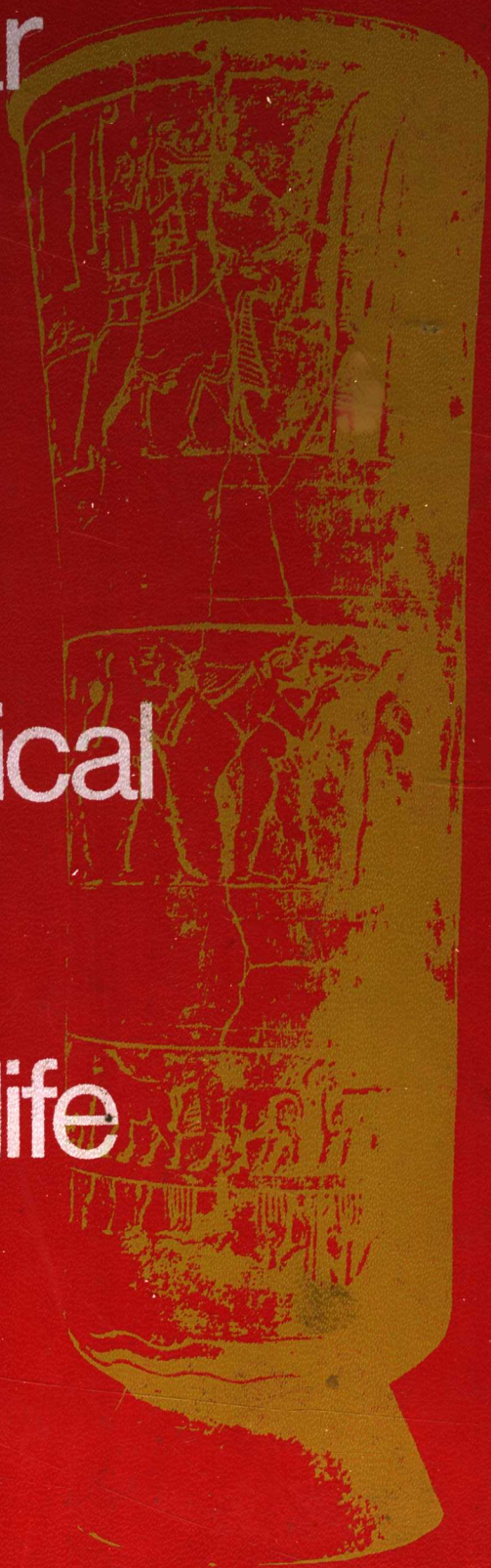
molecular evolution 2



biochemical evolution and the origin of life

edited by
e. schoffeniels

north-holland

Biochemical evolution and the origin of life

Proceedings of the International Conference on Biochemical Evolution

Edited by

E. Schoffeniels
Laboratoire de Biochimie,
Université de Liège
Liège Belgium

1971

NORTH-HOLLAND PUBLISHING COMPANY AMSTERDAM · LONDON

Library of Congress Catalog Card Number: 75-146189

ISBN North-Holland – series: 0 7204 4082 3
volume: 0 7204 4084 x

140 graphs and illustrations, 65 tables

Published by:
NORTH-HOLLAND PUBLISHING COMPANY – AMSTERDAM

Sole distributors for the U.S.A. and Canada:
AMERICAN ELSEVIER PUBLISHING COMPANY, INC.
52 VAN DER BILT AVENUE
NEW YORK, N.Y. 10017

Printed in the Netherlands

Marcel Florkin

EDITOR'S PREFACE

Since the publication of Marcel Florkin's great work *L'évolution biochimique'* *, it has become increasingly evident that the comparative biochemical approach to the interpretation of the adaptation of animals to their environment is a key factor in elucidating the irritating problem of evolution. If the concept stated, adaptation at the molecular scale, is not a new one, only in a few exceptionally favorable cases has it been possible to identify it. Adaptation is not necessarily recognizable by the study of single isolated molecules (adapted molecules), but may result rather from changes at the level of several molecular species involved in the complex polygenic adaptive mechanisms.

Under the enthusiastic guidance of Marcel Florkin, a systematic analysis of the molecular aspects underlying an adaptation recognizable at a higher level of organization, has been undertaken in the Department of Biochemistry of our University. This has led Marcel Florkin to define basic concepts in comparative biochemistry that have been widely accepted by those research workers directly involved in this field.

As a tribute to the outstanding contributions of Marcel Florkin, it seemed adequate to organize on the occasion of his 70th birthday, an International Conference on Biochemical Evolution.

This was very tempting owing to the fact that Prof. A. Buvet from the University of Paris was setting up the Third International Conference on the Origin of Life **, thus making it logical to couple this event, dealing with the prebiological evolution, with our meeting in Liège. Moreover, the symposium at Pont-à-Mousson was intended to honour Academician A.I. Oparin, rendering therefore even more attractive the whole idea of dedicating jointly these international manifestations to two of the most prominent personalities in the field of the formation and evolution of biological systems.

E. Schoffeniels

* Desoer, Liège; Masson, Paris, 1944; translated into English by S. Morgulis, Academic Press, New York, 1949.

** April 19–25, 1970, Pont-à-Mousson, France.

Biochemical Evolution and the Origin of Life, ed. E. Schoffeniels

ALLOCUTION of Professor M. WELSCH

Dean, Faculty of Medicine, University of Liège

Dear Colleagues, Ladies and Gentlemen,

As dean of the Medical School of Liege University, it is my privilege to open the present International Conference on Biochemical Evolution, and my great pleasure to extend a most heartfelt welcome to all participants.

We are very proud and happy to greet here a galaxy of talented biochemists who have notoriously illustrated their respective fields of investigation and we are grateful to them for visiting Liege after attending the Third International Conference on the Origin of Life held last week at Pont-à-Mousson.

We are especially glad of their presence here in view of the fact that this Conference is dedicated to one of them, our colleague Marcel Florkin, who will become professor emeritus within a few months. Marcel Florkin has gained the highest reputation in biochemistry and his name is associated with the development of several chapters of this science. However, it is mainly as one of the foremost proponents of Biochemical Evolution that he will appear this week.

His disciples and many friends are very pleased that his peers accepted to honour him with them by collaborating to the Conference and therefore setting it on the highest scientific level.

May I take this opportunity to tell Professor Florkin how much we admire his indefatigable activity, expressed not only by original investigations in biochemistry, but also by research in the History of Medicine, by devotedness to the organization of science at the international level and to the spreading of knowledge and love of arts in the Walloon community.

I wish Professor Florkin, on the eve of his seventieth birthday, many more years of continued activity and, thanking him warmly for everything he did and will do for Science and Arts, for the City of Liege and Wallonie, for our University and our Medical School, I beg him to accept the congratulations of his colleagues together with the expression of their deep respect and sincere affection.

Ladies and Gentlemen,

May I now ask your forbearance for a mere microbiologist who has to assume the presumptuous task of briefly introducing the Conference on Biochemical Evolution. I shall do it in calling to your attention a recent Symposium of the Society for General Microbiology devoted to Organization and Control in Prokaryotic and Eukaryotic Cells *. This subject might well have been dealt with by your group between the studies of prebiological evolution at Pont-à-Mousson and of biochemical evolution at Liège.

Revolutionary as it appeared when first introduced, the concept of biochemical evolution is now widely accepted. In fact, the rivalry which formerly opposed morphological to physiological sciences has now practically disappeared since biochemical thought has penetrated the whole field of biology. Everyone now agrees that structure is not merely to be described, but that it should be explained by its constituting molecules and by its biosynthesis. Conversely, it is also recognized that structural components of living beings explain their biochemical activities.

Thanks mainly to electron microscopy and to the study of isolated cellular organelles, living beings are now subdivided into two main groups, respectively designed as prokaryotic and eukaryotic organisms. This is a considerable generalization of a distinction proposed many years ago, among *Protista*, by the French protozoologist Chatton. Prokaryotic organisms comprize only unicellular beings, namely *Schizomycetes* (bacteria) and *Schizophyceae* (blue-green algae). All other living beings, whether unicellular or multicellular, are made of eukaryotic cells. Differences in cell organization justifying the distinction between those two large categories are many and well-known. They are fully discussed in the symposium to which I referred a few moments ago. The evolutionary relationships of the two groups, in particular, are reviewed in a remarkable assay by R.Y. Stanier.

In spite of important differences in structure and biochemical make-up, prokaryotic and eukaryotic cells have a number of important properties in common. Among them should be stressed: (1) the possession of an identical genetic code, together with common mechanisms for its translation and transcription; and (2) similar general pathways for a large number of biosyntheses, including photosynthesis which occurs in some organisms of both categories. It thus appears most likely that prokaryotes and eukaryotes derive from a common ancestor type of cell.

* Organization and control in prokaryotic and eukaryotic cells. H.P. Charles and B.C.J.G. Knight, editors (Cambridge University Press, 1970).

Prokaryotic cells are structurally much simpler than eukaryotic cells. But this would not be sufficient to assume that the former originated earlier than the latter. However, biochemical mechanisms considered as primitive are widely distributed among today prokaryotes, whereas they are absent or exceptional in eukaryotes. Thus, many prokaryotic species are obligate anaerobes and, furthermore, such species show a great diversity of the biochemical pathways used for energy production. On the other hand, the relatively few obligate anaerobes known among eukaryotic organisms appear as adapted to exceptional conditions of environment and exclusively use the glycolytic pathway. Among the phototrophic prokaryotes, some, the bacteria, are anaerobic and unable to utilize H_2O as H donor. Their photosynthetic apparatus comprizes but a single kind of reaction centre. All other prototrophs, whether prokaryotic, such as blue-green algae, or eukaryotic (other algae and green plants) use H_2O and evolve O_2 on the one hand, and require two kinds of reaction centres on the other hand. As it is generally admitted that life on earth occurred at a time when the atmosphere was devoid of oxygen and that this gas appeared largely as a result of its production through photosynthetic activity, it would appear that chemotropic anaerobic prokaryotes represented a very primitive form of life. From them derived, first, the phototrophic anaerobic prokaryotes, then the phototrophic aerobic prokaryotes, and later on, the chemotrophic aerobic prokaryotes.

As to the origin of present-day eukaryotic cells, a hypothesis advanced long ago seems to gain now much favour. It assumes that the eukaryotic cell originated from the acquisition, by a primitive type of cell, of prokaryotic endosymbionts. The theory was proposed in view of the frequent occurrence of symbionts, eukaryotic as well as prokaryotic, in the cells of many eukaryotic species belonging to a variety of groups. It became more attractive when electron-microscopic studies and biochemical observations stressed the many analogies existing between prokaryotic cells and characteristic organelles of eukaryotic cells such as mitochondria and chloroplasts. Additional arguments were found in the fact that such organelles do reproduce themselves, carry a part, small albeit functionally important, of the cell's genome and at least some of the biochemical armamentarium involved in its translation, transcription and replication.

The ancestral type of cell which, after associating with endosymbionts, became the origin of the eukaryotic line is visualized by some as a typical prokaryotic cell, by others as a cell which had already undergone important modifications toward the eukaryotic state, although lacking mitochondria and chloroplasts. According to Stanier, the crucial event might have been the

acquisition of a capacity to perform 'endocytosis' (phagocytosis and pinocytosis), necessarily coupled with the loss of ability to synthesize the cell wall. The prokaryotic cell thus modified would be able to obtain nutrients by the way of predation on other cells. In this line of predaceous cells, selection would not operate on new pathways replacing glycolysis to produce energy more efficiently, but would favour mutations improving predation. Thus could be explained the development of those structures highly characteristic of eukaryotes: the Golgi apparatus and the microtubular systems. The primitive eukaryotic cell thus evolved would, later on, when photosynthesis and respiration had appeared in prokaryotes, associate symbiotically with such organisms, and eventually retain them permanently as characteristic organelles, mitochondria and chloroplasts.

This is a wonderful and plausible, although by no means proven, story for the details of which I recommend reading the original papers of the published symposium. In referring to it, I simply wished to stimulate your interest and to provide a link between the topics discussed at Pont-à-Mousson and the lectures to be delivered here presently.

I thank you for your attention and declare now open the International Conference on Biochemical Evolution.

CONTENTS

Biochemical Evolution and the Origin of Life, ed. E. Schoffeniels

ELECTRONIC FACTORS IN BIOCHEMICAL EVOLUTION

B. PULLMAN
Institut de Biologie Physico-Chimique,
13 rue Pierre et Marie Curie, Paris 5e

The existence of chemical bonds can be accounted for exclusively by quantum-mechanics. It is the quantum theory, therefore, which contains in itself the possible explanation for the formation of the first, undoubtedly small, molecules in the universe and for their development into the numerous and large compounds known today.

The aim of this lecture is to outline the electronic factors which could possibly have played a substantial role in the chemical evolution leading to the selection of biomolecules, as indicated by the application to this problem of some basic ideas of quantum biochemistry.

Three aspects of the problem are particularly studied:
(1) the electronic factors associated with the presence of conjugated systems;
(2) the role of the electronic factors in intra- and intermolecular interactions;
(3) the role of the electronic factors in mutagenesis.

1. The significance of conjugated systems in biochemical evolution

A large number of essential biomolecules which are related to, or perform, the fundamental functions of the living matter (purines, pyrimidines, pteridines, flavins, quinones, carotenes, retinals, 'energy-rich' phosphates, practically all coenzymes etc...) are conjugated systems, generally heterocyclic, rich in π electrons, highly delocalizable. The possibility of utilizing this class of molecules must have represented for the process of biochemical evolution a number of substantial advantages susceptible to facilitate or and to accelerate the appearance and the development of life.

Among these advantages, we may quote:
(A) Thermodynamic stability;
(B) Radio- and photoresistance;
(C) Functional advantages.

2. Intermolecular associations and intramolecular interactions

The predominant role of the electronic factors in these phenomena, so essential for the production of the large organized structures, without which no elaborated form of life seems to be possible, is particularly well illustrated by the elucidation of the nature of *complementarity*, such as the one prevailing in the Watson–Crick coupling of purine and pyrimidine bases and of the conformational preferences of the amino acid residues of proteins.

3. The role of the electronic factors in mutagenesis

The role of the electronic factors in mutagenesis is particularly evident when considering the possible role in this respect of the tautomerization of the nucleic acid bases.

The detailed text of this lecture will be published in 'Exobiology' (C.Ponnamperuma ed.), North-Holland Publishing Company, Amsterdam.

Biochemical Evolution and the Origin of Life, ed. E. Schoffeniels

THE INFLUENCE OF THE GENETIC CODE ON PROTEIN EVOLUTION

Jack Lester KING
Department of Biological Sciences,
University of California, Santa Barbara

Evolution of the code

The origin of the genetic code has been the subject of much speculation; still almost nothing is known about pre-code evolution. There is an appearance of capriciousness about the genetic code. For no obvious reason, some amino acids are coded for by six codons, others by four, three, two or even only one codon, as in the cases of methionine and tryptophan. I should like to suggest that this apparent capriciousness is real – that the pattern of the code as such really makes no sense.

If the evolutionary origin of the genetic code was at all like other kinds of evolution, earlier versions of the code must have been variable, simply because one can have no evolution without variability. If variable, early codes must have also been ambiguous. It is hard to imagine how an organism might survive with an ambiguous genetic code, but there are many other aspects of early evolution that are also hard to imagine. Actually, the code is rendered slightly ambiguous in the presence of amber and ochre suppressors, and suppressor strains *do* survive.

In some fashion, however, some organism developed an unambiguous genetic code. What an enormous evolutionary advantage the first unambiguous genetic code must have had over the last ambiguous code! The organism prossessing it, and its descendents, must have quickly supplanted all other forms of life.

But the first unambiguous code must also have been the last code, the present code; no further changes could or can take place without the reintroduction (at least temporarily) of ambiguity and hence variability. Such a reintroduction would be far from likely to be evolutionarily advantageous. Thus the *first really workable* code is universal, and permanent.

The distribution of amino acid frequencies

There are some who prefer to believe that the present code is the best of all possible codes, and try to find a functional advantage in its every peculiarity. It has been observed by a number of investigators [6, 9, 12] that the code fits the distribution of amino acid frequencies very well – that is, by taking the known frequencies of nucleotide bases, arranging these bases in perfectly random permutations, and reading off the triplets with the amino acid code, one comes up with a pretty good prediction of amino acid frequencies actually seen in nature (fig. 1). Only arginine fails to come close to expectation, since its predicted frequency is more than twice its observed frequency.

There are two ways to interpret this correlation, as there are in most correlations. Either the code evolved to fit the observed amino acid frequencies, or the observed amino acid frequencies evolved to fit the code [13].

The first alternative takes us back to the unknown dark days when the

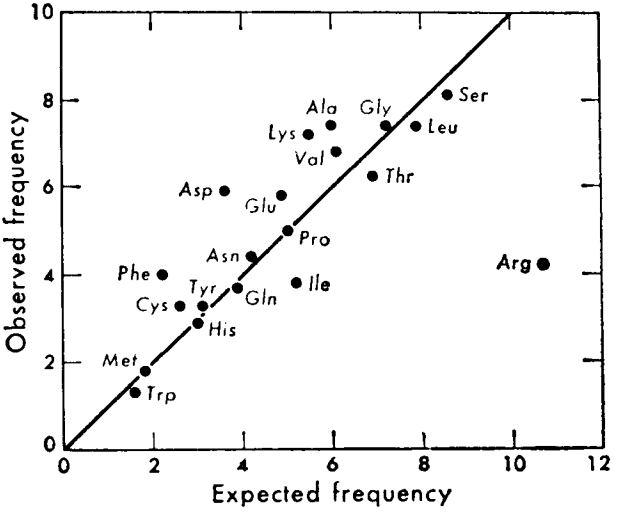


Fig. 1. Graph showing the similarity between the observed frequencies of amino acids in 53 completely sequenced mammalian proteins and the frequencies predicted by the genetic code and random permutations of DNA nucleotides. The frequencies are in percentages of total amino acid content. The straight line represents an idealized equality of expectation and observation.

code was variable, ambiguous, and evolving. The argument depends upon the assumption that each position in every protein is occupied by the amino acid that is functionally the best for that position, since it is adaptive evolution that placed it there; in the sum, therefore, the distribution of amino acid frequencies is the way it is because organisms function best this way. Implicit in the argument is the further assumption that the same distribution of amino acids was optimal back when the code was ambiguous and life existed in a highly reducing, ammonia-rich primordial sea. The code that could bring about the optimal amino acid frequency distribution with the least error would then evolve in preference to all other possible codes. This would be because a mutation in such a code would have a higher probability of producing a desirable replacement than an undesirable one. Gradually, then, the optimal code would be matched to the optimal amino acid frequency distribution. But this hypothesis requires that different codes should have been tried and rejected, and for reasons given above this seems unlikely.

The alternative is that the present-day distribution of amino acid frequencies reflects, more or less passively, present-day code, even though that code is largely arbitrary [9]. The frequency of each amino acid is proportional to the probability of its emergence through mutation.

If this is the case, however, then it must not be correct to assume that every position in every protein is occupied by the best out of twenty possible amino acids. On the contrary, one would have to assume a considerable degree of arbitrariness in protein composition. On other levels, the morphology and physiology of living organisms appear to be almost perfectly adapted – and these adaptations are mediated through proteins and other biological molecules. Can there then really be much latitude on the molecular levels?

I think that the answer to this question is that there need not be much latitude for arbitrary change in a given protein at a given time in order for random mutation to be an important and pervasive molding force over evolutionary time.

It is unlikely that a given protein is perfectly adapted to its function. Evolution is trial and error and successive approximations. There are more theoretically possible proteins of, say, 100 amino acids long than there are particles in the universe, and only an infinitesimal fraction have been tested. New proteins are derived from old by mutation, and *of course* mutation is going to be an important factor in evolution. This is true even if all evolutionary changes are due to adaptive response to selective pressures, and it is certainly true if a significant proportion of evolutionary changes is due to the pressure of mutation itself [5, 9].

How might a protein-coding gene change in evolutionary time? We do

know that most evolutionary changes in protein-coding genes are small single steps, the replacement of one DNA base pair by another. Other changes, such as the duplication of whole genes or parts of genes, minor deletions of one or a few codons, have occurred from time to time in the evolution of proteins, but single-base changes are by far the numerically most common kind of evolutionary change.

Darwinian evolution

We can be sure that Darwinian evolution – positive selection for beneficial mutations – must have occurred, because of the functional efficiency that we can observe. Furthermore, Darwinian evolution is unquestionably still going on, although not necessarily in every structural gene at any given time.

A structural gene coding for a polypeptide of 141 amino acid positions – such as the alpha chain of mammalian hemoglobin – is subject to 1269 possible single-base changes. About a quarter of these would not bring about any change in the primary structure of the polypeptide, because of the degeneracy of the code (it is conceivable that these synonymous changes might have some effect on the rate of translation of the messenger RNA, but this is pure speculation). There are about 1000 different polypeptides that could be derived from a given hemoglobin beta chain gene by single-base substitution.

Among these 1000 possibilities there may be some polypeptides that would confer upon the organism some slight competitive advantage. If one of these superior genes should occur by mutation, there is a reasonably good chance that it would survive and eventually replace the original allele. Then there would be a new and different spectrum of 1000 or so unique polypeptides available by single-step mutation. Some sites which could not previously undergo favorable changes before might now be open for change and *vice versa*. For instance, an allele change to a slightly more electropositive form might foreclose the possibility of another such change, but open up the possibility of changes in the opposite direction at any of a number of codons.

Eventually a gene might reach a point at which none of the alleles available by single-step mutation would be selectively superior. This point would constitute a 'local adaptive peak'. It would not necessarily code for the best of all possible polypeptides, but there would be no further progressive evolution possible so long as the situation remained as described. Of course, a change in the environment might change the relative selective values of the allele and some of its single-step alternatives, but there is no guarantee that this will always happen [11].