A FISSION-FUSION ORIGIN FOR LIFE

V. NORRIS
IFR 'Systems Integres' Laboratoire de Microbiologie Faculté des Sciences et Techniques de Rouen F76821 Mont Saint Aigan, France

D. J. RAINE*
Department of Physics and Astronomy, University of Leicester LE1 7 RH. U.K.

(Received 15 August, 1996)

Abstract. To develop a comprehensive 'cells-first' approach to the origin of life, we propose that protocells form spontaneously and that the fission and fusion of these protocells drives the dynamics of their evolution. The fitness criterion for this evolution is taken to be the the stability (conservation) of domains in the protocellular membrane as determined by non-covalent molecular associations between the amphiphiles of the membrane and a subset of the macromolecules in the protocell. In the presence of a source of free energy the macromolecular content of the protocell (co-)evolves as the result of (domain-dependent) membrane-catalysed polymerisation of the prebiotic constituents delivered to the protocell by fusion. The metabolism of the cell therefore (co-)evolves on a rugged fitness landscape. We indicate how domain evolution with the same fitness criterion can potentially give rise to coding. Membrane domains may therefore provide the link between protocells and the RNA/DNA-world.

1. Introduction

At what point in the evolution of life does the physical structure of precursor life-forms first appear and play a role? Is it, as in the RNA-first approach, at a point where the chemistry of replication is well advanced? Or is it, as in the 'cells-first' approach, at an earlier stage? And, if so, does the pre-cellular physical structure have only a passive role, as in Oparin's coercervate theory (Oparin, 1924), where the 'cell' is a colloidal aggregation which serves simply to confine the chemical evolution? Or, as Bernal (1949) and Needham (1968) were perhaps first to suggest, is the spontaneous formation of physical structure crucial to the development of life? The main argument in favour of a cells-first approach appears to be the efficiency of selection operating on spatially distinct systems (Koch, 1985) compared with the direct effect of selection on polymers in solution. Yet, whereas in the latter case the evolution can be understood in terms of a well-developed theory of optimisation on a fitness landscape (Kauffman, 1993), for the cells-first picture this remains to be worked out. The point in evolution at which cells emerged is unlikely to be determined by experiment. Therefore we must rely on theory to explain the emergence of template-directed synthesis and replication together with the minimal functionality of the cell. (To lift what Luisi (1993) calls the generous Darwinian fog.) In this paper we present an initial attempt to delineate some features of a possible theory.

* Author for all correspondence: e-mail jdr@leicester.ac.uk.

Origins of Life and Evolution of the Biosphere **28:** 523–537, 1998.

Several authors have considered the protocell as the starting point for the evolution of life. Deamer (1986), for example, invokes lipid vesicles as the primitive cell with diffusion into the cell and fission as the mechanisms of growth and evolution. This approach has to overcome several problems. First, it is not at all clear that lipids with sufficiently long fatty acid chains to form spontaneous structures can be obtained pre-biotically (Ferris, 1987). However, Deamer and Pashly (1989) argue that the amphiphilic components of the Murchison meteorite are surface-forming (hence that this is not a problem in Nature). Second, at low concentrations of lipids, micelle formation is preferred to that of vesicles. Perhaps this indicates a 'micellular period', prior to that to be considered here, but in which evolution would proceed in an analogous way. Alternatively, conditions of pH (via dissolved CO_2) or Ca^{++} concentration may favour the spontaneous assembly of vesicles. Bachman *et al*. (1992) start from a state in which the protocellular surfactant can catalyse the reactions necessary for its own spontaneous assembly and growth. A different approach is to take the first cellular 'organisms' to be proteinoid spheres (Fox, 1965).

The cell membrane acts as a barrier to diffusion, thereby retaining reaction products in the cell. But this implies that diffusion into the cell is also reduced. It may therefore be that the *external* surface of the protocell was the initial site of metabolism (Cavalier-Smith, 1987) and hence growth. An alternative mechanism of growth is through the fusion of vesicles (Baeza *et al*., 1994) which is what we invoke here.

A central problem in early evolution is the source of free energy. The free-energy requirements are much reduced if polymerisation reactions take place on surfaces (Cairns-Smith, 1982). Wachtershauser (1988) considers the first step to life to be surface 'organisms'. Much of his discussion is relevant to our theory. (But whether the particular reaction schemes suggested by Wachtershauser could or did work (Keefe *et al*., 1996) is not, of course, important here.) Such organisms may have been responsible for the production of the significant concentrations of amphiphiles, which we take as our starting point. (This would provide a further alternative solution to the problem of obtaining a high enough concentration.) We differ in that we suggest that the cell membrane rather than an inorganic surface plays a significant role in the subsequent evolution. Many of Wachtershauser's arguments for the inorganic surface over free solution can be advanced in favour of cell surface chemistry. It is at least conceivable then that prebiotic synthesis provides an input of sufficiently high free-energy 'food', provided the flow maintains the growing system, here the protocell, away from equilibrium. We shall see that here too the protocell surface plays a crucial role in preserving macromolecules from degradation and poisoning, hence shifting the equilibrium points of selected reactions.

Now, we know that compartmentalisation alone is not a useful condition for the origin of life (e.g. Eigen and Schuster, 1974). Our hypothesis is that the spontaneous

formation of closed compartments by surfaces *having a constituent-dependent domain structure* is a key evolutionary step.

In the case of a lipocell (see below) we know that in general a mixture of lipids will lead to such domain structures (e.g Mouritsen and Jørgensen, 1994). The domain structure will also be influenced by the presence of polymers in the cell interacting with the surface by, for example, electrostatic attraction. Conversely the domain structure will affect the catalysed polymerisation reactions on the cell surface and the changing polymer content of the cell will react back on the domain structure. (Comparison of the surface domain conformations to the structural conformations of modern enzymes is an intuitively compelling analogy.) In addition, the domain structure may control the fission (and fusion) characteristics of the cell.

We therefore arrive at a picture of the complex co-evolution of a polymerisation reaction network and the domain structure of the protocell surface which is determined by the stability of domains during fission (and fusion). The stability is brought about by the creation of (evolving) sets of polymers which mutually associate in the presence of the cell membrane. We shall argue that evolution occurs on a rugged fitness landscape and therefore leads to focusing of the main reaction products on to a small subset of bio-macromolecules. For the sake of argument we take this domain stabilisation to be derived from the mutual association of the lipid head groups, in the lipocell surface, with peptides, but it could equally involve other constituents of modern cells such as nucleotides polyhydrxybutyrates, polyamines and polyphosphates.

There is one further stage to be considered. A cells-first theory of the origin of life has been associated with a double-origin theory (Dyson, 1985). That is, the origin of cells is considered to be independent of the origin of their multiplication or self-replication: the RNA-world happens elsewhere and takes over the cell. This double origin seems to us to be a somewhat unsatisfactory theory. Rather, we should hope that replication emerges naturally from the evolutionary dynamics of the proto-cells. It is therefore of interest to examine whether there is any route by which this might *in principle* occur in our picture, independently of whether we believe this to be even a plausible route. This is the final stage of evolution to be considered here.

At this stage a new fitness criterion for the evolution emerges: any lipocells which can undergo fission to yield approximate copies of their domain structure have in effect a selective advantage for the accumulation of prebiotic molecules. Once this can happen therefore, the fitness criterion changes from a domain structure that affords stability to one that confers (approximate) *reproducibility* on the lipocell. This is a primitive form of coding because it preferentially produces those macromolecules that form stable associations down the generations. (These could be peptides and nucleotides.) Moreover, this operates in a hypercyclic structure and hence does not suffer from the error catastrophe (Eigen, 1971, see also Joyce 1991). In principle, therefore, in this theory, the first protocells may also provide the first step along the route to genetic evolution.

2. The Model

We shall use the term ‘protocell’ to apply indiscriminately to the first physically ordered structures; so this includes aggregates, micelles, liposomes and vesicles. We shall use the term ‘lipocell’ to refer to protocells in which the physical structure is a lipid vesicle with the bilayer stabilised by attachment of monomeric or polymeric compounds (e.g. amino acids or peptides).

We set up a simple model to illustrate first how macromolecular associations giving rise to relatively stable membrane domains can lead to focusing. We then consider the evolution of the lipocell metabolism and the evolution of the membrane domain structure in the lipocell population.

The system consists of

(i) amphiphilic, generally lipidic, molecules of various types (A, B...);
(ii) substrate, of possible types (u, v, ...) which may be monomers or short chain polymers and not all of which will be present initially. In referring to ‘chains’ we shall include 1-link chains (monomers) and short polymers. This class includes nucleotides and amino acids.

There may also be

(iii) metal ions.

Some subsets of the amphiphiles present spontaneously form protocells. The type of protocell (micelle or vesicle) will depend on amphiphilic concentration, abundance of other surfactants, temperature etc. The sizes and compositions of lipocells will determine their probabilities (or, equivalently their half-lives) for undergoing fission and fusion.

Crucial to the form of evolution is the notion of domains. A mixture of surfactants may spontaneously segregate according to the strengths of interaction amongst the different types (Tocanne *et al*., 1994). Such domains are known to be important in the membrane functions of contemporary cells where domain formation can depend on the presence not only of different lipid types but on other membrane-bound molecules such as steroids and proteins (Glaser, 1993). The domains of interest to us constitute chemically distinct regions that are relatively stabilised against disruption through fission of the protocell. Such structures are therefore relatively conserved by the dynamics of fission and fusion. Furthermore, we shall see that the domain structure is subject to change: thus the system is able to grow by the redistribution of molecular resources.

Our model therefore has the following minimal set of requirements:

(iv) primitive bilayers contain a domain-forming mixture of molecular species (A’, B’...) drawn from (i);

(v) domains can be further created and/or stabilised by non-covalent binding to macromolecules and their constituents (u', v'...) drawn from (ii);
(vi) such binding can stabilise domains against disruption by external mechanical forces (drying, mechanical agitation etc.) or, conversely, stabilise the associated molecules against degradation, or both. Domains can also be stabilised in this way against disruption by fission (and fusion).

Some chains polymerise. This is a crucial assumption and, as usual, raises the problem of the source of free energy. It is possible that one source could be the surface domain structure itself. However, the solution of this problem need not be specific to this theory. We might invoke prebiotic condensing agents (e.g. Oro *et al.*, 1984) or activated intermediates (Ferris *et al.*, 1990, 1996; Bohler *et al.*, 1996). Activated precursors have recently been used as part of a thioester-promoted condensation strategy to drive the self-replication of peptides (Lee *et al.*, 1996). The polymerisation may be weakly catalysed by the surface amphiphiles possibly with the help of associated chains and embedded metal ions. The catalysis of reactions on membranes may also proceed differently than in the bulk phase (Kinnunen *et al.*, 1994) and may be dependent on aggregation in the membrane (e.g. Melo *et al.*, 1992).

To be effective in producing a significant number of long chain polymers these must be removed from the catalytic site once formed (otherwise the only effect of catalysis is to bring the chains more rapidly into equilibrium with their constituent parts.) Thus, only polymers that can diffuse to an appropriate site on a protocell surface where they can be stabilised by association with the amphiphilic head groups, and possibly also with other chains, will be preserved from dissociation. So our further requirements are:

(vii) at least some domain structures are able to act as weak catalysts for the polymerisation of membrane-bound molecules. These polymers add to the set (u, v..). The exclusion from the domain of molecules that would poison nascent polymers also favours polymerisation.
(viii) at least some of the polymers in (vii) are able to bind to membrane domains. Once the concentration of a particular domain-forming polymer reaches a critical level, we assume that domain reorganisation can occur.

Fresh substrate is delivered to the protocell by fusion. This depends on the time for diffusion of free chains through the membrane being long. For example, if (u, v,...) are amino acids and the membrane is a lipid bilayer of thickness 5 nm (see below) the diffusion time is 2×10^{-5} s (Chakrabarti and Deamer, 1992). This may be long compared to a collision timescale even for a very dilute solution. For phospholipids the diffusion time across a bilayer can be days (see references in Norris, 1989). We can think of the vesicle with externally attached chain as a weakly bound intermediate state. The intermediate state will tend to collisionally dissociate

rather than allow the attached chain to be incorporated by diffusion. Thus, if fusion between colliding lipocells occurs efficiently then it can be assumed to be the way by which vesicles grow. If there is a strong barrier to fusion, so that it is a relatively rare event, then growth will be by absorption of molecules from solution. This may involve adherence to the outer monolayer, domain formation with alteration in the radius of curvature and endocytosis. In this case fusion would play only the role of forcing occasional large evolutionary steps on the population.

Protocells also undergo fission. Fission is relatively less disruptive to stabilised domains. (This is part of what we mean by stability.) These therefore form part of a new protocell, while material excluded from the stabilised domains is returned to the environment where it can be incorporated into new protocells. We assume that the formation of domain structure does not stabilise the protocell against fission, otherwise the system would evolve to a number of large inert vesicles. Thus, the population of protocells is dominated by a large population of small lipocells which can therefore be thought of as the 'food set' for the rarer larger cells.

3. Prebiotic Conditions

It is not part of our argument here to establish even a plausible prebiotic chemistry. The principles apply to any set of amphiphiles that can form micellar or vesicular structures and to monomers and short chain polymers that can form non-covalent associations with them. For the sake of argument, however, the following is a conceivable mapping between the model and prebiotic chemistry. This will provide a specific context in which we can discuss the evolution of the system.

(a) lipids: The abiotic synthesis of various lipids including one of the principle constituents of modern bacterial membranes, phosphatidylglycerol, has been obtained from mixtures of glycerol, phosphate, and fatty acids (with short, 12 carbon, chains) based on conditions possibly found in tidal pools (Hargreaves *et al.*, 1977). A wide range of alkanes, with carbon chains up to 23 have been obtained from both the Murchison meteorite and spark discharges (Kvenvolden, 1970). Using an evaporating pond model, one of the principal constituents of modern eukaryotic membranes, phosphatidylcholine, can be synthesised from phosphatidic acid and choline in the presence of cyanamide (Rao, 1982). Thus we assume that lipids are the amphiphiles that form the bilayer surface of the initial protocells.

A population of lipocells is subject to fusion and fission processes. Locally elevated calcium levels may provoke fusion between membranes containing anionic phospholipids whilst pH changes may induce fusion between membranes containing the neutral phospholipid, phosphatidyethanolamine (Lentz, 1994). Contact regions between separated phases induced by calcium have been proposed as the intermediate structures leading to calcium-induced fusion

of phosphatidylserine-containing bilayers (Lentz, 1994, Papahadjopoulos *et al.*, 1977). Although modern seawater contains 10mM calcium and the first seawater may have had concentrations three times higher (MacIntyre, 1970), it should be noted that these can produce 'hard water curds' of aggregated lipids.

(b) amino acids and peptides: we assume that all biologically important amino acids are produced prebiotically. There is no evidence that any of the set are later additions, although this would not affect our discussion. In addition, the evolution we propose allows for non-biological amino acids to be produced abiotically provided these are not incorporated into stabilised membrane structures. Long chain proteins require rather extreme conditions compared to those we envisage, although short chain peptides, long enough to have a weak catalytic function, might be produced under high temperatures in aqueous conditions. It is conceivable that the drying and heating of protocells led to condensation reactions between short peptides (Fox and Harada, 1958) and we assume that short peptides are present.

(c) nucleotides and polynucleotides: There is evidence that short nucleotidic sequences can form from activated nucleotides (Orgel, 1992). The source of such nucleotides may have been a mineral surface since clay has been shown to act as a catalyst in phosphate-ester formation (Ferris *et al.*, 1988). It should be noted, however, that there are stability problems with prebiotic nucleotides (Larralade *et al.*, 1995). However, we need something that is going to evolve into polynucleotides so, for the sake of argument, we assume nucleotides are present too.

There are two, not necessarily exclusive, possibilities for the site of further evolution. We should consider either the low concentrations of prebiotic molecules that can be obtained in the large volume of the primitive ocean or the higher concentrations that can occur in the much small volume of rock pools of the littoral by repeated drying. In the sea the maximum concentration of inorganically produced biotic macromolecules is 10^{-4} m (Miller, 1987), with degradation mainly through heating at volcanic vents giving a turnover time of 10^6 years. If the products include lipids (Ferris, 1987) this concentration is amply sufficient for micelle formation (Sarafan, 1994). Formation of liposomes may require a higher concentration and long unbranched fatty acid chains (Ferris, 1987). This might occur in pools where macromolecules may be protected from degradation by radiation, but the total quantities may not be significant compared with the lipid mass in sea-born micelles. Nevertheless, it may be that we should take as the next stage of evolution the spontaneous formation of lipid micelles in the primitive ocean. The micellar stage would then be a precursor stage to the formation of lipid vesicles. The evolution of micelles we imagine to occur according to the same principles of fission/fusion and domain formation.

We consider now how this system evolves. There are two aspects:

(i) metabolism; as discussed above, this includes both the creation of complex molecules from those synthesised prebiotically and the focusing on to a subset by mutual association in stable domains.
(ii) domain structure; the lipocell population will evolve in response to the accumulation of domain-stabilising polymers.

In reality these are not independent but, for simplicity of exposition, we shall initially treat them as such. In principle, the system could end up with a large ('unfocussed') set of macromolecules in dynamic equilibrium, or with a small set of inert vesicles in essentially static equilibrium (or with some cyclic or chaotic motion in state-space). Any of these outcomes would make vesicle formation inimical to the evolution of life. We shall argue quite the opposite: that the above properties are sufficient to yield a focused set of macromolecules which could reasonably be regarded as a possible precusor to cellular life.

4. Evolution of Metabolism

If there is a mixture of lipids present in a bilayer they will generally form domains. Now, we know that electrostatic interactions can attract small peptides to certain lipid head groups and hence attach them to the inner membrane of the lipid bilayer (Hammes and Schullery, 1970). The presence of the bilayer can influence the conformation of short peptides. These may adopt an alpha-helical form in the presence of the membrane and interact with the membrane lipids e.g. by insertion of the hydrophobic moiety (Hammes and Schullery, 1970). The first of these processes (the selective electrostatic interaction) represents a concentration of interacting peptides/lipids in the lipocell surface, a process which itself is conducive to catalytic polymerisation (Wachtershauser, 1988). We also know that even dipeptides can act as weak catalysts for polymerisation (Kauffman, 1993). Note that the arguments for enhanced polymerisation on inorganic surfaces translate immediately into arguments for polymerisation on the cell membrane. In particular, the entropic barrier depends on the dimensionality of the surface, not its composition.

In the contemporary cell, reactions are often catalysed by molecular associations that are not covalently bound. The transverse fluidity of the membrane surface is particularly appropriate as the site for analogous reactions in the lipocell. First, since a Brownian path is a 2-dimensional structure, a molecule diffusing freely in the surface will visit every site in a finite time. Second the surface hydrophobic effect tends to crowd macromolecules, further enhancing the probability of interaction. Thus the probability of the coming together of two parts of an enzyme with attached substrates is increased in the surface over the bulk phase. Thus, given

molecular association in surface domains, we expect that some polymerisation will be catalysed on the surface.

Note also we preserve not just those peptides that can be formed on the membrane, but the subset that can leave the formation site and attach elsewhere (in contrast to Wachtershauser). This (and fission-fusion) drives the evolutionary dynamics. The attachment of polymer to the membrane preserves the polymers from degradation, thereby driving the equilibrium point of the reaction for the production of *interacting* polymers towards polymer stability. We shall argue that this gives a mechanism for chemical focusing.

The products of catalysis in one region will stabilise other domains that will in turn exhibit weak catalysis. Provided we start with a sufficient number of different domains – this is probably guaranteed by the non-gaussian size and compositional distribution of vesicles – we obtain an autocatalytic network. Thus, following arguments for the evolution of a metabolism similar to Kauffman (1993), we end up with a catalytic network in which the set of domains catalyse their own formation. For the purpose of the metabolism we ignore for the moment the fact that the domains are in vesicles subject to fission and fusion and imagine the domains as separate entities (in some chemical-domain space). Then the polymer content of a domain determines its polymerising efficiency and its ability to stabilise polymers. The domain evolution can therefore be reduced to an evolution of polymers in polymer space. Kauffman has shown how such evolution can be discussed in terms of dynamics on a fitness landscape. Linear polymers of length equal to or less than N monomer units are assumed to have a catalytic efficiency determined by, on average, K monomers in the chain. The fitness of a polymer is determined in terms of its catalytic efficiency in the network of reactions that leads to its own formation. The only difference between us and Kauffman is that in our picture catalytic fitness includes the ability to form both catalytic and stabilising domains.

We expect the evolution to occur on a rugged fitness landscape. This is because we change the fitness of a polymer as a weak catalyst or a domain stabiliser only by making a large change in the polymer. Changing a single amino acid in one peptide is unlikely to alter very much the domain-forming properties. Let N be the number of monomers in the longest polymer and K the number that must be changed to alter the fitness of a polymer to stabilise a domain or participate in the catalytic properties of a domain. For example this might occur because the altered monomers bring about a conformational change in the polymer or increase its binding to the surface or to other surface bound polymers. Then our previous statement is equivalent to $1 << K << N$. This is the condition for evolution on a rugged landscape (Kauffman, 1993). Thus, the focusing occurs on to a range of polymers i.e. the autocatalytic network covers an extended region of polymer space even at late stages in the evolution. (Of course, the evolution of domain structure will act to change the fitness landscape, so this region will itself be subject to drift and dispersion on a longer timescale.)

Note that our system is locally in chemical equilibrium so a large flow of material is avoided. The presence and properties of the physical interface between domains is the cause of global non-equilibrium, not a flow. Thus, once the physical structure of vesicles and domains has spontaneously formed, we do not have to add any additional postulate concerning the flow of material through the system.

In summary, the principle is this: association with the membrane captures the new substrates that arrive when protocells fuse; this association both promotes polymerisation and preserves these polymers from hydrolysis (and, one hopes, poisoning by chain termination). Membrane association therefore leads to the accumulation of reaction products. Thus fusion leads to a linear growth of the cell and bulk reaction products, but, while the binding site lipids are in excess, the bound substrate grows approximately quadratically with time (measured in terms of the number of fissions). This is the dynamical basis of focusing.

Furthermore the ratio of bound lipids to surface area also increases. We assume this leads to fission before the growth of bound substrate enters the linear regime. This requires that the fission time should be sufficiently small, but also large enough to exceed the fusion time (i.e. the mean waiting time between fusion events).

5. Domain Evolution

In the previous section we concentrated on the evolution in polymer space subject to fitness to form stable domains without regard to the fission-fusion process (even though this was driving the evolution). We now focus on the evolution of the lipocell population in terms of the domain structure of the cells. The mechanical stresses created by growing domains will be relieved by fission with a better than random probability of leaving certain domains intact. (These are the relatively stabilised domains we have constructed.) We therefore have a population of lipocells (co-)evolving on the landscape of domain stability. This has the following correspondence with the genetic evolution of cells.

The ‘genetic content’ of a lipocell is its domain-forming polymer content. The ‘phenotype’ is defined by the domain structure. The fitness criterion is domain stability. Finally, N is here the number of polymers and K ($1 << K << N$) the number of polymers that on average interact with a given polymer to alter the domain-forming characteristics in the liposome. Thus we again arrive at evolution on a rugged landscape driven by fission and fusion. This implies that the product of evolution will be a limited range of domain structures representing local fitness peaks.

This is our promised conclusion: as a result of evolution the protocell population is focused but still varied. Without a varied population that can evolve on an evolving landscape, the system would become frozen, with the macromolecular content trapped on fitness peaks and unable to participate in further evolution (and more assumptions would be required for evolution to resume). It is important to

realise that the hypothesis of domain structure and stability has played a crucial role in this conclusion by providing an appropriate fitness criterion for the protocellular evolution.

6. Evolution of Coding

For the purpose of illustration, we now imagine the preceding evolution to occur under the condition that the system contains nucleotides (not necessarily pre-biotic) in addition to peptides. We assume that peptides and nucleotides can bind in the lipid bilayer. There is weak experimental evidence for this. For example, arginine is know to bind to the region of RNA which codes for it (Yarus, 1993) and DNA interacts with cationic lipids leading to aggregation of liposomes and ordering of the DNA (Radler *et al*., 1997). We show that this system has in principle the potential to evolve coding.

The polynucleotide-peptide evolution leads to a mutually associating or 'complementary' set, as defined in the following discussion.:

Suppose the lipocell makes nucleotides n_1, n_2, ... and peptides p_1, p_2, ... Let the subsets of nucleotides be labelled N_i, (i = 1, 2....) and the subsets of peptides P_i. If some p_k does not associate with an N_j for any j then it does not contribute to a stable domain. Therefore discard this p_k. Thus, only the p_i for which there exist a *complementary* N_i are selected by domain evolution. The argument can be reversed with N and P interchanged. Thus, lipocell evolution produces a complementary set.

Under what conditions is this complementary set $\{N, P\}$ a proper subset of all possible strings of monomers? Let p_{ij} be the probability that a nucleotide n_j is associated with a peptide p_i. Thus (p_{ij}) is (for all practical purposes) a matrix with entries 1 and 0. Then $\{N, P\}$ is a proper subset if this matrix is reducible i.e. can be partitioned into submatrices which are zero, except on the diagonal, by some permutation of the rows and columns. This is clearly a percolation problem. If all, or almost all, of the entries are 1s the matrix is irreducible. On the other hand reducible matrices exist. Thus there is a threshold for the probability of association below which the set $\{N, P\}$ is a proper subset.

It follows that evolutionary selection for domain stability favours lipocells that can *produce* complementary macromolecules (since these will have increased stability once they are formed and held in domains.) This is just a description of coding: the metabolic process (i.e. the macromolecules produced from the 'food set') is determined by molecular associations within the cell.

The complementary sets of $\{N, P\}$ need not be the same for all cells (and may depend on the lipid composition, hence size etc.). Call a set $\{N_i, P_i\}$ a class of lipocells. If cell division can occur in a class of vesicles (rather than fission to a single offspring vesicle) there is a finite probability that one set will yield daughter cells of the same set. Once this occurs that set will rapidly take over the population. Thus, coding in the model will lead to reproduction.

Finally we comment on the error catastrophe for this coded reproduction. The metabolism has a hypercyclic organisation: domain forming molecules catalyse the production of further domain forming catalysts in what, as we have just seen, is a closed system. We therefore expect the system to be secure from the error catastrophe. We note that our evolutionary picture also provides a way in which a hypercyclic type of functional organisation can evolve.

7. Discussion

For the purpose of characterising the various approaches to its origin we may attribute to life three types of organisation.

(a) chemical order in which a system is maintained far from chemical equilibrium by a flux of free energy. It is crucial for the origin of life that this serves to focus the system on a small fraction of the organic molecules that could in principle appear.
(b) spatial order such that the chemical organisation is maintained in spatial domains or cells.
(c) temporal order, by which we mean, for example, the process of cyclical replication as opposed to continuous repair.

In principle each of these types of order could occur independently. A chemically ordered but spatially homogenous system can function as a universal computing machine and therefore could exhibit the information processing properties associated with life. A spatially homogeneous system could be organised temporally, passing through a periodic sequence of states which might be thought of as a form of replication. Such a system need not evolve a mechanism of genetic coding. A spatially ordered metabolic system could, again in principle, be eternal. Thus, in accounting for the emergence of life we have to decide the sequence in which these aspects of structure arose.

Spontaneous spatial order of coacervate droplets in the pre-biotic medium is the initial step towards biological order in the early theory of Oparin (1924). Other authors have revived this theory in the context of membraneous structures as precursors of contemporary cells with a more active role for the membrane (e.g. Morowitz, Heinz and Deamer, 1987). At the opposite extreme of this class of models are the protenoid spheres of Fox (1965) where spatial and chemical structure, which arise at an earlier stage than replication, are different aspects of the same thing. The theory of evolution on clays in which inorganic ions carry the program for replication may be regarded as another member of this class of theories (Cairns-Smith, 1982).

Chemical and temporal order together are the first aspects of biological structure to emerge in ‘RNA first’ models. This is followed only later by the evolution of

cells, presumably because these have advantages for survival of the genetic code. Conversely (by definition) theories which invoke a metabolic origin for life put chemical order first. This may arise spontaneously in an autocatalytic network of chemical reactions. The genetic code may then enter as a second starting point (Dyson, 1985) or from the evolution of the catalytic network itself. In this picture the common cellular basis of life is (apparently) a historical accident.

There appears to be no way at present of deciding which of these pictures is correct. There are, however, several reasons why we favour the 'cells first' approach.

i) The initial phase of evolution involves redistribution and growth. Other theories obtain this by covalent bonding, a free-energy intensive approach. Given a standard array of prebiotic macromolecules, we use non-covalent bonding to initiate self-assembly and focusing.

ii) The cell contributes to focusing by preserving a subset of macromolecules from degradative and poisoning processes . We show that this is a focusing mechanism that can operate in a system that is locally in chemical equilibrium. The free-energy intensive operation of driving the system from equilibrium in its early phases is thereby avoided. (The evolution of a metabolism always requires some free-energy input that we do not address here.)

iii) The evolution necessarily focuses on a mutually associating set of macromolecules which will lead to effective catalysis and from which coding can, in principle, emerge.

In the particular form of 'cells-first' approach we adopt the three forms of order we have identified form spontaneously and together. Spatial structure arises from hydrophobic interactions giving rise to lipocells; dynamic order arises from fission and fusion of these structures; chemical order arises from the preferential stabilisation of mutually associating macromolecules within these structures.

Is our proposal subject to experimental test? We should distinguish here between the purely illustrative chemistry and the main structural features. Nevertheless it would be interesting to learn the extent to which mixtures of amino acids or nucleotides can give rise to, or stabilise, domain structures in lipid vesicles. Whether such structures have catalytic properties might also be tested. None of this is probably amenable to theory. On the other hand the mechanical properties of model membrane domains might be subject to both theoretical and experimental investigation (Norris and Manners, 1993). It is difficult to prove anything by the failure to obtain readily fissionable and fusible vesicles, but if fission and fusion play an important role it ought to be possible to find such model systems eventually. More immediately perhaps, the general arguments of the text concerning the evolutionary dynamics should be spelled out by detailed numerical simulation. We hope to return to this elsewhere.

Finally, membrane domains can be invoked to explain the regulation of initiation of chromosome replication (Norris, 1990, 1993, 1994), chromosome segregation (Norris, 1995), cell division, and differentiation (Norris and Madsen, 1995) in

modern bacteria. In the model proposed here, such regulation has its origins in the origin of life itself.

8. Conclusions

We have argued that the spontaneous formation and the fission and fusion of lipocells under pre-biotic conditions can provide a viable model for the evolution of life.

There are three main aspects to the model. These are the formation of domains in the lipocell membrane, the catalytic formation of polymers on the membrane surface and the evolutionary dynamics of the cell population driven by fusion and fission. The system evolves on a rugged fitness landscape determined by domain stability. The evolution of coding by domain structure provides a possible link to the RNA/DNA world.

Acknowledgements

VN acknowledges support from BBSRC while part of this work was carried out.

References

Bachmann, P. A., Luisi, P. L. and Lang, J.: 1992, *Nature* **357**, 57.
Baeza, I., Wong, C., Mondragon, R., Gonzalez, S., Ibanez, M., Farfan, N. and Arguetto, C.: 1994, *J. Mol. Evol.* **39**, 560.
Bernal, J. D.: 1949, *Proc. Phys. Soc.* **A62**, 537–558.
Bohler, C., Hill, A. R. and Orgel, L. E.: 1996, *Origins Life Evol. Biosphere* **26**, 1.
Cairns-Smith, A. G.: 1982, *Genetic Takeover and the Mineral Origins of Life* (Cambridge, University Press).
Cavalier-Smith, T.: 1987, *Cold Spring Harbor Symp. Quant. Biol. LII.*
Chakrabarti, A. C. and Deamer, D. W.: 1982, *Biochim. et Biophys. Acta* **1111**, 171–177.
Deamer, D. W.: 1986, *Origins Life Evol. Biosphere* **17**, 3–25.
Deamer, D. W. and Fleischaker, G. R.: 1994, *Origins of Life*, Jones and Bartlett.
Deamer, D. W. and Pashly, R. M.: 1989, in *Origins Life Evol. Biosphere* **19**, 21–38.
Dyson, F. J.: 1985, *Origins of Life*, Cambridge U.P., Cambridge, U.K.
Eigen, M.: 1971, *Naturwissenschaften* **58**, 465–523.
Eigen, M. and Schuster, P.: 1997, *Naturwissenschaften* **64**, 541–565.
Ferris, J. P.: 1987, *Cold Spring Harbor Symp. Quant. Biol* **52**, 29.
Ferris, J. P., Huang, C.-H. and Hagen, W. R.: 1988, *Origins Life Evol. Biosphere* **18**, 121.
Ferris, J. P., Kamaluddin, and Ertem, G.: 1990, *Origins Life Evol. Biosphere* **20**, 279.
Ferris, J. P., Hill, A. R., Liu, R. H. and Orgel, L. E.: 1996, *Nature* **381**, 59.
Fox, S. W.: 1965, *Nature* **205**, 328–340.
Fox, S. W. and Hadrada, K.: 1958, *Science* **128**, 1214.
Glaser, M.: 1993, *Curr. Opin. Struct. Biol.* **68**, 161–168.
Hammes, G. G. and Schullery, S. E.: 1970, *Biochemistry* **9**, 2555.
Hargreaves *et al*.: 1977, *Nature* **266**, 78.
Joyce, G. F.: 1991, *New Biologist* **3**, 399–407.
Koch, A. L.: 1985, *J. Mol. Evol.* **21**, 270–277.

Kauffman, S. A.: 1993, *The Origins of Order*, Oxford, U.P.
Kinnunen, P. K. J., Anu Kõiv, A., Lehtonen, J. Y. A., Rytömaa, M. and Mustonen, P.: 1994, *Chem. Phys. Lipids* **73**, 181–207.
Kvenvolden, K. *et al*.: 1970, *Nature* **228**, 923–926.
Larralade, R., Robertson, M. P. and Miller, S. L.: 1995, *Proc. Nat. Acad. Sci*. **92**, 8158.
Luisi, P. L.: 1993, in W. D. Stein and F. J. Varela (eds.), *Thinking about Biology; SFI Studies in the Science of Complexity Vol. III*, Addison-Wesley.
Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. and Ghadiri, M. R.: 1996, *Nature* **382**, 525.
Lentz, B. R.: 1994, *Chem. Phys. Lipids* **73**, 91.
MacIntyre, F.: 1970, in J. Walker and W. H. Freeman (eds.) *The Physics of Everyday Phenomena, Readings from Scientific American*, San Francisco.
Melo, E. C. C., Lourtie, I. M. G., Sankaram, M. B., Thompson, T. E. and Vaz, W. L. C.: 1992, *Biophys. J.* **63**, 1506–1512.
Miller, S. L.: 1987, *Cold Spring Harbor Symp. Quant. Biol*. **52**, 17.
Morowitz, H. J., Heinz, B. and Deamer, D. W.: 1988, in *Origins Life Evol. Biosphere* **18**, 281–187.
Mouritsen, O. G. and Jørgensen, K.: 1994, *Chem. Phys. Lipids* **73**, 3–25.
Needham, J.: 1968, *Order and Life*, Cambridge, Mass.
Norris, V.: 1989, *J. Theor. Biol*. **139**, 117–128.
Norris, V.: 1990, *J. Mol. Biol.* **215**, 67–71.
Norris, V.: 1993, *J. Theor. Biol.* **164**, 239–244.
Norris, V.: 1994, *Mol. Microbiol.* **15**, 985–987.
Norris, V.: 1995, *Mol. Microbiol.* **16**, 1051–1057.
Norris, V. and Madsen, M. S., 1995, *J. Mol. Biol.* **253**, 739–748.
Norris, V. and Manners, W.: 1993, *Biophys. J.* **64**, 1691.
Orgel, L. E.: 1992, *Nature* **358**, 203.
Oro, J., Basile, B., Cortes, S., Shen, C. and Yamron, T.: 1984, *Origins of Life* **14**, 237.
Papahadjopoulos, D., Vail, W. J., Newton, C., Nir, S., Jacobson, K., Poste, G. and Lazo, R.: 1977, *Biochim. Biophys. Acta* **465**, 579.
Rao, M., Eichberg, J. and Oro, J.: 1982, *J. Mol. Evol.* **18**, 196–202.
Radler, J. O., Koltover, I., Salditt, T. and Safinya, C. R.: 1997, *Science* **275**, 810.
Rokhsar, D. S., Anderson, P. W. and Stein, D. L.: 1986, *J. Mol. Evol.* **23**, 119–126.
Safran, S. A.: 1994, *Statistical Thermodynamics of Surfaces, Interfaces, and Membranes*, Addison-Wesley.
Tocanne, J.-F., Cezanne, L., Lopez, A., Piknova, B., Schram, V., Tournier, J.-F. and Welby, M.: 1994, *Chem. Phys. Lipids* **73**, 139–158.
Wachtershauser, G.: 1988, *Microbiol. Rev.* **52**, 452–484.
Yarus, M.: 1993, in *The RNA World*, Cold Harbour Spring Lab. Press.