Gaussian random energy model and Dyson's model for the origin of metabolism

Dominique J. Bicout* and Martin J. Field†
Laboratoire de Dynamique Moléculaire, Institut de Biologie Structurale—Jean-Pierre Ebel, 41 avenue des Martyrs, F-38027 Grenoble Cedex 01, France
(Received 24 April 1995)

The equivalence between Dyson's model for a population of prebiotic mutually catalytic molecules and the Gaussian random energy model is established. It is shown that, within the mean field approximation, the mathematical expression for the function that describes the autocatalytic capability of the whole system is completely determined by the nature of the interaction potential (or the force field) between the monomers. Our results along with those previously obtained by Abbott [J. Mol. Evol. **27**, 513 (1988)] on the replication of an autocatalytic system lead us to suggest that spin-glass theory is an appropriate model for the investigation of the origins of both metabolism and replication.

PACS number(s): 87.10.+e, 64.60.Cn, 87.15.−v

I. INTRODUCTION

The physics of disordered systems in condensed matter has become increasingly relevant to the study of biological systems. A representative example is the theory of spin glasses [1], which has found applications in as diverse domains as the investigation of neural nets [2] and in the study of protein dynamics [3–5]. A spin glass has a natural disorder because it contains frustrated interactions. Such systems exhibit a rugged potential energy surface, which is a feature that is found to be widespread in biological systems. It is believed that this complexity in the energy landscape is responsible for the great adaptability and the rich evolutionary dynamics of biological systems. It is thus possible to suppose that both the structure of the energy landscape and the resulting complexity that characterize contemporary biological systems are hereditary properties that emerged in the early evolution of biomolecules.

The Gaussian random energy model (REM) was introduced by Derrida [6] to study spin glasses. It is a simple model of a disordered system in which the energy levels are assumed to be independent Gaussian random variables. Bryngelson and Wolynes [7] first applied the model to study the folding of proteins and since then a number of groups have used it to study the folding of proteins [8–10] and other physical phenomena such as thermal properties [11] and relaxation processes [12,13]. Subsequently, Fernandez [14] has shown that the relaxation kinetics of RNA folding can be described with the REM. Recently, the model has also been used to model electron transfer reactions in biomolecules and in solvent environments [15].

In this paper we show how the REM can provide the theoretical framework for the model of the origin of metabolism that was developed by Dyson [16]. The outline of this paper is as follows. In Sec. II, we summarize Dyson's model and the problem that he was trying to address. In Sec. III, we describe the application of the REM to the same problem and derive all the equations which are necessary to show the equivalence of the two approaches, which is established in Sec. IV. Section V gives the conclusion.

II. DYSON'S MODEL FOR THE ORIGIN OF METABOLISM

Schrödinger, in his book *What is Life?* [17], first questioned our fundamental understanding of the physicochemical basis of life, but it was Von Neumann [18] who first used the computer metaphor for living organisms. He observed that there is a strong analogy between the way computers and living cells function. Two components are essential for the functioning of computers — the hardware, which processes information, and the software, which embodies the information. The analogues in living cells are the proteins and the nucleic acid, respectively. Proteins are the essential components for *metabolism* and the nucleic acid for *replication*, which constitute the two basic functions of life.

One of the principal problems when discussing the origin of life is the origin of each of these functions, metabolism and replication. Two hypotheses are possible. The first is *the single-origin hypothesis*, which assumes that life began only when the functions of metabolism and of replication, which were already present in a rudimentary form, were linked together. The second hypothesis, called by Dyson *the double-origin hypothesis*, postulates that life began with separate species of creature, one species capable of metabolism without replication, and the other one capable of replication without metabolism [16].

Dyson classifies theories of the origin of life into three main groups. Two of them, the theories of Oparin [19] and of Eigen *et al.* [20], are single-origin theories since

*Electronic address: bicout@ibs.fr
†Electronic address: mjfield@ibs.fr

both postulate a single process for the origin of life. Oparin places metabolism as the primary event with replication afterwards, while Eigen puts the primary emphasis on replication with metabolism being developed after replication had been established. It is only the third theory, that of Cairns-Smith [21], which is explicitly a double-origin theory. The first origin is the construction of a protein metabolic apparatus in conjunction with clays, which play a replicative role, while the second origin, which occurs after a period of biochemical evolution, is the replacement of the clay replicative apparatus with a nucleic acid one. (Recently, Morowitz [22] developed a model of life's origin in which cells originate first, proteins follow, and genes evolve last. We do not discuss this model here because it is reminiscent of Oparin's theory.)

The Eigen theory for the origin of replication has been extensively studied both experimentally and theoretically. In his work, which is discussed at length in the book *Origins of Life* [16], Dyson tried to establish a mathematical framework for the Oparin theory for the origin of metabolism in the same way that Eigen had done for his theory. To do this, he constructed a simple model (which he called his *toy model*) in which he considered the behavior of a population of molecules that are confined in a primitive cell or droplet. The essential feature of the model for the origin of metabolism are the transitions that occur for the population between quasistationary states, which are disorganized (*dead*), and those that are organized (*alive*).

He made a number of assumptions that we quote in full here:

1. (Oparin theory). Cells came first, enzymes second, genes much later.

2. A cell is an inert droplet containing a population of polymer molecules that are confined to the cell. The polymers are composed of monomer units that may be imagined to be similar to the amino acids that make up modern proteins. The polymers in the cell contain a fixed number N of monomers. In addition, there is an external supply of free monomers that can diffuse in and out of the cell, and there is an external supply of energy, which causes chemical reactions between polymers and monomers.

3. Cells do not die and do not interact with one another. There is no Darwinian selection. Evolution of the population of molecules within a cell proceeds by random drift.

4. Changes of population occur by discrete steps, each step consisting of a single substitution mutation. A mutation is a replacement of one monomer by another at one of the sites in a polymer.

5. At every step, each of the N sites in the polymer population mutates with equal probability $(1/N)$.

6. In a given population of polymers, the bound monomers can be divided into two classes, active and inactive.

7. The active monomers are in sites where they contribute to the ability of a polymer to act as an enzyme. To act as an enzyme means to catalyze the mutation of other polymers in a selective manner so that the correct species of monomer is chosen preferentially to move into a site that is active.

8. In a cell with a fraction x of monomers active, the probability that a monomer inserted by a fresh mutation will be active is $\phi(x)$. The function $\phi(x)$ represents the efficiency of the existing population of catalysts in promoting the formation of a new catalyst. The assumption that $\phi(x)$ depends on x means that the activity of catalysts is to some extent inherited from the parent population by the newly mutated daughter. The form of $\phi(x)$ expresses the law of inheritance from parent to daughter. The numerical value of $\phi(x)$ will be determined by the details of the chemistry of the catalysts.

9. The curve $y = \phi(x)$ is S shaped, crossing the line $y = x$ at three points, $x = \alpha, \beta, \gamma$ between zero and one.

10. Here we make a definite choice for the function $\phi(x)$ based on simple thermodynamics argument. It turns out that the function $\phi(x)$ derived from thermodynamics has the desired S shaped form to produce the three equilibrium states required by Assumption 9.

In the subsequent sections it will become apparent that some of these assumptions are unnecessary and are inherent in the REM version of the theory that we develop.

III. THE REM THEORY

The structure of a primitive cell or droplet consists of a mixture of free monomers and polymers of different sizes. According to assumption 6, each monomer has two nearly degenerate states that are separated by an energy that is different for different monomers. In our model, the monomers can be formally treated as Ising spins and we adopt the following notation to describe the state of the monomers — the ith monomer in its active form has an energy $-\epsilon_i$ and is in the state $\sigma_i = +1$; the ith monomer in its inactive form has an energy $+\epsilon_i$ and is in the state $\sigma_i = -1$. Within a droplet, the number and size of polymers are not fixed, but fluctuate among local minima under the influence of thermal effects and interactions among monomers. The interactions between the ith and jth monomers have a strength $-J_{ij}$ that we assume is a random variable whose distribution depends on the droplet. In this model, the energy expression for the droplet is

$$E = -\sum_i \epsilon_i \sigma_i - \sum_{\langle i,j \rangle} J_{ij} \sigma_i \sigma_j, \tag{1}$$

where the values of ϵ_i are Gaussian distributed with mean $\langle \epsilon_i \rangle = \bar{\epsilon}$ and fluctuations $\langle (\epsilon_i - \bar{\epsilon})^2 \rangle = \Delta\epsilon^2$. The fluctuations arise from droplet disorder. Similarly, we assume that the monomer interactions J_{ij} are also Gaussian distributed with mean $\langle J_{ij} \rangle = \overline{J}$ and fluctuations $\langle (J_{ij} - \overline{J})^2 \rangle = \Delta J^2$. Typically each monomer interacts on average with z other monomers.

The key element in this energy that makes the system spin-glass-like is the property of "frustration," which means that there are many unfavorable interactions because all favorable interactions cannot be satisfied simul-

taneously. As a consequence, the state becomes highly degenerate and there are a large number of different local energy minima, which may have similar energies even if they are not similar in structure. The number of such substates increase exponentially ($\simeq \exp\{0.2N\}$ for Ising systems) with the number of monomers.

Suppose that there is a droplet with N interacting monomers that is in a specific conformation with xN monomers in their active state. The mean energy for such a system is then

$$\overline{E}(x) = E_0 - (2\overline{\epsilon} - 4z\overline{J})xN - 4z\overline{J}x^2N; \qquad E_0 = N(\overline{\epsilon} - z\overline{J}). \tag{2}$$

For the standard deviation of the droplet energy, we assume the simplest model possible (in which the standard deviation is independent of x) and write it as

$$\Delta E^2 = N\left[\Delta\epsilon^2 + z\Delta J^2\right]. \tag{3}$$

The number of states with xN active monomers is given by

$$\Omega(xN) = \frac{N!}{(xN)![(1-x)N]!} \tag{4}$$

and so the configurational entropy, which is defined as $S^*(x) = Nk_B \ln[\Omega(xN)]$, can be written in the thermodynamic limit ($N \to \infty$) as

$$S^*(x) = -Nk_B\left\{x\ln(x) + (1-x)\ln(1-x)\right\}, \tag{5}$$

where k_B is the Boltzmann constant.

Using the thermodynamic relation for the temperature T, $\partial S/\partial E = 1/T$, it can be shown that the energy may be written as [15,11]

$$E(x) = \overline{E}(x) - \frac{\Delta E^2}{k_B T} \tag{6}$$

and the entropy as

$$S(x) = S^*(x) - \frac{k_B}{2}\left(\frac{\Delta E}{k_B T}\right)^2. \tag{7}$$

The critical temperature $T_c(x)$ is the temperature at which the entropy goes to zero. It is

$$T_c(x) = \frac{\Delta E}{[2k_B S^*(x)]^{1/2}} \tag{8}$$

and it attains its lowest value, $T_{c,\min} = \Delta E/k_B\sqrt{2N\ln 2}$, at $x = 1/2$. At high temperatures $T > T_c(x)$, the energy and the entropy are approximatively equal to the mean energy and the configurational entropy, respectively. For $T < T_c(x)$, the entropy of the droplet remains zero because it is trapped in one of its low-energy substates.

At high temperature $T > T_c(x)$, or for values of x such that $T > T_c(x)$, the free energy $F(x) = E(x) - TS(x)$ is given by the expression

$$F(x) = F_0 - (2\overline{\epsilon} - 4z\overline{J})xN - 4z\overline{J}x^2N - TS^*(x); \qquad F_0 = E_0 - \frac{\Delta E^2}{2k_B T}. \tag{9}$$

The free energy $F(x)$ may be analyzed in terms of an effective potential $U(x)$, which is defined as the sum of two terms:

$$U(x) = \frac{F(x) - F_0}{Nk_B T} = U_1(x) + U_2(x),$$

where

$$U_1(x) = x\ln(x) + (1-x)\ln(1-x), \tag{10}$$

$$U_2(x) = Ax - \frac{B}{2}x^2 = -\frac{B}{2}\left(x - \frac{A}{B}\right)^2 + \frac{A^2}{2B},$$

with

$$A = \frac{4z\overline{J} - 2\overline{\epsilon}}{k_B T} \quad ; \quad B = \frac{8z\overline{J}}{k_B T}. \tag{11}$$

The first term is the entropic potential, which is greatest as x approaches zero or one, and so favors large populations of active or of inactive monomers. The second term is a conventional potential energy and is the energy gained by activating xN interacting monomers. For our purposes, the importance of $U(x)$ lies in the fact that the equilibrium distribution of the population $P_e(x)$, in a droplet having xN monomers active, must be a Boltzmann distribution:

$$P_e(x) \sim \exp\left\{-\frac{F(x) - F_0}{k_B T}\right\} = \exp\{-NU(x)\}. \tag{12}$$

Now that we have defined an effective potential for the population of molecules in the droplet, we can investigate transitions between populations and in particular between disordered (*dead*) and ordered (*alive*) states, which is the idea that underlies the Dyson model and is contained in his assumption 9. To do this, we need to determine the conditions of bistability for $U(x)$.

A. Bistability conditions for the effective potential $U(x)$

Since $U_1(x) < 0$ on the whole interval $0 \le x \le 1$, $U(x)$ will have two minima at stable points and a maximum at an unstable one if $U_2(x) > 0$ on the same interval. This imposes the condition that $A > B/2$. Because of its two minima, $U(x)$ has to have in addition two points of inflection (points where the second derivatives are zero) provided $B > 4$. It follows that the conditions for bistability of $U(x)$ are:

$$A > A_c = 2 \text{ and } B > B_c = 4, \quad \text{or} \quad \overline{\epsilon} < \overline{\epsilon}_c = 2\overline{J} - k_B T \text{ and } \overline{J} > \overline{J}_c = k_B T/2z. \tag{13}$$

B. Extrema and points of inflection of $U(x)$

When the above conditions are fulfilled, the first derivative of $U(x)$ cancels at three points α, β, γ ($\alpha < \beta < \gamma$), which are solutions of the equation

$$\frac{1}{1+\exp\{A - Bx\}} = x. \tag{14}$$

α and γ are the stable points corresponding to the disordered and ordered states, respectively, while β is the unstable point that corresponds to the potential barrier separating the two stable points. The points of inflection, obtained by equating the second derivative of $U(x)$ to zero, are given by

$$x_{\pm} = \frac{1}{2}\left\{1 \pm \sqrt{1 - \frac{B_c}{B}}\right\}. \tag{15}$$

C. Marginally disordered and ordered states

Two interesting situations appear when the parameters (A, B) increase from the cusp at A_c, B_c where $x_+ = x_- = 1/2$. These marginal situations arise when one of the points of inflection, say x_- (x_+), coincides with one of stable points, say α (γ). In this case, $x_{\pm}$ are also solutions of Eq. (14) and the maximum of $U(x)$ coincides with one of its minima. Let

$$B = B_c \cosh^2 \frac{\theta}{2}. \tag{16}$$

The range of values of A which allow an order-disorder transition, is given by

$$A_- \leq A \leq A_+, \tag{17}$$

where

$$A_{\pm} = 1 \mp \theta + \exp(\pm\theta). \tag{18}$$

Therefore, we call *marginally disordered* a system having $A = A_-$ and corresponding to

$$\alpha = \beta = \frac{1}{1+\exp(\theta)}\,,\ \gamma. \tag{19}$$

The *marginally ordered* state corresponds to $A = A_+$ with

$$\alpha\,,\ \beta = \gamma = \frac{1}{1+\exp(-\theta)}. \tag{20}$$

The *symmetric* situation is obtained for $2A = B$. This correspond to

$$\alpha,\ \beta = \frac{1}{2}\,,\ \gamma = 1 - \alpha \tag{21}$$

and the potential $U(x)$ has equal minima at α and γ. For these reasons, the average energy $\bar{\epsilon}$ can be called the asymmetric energy since the symmetric case is obtained for $\bar{\epsilon} = 0$.

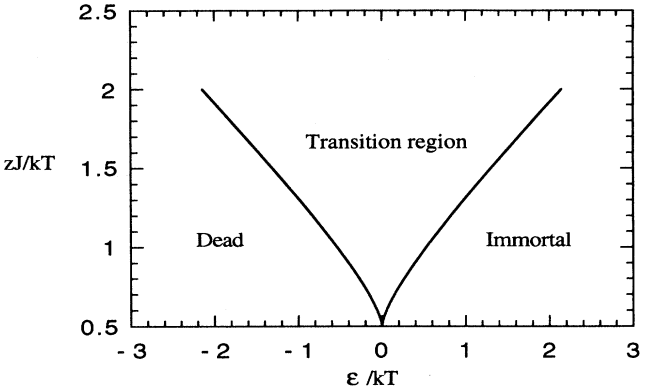

FIG. 1. Phase diagram in the space $(\bar{\epsilon}, z\bar{J})$ representing different possibilities for the population of a droplet. In the transition region (dead and alive coexisting), populations possess both ordered and disordered states. There are no ordered and disordered states in the dead and immortal regions, respectively.

Figures 1 and 2 display the phase diagram and the effective potential for the two marginal and the symmetric situations. The marginally ordered and disordered states are interesting because they are limit systems having the weakest and strongest capabilities for an order-disorder transition, respectively, with a given number of monomers species. In the marginally ordered system, there is no stable ordered state while the situation is reversed for the marginally disordered system. For these reasons, these states can be considered to be *dead* and *immortal,* respectively. All intermediate cases can be represented as a nonsymmetrical combination of dead and

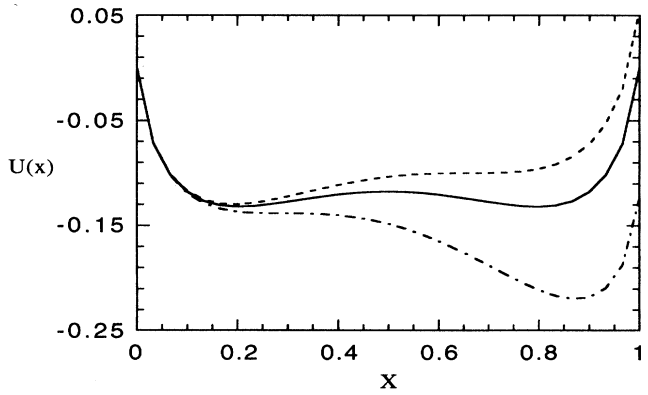

FIG. 2. The effective potential $U(x)$ for the three situations — marginally ordered (dead systems), symmetric, and marginally disordered (immortal systems). The corresponding values of the energies are $\bar{\epsilon} = -0.028 k_B T$, $z\bar{J} = 0.562 k_B T$ for dead systems (dashed line); $\bar{\epsilon} = 0.0$, $z\bar{J} = 0.576 k_B T$ for symmetric systems (solid line); and $\bar{\epsilon} = 0.062 k_B T$, $z\bar{J} = 0.607 k_B T$ for immortal systems (dot-dashed line).

immortal states and it is the interactions between them that allow order-disorder transitions.

We would like to determine how long it takes for a population to switch spontaneously from a dead state to an alive state. In other words, if we divide the 2^N different substates of a droplet into two disjoint sets: $\mathcal{D}$ the dead states, and $\mathcal{A}$ the alive states. We are interested in finding the time required for a polymer starting in a typical state in $\mathcal{D}$ to be transferred towards one in $\mathcal{A}$. At high temperature, $T > T_c$, the mean first passage time is given by the classical Arrhenius law

$$\tau \simeq \frac{2\pi}{\sqrt{F''(\alpha)F''(\beta)}} \exp\left\{\frac{F(\beta)-F(\alpha)}{k_BT}\right\}, \tag{22}$$

where $F''(\cdots)$ is the second derivative of F. By introducing the height of the barrier defined as

$$\Delta = U(\beta) - U(\alpha), \tag{23}$$

the average time required for a droplet to make the transition from dead to alive can be rewritten as

$$\begin{aligned}\tau &\simeq \frac{2\pi}{Nk_BT}\left\{\frac{\alpha(1-\alpha)\,\beta(1-\beta)}{[1-B\alpha(1-\alpha)]\,[1-B\beta(1-\beta)]}\right\}^{1/2} \\ &\quad\times \exp\{\Delta N\} \\ &= \tau_0 \frac{\exp\{\Delta N\}}{N},\end{aligned} \tag{24}$$

where τ_0 can be interpreted as the average transfer time between microstates at a given site. A critical population size N_c can now be defined within which a dead-alive transition can occur with a reasonable probability. N_c is such that the average number of transitions is equal to a certain threshold

$$\frac{N_c\tau}{\tau_0} = \exp\delta \iff N_c = \frac{\delta}{\Delta}. \tag{25}$$

The energy dependence of N_c is depicted in Fig. 3. It is apparent that N_c drastically increases as $\overline{J}$ tends to $\overline{J}_c$.

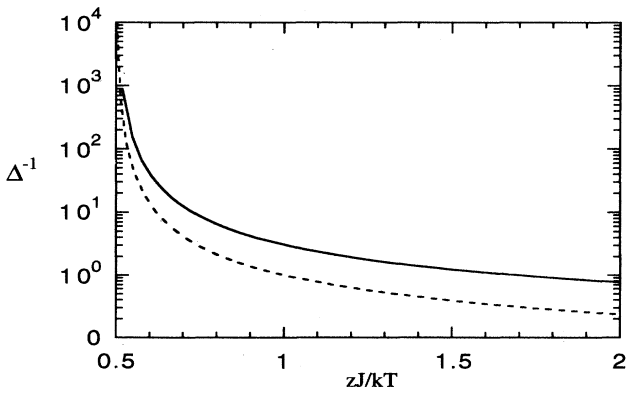

FIG. 3. The critical population size N_c ($\delta = 1$) as a function of the reduced interaction energy $z\overline{J}/k_BT$ for symmetric systems (solid line) and dead systems (dashed line). By definition, N_c is infinite for immortal systems.

IV. THE AUTOCATALYTIC PROBABILITY $\phi(x)$

To make the correspondence between the REM theory and the model of Dyson, we need to calculate the autocatalytic probability for the REM model. Let us consider a droplet containing N monomers. The state of our droplet system is given by the number $j = 0, 1, \ldots, N$ of monomers active in the droplet after k mutations. The variation of $j(k)$ corresponds to a discrete one-step Markov process $j \to j+1, j \to j-1$. As a consequence, the daughter population and its ancestors are not correlated. This fact is consistent with the REM, which assumes a fitness landscape constructed from independent random variables. Let $P(j,k)$ be the probability of finding j active monomers in the population after k mutation events have occurred. The master equation that describes the evolution of the population is

$$\begin{aligned}\partial_k P(j,k) &= t^+_{j-1}P(j-1,k) + t^-_{j+1}P(j+1,k) \\ &\quad -(t^+_j + t^-_j)P(j,k), \\ t^+_{-1} &= t^-_0 = 0.\end{aligned} \tag{26}$$

t^+_j and t^-_j are the transition probabilities per unit of mutation from j to $j+1$ and from j to $j-1$, respectively. We follow Dyson [23] and use the transition probabilities

$$\begin{aligned}j \to j+1 &: t^+_j = (1-x)\phi(x), \\ j \to j-1 &: t^-_j = x[1-\phi(x)],\end{aligned} \tag{27}$$

$x = j/N$, where we have introduced the probability $\phi(x)$ that the mutated unit be active in a droplet that already contains j monomers active. In other words, $\phi(x)$ describes the law of inheritance from parent to daughter. In this section, we determine the form of this probability without making any assumptions.

$t^+_j + t^-_j$ is the overall transition probability that the system will leave the state j. By choosing a random number in the interval $[0, t^+_j + t^-_j]$, we can realize the stochastic motion of the phase point $j(k)$, which undergoes a random walk through the phase space $(0, 1, \ldots,)$, which is governed by the transition probabilities. The mean number of mutations the population undergoes in the state j is

$$k_j = \frac{1}{t^+_j + t^-_j}. \tag{28}$$

The steady state is attained when the current flux $I(j)$ is zero:

$$I(j) = t^-_j P(j) - t^+_{j-1}P(j-1) = 0 \tag{29}$$

and so the stationary solution of the master equation is given by

$$P(j) = P(0)\prod_{l=1}^{j}\frac{t^+_{l-1}}{t^-_l} = P(0)\exp\left\{-\sum_{l=1}^{j}\ln\left(\frac{t^-_l}{t^+_{l-1}}\right)\right\}. \tag{30}$$

Using the Euler-MacLaurin summation formula,

$$\sum_{l=1}^{j} \ln\left(\frac{t_l^-}{t_{l-1}^+}\right) = \int_1^j \ln\left(\frac{t_l^-}{t_{l-1}^+}\right) dl + \frac{1}{2}\left[\ln\left(\frac{t_j^-}{t_{j-1}^+}\right) - \ln\left(\frac{t_1^-}{t_0^+}\right)\right] + \cdots \tag{31}$$

and in the limit of large N ($xN \gg 1$), we get

$$P(x) \simeq P(0)\,[Nu(x)]^{-1/2} \exp\{-NU(x)\}, \tag{32}$$

where

$$u(x) = \frac{x[1-\phi(x)]}{(1-x)\phi(x)} \frac{\phi(0)}{1-\phi(0)} \tag{33}$$

and the potential

$$U(x) = \int_0^x \ln\left(\frac{y}{1-y}\right) dy + \int_0^x \ln\left(\frac{1-\phi(y)}{\phi(y)}\right) dy. \tag{34}$$

At this stage we can compare the above expression for the effective potential with the one given in Eq. (10). The first term corresponds to the entropic potential and the second one is such that

$$\int_0^x \ln\left(\frac{1-\phi(y)}{\phi(y)}\right) dy = U_2(x). \tag{35}$$

The mathematical expression for the probability $\phi(x)$ is therefore given by

$$\phi(x) = \frac{1}{1+\exp\left\{\frac{\partial U_2(x)}{\partial x}\right\}} = \frac{1}{1+\exp\{A-Bx\}}. \tag{36}$$

This expression for $\phi(x)$ is exactly the same that Dyson uses in his model. Before going further, let us make three remarks:

(i) Within the mean field approximation, the potential $U(x)$ stems from the summation of two terms of entropic and energetic origins. The mean field approximation means that the efficiency of the activity of the catalyst, or the probability that the mutated unit be active, depends only on the total number of active monomers present and not on their detailed arrangement nor on the population before the mutation event.

(ii) As a consequence, the functional form of the autocatalytic probability $\phi(x)$ does not depend on the entropic term but is completely determined by the potential energy and is reminiscent of the transition probability for the importance sampling procedure used in Monte Carlo simulations:

$$\phi(x) = \frac{1}{1+\exp\left\{\frac{F(x+1/N)-F(x)}{k_BT}\right\}} \cong \frac{1}{1+\exp\left\{\frac{\partial U_2(x)}{\partial x}\right\}}. \tag{37}$$

Thus, the law of inheritance from parent to daughter is governed by the nature of the interactions between the monomers. It is remarkable to note that the probability $\phi(x)$ is typical of the sigmoidal functions that are used as input-output functions in neural networks, although, in this analogy, A represents the threshold and B the synaptic efficiency.

(iii) Assumptions 9 and 10, which require that the daughter population has the same average activity as the parent population at three points, is no longer necessary. Indeed, the S shape of $\phi(x)$ stems from the bistability of the potential $U(x)$ and the corresponding three crossing points of $\phi(x)$ are points where the first derivative of $U(x)$ is zero, i.e., where both the daughter and parent populations are equal on average and have the same characteristics.

To make the model more concrete, Dyson introduced two parameters a and b that specified the diversity of the population of monomers ($1+a$ is the number of monomer species) and the precision of the polymerizing catalysts, respectively. Dyson showed that when one assumes that every imperfect catalyst produces an energy-lowering proportional to x, the probability $\phi(x)$ takes the form

$$\phi(x) = \frac{1}{1+a\,b^{-x}}. \tag{38}$$

In terms of the parameters in the REM, it follows that

$$a = \sqrt{b}\,\exp\left\{-\frac{2\bar{\epsilon}}{k_BT}\right\}, \quad b = \exp\left\{\frac{8z\bar{J}}{k_BT}\right\}, \tag{39}$$

where $z\bar{J}$ is interpreted as the average energy by which the catalyst lowers the activation energy for the correct placement of a monomer. It is clear from these formulas that the diversity a of the monomer population and the discrimination factor b are intimately related. In the symmetric case, for example, just a single quantity, the average monomer-monomer interaction energy $z\bar{J}$ controls the system and a given number of monomer species fixes the precision of the polymerizing catalysts. This one-to-one relation between the diversity and the polymerization precision (sloppiness) implies a feedback control and is an important requirement for homeostatic equilibrium. For asymmetrical models ($\bar{\epsilon} \neq 0$), a system having a given sloppiness is able to tune its diversity in the allowed interval $A_- \leq A \leq A_+$ in order to increase or decrease its size N_c in the alive state.

Table I lists values for the parameters of some representative systems that were studied by Dyson [23]. It is to be noted that when $\bar{\epsilon} \ll z\bar{J}$, the monomers are strongly coupled and the model exhibits a large number of configurational substates, i.e., a rugged potential energy landscape. It should also be noted that values of $z\bar{J}$, which are consistent with having a reasonable probability for an order-disorder transition, are comparable to the values of the pairwise interactions between amino acids in proteins [24] and are of the order of a few percent of the thermal energy, k_BT. This means that such systems are flexible and adaptable, characters typical of metabolism, but less so of replication, which is less error tolerant. It would be instructive to investigate how the three systems dead, symmetric, and immortal individually relax to the

TABLE I. Typical values for the parameters of some representative models. The letters $\mathcal{D}$, $\mathcal{S}$, and $\mathcal{I}$ stand for dead, symmetric, and immortal systems, respectively.

	a	b	α	β	γ	$\bar{\epsilon}/k_BT$	$z\bar{J}/k_BT$	Δ^{-1}
$\mathcal{D}$	8	62.9	0.32	0.59	0.59	− 0.0043	0.518	371
$\mathcal{S}$	8	64	0.33	0.50	0.67	0.00	0.520	886
$\mathcal{I}$	8	65.7	0.39	0.39	0.70	0.0066	0.523	∞
$\mathcal{D}$	10	89.4	0.19	0.67	0.67	− 0.0280	0.562	34
$\mathcal{S}$	10	100.0	0.20	0.50	0.80	0.00	0.576	69
$\mathcal{I}$	10	128.0	0.29	0.29	0.87	0.0617	0.607	∞
$\mathcal{D}$	19	219.3	0.07	0.75	0.75	− 0.1246	0.674	5
$\mathcal{S}$	19	361.0	0.08	0.50	0.92	0.00	0.736	10
$\mathcal{I}$	19	3195	0.14	0.14	0.99	0.545	1.01	∞

equilibrium state as a function of the number of mutation events. Such dynamical studies have already been performed in other contexts by, for example, Shakknovich and Gutin [13], and Fernandez [14]. We will report the results of our application of these techniques elsewhere.

V. DISCUSSION

In this paper we have shown how the REM can be used as a theoretical framework for the model of the origin of metabolism developed by Dyson. This has been done by defining the effective potential of an interacting system of monomers and by rederiving the expression for the autocatalytic probability given by Dyson without any further assumptions.

We wish to conclude with a brief discussion on whether this same framework can be used to model replication as well as metabolism. Metabolism is based on homeostatic equilibrium. Homeostasis, which is the ability of a system to maintain a constant internal environment in a changing external environment, tends to maximize diversity of structure and flexibility of function. On the other hand, replication requires both diversity and stability for carrying macromolecular information. These characteristics can be modeled in spin-glass systems because they have diversity (many local minima), stability (barriers preventing escape from a single state), and flexibility (degenerate ground states). The origin of life is certainly a complicated process with conflicting situations, incidents, and adaptations working at different time scales so that the fitness function is far from a trivial one and must exhibit a rugged landscape. The concept of spin glass as a model of the transition to biological order has already been employed by Anderson [25], Stein and Anderson [26], Rokhsar *et al.* [27] and Amitrano *et al.* [28] to model the template replication of nucleic acids. Abbott [29] has used spin-glass systems to show that an autocatalytic system, as used by Dyson [16,23], is capable of replication without templating. Replication in such a system arises from the ability of the system to store and recover information as is done in neural networks. Therefore, it seems that the spin-glass analogy can give useful insights into both metabolism and replication and be used to incorporate them into the same theoretical framework. This is work we are currently engaged in.

ACKNOWLEDGMENTS

We thank the Institut de Biologie Structurale — Jean-Pierre Ebel, the Commissariat à l'Energie Atomique, and the Centre National de la Recherche Scientifique for support of this work.

[1] *Spin Glasses and Biology*, edited by D. L. Stein (World Scientific, Singapore, 1992).
[2] P. Peretto, *An Introduction to Modeling of Neural Networks.* (Cambridge University Press, Cambridge, 1992).
[3] A. Ansari, J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, and R. D. Young, Proc. Natl. Acad. Sci. U.S.A. **82**, 5000 (1985).
[4] D. L. Stein, Proc. Natl. Acad. Sci. U.S.A. **82**, 3670 (1985).
[5] H. Frauenfelder and P. G. Wolynes, Phys. Today **47**, 58 (1994).
[6] B. Derrida, Phys. Rev. B **24**, 2613 (1981).
[7] J. D. Bryngelson and P. G. Wolynes, J. Phys. Chem. **93**, 6902 (1989).
[8] E. I. Shakknovich and M. A. Gutin, Nature **346**, 773 (1990).
[9] J. Bascle, T. Garel, and H. Orland, J. Phys. (France) I **3**, 259 (1993).
[10] M.-H. Hao and H. A. Scheraga, J. Phys. Chem. **102**, 1334 (1995).
[11] R. D. Young and S. W. Powell, J. Chem. Phys. **101**, 9919 (1994).
[12] R. D. Young, J. Chem. Phys. **98**, 2488 (1991).
[13] E. I. Shakknovich and M. A. Gutin, Europhys. Lett. **9**, 569 (1989).
[14] A. Fernandez, Phys. Rev. Lett. **65**, 2259 (1990); Phys. Rev. A **45**, R8348 (1992).
[15] J. N. Onuchic and P. G. Wolynes, J. Chem. Phys. **98**, 2218 (1993).
[16] F. J. Dyson, *Origins of Life* (Cambridge University Press, Cambridge, 1985).
[17] E. Schrödinger, *What is Life? The Physical Aspect of the Living Cell* (Cambridge University Press, Cambridge, 1944); French edition; *Qu'est ce que la vie? De la*

Physique à la biologie (Edité par C. Bourgois, Sciences, Paris, 1986).

[18] J. Von Neumann, in *Cerebral Mechanisms in Behavior—The Hixon Symposium*, edited by L. A. Jeffress (John Wiley, New York, 1951), pp. 1–41.

[19] A. I. Oparin, *The Origin of Life on the Earth*, 3rd ed., translated by Ann Synge (Oliver and Boyd, Edinburg, 1957).

[20] M. Eigen, W. Gardiner, P. Schuster, and R. Winckler-Oswatitch, Sci. Am. **244**, 88 (1981).

[21] A. G. Cairns-Smith, *Genetic Takeover and the Mineral Origins of Life* (Cambridge University Press, New York, 1982).

[22] H. J. Morowitz, *Beginnings of Cellular Life: Metabolism Recapitulates Biogenesis* (Yale University Press, New Haven, CT, 1992).

[23] F. J. Dyson, J. Mol. Evol. **18**, 334 (1982).

[24] A. Kolinsky, A. Godzik, and J. Skolnick, J. Chem. Phys. **98**, 7420 (1993), see especially Table III.

[25] P. W. Anderson, Proc. Natl. Acad. Sci. U.S.A. **80**, 3386 (1983).

[26] D. L. Stein and P. W. Anderson, Proc. Natl. Acad. Sci. U.S.A. **81**, 1751 (1984).

[27] D. S. Rokhsar, P. W. Anderson, and D. L. Stein, J. Mol. Evol. **23**, 119 (1986).

[28] C. Amitrano, L. Peliti, and M. Saber, J. Mol. Evol. **29**, 114 (1989).

[29] L. F. Abbott, J. Mol. Evol. **27**, 513 (1988).