SPECULATIONS ON THE ORIGIN OF LIFE AND THERMOPHILY: REVIEW OF AVAILABLE INFORMATION ON REVERSE GYRASE SUGGESTS THAT HYPERTHERMOPHILIC PROCARYOTES ARE NOT SO PRIMITIVE

PATRICK FORTERRE[1],* FABRICE CONFALONIERI[2], FRANCK CHARBONNIER[1] and MICHEL DUGUET[2]

Laboratoire des Archéobactéries[1] et laboratoire d'Enzymologie des acides nucléiques[2]. Institut de Génétique et Microbiologie, CNRS, URA 1354 Université Paris-Sud, Bat. 409, 91405 Orsay Cdex, France

(Received November 5, 1993)

Abstract. All present-day hyperthermophiles studied so far (either *Bacteria* or *Archaea*) contain a unique DNA topoisomerase, reverse gyrase, which probably helps to stabilize genomic DNA at high temperature. Herein the data relating this enzyme is reviewed and discussed from the perspective of the nature of the last detectable common ancestor and the origin of life. The sequence of the gene encoding reverse gyrase from an archaeon, *Sulfolobus acidocaldarius*, suggests that this enzyme contains both a helicase and a topoisomerase domains (Confalonieri *et al., Proc. Natl. Acad. Sci.*, 1993, 90, 4735). Accordingly, it has been proposed that reverse gyrase originated by the fusion of DNA helicase and DNA topoisomerase genes. If reverse gyrase is essential for life at high temperature, its composite structure suggests that DNA helicases and topoisomerases appeared independently and first evolved in a mesophilic world. Such scenario contradicts the hypothesis that a direct link connects present day hyperthermophiles to a hot origin of life. We discuss different patterns for the early cellular evolution in which reverse gyrase appeared either before the emergence of the last common ancestor of *Archaea, Bacteria* and *Eucarya*, or in a lineage common to the two procaryotic domains. The latter scenario could explain why all today hyperthermophiles are procaryotes.

1. Introduction

Several authors have suggested the existence of a direct evolutionary link between the origin of life and a thermophilic common ancestor to all present day organisms (Woese, 1987; Wächtershauser, 1988; Pace, 1991; Stetter, 1992). The idea that life originated at high temperature is supported by some hypotheses about the primitive atmosphere involving a hot early Earth (for review see Kasting, 1993). On the other hand, the concept of a thermophilic common ancestor to the three domains of life (*Archaea, Bacteria* and *Eucarya*) has been inferred from the clustering of thermophilic procaryotes at the roots of the two procaryotic domains in 16S rRNA phylogenetic trees (Woese, 1987). Recently, the thermophilic nature of the last detectable common ancestor has been supported by the rooting of the universal tree of life in the bacterial branch (Woese *et al.*, 1990; Stetter, 1992), based on the analysis of composite phylogenetic trees of paralogous ATPases and elongation factors (Iwabe *et al.*, 1989; Gogarten *et al.*, 1989). However, the hot origin of life

* Corresponding author: Patrick Forterre, Tel: 33 1 69 41 74 89, Fax: 33 1 69 41 78 08, e-mail: forterre@igmors.u-psud.fr.

Origins of Life and Evolution of the Biosphere **25:** 235–249, 1995.

hypothesis has also been criticized, taking into account the instability of essential biological molecules, in particular RNA, at high temperature (Miller and Bada, 1988; Forterre, 1992). Furthermore, it has been recently shown that both new ATPase sequences and revaluation of the elongation factor phylogeny cast doubts on the rooting of the universal tree in the bacterial branch (Forterre *et al.*, 1993). At present, the correct rooting of the tree of life remains uncertain and therefore cannot be used to evaluate the hypothesis of a thermophilic common ancestor to the three domains. Furthermore, the weight of evidence available at present strongly imply that this universal ancestor was already a member of the DNA world (for reviews, see Lazcano *et al.*, 1992; Forterre *et al.*, 1994). Therefore, even if this ancestor was a hyperthermophile, the actual origin of life might have occurred either at high or low temperature.

Another way to test the hypothesis of a direct link between present day hyperthermophiles and a hot origin if life is to identify key molecular features required for life at high temperatures and ask the question: are there primitive traits or adaptative features?

One such feature is a unique DNA topoisomerase, reverse gyrase, which was discovered several years ago in the archaeon *Sulfolobus acidocaldarius* (Kikuchi and Asai, 1984). This enzyme is the only DNA topoisomerase which can introduces positive superturns into the DNA, i.e. it increases the number of topological links between the two strands of the double-helix (review in Forterre and Elie, 1993). From the beginning, it has been hypothesized that reverse gyrase could be essential to stabilize the DNA at high temperature in hyperthermophiles (Kikuchi *et al.*, 1986; Nadal *et al.*, 1986). Is reverse gyration of the DNA a primitive trait of life, directly connected to its putative hot origin, or else, is it an adaptative feature to high temperature? In other words, is reverse gyrase a new or an old enzyme ? Herein we discuss our recent results about reverse gyrase and their implications to the origin of this enzyme and the hot origin of life hypothesis.

2. Reverse Gyrase and Its Distribution in the Living World

Reverse gyrase is a type I DNA topoisomerase (Forterre *et al.*, 1985; Nakasu and Kikuchi, 1985), i.e. an enzyme which changes the DNA linking number in step of one unit, making a transient single-stranded break into one of the two DNA strands and passing the other strand through this break. This enzyme can use as substrate either a negatively supercoiled DNA (Kikuchi and Asai, 1984), or a relaxed DNA (Forterre *et al.*, 1985). In both case, reverse gyrase *increases* the linking number, ending up with a positively supercoiled product.

Reverse gyrases from different organisms are composed of a single polypeptide of apparent molecular mass around 120–130,000 (Nakasu and Kikuchi, 1985; Nadal *et al.*, 1988; Slezarev, 1988; Andere *et al.*, 1993). This monomeric structure is typical for type I DNA topoisomerases. However, reverse gyrase is unique among type I DNA topoisomerases since it is the only one which is ATP dependent,

TABLE I

Optimal growth temperatures of the archaeal and bacterial species in which reverse gyrase has been detected

Species	Optimal growth temperatures	References
ARCHAEA		
Pyrodictium occultum	105	Bouthier *et al.*, 1990
Methanopyrus kandlerii	100	Bouthier *et al.*, 1990
Pyrococcus furiosus	100	Bouthier *et al.*, 1990
Pyrococcus abyssi	100	Charbonnier *et al.*, 1992
Pyrobaculum Islandicum	100	Bouthier *et al.*, 1990
Thermococcus celer	92	Collin *et al.*, 1990
Staphylothermus marinus	92	Bouthier *et al.*, 1990
Desulfurococcus sacchavarovorans	92	Bouthier *et al.*, 1990
Acidianus infernus	90	Bouthier *et al.*, 1990
Thermoproteus neutrophilus	88	Bouthier *et al.*, 1990
Thermoproteus tenax	88	Collin *et al.*, 1988
Sulfolobus solfataricus	85	Collin *et al.*, 1988
Desulfurococcus mucosus	85	Collin *et al.*, 1988
Desulfurococcus mobilis	85	Collin *et al.*, 1988
Desulfurococcus amylolyticus	85	Slezarev, 1988
Sulfolobus shibatae	85	Bouthier *et al.*, 1990
Archaeoglobus fulgidus	83	Bouthier *et al.*, 1990
Methanothermus fervidus	83	Bouthier *et al.*, 1990
Sulfolobus acidocaldarius	80	Kikuchi and Asai, 1984
Desulfurolobus ambivalens	80	Charbonnier and Forterre, 1994
BACTERIA		
Thermotoga maritima	80	Bouthier *et al.*, 1991
Thermosipho africanus	75	Bouthier *et al.*, 1991
Calderobacterium hydrogenophilum	75	Andera *et al.*, 1993
Thermotoga thermarum	70	Bouthier *et al.*, 1991
Fervidobacterium islandicum	65	Bouthier *et al.*, 1991

ATP hydrolysis being required to drive the energetically unfavorable supercoiling reaction.

The discovery of reverse gyrase in a hyperthermophilec raises the question of whether this enzyme is indeed essential for life at high temperature. Until now, this has not been demonstrated directly, as a reverse gyrase mutant has yet to be isolated. Nevertheless, a strong indirect argument in favor of this hypothesis can be inferred from the distribution of reverse gyrase in the living world (Collin *et al.*, 1988; Bouthier de la Tour *et al.*, 1990, 1991). Reverse gyrase has never been

found in thrue mesophiles, in contrast this enzyme has been found in some species growing optimally between 60 and 80 °C and in *all* species growing optimally above 80 °C, either *Archaea* or *Bacteria*, tested so far (Table I). One of the most striking observation is that, in a set of closely related methanogenic archaea, the presence of reverse gyrase correlates with the optimal growth temperature of the organism (Bouthier de la Tour *et al.*, 1990). Reverse gyrase has been detected in some thermophilic bacteria which can even grow at mesophilic temperatures, for example the bacterion *Thermosipho africanus*, whereas other lack this activity, such as the bacterion *Thermus thermophilus*, indicating that reverse gyrase is facultative in thermophiles. However, its ubiquity in hyperthermophiles strongly suggests that the role of reverse gyrase is connected to high temperature and that this enzyme is essential for life at temperatures characteristic of hyperthermophiles.

Another indication of the importance of reverse gyrase *in vivo* came from the study of the DNA topological state in hyperthermophiles. It was discovered that the virus SSV1 from the hyperthermophilic archaeon *Sulfolobus shibatae* (formerly *Sulfolobus solfataricus* B12) harbors a positively supercoiled genome (Nadal *et al.*, 1986). More recently, it has been learned that plasmids isolated from hyperthermophilic archaea containing reverse gyrase have a higher linking number than plasmids of similar size isolated from mesophilic archaea or mesophilic and thermophilic bacteria (Forterre *et al.*, 1992; Charbonnier *et al.*, 1992; Charbonnier and Forterre, 1994). There is one known exception: a plasmid with a high linking number isolated from the moderately thermophilic methanogenic archaea *Methnobacterum thermoautotrophicum* in which we failed to detected a reverse gyrase activity (Charbonnier and Forterre, 1994). When the superhelical densities of these plasmids is extrapolated at the optimal growth temperature of their host organisms, using an equation that relates the DNA winding angle to the temperature up to 90 °C (Charbonnier *et al.*, 1992, Dugguet, 1993), the results indicate that plasmids from hyperthermophilic archaea containing reverse gyrase are relaxed, instead of being negatively supercoiled, as in mesophilic organisms. We also found that the linking number of the plasmid pGT5, isolated from the hyperthermophilic archaeon *Pyrococcus abyssi*, (formerly strain GE5) increases with the growth temperature of the organism (Charbonnier and Forterre, unpublished data). All there results suggest that reverse gyrase is required in hyperthermophiles to increase the DNA linking number *in vivo*.

It is not yet clear how reverse gyrase stabilizes the DNA in hyperthermophiles. A topologically closed DNA, even negatively supercoiled, is resistant to denaturation at high temperatures (Vinograd *et al.*, 1968; Marguet and Forterre, 1994) and a positively supercoiled DNA is no more resistant to depurination and hydrolysis of the phosphodiester bonds at high temperature than a negatively supercoiled one (Marguet and Forterre, 1994). A possible model is that the increase of the linking number mediated by reverse gyrase is required to compensate the passive untwisting of the DNA at high temperature, in order to maintain the helical repeat of the double helix at a constant length (Forterre *et al.*, 1994b).

It has not yet been determined if archaeal and bacterial reverse gyrases are homologous. Antibodies against *S. acidocaldarius* reverse gyrase do not recognize the enzyme in hyperthermophilic archaea other than Sulfolobales (Bouthier de la Tour *et al.*, 1990), accordingly, they cannot be used to check the putative homology of these enzymes in different domains. However, homology between archaeal and bacterial reverse gyrases is suggested by the fact that the purified reverse gyrase from the thermophilic bacteria *Calderobacterium hydrogenophilum* (Andera *et al.*, 1993) and *Fervidobacterium islandicum* (Bouthier de la Tour, unpublished data) have about the same size, enzymatic properties and molecular structure as the archaeal reverse gyrase. If reverse gyrases from the two procaryotic domains are indeed homologous, either this enzyme appeared in one of them and was later trasnferred horizontally into the other, or it should have been already present in the common ancestor of these two domains. In agreement with this second hypothesis, the short branches of the hyperthermophilic lineages in the archaeal and bacterial 16S rTNA trees and their positions near the roots of these trees have suggested that the common ancestor of both *Archaea* and *Bacteria* was a hyperthermophile (Woese, 1987; Stetter, 1992).

If the root of the universal tree of life is located in one of the two procaryotic domains, the presence of reverse gyrase in the common ancestor of both *Archaea* and *Bacteria* would imply that this enzyme was present in the last universal ancestor to the tree domains. In the hot origin of life hypothesis, reverse gyrase should have been even present in the very first cell with a genome composed of a double-stranded circular DNA. However, in the alternative hypothesis of a mesophilic origin of life, other scenarios can be imagined, for example, reverse gyrase might have appeared in the lineage leading to the common ancestor of the two procaryotic domains in a process of adaptation of these microorganisms to high temperature (see below). To determine which of these different possibilities is the most likely, it was important to get insights on the origin of reverse gyrase. This required first to determine its amino acid sequence in order to identify putative homologies with other proteins.

3. The Gene Encoding Reverse Gyrase and its Origin

We have recently cloned and sequenced the gene encoding the reverse gyrase from *S. acidocaldarius* (Confalonieri *et al.*, 1993). The sequence of this gene turns out to be unusual, being the first described combination of an helicase and a DNA topoisomerase gene.

As schematically shown in the Figure 1, the C-terminal half of reverse gyrase (from amino acid 630 to amino acid 1247) aligns with DNA topoisomerases I of a protein family whose prototype is the *E. coli* ω protein. These enzymes are typically present in *Bacteria*, however, an example has been detected in the eucaryon *Saccharomyces cerevisiae* (Wällis *et al.*, 1989). DNA topoisomerases of this family share with reverse gyrase the properties of binding transiently to the

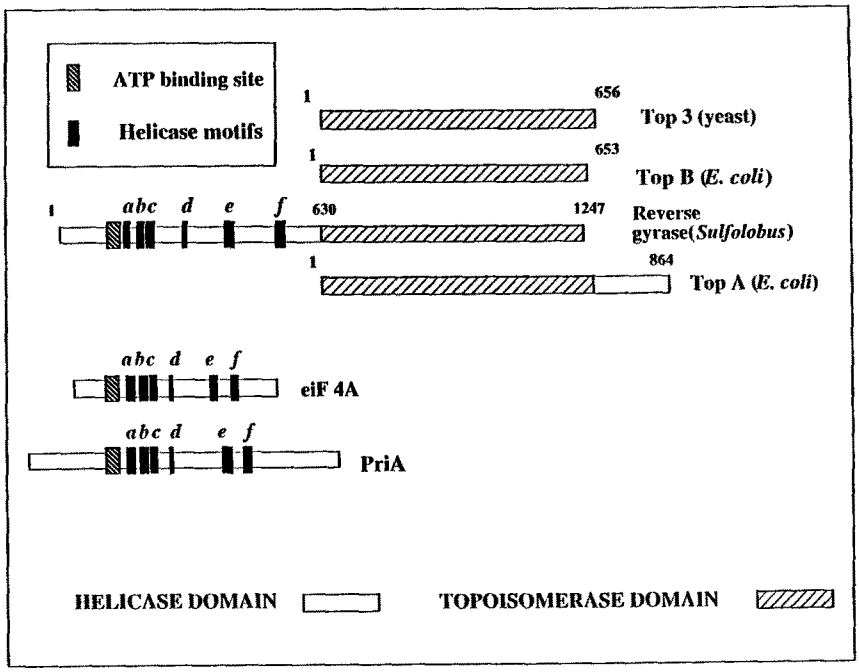

Fig. 1. Composite structure of the gene encoding the reverse gyrase of the archaeon *Sulfulobus acidocaldarius* (adapted from Confalonieri *et al.*, 1993; Forterre *et al.*, 1994). The helicase domain of reverse gyrase is characterized by a Walker box purine NTP-binding site (Walker *et al.*, 1982) and six (a to f) colinear 'helicase motifs'. It can be aligned with eiF4A, an eucaryotic translation initiation factor with an RNA helicase activity, or else with PriA, the DNA helicase of the *E. coli* primosome. The topoisomerase domain of reverse gyrase can be aligned with the type I DNA topoisomerases of the 5′ family: Top 3 (Wällis *et al.*, 1989), the yeast DNA topoisomerase I encoded by the gene TOP3, Top B (DiGate and Marians, 1989), the *E. coli* DNA topoisomerase III encoded by the gene *topB*, Top A (Tse-Dinh and Wang, 1986), the *E. coli* protein ω encoded by the gene *topA*.

5′ end of the DNA break in the course of the reaction, and relaxing negatively supercoiled DNA but not positively supercoiled DNA (i.e. they always increase the DNA linking number). However, in contrast to reverse gyrase, these enzymes do not use ATP for catalysis and, accordingly, none of them can push the reaction far enough to introduce positive supercoiling into the DNA.

The N-terminal half of the reverse gyrase gene, from amino acids 1 to 630, corresponds to an extension which cannot be aligned with any of the known DNA topoisomerase genes. This N-terminal extension contains an A motif of the Walker-type NTP binding pattern (Walker *et al.*, 1982), reading AxxGxGKT, that is most probably involved in the ATP dependent reverse gyration (Fig. 1). Furthermore, beside this ATP binding site, the N-terminal half of reverse gyrase contains several amino acid motifs which are typical of either RNA or DNA helicases (Confalonieri *et al.*, 1993), i.e. enzymes which separate the two strands of an RNA hairpin or a DNA duplex. These motifs are present in the same order as in typical RNA or DNA

helicases of the family SP2 (Gorbalenya and Koonin, 1993), strongly suggesting that the N-terminal domain of reverse gyrase confers to this enzyme an helicase activity.

This composite structure of reverse gyrase suggests a mechanism of action for this enzyme. It has been shown previously that the DNA tracking activity of a DNA helicase produces a wave of positive supercoiling in front of the moving enzyme and a wave of negative supercoiling behind it (Yang *et al.*, 1989). The same phenomenon occurs with an RNA polymerase and was first predicted by Liu and Wang (1988) in a theoretical paper as the 'twin-domains model'. If the helicase domain of reverse gyrase also produces two waves of opposite supercoiling, specific relaxation of the negatively supercoiled domain by the topoisomerase activity would produce a net accumulation of positive superturns (Confalonieri *et al.*, 1993).

Beyond these functional considerations, the composite structure of reverse gyrase has important implications for its origin. If reverse gyrase is an 'old' enzyme, i.e. if it appeared at the same time as the first DNA cell, one might hypothesize that its gene later split during the course of evolution to generate modern DNA helicase and topoisomerase genes. However, this would also imply that the ancestral reverse gyrase gave birth to all present day members of the superfamily of purine-binding proteins that are homologous to its helicase domain. This appears to be highly unlikely since, in that case, reverse gyrase would have appeared before proteins such as membrane transporters, V and F-ATPases, or else elongation and initiation factors for protein synthesis (Schulz, 1992). Accordingly, if reverse gyrase originated at the same time as double-stranded DNA, the most likely hypothesis is that it evolved from a RNA helicase gene endigenous to the RNA world. In that case, the origin of the topoisomerase domain remains unclear. It might have originated from a long C-terminal domain of the putative helicase gene or from the fusion of this RNA helicase with a putative 'RNA topoisomerase' of the RNA world. In both cases, this 'early' scenario for the emergence of reverse gyrase requires that at least all present-day 5′ type I DNA topoisomerases originated from reverse gyrase. Both of these scenatios viewing reverse gyrase as and 'old' enzyme appear quite speculative since they imply that reverse gyrase originated as the first DNA metabolizing enzyme!

In our opinion, a more likely hypothesis is that reverse gyrase originated from the 'late' fusion of a *bona fide* DNA helicase and DNA topoisomerase genes. Indeed many DNA helicases and topoisomerases probably evolved independently from different RNA-metabolizing enzymes of the RNA world during the period included between the appearence of the first DNA cell and the emergence of the last common ancestor to the three domains. This is suggested by the existence of several independent superfamilies of these enzymes (Gorbalenya and Koonin, 1993; Forterre *et al.*, 1994a). The existence of such a 'first age of the DNA world' in which different families of paralogous DNA metabolizing enzymes emerged has been proposed on the basis of an exhaustive analysis of DNA polymerase and topoisomerase evolution (Forterre, 1992; Forterre *et al.*, 1993, 1994b, see also

Laczano *et al.*, 1992). In this context, the late fusion hypothesis for the origin of reverse gyrase fits especially well with the modular evolution of helicases proposed by Gorbalenya and Koonin, (1993). For example, these authors reported that some proteins combine an helicase domain with a DNA site-specific endonuclease domain or a protease domain, either as a C-terminal or an N-terminal extension.

In this scenario, and if future work confirms that reverse gyrase is indeed essential for life at high temperature, organisms living during this first age of the DNA world could not have been hyperthermophiles before the appearance of this enzyme. Accordingly, the late fusion hypothesis for the origin of reverse gyrase contradicts *a prior* the existence of a direct link between present day hyperthermophiles and a hot origin of life. One can still speculate that ancient hyperthermophiles contained a different and primitive version of reverse gyrase, or that they don'd need this enzyme at all. However, a more attractive alternative is that reverse gyrase appeared in a mesophilic or moderately thermophilic lineage and that organisms in which this enzyme first originated gain the ability to live for the first time at temperature typical for hyperthermophiles.

4. Other Arguments Against a Direct Link Between Hyperthermophiles and a Putative Hot Origin of Life

It has been previously suggested that the hot origin of life scenario is in contradiction with the RNA world hypothesis because of the high instability of the RNA molecule at high temperature (Miller and Bada, 1988; Forterre, 1992). The sensitivity of RNA molecules to heat-induced hydrolysis is due to the reactive oxygen in the 2′ position of the ribose which could catalyze the intramolecular cleavage of phosphodiester bonds at high temperature (Ginoza *et al.*, 1964). A striking observation is that RNA thermodegradation is catalyzed by divalent salts, such as Mg^{++} or Zn^{++} (Lindalh, 1967; Butzow and Eichorn, 1975), which were certainly essential for numerous enzymatic activities of the RNA world. Pace (1992) suggested that RNA was stabilized in primitive cells by a very low water activity. One might also speculate that primordial RNA-like molecules were stable at high temperature in a two-dimensional metabolic network, such as the pyrite-based world postulated by Wächtershäuser (1988). However, the activity of water in the more advanced three-dimensional metabolism of RNA-based cells was probably not very different from its value in modern cells, in order to accomodate hydrolytic reactions required in a fully evolved ribozyme-catalyzed metabolism. Indeed, a huge number of the modified tRNAs isolated from modern hyperthermophiles exhibit methylation of the ribose 2′ oxygen (Edmonds *et al.*, 1991), which is precisely the Achille's heel of the RNA molecule at high temperature. In general, hyperthermophiles contain far more base modifications in their tRNA molecules than mesophiles, including many that are unique (Edmonds *et al.*, 1991). It has been shown in at least one case that these modifications are essential to prevent denaturation of the tRNA at temperature near the boiling point of water. In particular, whereas the unmodified

tRNA from the archaeon *Pyrodictium occultum*, produced from the *P. occultum* gene expressed in *E. coli*, has a Tm of around 80 °C, the natural modified tRNA isolated from the hyperthermophile itself had a melting temperature above 100 °C (Kuchino *et al.*, 1993).

Post transcriptional modifications of tRNA bases of ribose are certainly not primitive features since they are produced by very specialized enzymes (Björk *et al.*, 1987). The necessity of such additional components suggests again that hyperthermophiles appeared after a period of evolution in a mesophilic world. Of course, one might argue that modified ribonucleotides were once produced by the prebiotic chemistry and that the RNA world at high temperature was a world of modified RNA. However, one has to consider that methylation of the ribose, which is especially common at high temperature, precisely inactivates the reactive oxygen which plays the leading role in ribozyme catalysis. Pre-RNA molecules lacking the ribose 2′ OH (review by Joyce, 1989) might have been more stable than RNA itself at high temperature, but they would have also lack the catalytic properties required for a functional RNA world.

Since RNA modifications apparently play a role in the fine tuning of the translational mechanism and in many regulatory pathways (for review, see Persson, 1993), it might be simpler to imagine that tRNA modifications and tRNA modifying enzymes first appeared in the course of the refinement of the translation apparatus and were later recruited by hyperthermophiles for stabilization of the tRNA. Indeed, some of the tRNA modifications present extensively in hyperthermophilic archaea are also present in some mesophilic bacteria and eucaryotes (Edmonds *et al.*, 1991).

DNA is much more thermostable than RNA since it lacks the ribose 2′ OH which is responsible for RNA instability. Furthermore, instead to promote thermodegradation, as in the case of RNA, both Mg^{++} and Zn^{++} strongly protect DNA against thermodegradation (Lindahl, 1993; Marguet and Forterre, 1994 and unpublished data). Finally, the circularization of double-stranded DNA or the formation of topologically closed domains in a linear molecule introduce a strong barrier against denaturation of DNA duplex molecules (Vinograd *et al.*, 1967; Forterre *et al.*, 1992). Such a barrier does not exist in single-stranded RNA and this could explain why genomic DNA in hyperthermophiles are not especially GC rich, whereas the stable RNAs are (Marguet and Forterre, 1994). All these observations suggest that only organisms with double-stranded DNA genomes were able to adapt themselves to live in hyperthermophilic environments. Accordingly, even if life appeared at high temperature in a pre-RNA world, subsequent evolution of the RNA world probably occurred at lower temperature at least up to the invention of DNA.

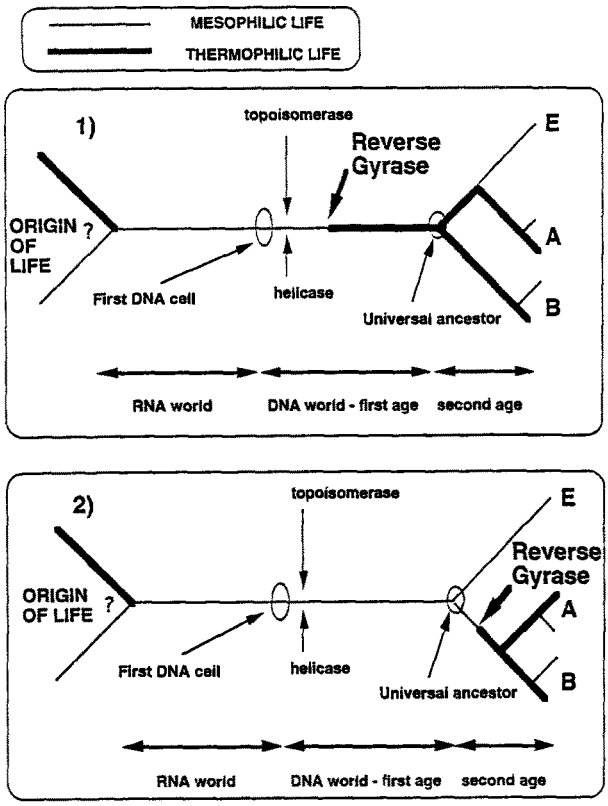

Fig. 2. Two scenarios for the emergence of reverse gyrase via a fusion event between helicase and topoisomerase genes. In the first (1), the fusion has occurred before the separation of the three domains of life, *Eucarya* (E), *Archaea* (A) and *Bacteria* (B). The universal tree has been tentatively rooted in the bacterial branch. In the second scenario (2), the universal tree has been rooted in the eucaryal branch and the fusion event has occurred in a lineage common to the two procaryotic domains. The thermophilic lineages are represented by bold lines. Two options (mesophile and thermophile) are proposed for the origin of life. The first and second ages of the DNA world are defined according to Forterre (1992).

5. Different Hypotheses for the Origin of Reverse Gyrase and Early Cellular Evolution

Figure 2 shows two possible evolutionary scenarios in which the direct link between the origin of life (either hot or cold) and the emergence of hyperthermophiles is eliminated. These models are both based on the 'late' fusion hypothesis for reverse gyrase described above. The first one (No. 1) implies that the last common ancestor of the three domains (the universal ancestor) was nevertheless a hyperthermophile. In that case, reverse gyrase should have appeared during the first age of the DNA world. However, since thermophiles cannot compete with mesophiles, it is not clear in such a scenario why primitive mesophilic cells which do not acquired reverse

gyrase completely disappeared, such that all present day organisms evolved from reverse gyrase containing cells? One possibility, suggested by Gogarten (1993), is that all primitive mesophiles disappeared after a giant meteoritic boiling impact which could have sterilized the Earth surface and the upper part of the Ocean (Maher and Stevenson, 1988), destroying all life forms except those thriving in abyssal hydrothermal systems. A weakness in this hypothesis is that today abyssal hydrothermal vents contain large number of mesophilic microorganisms which lives in the lower temperature part of the vents (for a review, see Prieur, 1992). Accordingly, if mesophiles already existed at the time of the putative giant impact, one has to make the additional assumption that this impact was sufficiently strong to even boil the bottom of the oceans. Alternatively, one would have to suppose that despite the existence of mesophiles elsewhere at the time of the impact, none were present near abyssal hydrothermal systems.

We presently favor the alternative scenario (No. 2) which is presented in Figure 2. In that case, the universal ancestor is seen as a mesophile, and reverse gyrase appears in an evolutionary lineage connecting the universal ancestor to the last common (hyperthermophilic) ancestor of the two procaryotic domains. In such a scenario, one has to explain why the common ancestor of procaryotes became a hyperthermophile. A possibility might be that the procaryotic phenotype itself emerged during the process of adaptation of this lineage to thermophily (Forterre, 1992). This hypothesis is based on the observation that the major problem faced by living organisms at very high temperatures is the heat-induced hydrolysis of macromolecules. The procaryotic phenotype thus appears well designed for hyperthermophily because of the high macromolecular turnover in procaryotic cells and the coupling between transcription and translation which avoids the crucial problem of mRNA degradation at very high temperature. One can speculate that the ancestors of procaryotes were subjected to a thermoreduction process which resulted in reduction in their genome size and speeding up of their metabolic rate. In that scenario, the gene fusion event leading to the reverse gyrase gene would have been a crucial step in the formation of organisms that resemble today's procaryotes.

It should also be noted that this hypothesis, linking the procaryotic phenotype to thermophily, would explain why only procaryotes thrive in the wide temperature range, 62 to 110 °C (Brock, 1978). Thermophilic or hyperthermophilic eucaryotes might have been overlooked because of specific cultivation problems, but an alternative hypothesis is that eucaryotes exhibit specific features incompatible with life at very high temperature. Since RNA is the most fragile macromolecule at high temperatures, one such feature might be the presence of more relics of the RNA world in eucaryotes than in procaryotes. The mRNA might be also unstable at high temperatures in eucaryotes because of the uncoupling of transcription and translation (Forterre, 1992).

The thermoreduction hypothesis for the origin of procaryotes implies that modern eucaryotes migh have preserved some features of the universal ancestor which have been lost in procaryotes. This idea is usually dismissed on the ground that

procaryotic microfossils appeared before eucaryotic ones. This is not a strong argument since primitive eucaryotes lack the cell wall structures that are required for fossilization (Sogin, 1991; Knoll, 1992). In fact, the most ancient microfossils of procaryotes (3.5 Giga years ago) (Schopf, 1992) and eucaryotes (2.1 Giga years ago) correspond to very evolved organisms, cyanobacteria and Grypania (green algae) (Han and Runnega, 1992), respectively, and cannot be representative of the first life forms on the Earth, at least if one consider their positions in the phylogeny deduced from rRNA analysis (Sogin, 1991; Knoll, 1992).

Although the scenario No. 2 and the thermoreduction hypothesis for the origin of procaryotes are attractive to us, one should remember that they are based on the current hypothesis that all procaryotes originated from an hyperthermophilic ancestor. The case for an hyperthermophilic ancestor of all *Archaea* is quite strong if one consider that: (i) hyperthermophilic lineages are slowly evolving in the archaeal 16S rRNA tree (Woese, 1987), (ii) this tree is rooted among different hyperthermophilic species (Achenbach-Richter *et al.*, 1988), (iii) reverse gyrase is found in the two archaeal kingdom (the *Crenarchaeota* and the *Euryarchaeota*, see Woese *et al.*, 1990), (iv) archaeal lipids, even from mesophilic species, are especially suitable for life at high temperature. In contrast, the case for an hyperthermophilic bacterial ancestor is not so strong. The 16S rRNA from bacterial hyperthermophiles are also slowly evolving and the root of the bacterial 16S rRNA tree has been located between the hyperthermophiles *Thermotoga* and *Aquifex* (Achenbach-Richter *et al.*, 1987; Burggraf *et al.*, 1992) but mesophilic species do not exhibit obvious moleular relics of their putative hyperthermophilic ancestor. Furthermore, it is not clear if the bacterial 16S rRNA tree can be safely extrapolated to the bacterial tree itself, since *Thermotoga* and *Aquifex* are not the most primitive lineages in the tree constructed from bacterial RNA polymerases sequence (Klenk *et al.*, 1994). The difficulty of extrapolating gene trees into species trees for very distantly related organisms is well illustrated by the existence of differences in the topology of the archaeal 16S and 23S rRNA trees (Philippe, 1993; Garrett *et al.*, 1994).

6. Conclusion

It is very important to determine the temperature range in which life appeared on the Earth. This is a prerequisite for the design of pertinent experiments in prebiotic chemistry. Unfortunately, this question is a difficult one, being characterized by a lot of uncertainties (Figure 3). In fact, we do not know what the temperature of our planet was at the time of the origin of life, the nature of the primordial biotope, or even the date, between 4.2 and 3.5 Giga years ago, when life appeared. The rooting of the universal tree is another challenge, which might even prove to be an impossible task if, as suggested by Meyer *et al.*, 1986, protein sequence comparisons cannot be used with confidence to identify phylogenetic relationships between very distantly related organisms. Taking into account all these difficulties, the study of mechanisms specifically designed for life at high temperature in hyperthermophilic

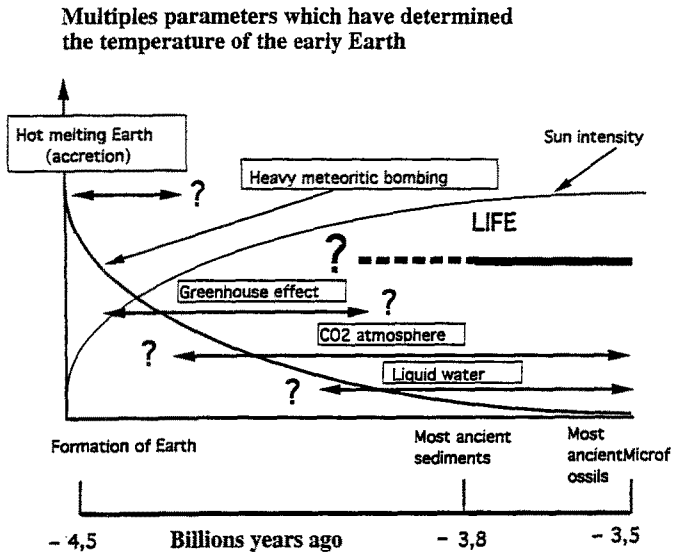

Fig. 3. Too many uncertainties prevent determination of the actual temperature of the early Earth. The rate of cooling of the Earth after its formation via accretion was dependent on the sun intensity, the decay in the heavy meteoritic bombing and the nature of the primitive atmosphere. At best, only estimates exist for these parameters (for reviews, see Lazcano *et al.*, 1992; Kasting, 1993). The actual time of the appearance of liquid water on Earth and the subsequent appearance of life is also unknown.

organisms can provide important new results and insights concerning the problems of the hot origin of life and the nature of the universal ancestor.

Here, we have not discussed the hypothesis of the hot origin of life *per se* (see Figure 2), but we have argued against a direct link between a hot origin of life and present day hyperthermophilic procaryotes. We have not discussed the problem of archaeal lipids, which could exhibit essential features for hyperthermophilic life, but further studies on their origin and relationships with bacterial and eucaryal lipids should be very informative. The study of tRNA modification enzymes is another avenue worth further exploration.

In the case of reverse gyrase, much remains to be done to clarify the role of this enzyme and its origin. A priority is to determine at the sequence level if the archaeal and the bacterial enzymes are indeed homologous and to localize precisely the position of reverse gyrase in the general phylogeny of DNA topoisomerases. This would require the sequencing of new reverse gyrase genes, as well as genes from related DNA topoisomerases, such as the ATP independent relaxing enzymes described both in moderately thermophilic archaea and in hyperthermophilic archaea and bacteria (Bouthier de la Tour *et al.*, 1990, 1991, 1993). One can be hopeful that a better knowledge of reverse gyrase and other key enzymes from hyperthermophiles will, in the end, give us clear-cut information about the history of these fascinating microorganisms.

Acknowledgements

We thank again our coworkers who have participated in the reverse gyrase saga: G. Mirambeau, M. Nadal and C. Jaxel. P.F. is especially grateful to A. Brack and A. Lazcano for promoting his participation to the 7th international ISSOL meeting. The work on DNA topoisomerases in our laboratories was supported by grants from the Association de la Recherche sur le Cancer (ARC). F. Confalonieri and F. Charbonnier were supported by fellowships from the ARC.

References

Achenbach-Richter, L., Gupta, R., Stetter, K. O., and Woese, C. R.: 1987, *Syst. Appl. Microbiol.* **9**, 34.

Achenbach-Richter, L., Gupta, R., Zillig, W., and Woese, C. R.: 1988, *Syst. Appl. Microbiol.* **10**, 231.

Andera, L., Mikulik, K., and Savelyeva, N. D.: 1993, *FEMS Microbial. Letters* **110**, 107.

Björk, G. R., Ericsson, J. U., Gustafsson, C. E. D., Hagervall, T. G., Jönsson, Y. H., and Wilkström, P. M.: 1987, *Ann. Rev. Biochem.* **56**, 263.

Bouthier de la Tour, C., Portemer, C., Nadal, M., Stetter, K. O., Forterre, P., and Duguet, M.: 1990, *J. Bacteriol.* **172**, 6803.

Bouthier de la Tour, C., Portemer, C., Huber, R., Forterre, P., and Duguet, M.: 1991, *J. Bacteriol.* **173**, 3921.

Bouthier de la Tour, C., Portemer, C., Forterre, P., Huber, R., and Duguet, M.: 1993, *Biochem. Biophys. Acta* **1216**, 213.

Brock, T. D.: 1978, *Thermophilic Microorganisms and Life at High Temperature*, Springer-Verlag, New-York, Heidelberg.

Burggraf, S., Olsen, G. J., Stetter, K. O., and Woese, C. R.: 1992, *Syst. Appl. Microbiol.* **15**, 352.

Butzow, J. J. and Eichorn, G. L.: 1975, *Nature* **254**, 358.

Charbonnier, F. and Forterre, P.: 1994, *J. Bacteriol.* **176**, 1251.

Charbonnier, F., Erauso, G., Barbeyron, T., Prieur, D., and Forterre, P.: 1992, *J. Bacteriol.* **174**, 6103.

Collin, R. G.,. Morgan, H. W., Musgrave, D. R., and Daniel, R. M.: 1988, *FEMS Micro. Letters* **55**, 235.

Confalonieri, F., Elie, C., Nadal, M., Bouthier de la Tour, C., Forterre, P., and Duguet, M.: 1993, *Proc. Natl. Acad. Sci. USA* **90**, 4753.

Di Gate, R. J. and Marians, K. J.: 1989, *J. Biol. Chem.* **264**, 17924–17930.

Duguet, M.: 1993, *Nucl. Acids Res.* **21**, 463–468.

Edmonds, C. G., Crain, P. F., Gupta, R., Hashizume, T., Hocart, C. H., Kowalak, J. A., Pomerantz, S. C., Stetter, K. O., and McCloskey, J. A.: 1991, *J. Bacteriol.* **173**, 3138.

Forterre, P.: 1992, in *'Frontiers of Life'*. Editions Frontières, Gif sur Yvette, France p. 221.

Forterre, P. and Elie, C.: 1993, in 'The Biochemistry of *Archaea*', *New Comprehensive Biochemistry* (Elsevier) **26**, 325.

Forterre, P., Mirambeau, G., Jaxel, C., Nadal, M., and Duguet, M.: 1985, *EMBO J.* **4**, 2123.

Forterre, P., Charbonnier, F., Marguet, E., Harper, F., and Hecnkes, G.: 1992, *Biochem. Soc. Symp.* **58**, 99.

Forterre, P., Benachenhou, N., Confalonieri, F., Duguet, M., Elie, C., and Labedan, B.: 1993, *BioSystems* **28**, 15.

Forterre, P., Bergerat, A., Gadelle, D., Elie, C., Confalonieri, F., Duguet, Holmes, M., and Dyall-Smith, m.: 1994a, *Syst. Appl. Microbiol.* **16**, 746.

Forterre, P., Marguet, E., Charbonnier, F., Mojica, F., and Bergerat, A.: 1994b, Abstract of the US/NZ Workshop on extreme thermophiles, December 8–11, Hamilton, New-Zealand.

Garrett, R., Aagaard, C., Andersen, M., Dalgaard, J., Lykke-Andersen, J., Phan, H. T. N., Trevisanato, S., Ostergaard, L., Larsen, N., and Leffers, H.: 1994, *Syst. Appl. Microbiol.* **16**, 680.

Ginoza, W., Carol, J. H., Vessey, K. B. and Camark, C.: 1964, *Nature* **203**, 606.

Gogarten, J. P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E. J., Bowman, B. J., Manolson, M. F., Poole, R. J., Date, T., Oshima, T., Konishi, J., Denda, K., and Yoshida, M.: 1989, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6661.
Gogarten, J. P.: 1993, in *Abstracts of the 7th Issol Meeting*, held in Barcelona (Spain) 4–9 July, p. 83.
Gorbalenya, A. E. and Koonin, E. V.: 1993, *Curr. Opini. in Struct. Biol.* **3**, 419.
Han, T.-M. and Runnega, B.: 1992, *Science* **257**, 232.
Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., amd Miyata, T.: 1989, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9355.
Joyce, G. F.: 1989, *Nature* **338**, 217.
Kasting, J. F.: 1993, *Science* **259**, 920.
Kikuchi, A. and Asai, K.: 1984, *Nature* **309**, 677.
Kikuchi, A., Shibata, T., and Nakasu, S.: 1986, *Syst. Appl. Microbiol.* **7**, 72.
Klenk, H. P., Palm, P., and Zillig, W.: 1994, *Syst. Appl. Microbiol.* **16**, 638–647.
Knoll, A. H.: 1992, *Science* **256**, 622.
Kuchino, Y., Ushida, C., Mizushima, H., Ueda, T., Watanabe, K., Stetter, K. O., Crain, P. F., and McCloskey, J. A.: 1993, in Abstract of the *International Workshop on Molecular Biology and Biotechnology of Extremophiles and Archaeabacteria*, held from 1–6 August 1993, Wako, Japan, p. 58.
Lazcano, A., Fox, G. E., and Oro, J. F.: 1992 in *The Evolution of Metabolism*, CRC press, Boca Raton, p. 237.
Lindalh, T.: 1967, *J. Biol. Chem.* **242**, 1970.
Liu, L. and Wang, J. C.: 1987, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7024.
Maher, K. A. and Stevenson, D. J.: 1988, *Nature* **331**, 612.
Marguet, E. and Forterre, P.: 1994, *Nucl. Acids Res.* **22**, 1681.
Meyer, T. E., Cusanovich, M. A., and Kamen, M. D.: 1986, *Proc. Natl. Aced. Sci. U.S.A.* **83**, 217.
Miller, S. L. and Bada, J. L.: 1988, *Nature* **334**, 609.
Nadal, M., Mirambeau, G., Forterre, P., Reiter, W-D. and Duguet, M.: 1986, *Nature* **321**, 256.
Nadal, M., Jaxel, C., Portemer, C., Forterre, P., Mirambeau, G., and Duguet, M.: 1988, *Biochemistry* **27**, 9102.
Nakasu, S. and Kikuchi, A.: 1985, *Embo J.* **4**, 2705.
Pace, N. R.: 1991, *Cell* **65**, 531.
Persson, B. C.: 1993, *Mol. Microbiol.* **8**, 1011.
Philippe, H.: 1993, *Nucl. Acids Res.* **21**, 5264.
Prieur, D.: 1992, in *Molecular Biology and Biotechnology of Extremophiles*, Blakie, Glasgow and London, p. 163.
Schopf, J. W.: 1992, in *The Proterozoic Biosphere, A Multidisciplinary Study*, J. W. Schopf and C. Klein (eds.), Cambridge University Press: New York, p. 583.
Schulz, G. E.: 1992, *Curr. Opin. Struct. Biol.* **2**, 61.
Slezarev, A. I.: 1988, *Eur. J. Biochem.* **173**, 395.
Sogin, M. L.: 1991, *Curr. Opin. Genet. and Dev.* **1**, 457.
Stetter, K. O.: 1992, in *'Frontiers of Life'*, Editions Frontières, Gif sur Yvette, France, p. 195.
Tse-Dinh, Y-C. and Wang, J. C.: 1986, *J. Mol. Biol.* **191**, 321.
Vinograd, J., Lebowitz, J., and Watson, R.: 1968, *J. Mol. Biol.* **33**, 173.
Wächtershäuser, G.: 1988, *Syst. Appl. Microbiol.* **10**, 207.
Walker, J. E., Saraste, M., Runswick, M. J., and Gay, J.: 1982, *EMBO J.* **8**, 945.
Wällis, J. W., Cherbet, G., Brodsky, G., Rolfe, M., and Rothstein, R.: 1989, *Cell* **58**, 409.
Woese, C. R.: 1987, *Microbiol. Rev.* **51**, 221.
Woese, C. R., Kandler, O., and Wheelis, M. L.: 1990, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576.
Yang, L., Jessee, C. B., Lau, K., Zhang, H., and Liu, L.: 1989, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6121.