UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL ATOMIC ENERGY AGENCY
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/916 -6

SEVENTH COLLEGE ON BIOPHYSICS:
Structure and Function of Biopolymers: Experimental and Theoretical Techniques.
4 - 29 March 1996

DNA Folding in Eukaryogenesis

Julian CHELA-FLORES
International Centre for Theoretical Physics
Miramare
P.O.Box 586
34100 Trieste
ITALY
&
Instituto Internacional de Estudios Avanzados
Universidad Simon Bolivar
Apartado 17606
Parque Central
Caracas 1015A
VENEZUELA

DNA FOLDING IN EUKARYOGENESIS (*)

Julian Chela-Flores (+)
International Centre for Theoretical Physics,
Trieste, Italy and
Instituto Internacional de Estudios Avanzados,
Caracas 1015A, Venezuela

Abstract. We examine properties that have contributed to raise the ancestral prokaryotic programmes to a level where we may appreciate a clear departure in eukaryotes from earlier themes in the evolution of the cell from the last common ancestor. We review the relevant aspects of eukaryogenesis within the wider scope of the origin and evolution of life. We shift our point of view from evolution of cell morphology to the point of view of the genes. In particular, we focus attention on possible physical bases for the way transmission of information has evolved in eukaryotes, namely, the inactivation of whole chromosomes. We review the main processes that take part in molecular genetics: transcription, DNA replication, and DNA packaging so as to be able to face tentative steps towards the as yet unknown principles underlying DNA bending and folding, which are the first stages of heterochromatization, the proposed hallmark of the eukaryotic cell. We discuss the possible relevance of deeper insights into eukaryogenesis for the next round of space missions that will search for life in both the inner as well as the outer Solar System.

1. Introduction

1.1. THE GROWTH OF EXOBIOLOGY

In the forthcoming planetary missions, particularly those in which landers will search for signs of life, it is of primary importance to decide which aspects of life to search for. Inevitably, our first seps in this direction will necessarily be modelled on our experience with the earth's biota.

The problem of the origin of the first nucleated cell (eukaryogenesis) occupies a central position in a wide range of problems which concern the origin and evolution of life ranging from chemical evolution to "exobiology" (i.e., the search for extraterrestrial intelligence, SETI, Drake and Sobel, 1992; Drake, 1996). SETI is relevant in the context of eukaryogenesis as the only form of intelligent life that we are familiar with is based on multicellular organisms of nucleated, or 'eukaryotic' cells (Domain Eukarya).

Hence, it is important to single out the factors that may have led to the transition from bacterial life to the nucleated cells and, ultimately, to multicellular organisms.

1.2. THE EVOLUTION OF ATMOSPHERIC OXYGEN IS CORRELATED WITH THE EVOLUTION OF THE LIVING CELL

Several lines of research suggest the absence of current values of O_2 for a major part of the Earth history. There are some arguments nevertheless that militate in favour of Archean atmospheres with values of the partial pressure of atmospheric oxygen O_2 (pO_2) about 10^{-12} of present atmospheric level (PAL).

(*) Lecture prepared for the Seventh College on Biophysics:*Structure and Function of Biopolymers: Experimental and Theoretical Techniques*. 4-26 March, 1996.

(+) Research Associate, *Dublin Institute for Advanced Studies*, 10, Burlington Road, Dublin 4, Ireland.

Some rocks, called 'igneous', have originated by solidification from a molten condition as they were poured out from volcanos. One example is provided by ***shale***, a rock that has played a role in our understanding of biological evolution. Shale is mainly clay that has hardened into rock. The onset of atmospheric oxygen is demonstrated by the presence in the geologic record of red shale coloured by ferric oxide. Such 'red beds' are estimated to be 2,000 million years old (corresponding to the Orosirian Period of the Paleoproterozoic), at a time when the oxygen levels may have reached 1-2% PAL, sufficient for the development of a moderate ozone protection for the microorganisms of the Proterozoic from ultraviolet (UV) radiation. In fact, UV radiation is able to split the O_2 molecule into the unstable O-atom, which in turn reacts with O_2 to produce ozone O_3, an efficient filter for the UV radiation.

The paleontological record suggests that the origin of the nucleated or eukaryotic cell (eukaryogenesis) occurred earlier than 1,500 million years before the present (My bp). Some algae may even date from 2,100 My bp (Han &Runnegar, 1992), a period comparable to the first onset of red beds. This is still rather late, compared to the earliest available prokaryotic fossils of some 3,500 My bp in (Schopf, 1993). However, if we keep in mind certain affinities between eukaryotes and archaebacteria (such as homologous factors in protein synthesis,'elongation factors'), we may argue that archaebacteria and the stem group of eukaryotes may have diverged at about the same time (Runnegar, 1994); this conjecture, combined with the lightest carbon isotope ratios from organic matter [$\delta^{13}C$ < -40 $^{o}/oo$ PDB], may imply that bacteria capable of oxidizing methane CH_4 ('methylotrops') may have been using methane produced by archebacteria that were able to produce it as a byproduct orf their metabolism. From the age of such fossils of organic matter a tentative date of 2,700 Mybp or earlier may be asigned to eukaryogenesis.

1.3. THE CONSTRAINTS IMPLIED BY THE BANDED IRON FORMATIONS

There are Archean rock formations (which may be found up to 2,000 My bp) that are significant in the evolution of life. These are laminated compounds of dioxide of silicon (silica) and iron. In reference to their laminated structure they are referred to as "banded iron formations" (BIFs). The period in which the BIFs were laid out ended some 1,800 My bp (Riding, 1992).

In the anoxic atmosphere of the Archean iron compounds could have been dispersed over the continental crust. They could have absorbed some oxygen thereby protecting photosynthesizers that could not tolerate oxygen. Such microorganisms in turn produced oxygen that combined with their environment to produce iron oxide (for example, hematite Fe_2O_3), which makes up the BIFs.

In strata dating prior to 2,300 My bp it has been observed that there is an abundance of the easily oxidized mineral form of uranium (IV) oxide (urininite, for example the well-known variety *pitchblende).* This supports the conclusion that we had to wait until about 2,000 My bp for a substantial presence of free O_2.

We may not exclude from the geochemical data earlier dates, in the lower Archean, for the first prokaryotic microflora (Schidlowski, 1995), although some critical considerations from the point of view of geochronology still do not rule out the possible origin of life immediately after the end of the Hadean subera (Moorbath, 1995). Once the eukaryotes enter the fossil record, its organization into multicellular organisms followed in a relatively short period (in a geological time scale).

Metazoans are hypothesized to have arisen as part of a major eukaryotic radiation in the Riphean Period, approximately 800-1,000 My bp (Knoll, 1994) when the level of atmospheric O_2 had reached 4-8% PAL. There is some evidence in the

Neoproterozoic, in Vendian time, for the existence of early diploblastic grades (Ediacaran faunas). These organisms were early metazoans with two germ layers, such as the modern coelenterates (jellyfishes, corals and sea anemones). Later on, when the level of atmospheric O_2 had reached values in excess of 10% PAL, these grades were overtaken in numbers by triploblastic phyla as the level of atmospheric O_2 had reached 40% PAL (Cambrian faunas, which were mainly metazoans with three germ layers) constituting at present the greater majority of multicellular animals.

We may obtain further insights from paleontology: acceleration in the evolutionary tempo is observed after the onset of eukaryogenesis, as it is clearly demonstrated by the microfossils of algae from the Neoproterozoic (Knoll, 1994) and by the macrofossils of the early Phanerozoic (Cambrian Period) (Conway-Morris, 1993). Such evolutionary changes within the first billion years of atmospheric oxygen rose the simple prokaryotic cell to eukaryotes, metazoans and metaphytes.

In this lecture we will maintain that it is possible that the accelerating evolutionary tempo may have had a counterpart in corresponding changes in the eukaryotic genome.

We propose a candidate for such a counterpart, namely, chromosome plasticity. Together with evolution in cellular morphology, there is corresponding evolution in structure, organization, and genetic regulation of the DNA in the nucleoid of prokaryotes. The simplest chromosomes, and possibly the earliest (Chela-Flores, 1994b), may be those of viroids and plasmids. Complexity, understood as increments in gene-expressing nucleic acid, increases from the RNA viroid level to the DNA prokaryotic chromosome (PC), found in the more evolved archaebacteria, eubacteria, chloroplasts, and mitochondria.

However, maximum complexity is only reached with the first appearance of the eukaryotic chromosome (EC). The consideration of the evolution from the PC to the EC is forced upon us when we look closer at properties of the contemporary genome of the living cell.

1.4. ARE THERE TRANSITIONAL MICROORGANISMS ?

We have considered recently the main problems concerning eukaryogenesis, a most significant event in the diversification of life (Chela-Flores, 1995b): Eukaryotes have their DNA linked in chromatin; the main organelles are normally in its cytoplasm. However, protozoans may provide examples of mitochondrion-less eukaryotes, there is even a whole phylum of amitochondrial protozoa, the microsporidia (Cavalier-Smith, 1987). Besides, in two out of three kingdoms of the Eukarya Domain (i.e., Animalia and Fungi) chloroplasts are absent. The origin of these two types of organelles in the eukaryotes is to be found, according to the Serial Endosymbiosis Hypothesis (Margulis, 1993), in separately evolved organisms.

Thus, at least the origin of the red algal chloroplast may be traced back to cyanobacteria. (This may be illustrated with the analysis of 16S ribosomal RNA of the unicellular marine red alga *Porphyridium*, (Bonen & Doolittle, 1976).) On the other hand, mitochondria may be linked with purple bacteria resemblig *Paracoccus denitrificans*. This prokaryote is suggested to be a plausible ancestor, because when all various biochemical parameters are taken into account, *P. denitrificans* resembles a mitochondrion much more closely than other aerobic bacteria (John & Whatley, 1975).

In fact, symbiosis implies that the ancestral prokaryote may have been taken up by a chloroplast-free amoeboid protoeukaryote; eventually the symbiont may have lost autonomy, possibly by horizontal gene transfer between the protomitochondrion and the host's nucleus. These events, which are different from natural selection, were factors that would have led to the evolution of a single-cell organism that had already integrated the metabolism of the partners in symbiosis.

Further support for symbiosis may be brought to our attention as both eukaryotic organelles we have just mentioned, have their own mechanisms of translation: we have seen that separate genetic codes are known for each organelle. On the other hand, the origin of the nucleus does not seem to be explained by symbiosis. This is the thesis of 'direct filiation' (i.e., differentiation), which sets some limitation on the extent to which symbiosis may have shaped the first eukaryotic cell, leaving the question of the nature of the earliest eukaryote as an open problem. Primitive eukaryotic organisms have been studied in detail. One such taxon is the family Cyanidiophyceae. These organisms are rhodophytes, commonly known as red algae (Seckbach, 1994). It has been argued that these acido-thermophilic algae may constitute a bridge between cyanobacteria and red algae (Seckbach, 1995).

In particular *Cyanidium caldarium* has a primitive eukaryotic cellular structure, which together with its typical biochemistry make it a possible candidate for a transitional cell between cyanobacteria and the earliest eukaryotic algae. Other remarkable properties of *C. caldarium* is that it may live at temperatures of up to 57 ^{o}C and shows better rates of growth and photosynthesis when cultured in an 'atmosphere' of pure CO_2 (Seckbach, 1972).

It is useful to know in which artificial atmospheres extreme organisms may survive. as we now know of several Solar-System planets and satellites that have atmospheres, not all similar to our own. For example, four large planetary satellites in the outer Solar System are known to have atmospheres (Hall *et al.*, 1995): Europa (cf., last paragraph in this section) and Io (Jupiter), Titan (Saturn) and Triton (Nepune). Both planets, Venus and Mars have CO_2 atmospheres. Although the surface temperature of Venus is much higher than what is suitable for known thermophilic organisms cooler underground temperatures remains to be discussed, as it should from what is known from the deep Earth biosphere (Gold, 1992).

The possibility of extending the biosphere deep into the silicate crust in at least another terrestrial planet (Mars) has some further implications. The values of the daily surface temperature of Mars lie in the range $190K < T < 240K$ (Kieffer *et al.*, 1976). The upper bound of the temperature range is not incompatible with cryophilic organisms that are known in the envronment. The present status of the search for life on Mars is reviewed in Table 1.

TABLE 1: The question of life on Mars (Based on Soffen, 1976).

Experiment	*Results*
Test for signs of photosynthesis or chemosynthesis induced by samples from the soil	Small incorporation of CO/CO_2 into organics
Measurement of any gaseous products from a soil sample	Initial rapid release of O_2; slow release of CO_2 , N_2
Search for the release of radioactive gas when the soil sample was exposed to a radioactive organic nutrient solution	Initial rapid release of labelled gas, followed up by slow release

However, UV radiation prevents this possibility, at least on the planetary surface devoid of any UV defense mechanism; but one possibility remains in life underground. This question seems pertinent to exobiology as we cannot exclude at present that the organisms that have been found to inhabit deep in the silicate crust of the Earth may have been deposited with the original sediment and survived over geologic time (Parkes and Maxwell, 1993).

Finally, another interesting possibility for future research concerns the satellite Europa, which has a uniform frozen surface, except for long crevices (Oro, 1996). Besides having oxygen in its atmosphere, as we mentioned above), calculations suggest that this satellite has an 80 km-deep ocean, under a 10 km-deep layer of ice. It is possible that organisms similar to our own archaebacteria may inhabit at liquid water temperatures that are estimated to be some 4 ^{o}C. Deep in our own oceans there are hot springs where such prokaryotes may live. We have seen in Sec. 1.2 that the divergence of archebacteria and the main stem of eukaryotes occurred at about the same time on Earth (Proterozoic eon).

1.5. A GENE-CENTERED APPROACH

The PC is a double-stranded DNA structure usually lacking:
1.4.1. Abundant packaging proteins (histones).
1.4.2. An enveloping membrane.
1.4.3. Different specialized regions, such as the nuclear organelle, associated with the site of ribosomal RNA-coding genes (nucleoli).
1.4.4. Ends formed by highly repeated sequences (telomeres).
1.4.5. Shut-down inhibition of gene expression (gene silencing). This may involve whole chromosomes, leaving some exceptional loci with the ability to transcribe pre-messenger RNAs (pre-mRNAs). PCs consist of a beaded structure, not unlike that of the EC (Griffith, 1976). Even histone-like proteins are known in some prokaryotes: *Escherichia coli* (Rouvier-Yaniv & Gros, 1975), Cyanobacteria (Haselkorn & Rouvier-Yaniv, 1976), the short rod-shaped human pathogen *Pseudomonas aeruginosa* (Kato *et al.*, 1990), and the obligate sexually-transmitted intracellular human parasite *Chlamydia trachomatics* (Hackstadt *et al.*, 1991)

We cannot argue in favour of a clear-cut difference between PCs and ECs from the point of view of genome size. Indeed, although PCs are normally smaller than ECs, some eukaryotes have very small chromosomes. One example is provided by the Rhodophyte *Cyanidioschyzon*. This seaweed has a genome of only 8 million (M) base pairs (bp) (Seckbach, 1995). This tiny genome is only about twice as long as the corresponding one in *E. coli* (3.5 Mbp).

On the other hand, the problem of eukaryogenesis is rendered still more difficult to define, as the EC has some characteristics which are not common to all eukaryotes. Some exceptions are particularly remarkable in lower eukaryotes, such as algal protists: in these cases we are faced with chromosomes lacking histones. For instance, in the dinoflagellates *Blastodinium* Chatton and *Amphidinium elegans* (Soyer, 1971) there are distinct chromosomes which are not associated with histones. *Prorocentrum micans* is a neurotoxin-producing marine dinoflagellate that occasionally may cause local outbreaks of extremely devastating red water; its chromosomes have no 'beaded' structure, which normally are due to sets of histones being complexed with DNA (such structures are called 'nucleosomes') (Herzog& Soyer, 1981).

Furthermore, the absence of histones is conspicuous in other eukaryotes, such as in three genera of fungi *Microsporum, Neurospora and Phycomyces* (Leighton *et al.*, 1971). For these reasons at least one exceptional group of eukaryotic chromosomes may be considered primitive (Maynard-Smith & Szathmary, 1995). Consequently, the origin and evolution of chromosomes becomes a relevant investigation in origin-of-life studies. The most likely cause for the evolution of

complex chromosome structure seems to be regulation of gene expression, a process which has reached its maximum expression in eukaryotes (Maynard-Smith, 1993).

1.6. CHROMATIN STRUCTURE

For a considerable time now, it has been evident that the integration of proteins complexed with DNA ('chromatin') has played a fundamental role in the regulation of gene expression (Littau *et al.*, 1965). Some chromatin replicates its DNA late in the S phase of the cell cycle (cf., Sec. 2.2); it is also dark-staining, due to the high degree of its DNA packaging. In order to differentiate such a special state of chromatin from its less dense counterpart ('euchromatin'), we refer to chromatin in the highly packed case as 'heterochromatin'.

However, it is convenient to introduce the concept of a dense form of chromatin which could be due to its specific DNA sequence. One such instance of chromatin contains highly repetitive DNA, which is associated with heterochromatization. This point will be considered below, in Sec. 3.2, in our discussion of 'satellite DNA'. A closely related state of chromatin, which is the result of regulation rather than structure, is sometimes found in a higher state of DNA packaging. Such chromatin is referred to as 'facultative heterochromatin'.

We reserve the term 'constitutive heterochromatin' to chromatin that finds itself in a dense state of packaging due to its permanent structure. In the fruitfly *Drosophila melanogaster*, a specific non-histone protein (HP-1) is known to influence directly chromatin structure (Eissenberg *et al.*, 1990); such protein may suppress the inactivation of gene expression in constitutive heterochromatin, demonstrating that such a protein could participate in a typically eukaryotic shut-down mechansim of chromosomes.

2. Evolution of DNA synthesis and gene regulation

2.1. ORIGINS OF EUKARYOTIC DNA REPLICATION AND TRANSCRIPTION

Two of the central pathways of biomolecular synthesis are relevant to our discussion, namely, DNA synthesis and transcription of pre-mRNA. These processes are well established in prokaryotes and were further elaborated by eukaryotes, generally increasing their complexity. In some cases some radical departures were initiated which, from the point of view of the genome, may be considered as true hallmarks of eukaryogenesis.

Some thirty years ago the 'replicon' model was introduced in an effort to understand bacterial DNA replication in terms of units of replication, the so-called 'replicons' (Jacob, 1993). The main themes of this model are:

2.1.1. A structural gene controls the synthesis of a specific protein, or 'initiator', which is involved in the initiation of DNA replication and,

2.1.2. A single origin of replication (i.e., the single target sequence recognised by the initiator) allows the starting of replication. In *E. coli*, for instance, the corresponding sequence 'ori-C' has 245 bp (Kornberg, 1988).

More complex eukaryotic DNA replication follows the guidelines identified in prokaryotes, but differs in some essential aspects:

2.1.3. A considerably richer repertoire of enzymes is needed for the generally larger eukaryotic genome (De Pamphilis, 1988).

2.1.4. Multiple origins of replication are spaced at an average of 50-100 thousand base pairs (kbp) (Hand, 1978; Liskens & Huberman, 1990).

2.1.5. Origins are activated at different times in the S phase of the cell cycle, but adjacent origins are activated at about the same time (Brewer & Fangman, 1993; Fangman & Brewer, 1992). Once again, in eukaryotes we find that the repertoire of enzymes required for RNA synthesis exceeds by far the set needed in bacterial transcription. In the process of the eukaryotic elaboration of earlier themes, nevertheless, the coupling between transcription and DNA replication is strictly preserved (Chela-Flores, 1992b). The main point we wish to emphasize here is that sets of adjacent genes transcribed collectively into a single pre-mRNA ('operons') may contain genes required for the initiation of transcription, as well as genes that may play a role in DNA replication. We may find the opposite situation, transcriptional factors may be components of eukaryotic origins of replication (Kornberg, 1988).

2.2. RELEVANCE OF DNA FOLDING IN AN EVOLUTIONARY CONTEXT

In this lecture we discuss the coupling of the processes of transcription, DNA replication, and DNA packaging, with the intention of approaching the question of DNA folding, the physical process underlying heterochromatization.. In the second lecture we will make a preliminary attempt to discuss the important topic of the effect of sequence-dependent DNA bendability on the early stages of the hierarchy of DNA folding, from the nucleosome filament to the metaphase chromosome.

From the point of view of Darwin's theory of evolution this problem is of some interest in the study of the origin of life, more precisely in the context of the origin of the first cell. The origin of the highest unnucleated taxa (the domains Eubacteria and Archaea) is very ancient; in fact, the earliest fossils are known to be from at least the Archean.

On the other hand, the taxon of the truly nucleated cells (Domain Eukarya), according to the fossil evidence may have preceded the evolutionary radiation of the early Cambrian, at the onset of the Phanerozoic, by over a billion years. Indeed, eukaryogenesis was perhaps the most significant event in the origin of biodiversity, which is today demonstrated by over 30 phyla of metazoans, encompassing some 30 million species.

We may raise the question of the physical basis of DNA folding, in an effort to explain early events in molecular genetics in terms of processes known to be going on in the contemporary eukaryotic genomes; such processes, if preserved by evolution, may represent aspects of the molecular mechanisms that led from the progenote to the first eukaryote. The biology of hot spring algae and other cells which have minimal eukaryotic characteristics are natural contemporary taxa in which to test these ideas.

For the purpose of this introductory section we begin by discussing the principal molecules we shall deal with in the remainder of these lectures. First of all we restrict our attention to the problem of polymerase dynamics. The progress of the enzyme complex- DNA polymerase- enhances the rate of RNA synthesis in DNA replication. This is an instance of a very general phenomenon which we shall now discuss in which the hypothesis of condensation has been used. We have discussed in the past (Chela-Flores, 1992b) the analogy between:

2.2.1. Chromatin structural changes, that may occur during gene expression (transcription and DNA replication) and

2.2.2. Certain physical phenomena-phase transitions- that normally occur in other forms of condensed matter.

This analogy may be presented in terms of the various polymerases involved (DNA as well as RNA). However, it is useful to highlight three aspects of polymerase dynamics:

2.2.3. RNA synthesis is required for the initiation of some replication origins. In fact, the concept of origin of replication is needed as in eukaryotes-our main concern in this lecture - every chromosome is composed of many origins *(oris)* that are activated at differnt times in the S phase of the cell cycle. The region itself that is served by one ori is called a *replicon.* The search for oris even in human chromosomes has progressed in recent times (Giacca *et al.,* 1994). In humans there are some 10^4 to 10^5 replicons (the replicon, besides having an *ori* has also a terminus at which replication stops). This strategy is needed for rapid replication, in view of the enormous length of the DNA fiber in each chromosome 7.8 x 10^9 bps / 24 $\approx$ 3 x 10^8 bps. For instance, in *Drosophila* one expects about one replication fork per 10 kb of DNA and with the multiple origins (some 3500) replication is completed in less than three minutes.

2.2.4. DNA segments used as cues for transcription can be part of the eukaryotic *oris.* These DNA segments used as part of the transcription mechanisms are of two types: promoters and enhanceres.

Promoters are nucleotide sequences in DNA at the beginning of a transcription unit , rather than strictly at the beginning of a gene A promoter is recognized by RNA pol II as the site to begin transcription. This aspect of transcription and replication, both promoters and their corresponding upstream sequences (enhancers) can be part of the eukaryotic origins. This may be illustrated with the *ori* of the circular tumor virus, the simian virus 40 (SV40).

2.2.5. The initiation of transcription and replication involve similar stages. In both processes initially a set of given enzymes is assembled at a particular site, so that synthesis starts at a certain time. Thus, the analogy between origins and promoters is persuasive. An origin can be regarded as a promoter sequence for the initiation of replication. There is a further analogy between the stage of assembly of enzymes and strand melting. This powerful analogy between the processes of transcription and replication suggests that in a model analysis of both processes, similar physical mechanisms should be responsible for the polymerase dynamics. This is particularly plausible in the approach of thermodynamics (Chela-Flores, 1992b).

2.3 BASES FOR A QUANTITATIVE STUDY OF POLYMERASE DYNAMICS

Certain aspects of transcription and DNA replication data are well established:

2.3.1. r_f *for prokaryotes is generally much faster than for eukaryotes,* where r_f denotes the rate of fork advancement.

2.3.2. r_t *for prokaryotes is generally much faster than for eukaryotes,* where where r_t denotes the transcription rate (or, the rate of advancement of the RNA polymerases).

2.3.3. The magnitude of r_f *is normally larger than that of* r_t .

2.3.4. *In prokaryotic cells the coupling of* r_f *and* r_t *is suggested by experimentally-induced inhibition of* r_t (Pato, 1975) *which, in turn, leads to a decrease in* r_f.

A linear relationship between r_f and r_t is suggested by Table 2:

TABLE 2: The rate of propagation of the replicating fork in selected organisms.

Organism	*Type of cell*	r_f *(kb/min/f)*	*T(C)*
Escherchia coli	Unicellular	50	-
E. coli	Unicellular	25	37
Saccharomyce cerevisiae (budding yeast)	Unicellular	7 - 20	-
Drosophila melanogaster (fruit fly)	Somatic	> 2.6	-
D. melanogaster	Embryonic	2.6	25
Xenopus laevis (Clawed toad)	Somatic	0.5	-
Triturus cristatus carnifex (Great-crested newt)	Somatic	1	25
Triturus vulgaris (smooth newt)	Spermatocyte	1	25
T. vulgaris	Spermatocyte	0.6	18
Cricetulus griseus (Chinese hamster)	Somatic	< 8.3	37
HeLa (Human)	Neoplastic	1.7	37

The corresponding values for transcription are given in Table 3:

TABLE 3: The rate of propagation of RNA polymerase for virus, bacteria and nucleated cells.

Type of cell	r_t *(kb/min*	*T (C)*
Bacteriophage T7	12.0	-
Escherchia coli	1.8 - 3.6	37
Eukaryotic	0.2	-

In view of this evidence we have conjectured that:

$$r_f \approx \mu r_t$$

The dimensionless parameter μ shall be assumed to be given by the ratio of two length parameters:

2.3.5. A characteristic length the replicon size (λ_f) which corresponds to a genetic element that replicates as a whole with a unique origin of replication.

2.3.6. A characteristic length λ_t the pre-messenger RNA (pre-mRNA).

The simplest hypothesis for constructing the dimensionless parameter μ in terms of the characteristic lengths λ_f and λ_t is that:

$$\mu = (\lambda_f / \lambda_t)$$

The hypothesis of the linear formula is suggested to take the form:

$$r_f \approx (\lambda_f / \lambda_t) \; r_t \tag{2.1}$$

2.4. PRELIMINARY THEORETICAL BASIS FOR CHROMATIN STRUCTURE

A possible way to rationalize formula (2.1) was given previously (Chela-Flores, 1992b). This theory was extended to include DNA already folded into chromatin (Chela-Flores, 1994): During interphase in the cell cycle it is this solenoidal arrangement that constitutes the most abundant form of chromatin. However, at later stages in the cell cycle this structure serves as the basis for further folding, ending up at the highest degree of folding observed at the metaphase chromosome.

In view of the basic role played by all the stages of the hierarchy, it is of considerable interest to find eventually a formalism by means of which we can anticipate the rather regular manner in which DNA folds in chromatin. This problem is of considerable difficulty.

We restrict our attention to the extreme cases of the lowest stages, or highest stages, of chromatin compaction. Some useful insights have already been gained from experiments (Widom, 1989). The biochemical basis for the difference between the more dense, compact structure (heterochromatin) and its less dense form (euchromatin) remains unknown.

Heterochromatin appears most frequently at the centromeres and telomeres of the chromosomes and is characterized by highly repeated sequences. The genetic function of highly repeated sequences has not been determined yet although some models attempt to account for impeding fork movement at a terminus, where such repeated sequences may be observed.

We have discussed these questions in the context of mean-field theory. In the simplest approximation the phenomenological equations are analytically identical at the various degrees of packaging, but their parameters are assumed to differ at the different chromatin packaging regimes. We have been led to a relationship between heterochromatin and late DNA replication and derived a formula correlating (during the S phase of the cell cycle) r_f (measured in nucleotides per minute) in a relation of inverse proportionality with the degree of DNA packaging:

$$r_f = \lambda \eta^{-1/2} \tag{2.2}$$

where the dimensional constant λ has been determined. This model suggests that in the heterochromatic regions of chromatin there is reduced activity of DNA polymerases. We discuss the possible relevance of our model to late replicating

telomeres in yeast and several higher eukaryotes. Eqns. (2.1) and (2.2) apply to DNA folded into chromatin. In this way we will be able to take heterochromatin into account, a form of chromatin whose importance will be underlined in the next section.

2.5. FROM CHROMATIN TO HETEROCHROMATIN

In the process of transcription more complexity is introduced by evolutionary mechanisms, as the RNA polymerase requires an array of activators, coactivators, and basal factors which, for instance, go well beyond the relatively simple set of sigma factors of *E. coli.* This set of enzymes collects on the sequence recognised by the RNA polymerase as the site to begin transcription, the so-called 'core promoter' (Tjian & Maniatis, 1994), and may be considered analogous to the *E. coli* sigma factors.

Unlike the simple bacterial strategy for synthesizing RNA transcripts, eukaryotes have neither simple adjacent controlling elements ('promoters'), nor DNA sequences that inhibit transcription ('operators'). Instead, RNA polymerases cannot work, in the case of eukaryotes, entirely with adjacent elements, but need to orchestrate their activity with distant segments (measured in kbp) called 'enhancers' and 'silencers' which, in turn, require their own set of transcription factors (Mitchell & Tjian, 1989). We return to the 'replicon model' in our search for typical mechanisms brought about in evolution by the requirements for DNA replication. In the numerous replication origins we may find a hint of such a typical eukaryotic mechanism. The temporal order for the initiation of origins is not controlled by a property of the origin itself; but a likely candidate for controlling this aspect of DNA replication in eukaryotes is control at the level of chromatin structure. In this context, as mentioned in Sec. 1.2 for over three decades it has been well known that genes on heterochromatin go through DNA replication late in the cell cycle (Lima-de-Faria, 1983). We return to this topic in Sec. 4, in order to rationalize the phenomenon of late DNA replication of heterochromatin.

3. Heterochromatin a hallmark of eukaryogenesis

3.1. CHROMOSOME PLASTICITY IN EUKARYOTES

Heredity, or transmission of qualities from ancestor to descendant, is reflected in fairly rigid chromosome organization of germ cells. In eukaryotes this may be illustrated, for instance, in genera of the same family of dicots, the Solanaceae, in the order Scrophulariales (Asteridae): the *Lycopersicon* (tomato) chromosome has a region between centromere and telomere which consists of a row of segments in which DNA is compacted into tight masses, largely inactive in transcription ('chromomeres'); in *Petunia*, in spite of being another genera of the same family, the abundance of chromomeres is not preserved, since larger blocks of heterochromatin are observed A tiny bit of heterochromatin may be superficially indistinguishable from a eukaryotic chromomere, Brown, 1966). These two genera of the Solanaceae Family illustrate how quickly the evolutionary process can induce rearrangements of heterochromatin, while preserving general chromosome structure. This capacity for chromosomes to be molded (their 'plasticity') preserves general chromosome organization. This property may be achieved through several mechanisms including:

3.1.1. Chromosomal mutations. These may consist of translocations of discrete DNA segments ('transposable elements') between non-homologous DNA sites. This process may occur not only in eukaryotes (Spradling, 1994), but also in prokaryotes (Shapiro, 1982).

3.1.2. Horizontal gene transfer (HGT). DNA segments from one species may be transferred to another, where it may be integrated into the genome of the recipient cell. We have reviewed recently HGT, a process which may affect both eukaryotes and prokaryotes (Chela-Flores, 1995b; 1996).

The examples mentioned in Sec. 1.2 demonstrate that some of the themes developed by eukaryotes are already present in prokaryotes, in spite of the small size of the bacterial genome. Condensation of whole chromosomes is also anticipated in the small genomes of some prokaryotes. Gene silencing has reached a central position in eukaryotic gene expression in the context of sexual reproduction (Brown, 1966). However, bacterial shut-down processes in whole chromosomes have been observed in the cycle of the two developmental phases of the blindness-inducing parasite *C. trachomatics* (cf., Sec. 1.2). Infection of susceptible cells begins by a metabolically inert *C. trachomatics*, in which its core consists of apparently condensed chromatin (Barry *et al.*, 1992). Global regulation of gene expression, as exemplified by the mechanism for controlling bacterial virulence in *C. trachomatics*, is considerably developed in the much larger eukaryotic genome of metazoans and metaphytes.

3.2. ORIGINS OF FACULTATIVE HETEROCHROMATIN

Eukaryotes base their global regulation of gene expression on their ability to manipulate facultative heterochromatin. This process had to await the evolution of heterochromatin in the lower eukaryotes. The nuclei of primitive single-celled protists, such as the flagellated green alga *Clamydomonas reinhardii*, have some repeated DNA, a single nucleolus, but no heterochromatin (Britten & Kohne, 1968).

Early work has been reported on mammalian DNA of different density composed of relatively short, highly repetitive polynucleotide sequences ('satellite DNA'). This fraction is about 10% of all the DNA. Satellite DNA has been observed in the colourless alga *Polytoma* (Yunis & Yanismeh, 1971), in the euglenoid *Euglena gracilis*. The parazoan *Microcyona* (a sponge) also has satellite DNAs (Britten & Kohne, 1968); at this early stage in evolution in a branch separate from the metazoans, DNA may have developed repeated sequences by mechanisms analogous to gene amplification (Lewin, 1994). This early repetition of DNA sequences may have been preserved because they may have served some advantageous structural role. All satellite DNAs have the property of heterochromatization in common, in spite of being species-specific. Once heterochromatin had been established in higher eukaryotes the phenomenon of gene silencing was possible. In higher eukaryotes constitutive-heterochromatic, repetitive-DNA is well documented in a wide range of taxa, for instance:

3.2.1. Metazoans of diploblastic phyla (coelenterates) and triploblastic phyla (arthropods, mollusks, and chordates) (Britten & Kohne, 1968).

3.2.2. Metaphytes from monocots of the subclass Commelindae (*Secale cereale*, rye) to dicots of the subclass Rosidae (*Phaseolus vulgaris*, bean) (Britten & Kohne, 1968).

On the other hand, facultative heterochromatin may also be documented in a wide range of taxa (including invertebrates, vertebrates and plants):

3.2.3. In the phylum Arthropoda. The homopteran *Pseudococcus obscurus* (mealy bugs, or coccid in the Cicadidae Family) is able to silence a whole set of paternal chromosomes early in its development. In the dipteran genus *Miastor* (gall midges), it has been observed that the germ cells of these plant pathogens of the Cecidomyiidae Family have twenty-nine chromosomes, while the soma have only six. In this case 23 'E-chromosomes' are inactivated (Painter, 1966).

3.2.4. In the phylum Nematoda. A parasite of both vertebrates and invertebrates, *Ascaris megalocephala*, loses some chromosome segments early in its development (Yunis & Yanismeh, 1971).

3.2.5. In the phylum Chordata. Prototherian mammals, both marsupials and monotremes, are able to silence part of their sex chromosomes, while eutherian mammals are able to silence the full chromosome, only leaving a few genes active (Lyon, 1990), a point to be discussed more fully in Sec.4.

3.2.6. In the class Monocot. The perennial rye grass *Lolium perenne* of the Poaceae Family is characterized by two-ranked many-flowered spikelets; in this genus whole inactivated chromosomes may be concerned in the process of cell division (Cameron & Rees, 1967). In *S. cereale*, a species of hardy annual cereal grass, a number of supernumerary B-chromosomes are abundant in Asian populations, but are rare in Europe. These chromosomes are heterochromatic and have been shown to affect the duration of the mitotic cycle (Ayonoadu &Rees, 1967), as well as meiosis (Cameron & Rees, 1967).

3.3. PHYSICAL ASPECTS OF FACULTATIVE HETEROCHROMATIN

Mammals are probably a good taxon where we should focus attention on the physical aspects of shut-down processes of gene expression occurring in large sections, or even in whole chromosomes. The rapid evolution of this class of chordates may be exemplified, for instance, by the evolution of cetacean swimming (Novacek, 1994). This suggests that mammalian genomes are particularly dynamic and plastic. Indeed, mammalian genomic plasticity provides us with examples of novel physical phenomena:

3.3.1. Spread of X inactivation (Mohandas *et al.*, 1987), a phenomenon which may be identified through chromosomal translocations.

3.3.2. Initiation of inactivation at a specific locus, called the X chromosome inactivation centre (XIC). This locus is identified through the expression of RNA transcripts, which are specific to the X-inactivated chromosome but remain untranslated into proteins (Goldman, 1992).

3.3.3. Gene escape is a third phenomenon in which exceptional genes remain active. This may imply that X-inactivation is brought about by the spread of heterochromatization along the chromosome from the XIC.

The extent of the spread of X inactivation is clearly variable in ontogeny: in early embryogenesis the time of first appearance of X inactivation ranges from the blastocyst in some species to early neurulation in other species. On the other hand, late in ontogeny (in aging female mammals) inactivation may disappear (Wareham *et al.*, 1987). It is remarkable that the extent of the spread of inactivation is also variable in phylogeny. In order to understand this property we should approach

Ohno's Law that rationalizes some genetic experiments on monotremes. We first recall that this taxon, Monotremata, comprises the duck-billed platypus (*Ornithorhynchus anatinus*), as well as echidnas (*Tachyglossus* and *Zaglossus*).

3.4. OHNO'S LAW

In a classical paper (Ohno, 1973) it is postulated, as a fact, that the X chromosome of any eutherian mammal species, regardless of whether it be placental or marsupial, is the exact genetic equivalent of the human X chromosome. There is wide support for this postulate from human X-linked genes, since they are also X-linked in other mammals. Ohno's Law is supported by data from about 20 species included in several orders.

Ohno interprets this striking phenomenon as a case of a "frozen accident". Housekeeping genes such as phosphoglycerate kinase (X-linked in man, horse, and kangaroo), have no direct connexion with sex determination. Yet, they remained X-linked in other mammalian species, as the X-chromosome happened to be selected for gene silencing.

Monotremes, on the other hand, display incipient inactivation, which is observed to spread along the X-chromosome as evolutionary processes raise early mammals to the level of therian mammals. In fact, in monotremes, inactivation occurs only in the short arm of the X chromosome (X_p) in some tissues (Watson *et al.*, 1990). The most

recent deviations of Ohno's law in two mouse species do not change our main conclusions (Palmer *et al.*, 1995).

The argument for inserting these facts into the evolution of prototherian mammals is as follows: inactivation begins at an XIC. Its spreading along the chromosome is suggested by the inactivation of attached autosomal material (i.e., from chromosomes other than the sex chromosomes) in X-autosome translocations. Yet, one may argue that this autosomal inactivity could be due to a position-effect, in view of the proximity of the autosomal material concerned to X-chromosome heterochromatin.

However, both the spreading of inactivation, as well as the data from X-autosome translocations, may be compatible. As we have seen above, in Sec. 1.2, there are proteins specifically associated with position-effect variegation. These proteins could be part of the molecular mechanism for facultative heterochromatin (Lyon *et al.*, 1982). In the course of evolution the length of the inactivated region ξ (the 'distance of spread') has increased gradually, from monotremes to eutherians (Watson *et al.*, 1990).

4. Biophysical aspects of chromosome dynamics

4.1. THE VALUE OF r_f AS THE FORK APPROACHES HETEROCHROMATIN

The present theoretical approach suggests that there is a decrease in r_f as the replicating fork approaches heterochromatin from a euchromatic region of the chromosome. This may account for the experimental observation that heterochromatin synthesizes its DNA later than the euchromatic regions, for which there has been experimental support for over forty years (Lima-De-Faria and Jaworska, 1968).

However, although the molecular mechanism correlating heterochromatin and late replication remains unknown, such a relationship is nevertheless expressed in the present approach by the formula in Eqn. (2.2).

In addition, the r_f parameter does not always have the same constant value. This remark may be further illustrated with the observation that there are instances both in prokaryotes (*Escherichia coli*), as well as in eukaryotes (*Saccharomyces cerevisiae*), in which the replicating fork slows down under certain conditions:

4.1.1. In *Escherichia coli* the slowing down of the r_f parameter has been observed when replication and transcription occur in opposite directions. An earlier interpretation of such phenomena is that the slowed down movement of the fork may be understood in terms of its obstruction by transcription complexes temporarily immobilized on the DNA template.

4.1.2. In *Saccharomyces cerevisiae* there is an origin of replication 40 kilobase pairs (kbps) from the end of chromosome III (the telomere, ending with approximately 100 bps of irregularly repeated sequences). Bidirectional replication begins early in the S phase; one fork moves outward toward the telomere, initially at r_f = 4 kb /min/fork. In the terminal 15 kbps the fork slows down to r_f = 1.3 kb /min/fork, as found in experiments (Fangman and Brewer, 1991).

In Table 4 we have shown the theoretical expectation for the slowing down of the r_f parameter. In these cases the fork moving outwards towards the telomere finds an impediment that slows down its movement:

TABLE 4: The slowing down of the r_f parameter.

Organism	***Experimental value of r_f in euchromatin (kb/min/f)***	***Experimental value of r_f in the telomere (kb/min/f)***	***Theoretical value of r_f in the telomere (kb/min/f)***
Saccharomyces cerevisiae	4	1.3	1.3
Drosophila melanogaster (fruit fly)	> 2.6	-	> 0.8
Xenopus laevis (South African clawed toad)	0.5	-	0.16
Cricetulus griseus (Chinese hamster)	< 8.3	-	< 2.7
HeLa (Human)	1.7	-	0.6

The velocities r_f of the replicating forks are given in kilobase pairs/minute/fork (kb/min/f). The theoretical value of the r_f parameter in the telomere in the case of *Saccharomyces cerevisiae* has been fixed using the given experimental value. This helps us to infer the appropriate value of the η parameter for the degree of packing at the telomere:

$$\eta \text{ (telomere)} = 4.5 \times 10^3 \tag{4.1}$$

(a reasonable value for the heterochromatic telomere in the appropriate range of the heterocromatic metaphase chromosome of Table 4). We have used in our formula (4.1) the values r_f (euchromatin) = 4 kb/min/f and r_f (telomere) = 1.3 kb/min/f, taken from various experiments reported above (all the other values given in the last column of Table 4 are predictions); we also took

$$\eta \text{ (euchromatin)} = 5 \times 10^2, \tag{4.2}$$

in agreement with the value given in Table 5.

TABLE 5: Some typicaal values of the packing density.

DNA Organization	*Packing density* η	*Density (mg/ml)*
The 100-Å beads-on-string fibre	10	
The 300-Å beads-on-string fibre	(25-40) x 10	
Interphase chromatin	$10^2 - 10^3$	
Metaphase chromosome (heterochromatic)	10^4	> 200

We remind the reader that at the ends of all linear eukaryotic chromosomes there are unique structures that have been called telomeres, as we have already mentioned telomeres may be characterized by heterochromatin.

Our work suggests that the packaging of chromatin in higher order folding may account, in part, for such regulation of the values of the r_f parameter. In this context it should be stressed that heterochromatin appears at the telomeres. In such conditions we expect from our formula (2.2) that the value of the r_f parameter will decrease; this is in agreement with the observations of Fangman and Brewers in *S. cerevisiae.*

4.2. PHYSICAL PARAMETERS IN X-CHROMOSOME INACTIVATION

We wish to discuss four physical parameters in relation with heterochromatization:

4.2.1. The distance of spread ξ.

4.2.2. The degree of packaging in chromatin given by the parameter $\eta = L_1 / L_2$ where L_1 denotes DNA length in the fully extended state, and L_2 denotes the length of the folded state of condensation. Heterochromatin is characterized by high values of $\eta \approx 10^4$, while the 100-Å DNA fiber of eukaryotes has a value of $\eta \approx 10$.

4.2.3. The rate of advancement of the replication fork through euchromatin (r_f). The interest in this parameter is justified by the fact that r_f is a variable that is subject to measurement. For instance, in yeast chromosome 3 the replication fork normally advances through euchromatin at a rate of r_f = 4kbp/min (Fangman & Brewers, 1991). The fork, a multienzyme complex slows down by a factor of 4, as it enters the telomere, whose chromatin is in the heterochromatic state (10^4), while its spread is limited to a few kbp.

4.2.4. In first approximation late replication of heterochromatin implies the existence of a fourth parameter (λ), the direct proportionality factor between r_f and η.

Our previous analysis of this problem was based on the assumption that biological function, such as transcription, DNA replication, and compaction may be viewed as correlates of collective phenomena, rather than chemical detail (Chela-Flores, 1987). This postulate received more formal bases in a preliminary analytical approach in terms of mean-field theory. Questions discussed previously included the coupling of transcription and DNA replication in eukaryotes (Chela-Flores, 1992b), and DNA folding (Chela-Flores, 1994), in which the parameters r_f and η are functionally related by our formula (2.2), namely, $r_f = (\lambda\eta)^{-1/2}$.

If the underlying idea is correct that collective effects are appropriate correlates of some biological phenomena, then certain interesting consequences may be expected: as a chromosome early in embryogenesis is active, at a certain time the XIC triggers off a signal that spreads the heterochromatic state up to a distance ξ.

In lower mammals ξ is smaller than the length of the small arm of the X chromosome $L(X_p)$, i.e., $\xi < L(X_p)$. Then, for distances $d > \xi$, the chromosome is euchromatic. From the above relationship between distance of spread and packaging density, it follows that there should be a corresponding slowing down of the replication fork in proportion to the extension of its spread of X inactivation. We recall that a reduction in the r_f parameter has been observed in yeast. However, in prototherians a phenomenon of slowing down of the replication fork is expected to occur as this multienzyme complex enters the telomere region. According to the present analysis, there should be additional (still to be detected) increments of the r_f parameter: as the fork, starting from the centromeres, covers a distance ξ and enters the euchromatic region, the rate of fork propagation should increase.

5. Discussion

We have attempted to define some aspects of the problem of the origin and evolution of the first nucleated cell. We have emphasized the role of DNA packaging in a gene-centred approach to eukaryogenesis. The underlying physical mechanism is a two-stage process, first DNA bending and secondly, DNA folding through all the stages leading up to the heterochroamtic metaphase chromosome. Further, we have attempted to demonstrate that in the problem of the origin of the eukaryotic cell some physical aspects of the evolution of animals (Class Mammalia), at the chromosome level, may be approached with some physical methods used in the study of non-living condensed matter.

One pressing question in exobiology is whether, by adopting an approach such as the present one (a gene-centred appraoch) rather than the traditional cytological criterion for an eukaryotic cell, we attempt to devise assays for acertaining whether what we encounter in a search for life by, for instance a Mars lander, is a prokaryotic or a eukaryotic organism. In the positive case of identifying a fossil of a certain microorganism, it is not yet clear how to identify in an unequivocal manner in a given microfossil, whether there are traces of the hallmark of eukaryogenesis (heterochromatin).

On the other hand, if *living microorganisms* are found in future space missions the present approach has definite advantages, as it does not base its identification on morphological properties, such as the presence of organelles, which may be missing in a whole phylum of eukaryotes (the microsporidia). The present work suggests one possibility for such an assay: one may search for cellular replication and investigate whether there is a delay in replication of chromosome segments, which would characterize them as heterochromatic.

It should be noted that the straightforward search for a membrane-bounded set of chromosomes (in a single- cell organism that might be encountered) evidently does not answer the question, as there are prokaryotes, such as *Gemmata oscuriglobus* that do have a membrane-bounded nucleoid (Fuerst and Webb, 1991). The case of the origin of multicellularity is discussed in Sec. 6.

As stated in Sec. 1.3, we cannot exclude at present organisms that may inhabit deep in the silicate crust of the rocky planets. The case of the Earth suggests this possibility. Such 'extremophiles' may have been deposited with the original sediment and survived over geologic time (Parkes and Maxwell, 1993). Hence, if such organisms are alive deep in the Martian crust, then assays should be formulated to acertain whether the first steps towards eukaryogenesis have been taken (as suggested above).

6. Conclusions.

In the present scenario it is important to acertain whether the first steps towards *multicellularity* could also have been taken. However, the origin of metazoans cannot be inferred from the fossil record. Indeed, metazoans are too well developed when they make the first fossil appearance (Lipps *et al.*, 1992). However, it is not possible to exclude a Precambrian metazoan history of several hundred million years from the best available data set (Field *et al.*, 1988).

We recall that in molecular evolution we may define a *genetic distance* as a measure of the number of nucleotides substitutions per nucleotide site between two homologous DNA sequences since the division between the sequences. It has been found that the distance between the phyla Echinodermata and Chordata turns out to be substantially less than the distances between these phyla and the cnidarians. But the echinoderm-chordate divergence cannot be earlier than the Cambrian period. These arguments of molecular evolution suggest that metazoans may have had a long Precambrian history (Runnegar, 1992). But from Sec. 1.2, we know that the fossil evidence suggests that eukaryogenesis may have been more recent that 2,000 My bp. These results demonstrate how intimately related are the questions of late onset of eukaryogenesis and early origin of metazoans.

We may conclude that on Earth these events may have been separated by an abiotic limiting factor, possibly the eventual increment of atmospheric oxygen (Sec. 1.2). In any exobiological consideration in which detailed atmospheric evolution remains missing, the search for life should not be entirely constrained to assays that are envisaging microorganisms that are a priori assumed to resemble organisms restricted to the domain Bacteria.

To summarize we have drawn two main conclusions from our analysis of data from molecular evolution, genetics, geology and paleontology:

6.1 1. In the search for unicellular organisms in the assays that might be considered in the futurre space missions, the hallmark of eukaryotes should not be overlooked. This may be implemented, for instance, by monitoring cell division, as we could attempt to identify *heterochromatic* sectors of the chromosomes that may replicate late.

6.1.2. We cannot exclude at present that in planetary or satellite atmospheric evolution that may have differed from the Earth, precautious multicellularity cannot be excluded (both of prokaryotes but particularly of eukaryotes. In those cases are possible, assays should not be ruled out that may test whether *micrometazoan organisms* [possibly similar to modern marine larvae (Davidson *et al.*, 1995)]may be present deep in the crust of terrestrial planets such as Mars,

6.1.3. We cannot rule out at present the existence of organisms uniquely adapted to the alien chemistry that characterizes some extraterrestrial environments, in particular the Martian soil that was carefully explored in two widely separated locations in the Viking missions (Murray, 1981).

References

Ayonoadu, U.W.U. and Rees, H. (1968). The regulation of mitosis by B chromosomes in rye, Heredity **23**, 164.

Barry III, C.E., Hayes, S.F., and Hackstadt, T. (1992). Nucleoid condensation in *Escherichia coli* that express a *Chlamydial* histone homolog, Science **256** 377-379.

Blumenthal, A.B., Kriegstein, H.J. & Hogness, D.S. (1974). Cold Spring Harbor Symp. Quant. Biol. **38,** 205-223 c1.

Bonen, L and Doolittle, W.F. (1976). Partial sequences of 16S rRNA and the phylogeny of blue-green algae and chloroplasts, Nature **261,** 669-673.

Brewer, B.J. and Fangman, W.L. (1993). Initiation at closely spaced replication origins in a yeast chromosome, Science **262,** 1728-1731.

Brown, S.W.(1966). Heterochromatin, Science **151,** 417-425.

Britten R.J. and Kohne, D.E.(1968). Repeated sequences in DNA, Science **161**, 529-540.
Callan, H.G. (1972). Replication of DNA in the chromosomes of eukaryotes. Proc.R. Soc. Lond. **B 181,** 19-41.

Cameron, F.M. and Rees, H. (1967). The influence of B chromosomes on meiosis in *Lolium*, Heredity **22,** 446-450.

Cavalier-Smith, T.(1987). Eukaryotes with no mitochondria, Nature **326** , 332-333.

Chandra, H.S. and Brown, S. (1975). Chromosome imprinting and the mammalian X chromosome, Nature **253**, 165-168.

Chela-Flores, J. (1987). Towards a collective biology of the gene. J. Theor. Biol. **126,** 127-136.

Chela-Flores, J. (1992a). Influence of Chromatin Molecular Changes on RNA Synthesis during Embryonic Development, Acta Biotheoretica **40,** 41-49.

Chela-Flores, J. (1992b). Towards the Molecular Bases of Polymerase Dynamics. J. Theor. Biol. **154,** 519-539 and Erratum: J. Theor. Biol. **157** (1992), 269.

Chela-Flores, J.(1994a). Towards the theoretical bases of the folding of the 100-A nucleosome filament, J. Theor. Biol. **168** , 65-73.

Chela-Flores, J.(1995a). Are there molecular relics from the origin of life? In: J. Chela-Flores, M. Chadha, A. Negron-Mendoza, and T. Oshima. A. (1995). pp.185-199.

Chela-Flores, J. (1995b). Some physical problems in biology: Aspects of the origin and structure of the first cell. In: Ponnamperuma, .C. and Chela-Flores, (Eds.) (1995). pp. 315-330.

Chela-Flores, J. (1996). Preservation of relics from the RNA world through natural selection, symbiosis and horizontal gene transfer. Acta Biotheoretica (in press).

Chela-Flores, J. Chadha, M., Negron-Mendoza, A. and Oshima, T. (Eds.).(1995). *Chemical Evolution: Self-Organization of the Macromolecules of Life.* A. Deepak Publishing: Hampton, Virginia, USA

Chela-Flores, J. and Raulin, F. (Eds.). (1996). *Chemical Evolution: Phys:cs of the Origin and Evolution of Life.* To be published by Kluwer Academic Publishers, Dordrecht, The Netherlands.

Conway-Morris, S.(1993). The fossil record and the early evolution of the Metazoa, Nature **361**, 219-225.

Darnell, J., Lodish, H. & Baltimore, D. (1990). *Molecular Cell Biology*. 2nd ed.New York: W.H. Freeman & Co. p. 538.

Davidson , E.H., Peterson, K.J., and Cameron, R.A. (1995). Origin of bilaterian body plans: evolution of developmental regulatory mechanisms. Science **270**, 1319-1325.

Dawkins, R. (1989). *The selfish gene*. Second Edition. Oxford University Press.

De Duve, C. (1984). *A guided tour of the living cell.* Vol **2.** Scientific American Books: New York. p. 370

De Pamphilis, M.L. (1988). Transcriptional elements as components of eukaryotic origins of DNA replication, Cell **52**, 635-638.

Douglas, S.E., Murphy, C. ' ., Spenser, D.F., and Gray, M.W. (1991). Cryptomonad algae areevolutionary chimaeras of two phylogenetically distinct unicellular eukaryotes, Nature **350**, 148-151.

Drake, F. (1996). Ongoing and future searches for extraterrestrial intelligent life. In: Chela-Flores, J. and Raulin, F. (Eds.). (1996).

Drake, F. and Sobel, D. (1992). *Is there anyone out there? The scientific search for Extraterrestrial Intelligence.* Delacorte Press: New York.

Eissenberg, J.C., James, C., Foster-Harnett, D.M., Harnett, T., Ngan, V., Elgin, S.C.R. (1990). Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*, Proc. Natl. Acad. Sci. USA **87**, 9923-9927.

Fangman, W.L. and Brewer,B.J. (1991). Activation of replication origins within yeast chromosomes. Ann. Rev. Cell Biol. **7**, 375-402.

Fangman, W.L. and Brewer, B.J. (1992). A question of time: replication origins of eukaryotic chromosomes, Cell **71**, 363-366.

Field, K.G., Olsen, G.J., Lane L.D., Gionannono, S.J., Ghiselin, M., T., Raff, E.C., Pace, N.R., and Raff, R.A. (1988). Molecular phylogeny of the animal kingdom. Science **239**, 748-753.

Finch, J.T. and Klug, A. (1976). Solenoidal model for superstructure in chromatin. Proc. Natl. Acad. Sci. USA **73**, 1897-1901.

Fuerst, J.A. and Webb, R.J. (1991). Membrane-bounded nucleoid in the eubacterium *Gemmata oscuriglobus*, Proc. Natl. Acad. Sci. USA **88**, 8184-8188.

Giacca, M., Zentilin, L., Norio, P., Diviacco, S., Dimitrova, D., Contreas, G., Biamonti, G., Perini, G., Weighardt, F., Riva, S., and Falaschi, A. (1994). Fine mapping of a replication origin of human DNA. Proc. Natl. Acad. Sci. USA **91**, 7119-7123.

Gold, T. (1992). The deep, hot biosphere. Proc. Natl. Acad. Sci. USA **89**, 6045-6049.

Goldman, M.A. (1992). The silence of the X, Nature Genetics **2**, 169-170.

Gould, S.J. (1989). *Wonderful Life*. The Burgess Shale and the Nature of History. Penguin Books: London.

Griffith, J.D.(1976) Visualization of prokaryotic DNA in a regularly condensed chromatin-like fiber, Proc. Natl. Acad. Sci. USA **73**, 563-567.

Hackstadt, T., Baer, W., and Ying, Y.(1991).*Chlamydia trachomatis* developmentally regulated protein is homologous to eukaryotic histone H1, Proc. Natl. Acad. Sci. USA **88**, 3937-3941.

Hall, D.T., Strobel, D.F., Feldman, P.D., McGrath, M.A. and Weaver, H.A. (1995). Detection of an oxygen atmosphere on Jupiter's moon Europa. Nature **373**, 677-679.

Han, T.-M. and Runnegar, B. (1992). Megascopic eukaryotic algae from the 2.1-billion-year-old Negaunee iron-formation, Michigan, Science **257**, 232-235.

Hand, R. (1978). Eukaryotic DNA: organization of the genome for replication, Cell **15**, 317-325.

Haselkorn, R. and Rouvier-Yaniv, J. (1976). Cyanobacteria DNA-binding protein-related *Escherichia coli* HU protein, Proc. Natl. Acad. Sci. USA **73**, 1917-1920.

Herzog, M. and Soyer, M.-O. (1981). Distinctive features of dinoflagellate chromatin. Absence of nucleosomes in a primitive species *Prorocentrum micans* E., Eur. J. Cell Biol. **23**, 295-302.

Hulton, C.S.J., Seirafi, J., Hinton, J.C.D., Sidebotham, J.M., Waddell, L., Pavitt, G.D., Owen-Hughes, T., Spassky, A., Buc, H., and Higgins, C.F. (1990). Histone-like protein H1 (HN-S), DNA supercoiling, and gene expression in bacteria. Cell **63**, 631-642.
Jacob, F. (1993) The replicon: thirty years later, Cold Spring Harbor Symposia on Qual. Biol. **58**, 383-387.

John, P. and Whatley, F.R. (1975). *Paracoccus denitrificans* and the evolutionary origin of the mitochondrion, *Nature* **254**, 495-498.

Jones, N. and Rees, H. (1967). Genotypic control of chromosomal behaviour in rye. XI. The influence of B chromosomes on meiosis, Heredity **22**, 333-347.

Kato, J., Misra, T.K., and Chakrabarty, A.M. (1990). AlhR3, a protein resembling eukaryotic histone H1, regulates alginate synthesis in *Pseudomonas aeruginosa*, Proc. Natl. Acad. Sci. USA **87**, 2887-2891.

Kieffer, H.H., Christensen, P.R., Martin, T.Z., Miner, E.D. and Palluconi, F.D. (1976). Temperatures of the Martian Surface and Atmosphere: Viking Observation of Diurnal and Geometric Variations. Science **194**, 1346-1351.

Knoll, A.H. (1994). Proterozoic and Early Cambrian protists: Evidence for accelerating evolutionary tempo, Proc. Natl. Acad. Sci. USA **91**, 6743-6750.

Kornberg, A. (1988): DNA replication, J. Biol. Chem. **263**, 1-4.

Kornberg, A. and Baker, T.A. (1992). *DNA Replication.* Second edition. New York: W.H. Freeman and Co.

Leighton, T.J., Dill, B.C., Stock, J.J., Phillips, C. (1971). Absence of histones from the chromosomal proteins of fungi, Proc. Natl. Acad. Sci.USA **68**, 667-680.

Lewin, B. (1994). *Genes V.* Oxford University Press, pp.1087-1091 .

Lima-de-Faria, A. 1983). *Molecular evolution and organization of the chromosome.* Elsevier: Amsterdam, pp. 1186.

Lima-De-Faria, A. and Jaworska, H. (1968). Late DNA synthesis in heterochromatin. Nature **217**, 138-142.

Lipps, J.H., Bengston, S. and Farmer, J.D. (1992). The Precambrian-Cambrian Evolutionary Transition. In: Schopf, J.W. and Klein, C. (1992). pp. 453-457.

Liskens, M.H.K. and Huberman, J.A. (1990). The two faces of higher eukaryotic DNA replication origins, Cell **62,** 845-847.

Littau, V.C., Burdick, C.J., Allfry, V.G., and Mirsky, A.E. (1965). The role of histones in the maintenance of chromatin structure, Proc. Natl. Acad. Sci. USA **54,** 1204-1212.

Lyon, M.(1968). Chromosomal and subchromosomal inactivation, Ann. Rev. Genet. **2**, 31-52.

Lyon, M.F. (1990). Evolution of the X chromosome, Nature **348**, 585-586.

Lyon, M.F., Zenthon, J., Evans, E.P., Burtenshaw, MD., Wareham, K.A., and Williams, E.D. (1986). Lack of inactivation of a mouse X-linked gene physically separated from the inactivation centre, J. Embyol. exp. Morph. **97**, 75-85.

Margulis, L. (1993). *Symbiosis in Cell Evolution.* Second Ed. New York: W.H. Freeman and Co.

Maynard Smith, J. (1993). *The theory of evolution.* Canto Edition, Cambridge University Press, Cambridge (UK) p. 122.

Maynard Smith, J. and Szathmary, E. (1995). The major evolutionary transitions,Nature **374**, 227-232.
Moorbath, S. (1995). Age of the oldest rocks with biogenic components. In: Ponnamperuma, C. and Chela-Flores, J. (Eds.). (1995). pp. 85-94.

Mitchell, P.J. and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA-binding proteins, Science **245**, 371-378.

Mohandas, T., Geller, R.L., Yen, P.H., Rosendorf, J., Bernstein, R., Yoshida, A., and Shapiro, L.J. (1987). Cytogenic and molecular studies on a recombinant human X chromosome: implications for the spreading of X chromosome inactivation, Proc. Natl. Acad. Sci. USA **84**, 4954-4958.

Murray, B., Malin, M.C., and Greely, R. (1981). Earthlike Planets. Surfaces of Mercury, Venus, Earth, Moon, Mars. W.H. Freeman& Co.: San Francisco. p. 317.

Novacek, M.J. (1994). Whales leave the beach, Nature **368,** 807.

Ohno, S. (1973). Ancient linkage groups and frozen accidents, Nature **244**, 259-262.

Oro, J. (1996). Cosmic Evolution, Life and Man. In: Chela-Flores, J. and Raulin, F. (Eds.). (1996).

Painter, T.S. (1966). The role of the E-chromosomes in Cecidomyiidae, Proc. Natl. Acad. Sci. USA **56**, 853-855.

Palmer, S.,Perry, J. and Ashworth, A. (1995). A contravension of Ohno's law in mice, Nature Genetics **10**, 472-476.

Parkes, J. and Maxwell, J. (1993). Some like it hot (and oily). Nature **365**, 694-695.

Pato, M.L. (1975). Alteration of the rate of movement of deoxyribonucleic acid replication forks. J. Bacteriol. **123,** 272-277.

Ponnamperuma, C. and Chela-Flores, J. (Eds.). (1995). *Chemical Evolution: The Structure and Model of the First Cell.* Kluwer Academic Publishers, Dordrecht, The Netherlands.

Riding, R. (1992). The algal breath of life, Nature **359,** 13-14.

Rouvier-Yaniv, J. and Gros, F. 1975). Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli*, Proc. Natl. Acad. Sci. USA **72**, 3428-3432.

Runnegar, B. (1992). Origin and Diversification of the Metazoa. In. Schopf, J.W. and Klein, C. (1992). pp. 485.

Runnegar, B. (1994). Proterozoic eukaryotes: Evidence from biology and geology. In: *Early Life on Earth*. Proc. Nobel Symposium No. 84. Ed. Bengston, S. New York: Columbia University Press. pp.287-297.

Schidlowski, M. (1995). Early Terrestrial Life: Problems of the oldest record. In: Chela-Flores, J., M. Chadha, A. Negron-Mendoza, and T. Oshima (Eds.). (1995). pp. 65-80.

Schopf, J.W. (1993). Microfossils of the Early Archean Apex Chert: New Evidence of the Antiquity of Life, Science **260**, 640-646.

Schopf, J.W. and Klein, C. (1992). The Proterozoic Biosphere. A Multidisciplinary Study. Cambridge: Cambridge University Press.

Seckbach, J. (1972). On the fine structure of the acidophilic hot-spring alga *Cyanidium caldarium:* a taxonomic approach. *Microbios* **5**, 133-142.

Seckbach, J. (1994). The natural history of *Cyanidium* (Geitler, 1933): past and present perspectives. In: Evolutionary pathways and enigmatic algae: *Cyanidium caldarium* (Rhodophyta) and related cells. Ed. Seckbach, J. Kluwer Academic Publishers: Dordrecht, The Netherlands. pp. 99-112.

Seckbach, J. (1995). The first eukaryotic cells-Acid hot-spring algae. In: Ponnamperuma, C. and Chela-Flores, J. (Eds.). (1995). pp. 335-345.

Shapiro, J.A. (1982). Variation as a genetic engineering process, In: Bendall, D.S. *Evolution from molecules to man.* Cambridge University Press. pp. 253-270.

Soffen, G.A.(1976). Scientific results from the Viking Mission. Science 194, 1274-1276.

Sogin, M., Gunderson, J., Elwood, H.., Alonso, R.A., Peattie, D.A. (1989).Phylogenetic meaning of the kingdom concept:an unusal ribosomal RNA from *Giarda Lamblia,* Science **243**, 75-77.

Soyer, M.-O. (1971). Structure du noyau des *Blastodinium* (Dinoflagelles parasites), Chromosoma **33**, 70-114.

Spradling, A.C. (1994). Transposable elements and the evolution of heterochromatin. In: Molecular evolution of physiological processes, Ed. D.M. Famborough. Rockefeller University Press: New York.. pp. 69-83.

Stryer, L. (1988). *Biochemistry*. 3rd Ed. New York: W.H. Freeman and Co. p. 831.

Tjian, R. and Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces, Cell **77**, 5-8.

Wareham, K.A., Lyon, M.F., Glenister, P.H.. and Willams, E.D. (1987). Age related reactivation of an X-linked gene, Nature **327**, 725-727.

Watson, J.M., Spenser, J.A., Riggs, A.D., and Graves, J.A.M. (1990). The X chromosome of morotremes shares a highly conserved region with the eutherian and marsupial in spite of the absence of X chromosome inactivation, Proc. Natl. Acad. Sci. USA **87**, 7125-7129.

Widom, J. (1989). Toward a unified model of chromatin folding. Annu. Rev. Biophys. Biophys. Chem. **18**, 365-395.

Yunis, J. and Yanismeh, W.G. (1971). Heterochromatin, satellite DNA, and cell function, Science **174**, 1200-1209.