefinite duration during which the properties of the system remain almost constant.

[5] To assign relative survival probabilities, for the simulations shown here we weighted graphs according to the product of the production rates of their vertex elements. As we now know, this overemphasizes the importance of short graphs. We intend to repeat these simulations with the proper weighting function as described in section 4.3 in the future.

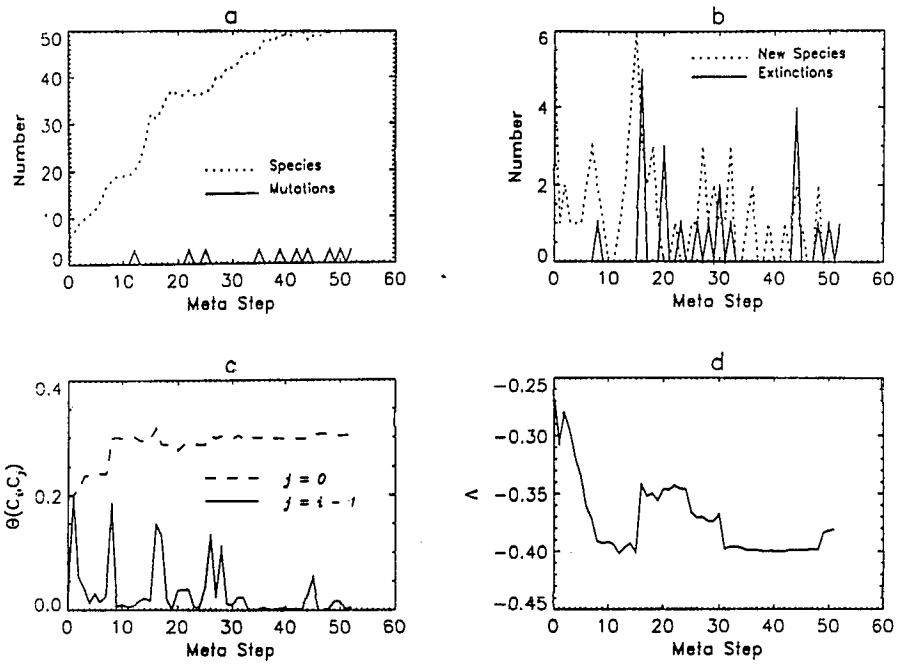

FIGURE 3 *The evolution of an autocatalytic metabolism.* (a) plots the number of polymer species as a function of time, measured in "meta-steps" (see text). The triangles on the horizontal axis indicate mutations that generate evolutionary modifications. (b) shows the number of new species created (dotted) and the number of old species that become extinct. (c) plots the difference in the angle of the concentration vector (see text) between the present step and the previous step (solid) or the initial condition (dashed). (d) The slope Λ is one indication of the deviation from equilibrium properties. The forward rate constant k_f, the reverse rate constant k_r, and the catalytic efficiency ν are randomly varied in the range $k_f \in [10^1, 10^2]$, $k_r \in [1, 10]$, $\nu \in [10^3, 10^7]$. The unbinding constant $k_u = 10^5$, the driving $\delta = 10^2$, the mass concentration $m_0 = 10^{-1}$, the concentration threshold $= 10^{-4}$, and the probability of catalysis $p = 4.5 \times 10^{-3}$. An upper limit on the number of polymers was set at 50; autocatalytic subgraphs were chosen from those with one or two nodes.

In order to measure the change of the state of the network in quantitative terms, in Figure 3(c) we use the angle $\Theta(C_i, C_j)$ between the concentration vectors C_i and C_j. The *concentration vector* C_i is defined as the infinite-dimensional vector whose coordinates are the concentrations of each possible species. C_i uniquely specifies the state of the system at meta-step i. Figure 3(c) demonstrates that the angle between concentration vectors changes more rapidly at the beginning, indicating that evolutionary modifications at later stages have a smaller effect on the autocatalytic metabolism.

In the companion paper we introduced Λ, the slope of the concentration profile, as a measure of the deviation of the properties of the system from those at equilibrium. Larger (less negative) values of Λ indicate a larger deviation from equilibrium. Λ decreases during the initial deterministic evolution; an evolutionary modification triggered by the first mutation increases Λ significantly for a while, but subsequent mutations cause it to decrease again. The values of Λ in this simulation indicate that none of these metabolisms are very robust; however, it is interesting to see that an evolutionary modification can make a significant change in Λ.

Figure 4 shows another sequence of evolutionary modifications. In this case we admit mutations involving subgraphs with up to three (rather than two) vertices, and use slightly different kinetic parameters, as well as a higher ceiling for the number of possible species.

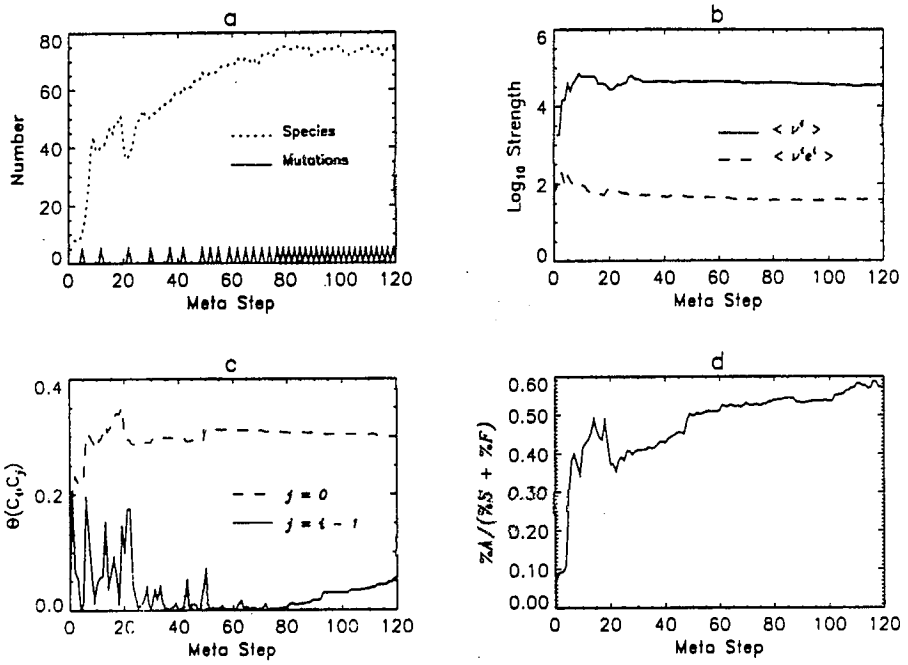

FIGURE 4 *A second simulation of an autocatalytic metabolism.* Figures (a) and (c) are similar to those of from those of Figure 3. (b) plots the mean catalytic efficiency efficiency $\langle \nu \rangle$ of the autocatalytic metabolism (solid) and the mean catalytic strength $\langle \nu e_i \rangle$, where e_i is the concentration of catalyst e. (d) plots the mass ratio, i.e., the ratio of the mass of the autocatalytic metabolism (without the food set) to that of the sum of the food set and the background. The parameters are the same as those of Figure 3, except that $k_u = 10^6$, the limit on the number of polymers was set at 75, and autocatalytic subgraphs were chosen from those with three vertices or less.

Figure 4(a) shows the number of species as a function of time; interestingly, some mutations cause the number of species (and hence the diversity) to decrease. Early mutations typically have a large effect, in that they require several steps of the metadynamics algorithm to reach a metadynamical fixed point and therefore must have generated several new internally catalyzed pathways and several new chemical species. Later mutations, however, only trigger a single metadynamical step. (They may generate multiple species and internal pathways, but the number is probably fewer). Somewhat surprisingly, in spite of this Figure 4(c) shows that the later mutations cause a larger change in the angle of the concentration vector. At this point we do not know how to interpret this.[6]

Figure 4(d) plots another measure of the deviation from equilibrium, the mass ratio of the autocatalytic metabolism relative to the sum of the food set and the background. Although there are several periods where the mass ratio decreases, there is an overall tendency for it to increase with time. Figure 4(b) shows the mean catalytic efficiency as a function of time. It rises to a maximum and then tends to decrease slightly; we do not understand why.

These numerical experiments are admittedly very preliminary. The parameters used were somewhat arbitrary, and did not generate very robust autocatalytic metabolisms. Nonetheless, they do demonstrate the basic principle that a metabolism can evolve well past its initial pinned state, and that the resulting set of evolutionary modifications can generate new molecular species, and cause old species to become extinct.

6. DISCUSSION

We have illustrated a process of "chemical evolution" that bears many similarities to biological evolution, at the same time that it is distinctly different. In this section we discuss this analogy in more detail.

There is a clear notion of variation and selection within autocatalytic metabolisms, and we feel that the term "evolution" is well deserved. Random variations, which play the role of mutations, are generated by spontaneous reactions. Some of these variations have no effect, and simply die out. Others have large effects, generating several new chemical species and perhaps causing others to die out, substantially altering the composition of the autocatalytic metabolism. As in evolutionary biology, "favorable" variations are by definition those that propagate themselves. Spontaneous fluctuations provide random variation, and chemical kinetics provide selection.

[6] Similar behavior has been observed by Rasmussen et al.[12]

We now develop the analogy with the evolution of biological organisms in more detail. A possible point of confusion concerns the level at which to make the identification. In the self-reproducing reactions studied by Eigen, the notion of "chemical species" can be roughly identified with "biological species." This is literally the case in the experiments with the $Q\beta$ virus. For autocatalytic metabolisms, however, the individual chemical species are only elementary building blocks; they are not in any sense alive on their own. The "organism" is the entire metabolism. The "phenotype" of the organism is the concentration vector C_i, i.e., the set of species and concentrations that comprise the autocatalytic metabolism.

We have demonstrated that the time history of an autocatalytic metabolisms contains periods where the system is pinned and change is very slow, and other periods where a fluctuation triggers the exploration of new internally catalyzed pathways and change is quite rapid. We propose that the chemical kinetics of internally catalyzed pathways can be regarded as a "developmental algorithm." The "genotype" is any list of chemical species that produce a given pinned state through purely internally catalyzed pathways. Note that there are many possible concentration levels and lists of chemical species that are connected by internally catalyzed pathways leading to the same pinned state. A subset of the metabolism can thus regenerate the entire metabolism. Such a subset has been called a *seeding set* by Fontana.[9] We regard the genotype of a given pinned state as the equivalence class of its seeding sets. Since the seeding sets will all regenerate the metabolism regardless of the concentrations of their elements, the genotype is given by a simple list, stating which species are present. It thus encodes information about the phenotype in symbolic form, in a manner analogous to the sequence of base pairs in a contemporary organism. The chemical kinetics produces the phenotype from the genotype, in a highly simplified but analogous manner to that of a contemporary organism.

As we have demonstrated here, autocatalytic metabolisms go through a progression of events, involving the alteration of the genotype and the generation of new phenotypes. Since the expression of the directly catalyzed pathways is rapid, while the time intervals between evolutionary modifications may be large, when viewed over an expanse of time, these have the appearance of "punctuated equilibria." This illustrates a difference between the evolution of autocatalytic metabolisms and that of biological organisms: the evolutionary process and the developmental process are one and the same. In this context a single organism can evolve—the selection comes from the laws of chemistry, without any obvious need for competition with other organisms.

This brings up the important issue of "identity." If an autocatalytic metabolism is clearly confined to a given container, then the genotype and phenotype of the autocatalytic metabolism make it possible to distinguish it from other metabolisms in other containers. The number of possible metabolisms is quite high, and the probability that two metabolisms in different containers will be in different pinned states is extremely high. Each metabolism might evolve on its own, without any need for competition with others.

However, one can also imagine that the vessels containing the metabolisms occasionally come into physical contact, so that some material diffuses from one

to the other. In this case we can imagine a population of evolving metabolisms, each occasionally infecting the others with pieces of its own genotype. An infection may trigger new catalyzed pathways, and substantially change the evolution of the infected host.

One can also imagine a situation in which autocatalytic metabolisms are more continuously distributed through space, with ongoing diffusive coupling. In this case it becomes difficult to assign a notion of "identity." Nonetheless, there might be some interesting spatial inhomogeneity that fosters evolutionary behavior that is qualitatively different from that within a spatially isolated well-stirred reactor.

We have tried to suggest that autocatalytic metabolisms can be viewed as proto-organisms with a crude "genetic code," consisting of a list of polymers. The code is "interpreted" by chemical kinetics. We also suggest that autocatalytic metabolisms "evolve"; they certainly meet the basic requirements of information storage of a genotype and of a selection mechanism to amplify random variations. Nonetheless their evolution is distinctly "chemical" rather than biological. There are substantial differences, most of which stem from the fact that the identity of an individual autocatalytic metabolism is not as well defined as that of a biological organism or that of a templating RNA string; there is not necessarily a clear distinction between an organism and a population of organisms. However, we feel that these differences make this system more rather than less interesting. Many of the same questions that have eternally plagued evolutionary biology also surface here. In particular, how does the evolution of autocatalytic metabolisms differ from a random walk? Is there any sense in which they "make progress" as they evolve? We hope that future studies of autocatalytic metabolisms may shed some light on these questions, and help clarify how evolution applies in a broader sense.

ACKNOWLEDGMENTS

We would like to thank Steen Rasmussen and Chris Langton for critical readings of the manuscript.

REFERENCES

1. Bagley, R. J., and J. D. Farmer. "Spontaneous Emergence of a Metabolism." This volume.
2. Bagley, R. J., J. D. Farmer, S. A. Kauffman, N. H. Packard, A. S. Perelson, and I. M. Stadnyk. "Modeling Adaptive Biological Systems." *Biosystems* **23** (1989): 113–138.
3. Calvin, M. "Chemical Evolution and the Origin of Life." *Amer. Sci.* **44** (1956): 248–263.
4. Calvin, M. *Chemical Evolution.* Oxford: Oxford University Press, 1969.
5. Eigen, M. "Self-Organization of Matter and the Evolution of Biological Macromolecules." *Naturwissenschaften* **58** (1971): 465–523.
6. Eigen, M., and P. Schuster. "The Hypercycle: A Principle of Natural Self-Organization. Part A: The Emergence of the Hypercycle." *Naturwissenschaften* **64** (1977): 541–565.
7. Farmer, J. D., S. A. Kauffman, and N. H. Packard. "Autocatalytic Replication of Polymers." *Physica D* **22** (1986): 50–67.
8. Farmer, J. D., N. H. Packard, and A. S. Perelson. "The Immune System, Adaptation, and Machine Learning." *Physica D* **22** (1986): 187–204.
9. Fontana, W. "Algorithmic Chemistry: A Model for Functional Self-Organization." This volume.
10. Karlin, S., and H. M. Taylor. *A First Course in Stochastic Processes.* New York: Academic Press, 1975.
11. Miller, J. G. *Living Systems.* New York: McGraw Hill, 1978.
12. Rasmussen, S., C. Knudsen, and R. Feldberg. "Dynamics of Programmable Matter." This volume.
13. Rössler, O. "A System Theoretic Model of Biogenesis." *Z. Naturforschung* **B26b** (1971).
14. Rössler, O. "Design for Autonomous Chemical Growth Under Different Environmental Constraints." In *Progress in Theoretical Biology*, edited by R. Rosen and F.M. Snell. 1972.
15. Rössler, O. "Chemical Automata in Homogeneous and Reaction-Diffusion Kinetics." *Notes in Biomathematics* **B4** (1974).
16. Spencer, H. *First Principles.* 1862.

Artificial Life II

PROCEEDINGS OF THE SECOND
INTERDISCIPLINARY WORKSHOP ON
THE SYNTHESIS AND SIMULATION OF
LIVING SYSTEMS,
HELD FEBRUARY, 1990
IN SANTA FE, NEW MEXICO

Editors

Christopher G. Langton
Complex Systems Group, Theoretical Division
Los Alamos National Laboratory

Charles E. Taylor
Department of Biology
University of California, Los Angeles

J. Doyne Farmer
Complex Systems Group, Theoretical Division
Los Alamos National Laboratory

Steen Rasmussen
Complex Systems Group, Theoretical Division
Los Alamos National Laboratory

Proceedings Volume X

Santa Fe Institute
Studies in the Sciences of Complexity

[Addison-Wesley Information]

Publisher: *Allan M. Wylde*
Production Manager: *Jan Benes*
Marketing Manager: *Laura Likely*

Director of Publications, Santa Fe Institute: *Ronda K. Butler-Villa*
Technical Assistant, Santa Fe Institute: *Della L. Ulibarri*

Library of Congress Cataloging-in-Publication Data

This volume was typeset using TEXtures on a Macintosh II computer. Camera-ready output from a NEC Silentwriter 2 printer.

ABCDEFGHIJ-???????????