Molecular Microbiology (2003) **47**(1), 13–22

MicroHypothesis

From pre-cells to Eukarya – a tale of two lipids

G. Wächtershäuser
Tal 29, D-80331 Munich, Germany.

Summary

A mechanistic hypothesis for the origin of the three domains of life is proposed. A population of evolving pre-cells is suggested to have had a membrane of a racemate of chiral lipids that continuously underwent spontaneous symmetry breaking by spatial phase segregation into two enantiomerically enriched membrane domains. By frequent pre-cell fusions and fissions these membrane domains became partitioned between two pre-cell subpopulations having predominantly one lipid enantiomer or the other. The origin of the Bacteria and Archaea is explained by divergence of first a population of proto-bacteria and later a population of proto-archaea from the evolving pre-cells, each by the emergence of an enantio-selective lipid biosynthesis within the corresponding pre-cell subtype. The origin of the Eukarya is explained by symbiosis between a population of Bacteria and a subpopulation of pre-cells with a predominance of the bacteria-type lipid enantiomers.

Introduction

The problem of the early evolution of life is now, after the completion of so many genome sequencing projects, as vexing and as tantalizing as ever. Why, after its inception did the evolution of life remain a unitary process for so long? What was it that finally drove life apart forging three domains of fundamentally different organization, the Bacteria, the Archaea and the Eukarya? and why, through billions of years of subsequent evolution was life stuck with these three, and only these three, domains?

A theory is here suggested, which attempts a coherent answer to these questions. It is built on three foundations: the pre-cell theory of Otto Kandler, comparative genomics and the physical chemistry of lipid membranes. It can be best conceptualized with the understanding that the overall process of evolution is divided into three phases. The first phase begins with the origin of life. The second phase is the phase of divergence into the three domains of Bacteria, Archaea and Eukarya (Woese, 1987; Iwabe *et al*., 1989; Woese *et al*., 1990). The third phase covers the parallel evolution of these three co-existing domains up to the present day. The theory here proposed is concerned mainly with the second phase.

Accepted 18 September, 2002. *For correspondence. E-mail info@patent.de; Tel. (+49) 89 2199760; Fax (+49) 89 223759.

The pre-cell theory

The pre-cell theory of Kandler (1994a,b; 1998), which has recently been adopted by Woese (1998; 2002) provides the framework for understanding the second phase of evolution. Kandler (1994a,b) has defined pre-cells as 'metabolizing self-reproducing entities exhibiting most of the basic properties of a cell, but unable to limit the frequent mutual exchange of genetic information'. Therefore, the pre-cells so defined are, from a biological point of view, incomplete cells even though from a geometric point of view they are complete cellular entities. Because all pre-cells are defined as in mutual exchange of genetic material they must have formed in Kandler's view one coherent population (Kandler, 1994a,b). However, not all the pre-cells within this coherent population were genetically identical. Rather, in Kandler's (1994a,b) terminology the total population of pre-cells was 'multiphenotypical' and distributed over a variety of habitats each habitat harbouring a different subpopulation. (This contrasts to Woese's (1998) terminology whereby the ancestral state was a 'communal diverse conglomerate'). Some subpopulations may have been autotrophic and others heterotrophic. Some may have been anaerobic and others microaerophilic; some may have been H_2-consumers and others H_2-producers. As we shall see this phenotypical diversity of the pre-cells is important for the domain divergence in the second phase of evolution.

Kandler's pre-cell theory assumes a trunk evolution of pre-cells, which began at some point during the first phase of evolution and from which the three separate domains eventually diverged one after the other in the second phase. At each point of divergence a founder population

branched off from a phenotypically distinct subpopulation of the pre-cells, while the trunk evolution of the whole pre-cell population continued (Fig. 1). Specifically, Kandler assumes the bacterial lineage to have diverged first at a pre-cell stage PC-1 to embark on its own separate evolutionary path, while the trunk evolution of pre-cells continued. Thereafter, at a more evolved pre-cell stage PC-2 the archaeal lineage diverged to embark on yet another separate evolutionary path, while the trunk evolution of pre-cells still continued. The eukaryal lineage diverged from the pre-cells at a still more evolved stage PC-3. On the subsequent fate of the pre-cells Kandler's theory remains undecided.

Let us now look at the pre-cells in greater detail. As empirical base for this undertaking we will employ comparative genomics. It has been found that in all bacterial and archaeal genomes sequenced so far certain orthologous genes, mainly for transcription and translation, are organized in conserved gene clusters of different lengths and that alignments of these clusters may be fitted together like pieces of a jigsaw puzzle (Wächtershäuser, 1998). This produces a composite alignment with overlapping gene clusters. This composite alignment allows us to hypothetically reconstruct extended genome segments of ancestral life forms including the bacterial common ancestors (BCA) and the archaeal common ancestors (ACA) shown in Fig. 2. These reconstructed genome segments are amazingly similar and even the few deviations are again constant throughout the domain of Bacteria and throughout the domain of Archaea, respectively. Lateral transfer of the whole gene cluster cannot be invoked for explaining this extreme conformity within and between the domains of Bacteria and Archaea. Lateral transfers of such large gene clusters between distantly related groups of organisms would cause havoc in the highly complex machineries for transcription and translation. Quite generally, the lateral transfer of any long cluster of genes coding for components of a vital and complex cell machinery would be lethal for the host organism (Wächtershäuser, 1998). The only reasonable explanation for this universality would be that the pre-cells prior to the divergence of the bacterial lineage, i.e. immediately before stage PC-1, must have had already a genome with a remarkably stable cluster of more than forty genes mainly for transcription and translation, which underwent little change over billions of years. This means that the overall population of pre-cells, whereas distinctly non-unitary with regard to the metabolic genes for survival in the various habitats, must have been a unitary form of life with regard to the genes for the genetic machinery. Thus the population of pre-cells prior to the divergence of the domain Bacteria constitutes a last universal common ancestor (LUCA) with regard to the genetic machinery, but not with regard to the metabolic genes.

Now we can take a closer look at the genes contained in the archaic gene cluster in the hope that this will give us important insights into the mechanistic forces that forged the three domains of life and specifically their genetic machinery and their cell envelope. Turning first to the genetic machinery of the pre-cells at stage PC-1 we can now conclude from the apparent stability of the corresponding portion of the genome of these pre-cells that they must have been circular double-stranded DNA chromosomes, perhaps of the size of extant plasmids. It has been suggested that the earliest genetic elements were open-ended linear gene-sized elements (Woese, 1998). But these fray easily at their ends and therefore are not suitable as information carriers at this stage. Linear chromosomes with a variety of terminal stabilization structures are found in many extant organism (Redenbach, 2002), but they must have appeared much later in the third phase of evolution and polyphyletically. Moreover, the archaic gene cluster comprises the gene for the universal antitermination factor NusG (Fig. 2). This indicates that transcription may have continued over long DNA stretches and occasionally perhaps around the entire circle of the chromosome as a form of rolling circle transcription. Given the lack of universality of transcription initiation we may well assume that pre-cell transcription operated without promoters and without controlled termination, generating long multicistronic transcripts of varying lengths. This immediately suggests that pre-cell replication could have been a rolling circle process as well. A mechanistic transition from rolling circle transcription to rolling circle replication would have required nothing more than a replacement of ribonucleotides by deoxyribonucleotides and it would readily explain why RNA primers are needed in DNA replication to this day. Incidentally, if these rolling circle processes would have begun with a relaxing nick, the pre-cells would have needed neither helicase nor topoisomerase.

Leaving now the genetic machinery and turning to the cell envelope of the pre-cells prior to PC-1 we note that the canonical genes in the archaic gene cluster (Fig. 2) comprise the genes coding for SecE and SecY. These are the two essential subunits of membrane-bound protein translocase (Van Wely *et al*., 2001). Their presence as canonical genes in the archaic gene cluster strongly suggests that the cell envelope of the pre-cells immediately before stage PC-1 was a stable closed lipid membrane with a membrane-anchored protein translocase.

The pre-cell lipid membrane

The results obtained so far are in harmony with each other. The inquiry began with the definition of the pre-cells as exhibiting most of the basic properties of a cell. It was then argued that the pre-cells must have had already a

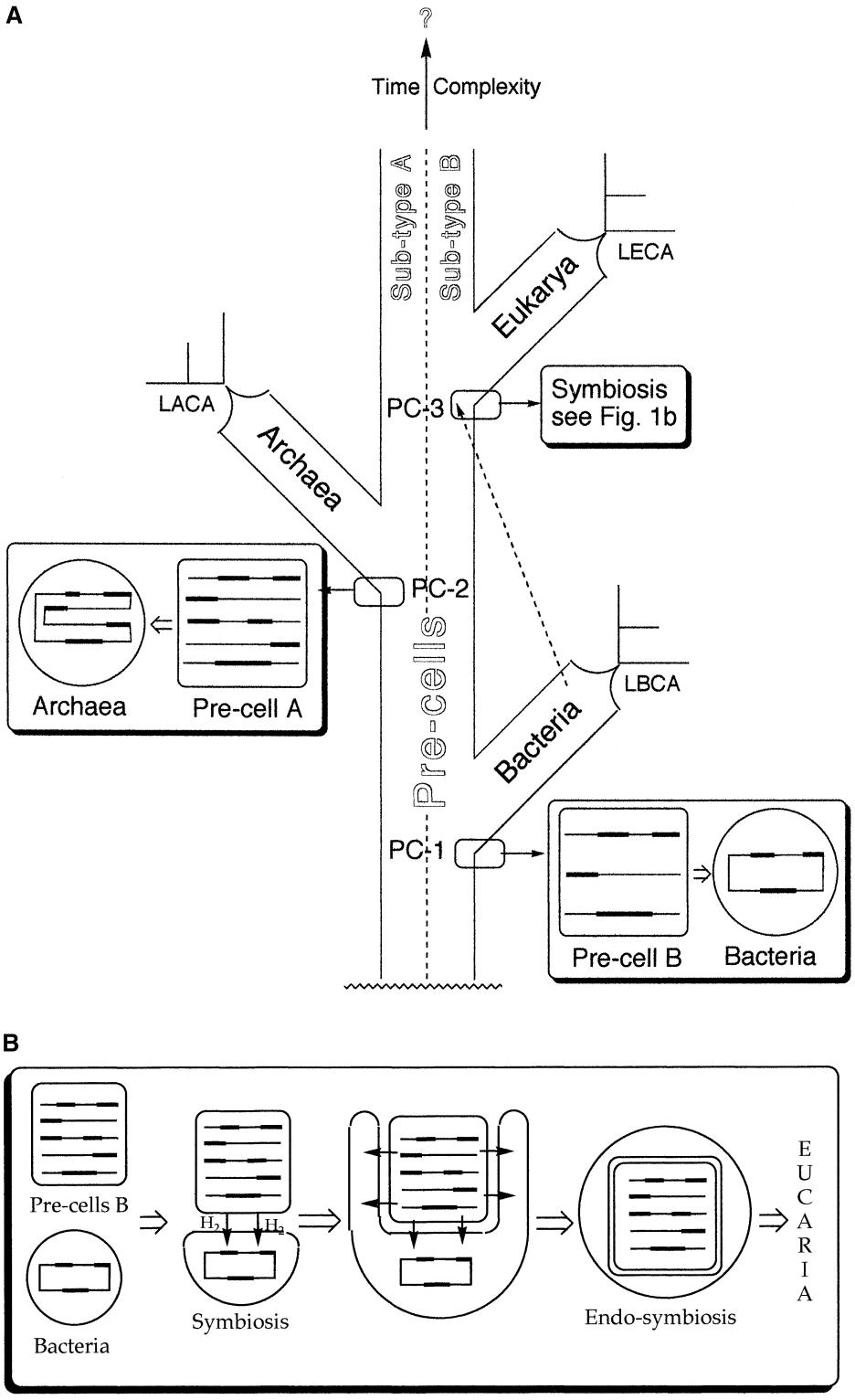

Fig. 1. Illustration of the divergence of the domains Bacteria, Archaea and Eukarya from stages PC-1, PC-2 and PC-3 in the second phase of evolution (modified from Kandler, 1994a,b).
A. The overall process. (B is inset in Fig. 1A). The wavy line at the bottom signifies that the shown second phase of evolution continues an older first phase of evolution beginning with the origin of life. The widths of the trunk and branches signify the diversity of subpopulation in different habitats and the massiveness of parallel evolution. Pre-cells are symbolized by squares, the horizontal lines with thin and fat segments signifying the linear intermediates of rolling circle plasmid replication. Archaea and Bacteria are symbolized by circles with circular chromosomes.
B. Symbiosis.

A	BCA	SecE	NusG	L11	L1	L10	L12	RpoB	RpoC	L30	—	S12	S7	EF-G	EF-Tu	S10	L3	L4	L23	L2 of Fig. 2 B
	ACA	SecE	NusG	L11	L1	L10	L12	RpoHB"B'	RpoA'A"	L30	NusA	S12	S7	EF-G	EF-Tu	S10	L3	L4	L23	L2 of Fig. 2 B
	PC-I	SecE	NusG	L11	L1	L10	L12	RpoHB"B'	RpoA'A"	L30	—	S12	S7	EF-G	EF-Tu	S10	L3	L4	L23	L2 of Fig. 2 B
B	BCA	L2	S19	L22	S3	L16	L29	S17	L14	L24	—	L5	S14	S8	L6	—	—	L18	S5	L30 of Fig. 2 C
	ACA	L2	S19	L22	S3	—	L29	S17	L14	L24	S4E	L5	S14	S8	L6	L32	L19	L18	S5	L30 of Fig. 2 C
	PC-1	L2	S19	L22	S3	—	L29	S17	L14	L24	—	L5	S14	S8	L6	—	—	L18	S5	L30 of Fig. 2 C
C	BCA	L30	L15	SecY	Adk	Map	IF-1	L36	—	S13	—	S11	S4	RpoA	L17	L13	S9	—	S2	
	ACA	L30	L15	SecY	Adk	—	—	L36	Cmk	S13	S4	S11	—	RpoA	L17	L13	S9	RpoN	S2	
	PC-I	L30	L15	SecY	Adk	—	—	L36	—	S13	—	S11	S4	RpoA	L17	L13	S9	—	S2	

Fig. 2. Ancestral gene clusters (modified from Wächtershäuser, 1998), reconstructed for the bacterial common ancestor (BCA), the archaeal common ancestor (ACA) and the pre-cells of stage PC-1. SecE, SecY, translocase units; NusG, NusA, anti-termination factors; L11, etc., large subunit ribosomal proteins; S12, etc., small subunit ribosomal proteins; RpoB, etc., RNA polymerase units; EF-G, EF-Tu, elongation factors; Adk, adenylate kinase; Map, methionine aminopeptidase; IF-1, initiation factor; Cmk, cytidylate kinase.

complex multicomponent genetic machinery, which certainly must have required containment inside a stable cellular structure. It was then argued that the pre-cells must have been bounded by a stable lipid membrane. This leads to the following key question: what were the lipids of this pre-cell membrane?

All extant species have lipids with basically the same structure: a glycerol-phosphate head group to which two long-chain hydrophobic rests are attached. It is this structure that gives the lipids the wedge-shape required for the formation of stable closed vesicles (Ackermann, 1992). It therefore seems compelling that the stable membrane of the pre-cells had lipids with this kind of basic structure. But then we come across a peculiar fact. All species of the domains Bacteria and Eukarya have glycerol-3-phosphate lipids, while all species of the domain Archaea have glycerol-1-phosphate lipids (Kates, 1993; Nishihara *et al.*, 1998). So far not a single exception has been reported, nor a single case where both these two enantiomers of the chiral glycerol-phosphate lipids co-exist in a hetero-chiral membrane. How do we account for this situation.

Two possibilities come immediately to mind. The first possibility is that the membrane of the pre-cells prior to PC-1 consisted of lipids of the bacterial enantiomeric type and that these evolved into the archaeal enantiomeric type. The second possibility is that the membrane of the pre-cells prior to PC-1 consisted of lipids of the archaeal enantiomeric type and that these evolved into the bacterial enantiomeric type. Both possibilities require that evolution would have had to proceed through intermediate organisms with heterochiral hybrid membranes of a mixture of lipids of the bacterial and archaeal enantiomeric types. It is here proposed that as a result of lipid incompatibility such heterochiral hybrid membranes are less stable than the homochiral membranes of the Bacteria and Archaea. This means that a gradual evolution through intermediate organisms with heterochiral hybrid membranes is selectively disfavoured. Therefore, a transformation of a bacterial membrane into an archaeal membrane or *vice versa* through a hybrid heterochiral membrane would have been counter-selective. This contradicts all previous theories, which explicitly postulate or implicitly entail such transformation (Cavalier-Smith, 1987; Sogin, 1991; Gupta and Golding, 1996; Martin and Müller, 1998; Koga *et al.*, 1998; Horiike *et al.,* 2001; Hartman and Fedorov 2002; Cavalier-Smith 2002). The problem of lipid membrane transformation has been discussed by Forterre (2001). A proposal by Zillig *et al.* (1992) avoids this transformation by postulating that the bacterial lipid membrane and the archaeal lipid membrane emerged independently from each other by replacing an ancestral non-lipid cell membrane made of proteins. Incidentally, lipid incompatibilities in cases of chiral discrimination are well established (Sackmann, 1982; Nassoy *et al.*, 1995; Vollhardt *et al.*, 1996; Kaganer *et al.*, 1999; Uragami *et al.*, 2000).

A solution of this problem is now proposed. Generally speaking, it is proposed that whereas evolution from a homochiral stable membrane through a less stable heterochiral membrane into the antipodal stable membrane would have been counter-selective and thus of low probability, an evolution could have proceeded with much higher probability from a heterochiral membrane of lesser stability separately into the two antipodal more stable heterochiral membranes. Specifically, according to this proposal the pre-cells before the divergence of the domain Bacteria did have a stable bilayer membrane of a racemate of chiral lipids, which may well have been glycerol-phosphate lipids. These lipids were synthesized inside the pre-cells. The metabolism of the pre-cells was still in the stage of a conversion of non-stereospecific reaction steps to reaction steps catalysed by stereo-specific enzymes (Wächtershäuser, 1988; 1992). In this evolutionary phase the glycerol moiety of the lipids is assumed to have

been synthesized as a racemate with 50% glycerol-1-phosphate units and 50% glycerol-3-phosphate units, perhaps non-enzymatically by inorganic transition metal catalyst or by the kind of primitive non-stereo-specific dihydroxy-acetone phosphate reduction catalyst or enzyme envisioned by Koga *et al*. (1998). Any downstream enzymes must have worked well with both enantiomers (D + L) or diastereomers (e.g. D-L + L-L or D-D + L-D). This is still true for a cytidyltransferase in the archaeal lipid biosynthesis, that works equally well for activating both enantiomers of Di-*O*-geranylgeranoic-glycero-phosphate as well as both enantiomers of Di-*O*-geranylgeranyl-glycero-phosphate (Morii *et al*., 2000).

As a result of chiral discrimination the racemate of lipids of the pre-cell membrane is assumed to have undergone spontaneous symmetry breaking by spatial segregation into a micropattern of two membrane domains each with a predominance of one enantiomer (or diastereomer) or the other. This racemic pre-cell membrane, although not as stable as a homochiral membrane, was presumably stable enough for generating fully functional stable organisms and a definite organism-environment dichotomy over a long period of evolution, perhaps over hundreds of million years, and even under assumed hot conditions of the early Earth. This assumption is supported by the experimental finding that vesicles of a fraction of archaeal lipids and vesicles of bacterial lipids can be forced to fuse by freezing with liquid nitrogen and thawing, followed by sonication (Elferink *et al*., 1992; In't Veld *et al*., 1992).

Now, we make the important assumption that the pre-cells must have undergone frequent fusions and fissions (Wächtershäuser, 1988; 1992), notably at high environmental temperatures. This had the consequence that in the course of these fusions and fissions the pre-cells segregated into one subset A with archaea-type lipid enantiomers or diastereomers predominating and into another subset B with bacteria-type lipid enantiomers or diastereomers predominating. These two subsets of pre-cells, segregated strictly by physical-chemical forces (Wächtershäuser, 1988; 1992), turned out to be the placeholders for the later emergence of the phylogenetic domains of Archaea and Bacteria.

Closer structural details of the lipids of the pre-cells must remain undecided at present. They may have been bis-esters of fatty acids (as in Bacteria) or bis-ethers of isoprenoid alcohols (as in Archaea) or bis-esters of isoprenoid acids (Wächtershäuser, 1992). Moreover, the lipids of the pre-cell membranes may well have had a *combination* of head group modifications. For it has been shown that the head group interactions of such lipid mixtures produce increased membrane stability (Safran *et al*., 1990), notably if this also means an avoidance of undue accumulations of negative charges. The removal of these theoretical ambiguities may have to wait for further metabolic and phylogenetic studies as well as lipid experiments. Experimental tests could be carried out *in vivo* by genetic engineering of lipid biosynthesis and *in vitro* by membrane or vesicle formation from lipid mixtures or by sonicating two types of vesicles separately obtained from the two enantiomeric or diastereomeric lipid types for inducing fusions and fissions. This test programme may be aided by theoretical calculations of chiral discrimination (Pelizzola *et al*., 2000) in curved lipid membranes. This means that the theory is mature enough to expect improvements from conclusive experimental falsifications.

The proposal of frequent fusions and fissions provides a simple explanation for the assumed frequent mutual exchange of genetic information of the pre-cells. For such frequent fusions and fissions lead to a wholesale exchange and assortment of chromosomes (Wächtershäuser, 1988) between different pre-cells. For fully appreciating this effect it is helpful to distinguish between processes of lateral transfer of individual genes (or operons of genes) and processes of fusion and fission of pre-cells. A process of lateral transfer of discrete genes or operons of genes (Woese, 1998) occurs from one cell of one species to another cell of another species. But the fusion of two pre-cells leads to a new, enlarged pre-cell with a combined 'polyploidal' genome, which is partitioned (cf. Woese, 1998) between two daughter pre-cells during a subsequent fission. This means that the process of fusion and fission of pre-cells is a highly promiscuous quasi-sexual process. It generates a huge gene pool undergoing a massively parallel evolution of the pre-cell population due to large-scale reshuffling of the chromosomes, which themselves must have replicated with high accuracy. In this large-scale chromosome reshuffling the genetic endowments for uptake and utilization of food and energy became differently assorted under the physical and chemical selection pressures in different habitats. But the chromosomes with genes for transcription and translation (and perhaps replication) stayed universal over all habitats. This explains how the pre-cell population could have been at the same time multiphenotypical as well as unitary.

Divergence of the domains Bacteria and Archaea

Turning now to the problem of the divergence of the domain Bacteria the view (Koga *et al*., 1998) is here adopted that the emergence of an enzyme for the stereospecific formation of glycerol-3-phosphate units (either at the level of dihydroxyacetone phosphate reductase or at a later level) was decisive. It had the immediate effect that the lipids became homochiral in their glycerol moiety. They were chirally identical to the dominant enantiomer of the lipids of pre-cells B at stage PC-1. Therefore, this subpopulation of pre-cells B became automatically fixed biosyn-

thetically as the first population of true bacterial cells, i.e. the first bacterial common ancestor (Fig. 1A). In a similar fashion at the later stage PC-2 the emergence of an enzyme for the stereo-specific formation of glycerol-1-phosphate units (Koga *et al*., 1998) led to lipids that were chirally identical to the dominant enantiomer of the lipids of pre-cells A. Therefore, this subpopulation of pre-cells A became automatically fixed biosynthetically as the first population of true archaeal cells, i.e. the first archaeal common ancestor (Fig. 1A). In accordance with the proposal by Koga *et al.* (1998) the two antipodal enantioselective enzymes must have originated from different ancestor enzymes, perhaps by recruitment. Incidentally, the two enantioselective pioneer enzymes of the two antipodal lipid pathways may have appeared at the same pathway step or at different pathway steps by enzyme recruitment from different ancestral enzymes.

This account satisfies the principle of continuity. It corresponds neatly to the tree topology according to the pre-cell theory: a population of Bacteria with a homo-chiral lipid membrane and a population of Archaea with the antipodal homo-chiral lipid membrane diverged consecutively from the evolving population of pre-cells. The Bacteria are assumed to have diverged first and at a time, when the evolving pre-cells at stage PC-1 had a relatively simple information processing machinery. By the time the Archaea diverged, the universal pre-cells at stage PC-2 had a significantly more complex information processing machinery. Thus Kandler's pre-cell theory readily explains this fundamental complexity difference between the domains Bacteria and Archaea.

Now, it is important to note that fusions must have occurred frequently amongst the bacterial cells creating a unitary population of Bacteria. Fusions between Bacteria and pre-cells while still possible occurred less frequently due to the differences of the lipids. Therefore, the unitary population of Bacteria embarked on its own distinct process of evolution, which caused a progressive alienation of the unitary population of the Bacteria from the pre-cells. Given this progressive alienation chimaeric fusions between a bacterial cell and a pre-cell must have become progressively deleterious or even lethal because of increasing incompatibilities of the cellular machineries. This means that any mutation within a certain subpopulation of Bacteria further inhibiting chimaeric fusions must have given a selective advantage to this subpopulation over the others, which would have been swallowed up by the pre-cells. This means that the lipids of the bacteria would have evolved away from the lipids of the pre-cells. The same holds for the Archaea. In this way the bacterial lipids underwent head group modification and restriction mainly to saturated ester lipids of linear fatty acids. The archaeal lipids underwent head group modification and restriction mainly to ether lipids of methyl-branched isoprenoid alcohols. As this hypothesis means that the compatibility between lipid membranes decreases with progressive lipid modification it can be readily tested experimentally by sonicating a mixture of vesicles of different lipid membranes.

Turning now to the evolutionary fate of the genetic machinery we note that the plasmid-type chromosomes of the unitary population of Bacteria combined to one large circular chromosome. This had the advantage of facilitating linkage of replication with cell division. Some plasmids remained and their replication became synchronized with cell division. Others turned into rolling circle viruses. At the same time replication and transcription became more and more separated. The same sequence of events occurred independently within the unitary population of Archaea. Within the unitary populations of the Bacteria and Archaea enzyme complexes for modern DNA replication (Leippe *et al*., 1999) with co-ordinated leading strand and lagging strand synthesis departed independently from the much simpler system of rolling circle replication of the pre-cells. This explains the pronounced differences between replication of the Bacteria and replication of the Archaea (Leippe *et al*., 1999). In addition the mechanisms of transcription and translation were refined independently in both domains (Woese, 1998).

With the separate appearance of biosynthetic pathways for both types of chiral lipids the divergence of two stable lineages of domain evolution with one or the other chiral type would become highly favoured to the point of becoming inescapable. The Bacteria continued to evolve as a unitary population until the emergence of fusion-prohibiting cell walls (Woese, 1983; Kandler, 1994a,b; 1998). The last unitary population of Bacteria, within which this happened is here denoted as 'last bacterial common ancestor' (LBCA). Thereafter the domain Bacteria diverged into a number of phyla (see Fig. 1A). Similarly the Archaea continued to evolve as a unitary population until the emergence of fusion-prohibiting cell walls. The last unitary population of Archaea, within which this happened is here denoted as 'last archaeal common ancestor' (LACA). Thereafter the domain Archaea diverged into a number of phyla (see Fig. 1A).

The above proposal appears presently to be logically independent from and hence compatible with any one of the various hypotheses concerning the thermal evolution of life. Specifically, at one extreme it is compatible with an irreversible thermally downward evolution. In this case the relatively high temperature of the primordial habitats would have posed a burden for the emergence of a self-supporting lipid membrane. The lipids would have to have been already rather evolved. But the rates of fusion and fission of the pre-cells would have been increased by the high temperature. The emergent fatty acid ester lipids

would have to be stable enough at the high temperature. The appearance of isoprenoid ether lipid structures would have been driven not by thermal stability requirement but rather by domain segregation. At the other extreme the present theory is also compatible with a mesophilic origin of life and a thermally upward evolution (Forterre, 1996). In this case the appearance of the isoprenoid ether lipid structures would have been driven by thermal adaptation, even though it might seem odd that the relatively stable carboxylate ester bonds would have been replaced by more stable ether bonds while the less stable phosphate ester bonds would have been maintained. The above-outlined experimental programme could well discriminate between these alternatives.

Divergence of the domain Eukarya

The postulate of chiral segregation explains by physical necessity the existence of two mutually exclusive and jointly exhaustive primary domains: Bacteria and Archaea. But why is there the third domain Eukarya with bacteria-type lipids and what is its mechanistic origin?

The following answer is based on Kandler's assumption that after the divergence of the Bacteria and the Archaea, the population of pre-cells continued to evolve up to stage PC-3. At this stage eukaryotism began with the formation of a nucleus by endosymbiosis. This endosymbiosis is suggested to have occurred between the now more evolved pre-cells type B (endosymbionts) and bacteria (hosts). This endosymbiosis may well have been forged by a mechanism according to an important proposal by Martin and Müller (1998), which has been inspirational for the present theory. Here, however, a symbiotic association is assumed between a subpopulation of heterotrophic, H_2-producing pre-cells B and a population of autotrophic, H_2-consuming bacteria with a progressive biochemical dependence and physical contact until the pre-cells became completely enclosed by the bacteria (Fig. 1B). One of the revolutionary aspects of the proposal by Martin and Müller, which is incorporated into the present theory, must be seen in its independence from phagocytosis or even a pre-existent cytoskeleton. The special symbiosis here proposed generated a population of 'protokarya' (Kandler, 1994b) with an external bacteria-type membrane and a poly chromosomal pre-cell-derived nucleus having a double membrane. The outer nuclear membrane was of bacterial type while the inner nuclear membrane was that of pre-cells B. Under the influence of the bacterial (host) enzymes for lipid biosynthesis the inner nuclear membrane was soon fixed biosynthetically to a strictly bacterial type.

As a result of the composite cell structure the protokarya were forced to evolve in a fundamentally different way than the Bacteria and Archaea. We begin with the problem that nutrients as well as waste products of the metabolism of the endosymbionts had to pass through three membranes. This created a selection pressure for relocating the metabolism of the endosymbiont into the outer cytoplasm thus irreversibly forming an integrated metabolism with pathways derived from the pre-cell endosymbiont and pathways derived from the bacterial host. Concomitantly the endosymbiont was converted into a nucleus. Whereas some translation appears to have remained to this day inside the nucleus (Brogna, 2001; Ibora *et al*., 2001) most protein synthesis was relocated into the outer cytoplasm. This explains the extant roundabout biosynthesis of the ribosome, some ribosomal proteins being imported into the nucleus for the assembly of incomplete ribosomal subunits to be exported and completed outside the nucleus. All mRNAs generated inside the nucleus also had to be transported into the outer cytoplasm. Therefore, nuclear pores were invented. This means that the organization of genes in gene clusters or operons lost its benefit, which previously consisted in the conjunction of transcription and translation of polycistronic mRNAs. The genes became solitary and the mRNAs became short, capped and polyadenylated. This theory explains why the Eukarya are most closely related to Archaea with regard to information processing while their metabolism and lipids are bacterial. Further, it explains the apparent primitive nature of the Eukarya by the fact that the nucleus derived mainly by fixation of the genome of the primitive non-streamlined pre-cells of stage PC-3. The further simple assumption that the bacteria-derived cell membrane and the bacteria-derived outer nuclear membrane grew at a higher rate than the pre-cell-derived inner nuclear membrane is sufficient to explain the generation of an endoplasmatic reticulum within an enlarged protokaryal cell. The chromosomes of the host were abandoned with individual genes being relocated into the nucleus. Linear chromosomes were derived from the linear intermediates of the rolling circle replication of the pre-cell chromosomes, perhaps with a multiplication in length.

Because of separation of the chromosomes from the outer cell membrane by the nuclear double membrane there was no selective advantage in their unification into one large circular chromosome, as its replication could not have become linked to the outer cell membrane during cell division. Instead, a microtubule spindle apparatus emerged for mitosis. Mitosis was pioneered as closed mitosis within a closed nucleus as shown by most eukaryal phyla. Open mitosis with a fragmentation of the nuclear membrane that is found mainly in animals and higher plants (Hausmann, 1985) must have appeared much later. Only after the appearance of microtubules did phagocytosis become possible.

The origin of the joint precursor of mitochondria and hydrogenosomes by endosymbiosis of a hydrogen-

generating species of alpha-proteobacteria and hydrogen-consuming protokarya would have occurred later after the appearance of the alpha-proteobacteria much like proposed by Martin and Müller (1998).

Within the above theory it has been assumed that the Eukarya diverged after the divergence of the Archaea and that the pre-cells PC-3 from which the Eukarya diverged were more evolved than the pre-cells PC-2 from which the Archaea diverged. Alternatively, however, it may be assumed that the Eukarya diverged at about the same time as the Archaea or even before the Archaea. The empirical evidence presently does not allow to discriminate between these possibilities. In closing it should be emphasized that in contrast to autogenous theories of the origin of the nucleus (Martin, 1999) the above explanation has the advantage of being based on a concatenation of specific selection pressures triggered by the original symbiosis.

Origin of the pre-cells – an afterthought

The above theory of domain emergence does not depend on any particular theory on the origin and early evolution of lipids. It is however, most readily agreeable with the theory of a primordial chemo-autotrophic anabolism on surfaces of colloidal or microcrystalline particles with transition metal/sulphur structure (Wächtershäuser, 1988, 1992) giving rise to long-chain carboxylic acid lipids, notably fatty acid lipids or unsaturated isoprenoid acid lipids like geranylgeranoic acid. The formation of acetyl-thioester (CH_3-CO-S-CH_3) from CO and H_2S has been demonstrated experimentally in the presence of co-precipitated (Fe,Ni)S under conditions of volcanic exhalations (Huber and Wächtershäuser, 1997). Today both the fatty acid lipid pathway and the mevalonate isoprenoid lipid pathway begin with acetyl-thioester via malonic acid thioester. Therefore, primordial lipid biosynthesis may well have been based on thioester condensations. Such condensations may have proceeded via 3-methyl-glutaconic acid thioester (RSOC-CH = C(CH_3)-CH_2-COOH), a vinyloge of malonic acid thioester, to unsaturated achiral isoprenoid acids, like geranylgeranoic acid (Wächtershäuser, 1992).

What was the first function of these long-chain carboxylic acid lipids? In the presence of mineral catalysts an accumulation of lipids has the inevitable consequence of surface lipophilization (Wächtershäuser, 1988; 1992). This means a locally lowered water and proton activity and a suppression of hydrolytic reactions. Thereby, numerous condensation reactions including lipid synthesis itself are promoted (Wächtershäuser, 1988; 1992). At a certain surface concentration achiral carboxylic acid lipids would self-organize to mineral-supported membranes. These lipids would have evolved in the direction of better and better surface lipophilization. Ultimately the surface-bonded lipids would have self-organized into self-supporting closed vesicle membranes around the colloidal or microcrystalline mineral particles. Such early vesicles would have been unstable, notably in a hot environment. Nevertheless, they afforded already an effective protection and isolation of the constituents of the metabolism and of the unfolding genetic machinery from adverse chemical conditions of the environment. This had an immediate selective advantage. The further evolution proceeded therefore automatically in the direction of increasing stability of the self-supporting membrane vesicles. Specifically, with the emergence of racemic glycerol-1-phosphate the achiral carboxylic acid lipids were replaced by racemic lipids of bis-esters of long-chain carboxylic acids with glycerol-phosphate while still bonded to surfaces. Subsequently head group modification provided for greater and greater stability of the self-supporting membranes. This is seen as the origin of the pre-cells in the first phase of evolution.

Turning finally to the evolution of isoprenoid lipids we note that the extant isoprenoid pathway proceeds exclusively through isoprenoid alcohol diphosphates. Early on, during pre-cell evolution prior to PC-1, these may have been formed by reduction of the corresponding isoprenoid acid thioesters and subsequent phosphorylation. This would explain the universality of the isoprenoid pathway downstream from isopentenyl pyrophosphate. Later a reductive conversion of 3-hydroxy-3-methylglutaroyl thioester (RSOC-CH_2-C(OH)(CH_3)-CH_2-COOH), to mevalonic acid (HO-CH_2-CH_2-C(OH)(CH_3)-CH_2-COOH) would have emerged in pre-cell evolution between PC-1 and PC-2, which ushered in the mevalonate pathway that is found in all Archaea and Eukarya. In the bacterial domain by contrast the non-mevalonate pathway to IPP (Rohdich *et al*., 2002) via 2C-methyl-D-erythritol 4-phosphate emerged between PC-1 and LBCA and it remained restricted to the domain Bacteria and to the bacteria-derived plastids.

Acknowledgements

I would like to express my gratitude first and foremost to Otto Kandler for discussions regarding his pre-cell theory and for reading several drafts of the manuscript. Experimental demonstration of thioester formation by Claudia Huber, funded by the Deutsche Forschungsgemeinschaft, has been a critical support. I express my gratitude to Dorothy Wächtershäuser for decisive help in the presentation of the theory and to Adelbert Bacher and Nicolas Glansdorff for helpful discussions.

References

Ackermann, T. (1992) *Physikalische Biochemie*. Berlin: Springer Verlag, pp. 161–165.

Brogna, S. (2001) Pre-mRNA processing: insights from nonsense. *Curr Biol* **11:** R838–R841.

Cavalier-Smith, T. (1987) The origin of eukaryotic and archaebacterial cells. *Ann New York Acad Sci* **503:** 17–54.

Cavalier-Smith, T. (2002) The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int J Syst Evol Microbiol* **52:** 7–76.

Elferink, M.G., de Wit, J.G., Demel, R., Driessen, A.J., and Konings, W.N. (1992) Functional reconstitution of membrane proteins in monolayer liposomes from bipolar lipids of *Sulfolobus acidocaldarius*. *J Biol Chem* **267:** 1375–1381.

Forterre, P. (1996) A hot topic: the origin of hyperthermophiles. *Cell* **85:** 789–792.

Forterre, P. (2001) Genomics and early cellular evolution. The origin of the DNA world. *C R Acad Sci Paris, Sci la Vie/Life Sci* **324:** 1067–1076.

Gupta, R.S., and Golding, G.B. (1996) The origin of the eukaryotic cell. *Trends Biochem Sci* **21:** 166–171.

Hartman, H., and Fedorov, A. (2002) The origin of the eukaryotic cell: a genomic investigation. *Proc Natl Acad Sci USA* **99:** 1420–1425.

Hausmann, K. (1985) *Protozoologie.* Stuttgart: Georg Thieme Verlag.

Horiike, T., Hamada, K., Kanaya, S., and Shinozawa, T. (2001) Origin of nucleic acid cell nuclei by symbiosis of Archaea in Bacteria is revealed by homology-hit analysis. *Nat Cell Biol* **3:** 210–214.

Huber, C., and Wächtershäuser, G. (1997) Activated acetic acid by carbon fixation on (Fe,Ni) S under primordial conditions. *Science* **276:** 245–247.

Ibora, F.J., Jackson, D.A., and Cook, P.R. (2001) Coupled transcription and translation within the nuclei of mammalian cells. *Science* **293:** 1139–1142.

In't Veld, G., Elferink, M.G., Driessen, A.J., and Konings, W.N. (1992) Reconstitution of the leucine transport system of Lactococcus lactis into liposomes composed of membrane-spanning lipids from *Sulfolobus acidocaldarius*. *Biochemistry* **31:** 12493–12499.

Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., and Miyata, T. (1989) Evolutionary relationship of Archaebacteria, Eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc Natl Acad Sci USA* **86:** 9355–9359.

Kaganer, V.M., Möhwald, H., and Dutta, P. (1999) Structure and phase transitions in Langmuir monolayers. *Rev Mod Phys* **71:** 779–819.

Kandler, O. (1994a) The early diversification of life. In *Nobel Symposium No. 84.Early Life on Earth*. Bengtson, S. (ed.). New York: Columbia University Press, pp. 152–160.

Kandler, O. (1994b) Cell wall biochemistry in Archaea and its phylogenetic implications. *J Biol Phys* **20:** 165–169.

Kandler, O. (1998) The early diversification of life and the origin of the three domains: a proposal. In *Thermophiles: the Keys to Molecular Evolution and the Origin of Life*. Wiegel, J., and Adams, M.W.W. (eds). London: Taylor & Francis, pp. 19–28.

Kates, M. (1993) Membrane lipids in Archaea. In *The Biochemistry of Archaea (Archaebacteria).* Kates, M., Kushner, D.J., and Matheson, A.T. (eds). Amsterdam: Elsevier, pp. 261–295.

Koga, Y., Kyuragi, T., Nishihara, M., and Sone, N. (1998) Did archaeal and bacterial cells arise independently from noncellular precursors? A hypothesis stating that the advent of membrane phospholipid with enantiomeric glycerophosphate backbones caused the separation of the two lines of descent. *J Mol Evol* **46:** 54–63.

Leippe, D.D., Aravind, L., and Koonin, E.V. (1999) Did DNA replication evolve twice independently? *Nucleic Acids Res* **27:** 3389–3401.

Martin, W. (1999) A briefly argued case that mitochondria and plastids are descendents of endosymbionts, but that the nuclear compartment is not. *Proc R Soc Lond B* **266:** 1387–1395.

Martin, W., and Müller, M. (1998) The hydrogen hypothesis for the first eukaryote. *Nature* **392:** 37–41.

Morii, H., Nishihara, M., and Koga, Y. (2000) CTP. 2,3-di-*O*-geranylgeranyl-*sn*-glycero-1-phosphate cytidyltransferase in the methanogenic archaeon *Methanothermobacter thermoautotrophicus*. *J Biol Chem* **275:** 36568–36574.

Nassoy, P., Goldman, M., Bouloussa, O., and Rondolez, F. (1995) Spontaneous chiral segregation in bidimensional films. *Phys Rev Lett* **75:** 457–460.

Nishihara, M., Kyuragi, T., Sone, N., and Koga, Y. (1998) *sn*-Glycerol-1-phosphate dehydrogenase: a key enzyme in the biosynthesis of ether phospholipids in Archaea. In *Thermophiles: the Keys to Molecular Evolution and the Origin of Life*. Wiegel, J., and Adams, M.W.W. (eds). London: Taylor & Francis, pp. 19–28.

Pelizzola, A., Pretti, M., and Scalas, E. (2000) Heterochirality in Langmuir layers and antiferromagnetic Blume-Emery-Griffiths model. *J Chem Phys* **112:** 8126–8136.

Redenbach, M. (2002) Warum haben einige Bakterien lineare Chromosomen und Plasmide? *Biospektrum* **2:** 158–163.

Rohdich, F., Hecht, S., Gärtner, K., Adam, P., Krieger, C., Amslinger, S., *et al.* (2002) Studies on the nonmevalonate terpene biosynthetic pathway: metabolic role of IspH (LytB) protein. *Proc Natl Acad Sci USA* **99:** 1158–1163.

Sackmann, E. (1982) Physikalische grundlagen der molekularen organisation und dynamik von membranen. In *Biophysik*. Hoppe, W., Lohmann, W., Markl, H., and Ziegler, H. (eds). Berlin: Springer, pp. 439–471.

Safran, S.A., Pincus, P., and Andelman, D. (1990) Theory of spontaneous vesicle formation in surfactant mixtures. *Science* **248:** 354–355.

Sogin, M.L. (1991) Early evolution and the origin of eukaryotes. *Curr Opin Gen Dev* **1:** 457–463.

Uragami, M., Miyake, Y., and Regen, S.L. (2000) Influence of headgroup chirality on the mixing behavior of phosphatidylglycerol mimics in fluid bilayers. *Langmuir* **16:** 3491–3496.

Van Wely, K.H.M., Swaving, J., Freudl, R., and Driessen, A.J.M. (2001) Translocation of proteins across the cell envelope of Gram-positive bacteria. *FEMS Microbiol Rev* **25:** 437–454.

Vollhardt, D., Emrich, G., Gutberlet, T., and Fuhrhop, J.H. (1996) Chiral discrimination and pattern formation in N-dodecylmannonamide monolayers at the air-water interphase. *Langmuir* **12:** 5659–5663.

Wächtershäuser, G. (1988) Before enzymes and templates: Theory of surface metabolism. *Microbiol Rev* **52:** 452–484.

Wächtershäuser, G. (1992) Groundworks for an evolutionary biochemistry – the iron-sulfur world. *Prog Biophys Mol Biol* **58:** 85–201.

Wächtershäuser, G. (1998) Towards a reconstruction of ancestral genomes by gene cluster alignment. *System Appl Microbiol* **21:** 473–477.

Woese, C.R. (1983) The primary lines of descent and the universal ancestor. In *Evolution from Molecules to Men*. Bendall, D.S., (ed.). Cambridge: Cambridge University Press, pp. 209–233.

Woese, C.R. (1987) Bacterial evolution. *Microbiol Rev* **51:** 221–271.

Woese, C.R. (1998) The universal ancestor. *Proc Natl Acad Sci USA* **95:** 6854–6859.

Woese, C.R. (2002) On the evolution of cells. *Proc Natl Acad Sci USA* **99:** 8742–8747.

Woese, C.R., Kandler, O., and Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria and Eucarya. *Proc Natl Acad Sci USA* **87:** 4576–4579.

Zillig, W., Palm, P., and Klenk, H.-P. (1992) The nature of the common ancestor of the three domains of life and the origin of the Eucarya. In *Frontiers of Life.* Tran Thanh Van, J., Mounolou, J.C., Schneider, J., and McKay, C. (eds). Gif-sur-Yvette: Editions Frontiers, pp. 181–193.