

**MOLECULAR BIOLOGY  
INTELLIGENCE  
UNIT**

# The Genetic Code and the Origin of Life

Lluís Ribas de Pouplana, Ph.D.

The Scripps Research Institute  
La Jolla, California, U.S.A.

and

ICREA and Barcelona Institute for Biomedical Research  
Barcelona Science Park, Barcelona, Spain

LANDES BIOSCIENCE / EUREKAH.COM  
GEORGETOWN, TEXAS  
U.S.A.

KLUWER ACADEMIC / PLENUM PUBLISHERS  
NEW YORK, NEW YORK  
U.S.A.

# THE GENETIC CODE AND THE ORIGIN OF LIFE

## Molecular Biology Intelligence Unit

Landes Bioscience / Eurekah.com  
Kluwer Academic / Plenum Publishers

Copyright ©2004 Eurekah.com and Kluwer Academic / Plenum Publishers  
All rights reserved.

No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Printed in the U.S.A.

Kluwer Academic / Plenum Publishers, 233 Spring Street, New York, New York, U.S.A. 10013  
<http://www.wkap.nl/>

Please address all inquiries to the Publishers:

Eurekah.com / Landes Bioscience, 810 South Church Street, Georgetown, Texas, U.S.A. 78626

Phone: 512/ 863 7762; FAX: 512/ 863 0081

<http://www.eurekah.com>

<http://www.landesbioscience.com>

ISBN 0-306-47843-9

The Genetic Code and the Origin of Life edited by Lluís Ribas de Pouplana. Landes / Kluwer dual imprint. Landes series: Molecular Biology Intelligence Unit.

While the authors, editors and publisher believe that drug selection and dosage and the specifications and usage of equipment and devices, as set forth in this book, are in accord with current recommendations and practice at the time of publication, they make no warranty, expressed or implied, with respect to material described in this book. In view of the ongoing research, equipment development, changes in governmental regulations and the rapid accumulation of information relating to the biomedical sciences, the reader is urged to carefully review and evaluate the information provided herein.

### Library of Congress Cataloging-in-Publication Data

The genetic code and the origin of life / [edited by] Lluís Ribas de Pouplana.

p. ; cm. -- (Molecular biology intelligence unit)

Includes bibliographical references and index.

ISBN 0-306-47843-9

1. Evolutionary genetics. 2. Genetic code. I. Ribas de Pouplana, Lluís. II. Series: Molecular biology intelligence unit (Unnumbered)

[DNLM: 1. Genetic Code. 2. Evolution, Molecular. QH 450.2 G3275 2004]

QH390.G445 2004

572.8'633--dc22

2004015694

To Berta, Bernat, and Irene.  
The future of the origin.

# CONTENTS

Foreword .....	xi
Preface .....	xv
<b>1. The Early Earth .....</b>	<b>1</b>
<i>Oliver Botta and Jeffrey L. Bada</i>	
Raw Materials .....	1
Formation of the Solar System .....	2
Formation of the Earth .....	3
The Early Atmosphere, Ocean and Climate .....	4
Organic Compounds on the Early Earth .....	6
Sources of Prebiotic Organic Compounds .....	6
Prebiotic Organic Compounds from Space .....	9
The Prebiotic Soup and the First Living Entities .....	9
Life as We Do Not Know It: Nonheterotrophic Hypotheses for the Origin of Life .....	11
<b>2. Reconstructing the Universal Tree of Life .....</b>	<b>15</b>
<i>James R. Brown</i>	
Topology of the Universal Tree .....	16
Uprooting the Universal Tree .....	19
Genomes and HGT .....	19
Possible HGT Patterns and Processes .....	22
Universal Trees Based on Multiple Datasets .....	23
<b>3. The Nature of the Last Common Ancestor .....</b>	<b>34</b>
<i>Luis Delaye, Arturo Becerra and Antonio Lazcano</i>	
Universal Phylogenies and the Search for the Cenancestor .....	35
Progenote Swarms or Prokaryote-Like Cenancestors .....	37
The Nature of the Cenancestral Genome: DNA or RNA .....	38
Some Like It Very, Very Hot .....	41
Trimming the rRNA-Based Universal Trees .....	42
<b>4. Ribozyme-Catalyzed Genetics .....</b>	<b>48</b>
<i>Donald H. Burke</i>	
Two RNA World Views .....	48
RNA-Catalyzed Genetics I: Nucleotide Polymerization .....	49
RNA-Catalyzed Genetics II: Protein Synthesis .....	55
Towards an RNA-Catalyzed Metabolism: What's Missing? .....	64
<b>5. The Scope of Selection .....</b>	<b>75</b>
<i>Michael Yarus and Rob D. Knight</i>	
Calculations .....	76
Results .....	76
Summing Up .....	79
Appendix .....	84

<b>6. The Evolutionary History of the Translation Machinery .....</b>	<b>92</b>
<i>George E. Fox and Ashwinikumar K. Naik</i>	
Translation and the Origin of Life .....	92
Origins of Translation: What Can We Hope to Learn in the Near Future? .....	93
Timing Information .....	94
Insights to Ribosomal History from tRNA Structure .....	95
Individual Protein History .....	96
Ribosomal Protein S1 .....	96
Ribosome Subunit Evolution: Does Assembly Recapitulate History? .....	97
Order of Events Model for the Development of the Translation Machinery .....	100
Implications and Future Work .....	102
 <b>7. Functional Evolution of Ribosomes .....</b>	 <b>106</b>
<i>Carlos Briones and Ricardo Amils</i>	
Historical Perspective .....	106
Ribosomes and Translation .....	107
Ribosomal RNA .....	107
In Vitro Reconstitution of Ribosomes .....	108
Functional Ribosomal Neighborhoods .....	108
Protein Synthesis Inhibitors as Functional Markers .....	109
Evolutionary Clocks and Molecular Phylogeny .....	109
Functional Phylogeny of Ribosomes .....	110
Functional Analysis of Archaeal Ribosomes .....	111
Phylogenetic Value of Ribosomal Functional Analysis .....	111
Phylogenetic Bases of Ribosomal Functiotype .....	113
 <b>8. Aminoacyl-tRNA Synthetases as Clues to Establishment of the Genetic Code .....</b>	 <b>119</b>
<i>Lluís Ribas de Pouplana and Paul Schimmel</i>	
The Two Classes of Aminoacyl-tRNA Synthetases .....	120
Evolution of Aminoacyl-tRNA Synthetases from Phylogenetic Studies .....	122
Pairs of Subclasses and Their Significance .....	124
Further Support for the ARS-Pair Theory from the Editing Domains .....	125
A Model for the Emergence of Extant ARS and Establishment of the Genetic Code .....	127
The ARS-Pairs in the Context of Theories About the Origin of the Genetic Code .....	128

<b>9. The Relation between Function, Structure and Evolution of Elongation Factors Tu .....</b>	<b>134</b>
<i>Mathias Sprinzl</i>	
Functions of Elongation Factor Tu .....	134
The Functional Cycle of EF-Tu .....	135
The Structure of EF-Tu .....	137
Mechanism of GTPase Activation .....	138
Interaction of EF-Tu GTP with Aminoacyl-tRNA .....	139
Evolution of EF-Tu and tRNA .....	141
<b>10. Origin and Evolution of DNA and DNA Replication Machineries ....</b>	<b>145</b>
<i>Patrick Forterre, Jonathan Filé and Hannu Myllykallio</i>	
Origin of DNA .....	146
Origin and Evolution of DNA Replication Mechanism .....	153
Evolution of Specific Mechanisms Associated to Cellular DNA Replication: Two Case Studies .....	162
<b>11. Early Evolution of DNA Repair Mechanisms .....</b>	<b>169</b>
<i>Jocelyne DiRuggiero and Frank T. Robb</i>	
Life in Extremis .....	169
Experimental Evidence of DNA Repair Mechanisms .....	171
Molecular Mechanisms .....	171
Did Basic DNA Repair Mechanisms Evolve More Than Once .....	177
<b>12. Extant Variations in the Genetic Code .....</b>	<b>183</b>
<i>Manuel A.S. Santos and Mick F. Tuite</i>	
Mechanisms of Codon Reassignment .....	186
The Selective Forces Driving Evolution of Alternative Genetic Codes .....	188
Structural Alterations in the Translation Machinery Are Required for Codon Reassignment .....	191
Future Prospects and Implications for Functional Genomics .....	196
<b>13. Adaptive Evolution of the Genetic Code .....</b>	<b>201</b>
<i>Rob D. Knight, Stephen J. Freeland and Laura F. Landweber</i>	
Framing the Questions .....	201
Is the Choice of Coding Components Optimal? .....	201
An Optimal Pattern of Code Degeneracy? .....	204
An Optimal Pattern of Codon Assignments? .....	207

<b>14. Expanding the Genetic Code in Vitro and in Vivo .....</b>	<b>221</b>
<i>Thomas J. Magliery and David R. Liu</i>	
New Codes in Vivo .....	230
Alternate Substrates for Aminoacyl-tRNA Synthetases .....	231
Selections and Screens for Altered Amino Acid Specificity of aaRSs .....	238
<b>Index .....</b>	<b>251</b>

# EDITOR

**Lluís Ribas de Pouplana, Ph.D.**

The Scripps Research Institute

La Jolla, California, U.S.A.

and

ICREA and Barcelona Institute for Biomedical Research

Barcelona Science Park

Barcelona, Spain

*Chapter 8*

# CONTRIBUTORS

Ricardo Amils

Centro de Biología Molecular

UAM-CSIC

Universidad Autónoma de Madrid,

Cantoblanco

Madrid, Spain

*Chapter 7*

Jeffrey L. Bada

Scripps Institution of Oceanography

University of California at San Diego

San Diego, California, U.S.A.

*Chapter 1*

Arturo Becerra

UNAM

Cd. Universitaria

México, D.F., México

*Chapter 3*

Oliver Botta

Scripps Institution of Oceanography

University of California at San Diego

San Diego, California, U.S.A.

*Chapter 1*

Carlos Briones

Centro de Astrobiología

Torrejón de Ardoz

Madrid, Spain

*Chapter 7*

James R. Brown

Bioinformatics Division

GlaxoSmithKline

Collegeville, Pennsylvania, U.S.A.

*Chapter 2*

Donald H. Burke

Indiana University

Bloomington, Indiana, U.S.A.

*Chapter 4*

Luis Delage

UNAM

Cd. Universitaria

México, D.F., México

*Chapter 3*

Jocelyne DiRuggiero

Department of Cell Biology

and Molecular Genetics

University of Maryland

College Park, Maryland, U.S.A.

*Chapter 11*

Jonathan Filée

Institut de Génétique et Microbiologie

Université Paris-Sud XI

Orsay, Cedex, France

*Chapter 10*

Patrick Forterre

Institut de Génétique et Microbiologie

Université Paris-Sud XI

Orsay, Cedex, France

*Chapter 10*

George E. Fox

Department of Biology and Biochemistry

University of Houston

Houston, Texas, U.S.A.

*Chapter 6*



Stephen J. Freeland  
Department of Ecology  
and Evolutionary Biology  
Princeton University  
Princeton, New Jersey, U.S.A.  
*Chapter 13*

Rob D. Knight  
Department of Molecular, Cellular  
and Developmental Biology  
University of Colorado  
Boulder, Colorado, U.S.A.  
*Chapter 5, 13*

Laura F. Landweber  
Department of Ecology  
and Evolutionary Biology  
Princeton University  
Princeton, New Jersey, U.S.A.  
*Chapter 13*

Antonio Lazcano  
UNAM  
Cd. Universitaria  
México, D.F., México  
*Chapter 3*

David R. Liu  
Department of Chemistry  
Harvard University  
Cambridge, Massachusetts, U.S.A.  
*Chapter 14*

Thomas J. Magliery  
Yale University  
Molecular Biophysics and Biochemistry  
New Haven, Connecticut, U.S.A.  
*Chapter 14*

Hannu Myllykallio  
Institut de Génétique et Microbiologie  
Université Paris-Sud XI  
Orsay, Cedex, France  
*Chapter 10*

Ashwinikumar K. Naik  
Department of Biology and Biochemistry  
University of Houston  
Houston, Texas, U.S.A.  
*Chapter 6*

Frank T. Robb  
Center of Marine Biotechnology  
University of Maryland Biotechnology  
Institute  
Baltimore, Maryland, U.S.A.  
*Chapter 11*

Manuel A.S. Santos  
Department of Biology  
University of Aveiro  
Aveiro, Portugal  
*Chapter 12*

Paul Schimmel  
The Scripps Institute  
La Jolla, California, U.S.A.  
*Chapter 8*

Mathias Sprinzl  
Laboratorium für Biochemie  
Universität Bayreuth  
Bayreuth, Germany  
*Chapter 9*

Mick F. Tuite  
Department of Biosciences  
University of Kent  
Kent, United Kingdom  
*Chapter 12*

Michael Yarus  
Department of Molecular, Cellular  
and Developmental Biology  
University of Colorado  
Boulder, Colorado, U.S.A.  
*Chapter 5*

---

---

# FOREWORD

---

---

## Early Thoughts on RNA and the Origin of Life

The full impact of the essential role of the nucleic acids in biological systems was forcefully demonstrated by the research community in the 1950s. Although Avery and his collaborators had identified DNA as the genetic material responsible for the transformation of bacteria in 1944, it was not until the early 1950s that the Hershey-Chase experiments provided a more direct demonstration of this role. Finally, the structural DNA double helix proposed by Watson and Crick in 1953 clearly created a structural framework for the role of DNA as both information carrier and as a molecule that could undergo the necessary replication needed for daughter cells.

Research continued by Kornberg and his colleagues in the mid-1950s emphasized the biochemistry and enzymology of DNA replication. At the same time, there was a growing interest in the role of RNA. The 1956 discovery by David Davies and myself showed that polyadenylic acid and polyuridylic acid could form a double-helical RNA molecule but that it differed somewhat from DNA. A large number of experiments were subsequently carried out with synthetic polyribonucleotides which illustrated that RNA could form even more complicated helical structures in which the specificity of hydrogen bonding was the key element in determining the molecular conformation. Finally, in 1960, I could show that it was possible to make a hybrid helix. The RNA molecule polyadenylic acid and the DNA molecule polydeoxythymidylic acid could form a double helix, even though it was known that the conformation of the DNA backbone and the RNA backbone were different from each other. This suggested a molecular model for the production of an RNA strand on a DNA strand. Work by Hurwitz, Stevens and Weiss on purification of RNA polymerase eventually led to identification of the system in which DNA-dependent RNA could be synthesized. This set the stage early in 1961 for experiments by several groups that suggested that a rapidly-turning-over RNA molecule called messenger RNA was active in directing the synthesis of proteins on ribosomes, a result that was confirmed by experiments of Nirenberg and associates which showed that the RNA molecule polyuridylic acid directed the synthesis of polyphenylalanine on ribosomes.

In the early 1960s I was pondering the impact of all of this newly acquired knowledge on problems of evolution and the origin of life. The most popular ideas about the origin of life in the 1940s was that espoused by the Russian scientist Oparin, who expressed the belief that life began in coaccervates of polypeptide chains that formed specialized environments leading to the production of enzymatic activity and eventually to living systems. This was a view that I felt was likely to be incorrect since it did not explain the fundamental role of the nucleic acids in providing the information needed for specifying biological systems.

When the chemist and spectroscopist Michael Kasha was a Visiting Professor at Harvard for the 1960-1961 term, we spent some time doing experiments together. It was toward the end of his stay that he approached me to ask if I would contribute a chapter to a book that he and Bernard Pullman were editing as a Festschrift for Albert Szent-Gyorgyi. This book, with the title *Horizons in Biochemistry* (Academic Press, New York, 1962), represented an opportunity for me to put down a number of thoughts that I had about the origin of living systems. I wrote an article with the title "On the Problems of Evolution and Biochemical Information Transfer" (pp. 103-126). It presented a brief overview of the way information was transmitted from DNA to RNA and eventually to directing the synthesis of proteins through the interaction of transfer RNA molecules with messenger RNA in ribosomes. In this essay, I stressed the fact that life could not have originated with protein molecules since it could not explain how nucleic acids came to control protein synthesis. I stressed that it was more likely that polynucleotides were the origin of living systems. I postulated a primitive environment in which polynucleotide chains are able to act as a template or as a somewhat inefficient catalyst for promoting the polymerization of the complementary nucleotide residues to build up an initial two-stranded molecule. Such an inefficient system could be followed by denaturation of the nucleic acid duplex and continuation of the process which would ultimately lead to an increasingly larger number of nucleic acid polymers. I then outlined various ways in which the polymerization of nucleic acids might be coupled with an inefficient polymerization of amino acids. Of key importance in this view was the development of primitive activating enzymes that would begin to relate a specific nucleic acid sequence to the assembly of specific amino acids. Thus, "life" was viewed as starting with a coupling of nucleic acid polymerization and amino acid polymerization although it was stressed that the prototype of this reaction may have been in a form quite different from that which we observe today.

In another section I asked the question, why are there two nucleic acids in contemporary biological systems? It seemed reasonable to believe that both RNA and DNA stemmed from a common precursor. However, it was apparent in contemporary biological systems that DNA seems to act as a major carrier of genetic information, while the RNA molecule is used to convert the genetic information into actual protein molecules. However, I noted that RNA molecules are also able to carry genetic information as in RNA-containing viruses. Thus, it seemed reasonable to speculate that the first polynucleotide molecule was initially an RNA polymer that was able to convey genetic information as well as organize amino acids into specific sequences to make proteins.

This article, published in 1962, was probably the first statement to suggest that RNA was the fundamental nucleic acid involved in the origin of living systems.

It is interesting to note that, in this same essay, I discussed the method by which the newly discovered messenger RNA was made. I suggested the possibility that messenger RNA may be made in vivo as complementary copies of one or both strands of DNA. If both strands are active, then the DNA would produce two RNA strands, and only one of these might be active as messenger RNA in protein synthesis. The other strand, I speculated, might be a component of a control or regulatory system. This is probably the first statement of an anti-sense function for RNA molecules. It also suggests the possibility that RNA could have other regulatory functions.

These statements were published over 40 years ago. Today we have a wealth of information that strengthens the role of RNA in the early evolution of life. The discovery of ribozymes by Cech and the more recent discovery of micro-RNAs that have a variety of functions in controlling the development of biological systems suggests that these may be trace evidence of what has been called the “RNA world”, meaning an era in early evolution in which RNA played a dominant role in both replication and in carrying out a number of chemical modifications leading to the organization of present-day biological systems.

Given our present much more extensive knowledge base concerning the role of nucleic acids in biological systems, it leaves open the issue of when “life” actually began. Preceding the arrival of RNA, there must have been an enormous complexity of chemical reactions. The recent research by Eschenmoser and colleagues points out the possible participation of the four carbon threose nucleic acid polymers as precursors of present-day RNA molecules. It is likely that we will not be able to define a precise event that led to the origin of life. Rather, we are likely to view a growing level of molecular complexity that eventually yields a system that we would call “living” but for which it would be very hard to define a unique point at which we can say that life began. Of course, a key element is the extent to which all of these processes in early evolutionary history were error prone. These errors provided the substrate for Darwinian selection since, among the errors in the system, some will create efficiencies that form the basis for selection that eventually provides the direction for molecular evolution.

*Alexander Rich  
Department of Biology  
Massachusetts Institute of Technology  
Cambridge, Massachusetts, U.S.A.*

---

---

## PREFACE

---

---

It has been said that everybody thinks that he understands evolution but nobody really does. A key word in the previous phrase is 'everybody'. Unlike many other disciplines of biology that are rarely pondered upon by laymen, interest in evolution far precedes the application of scientific method to the problem. The question 'where do we come from?' is the cradle of all forms of pantheism and continues to be one of the central questions of modern biology. The complexity of the problem is such that pessimistic predictions abound on our chances of ever understanding the origin of life. And yet, we are making extraordinary progress.

The field advances mainly through a punctuated cycle of large-scale theoretical predictions and breakthroughs followed by the slow gathering of experimental evidence. Among the most spectacular successes of the field one must surely count the prediction and exploration of the role of RNA, which has provided a satisfactory background to the origin of modern species and a crucial insight into extant cell metabolism. On the other hand, we are still far from understanding the transition from a chemical world to this RNA world that gave rise to life on earth.

This book deals mainly with the processes that led from an established RNA world to the modern rule of the genetic code. The first three chapters of the book describe general concepts regarding the origin of life. These are followed by two reviews of theoretical and experimental studies of the RNA world that illustrate our current understanding of this crucial phase of life evolution. Finally, the biochemistry and the evolution of the central components of the modern genetic code are dissected individually. The goal of this structure is to provide the reader with relevant and detailed information on each of these aspects of the evolution of the code within a temporal perspective of the general evolution of life.

I need to acknowledge the generosity of the authors that have contributed to this volume. They have written excellent reviews and have endured my demanding messages with considerable patience. Many people have read and commented chapters of the book or its general organization. I need to thank especially Ricard Amils, Gustaf Arrenhius, Hugues Bedouelle, Kirk Beebe, Oliver Botta, Jim Brown, Stephen Cusack, Patrick Forterre, Magali Frugier, Rezha Gadhiri, Jerry Joyce, Roshan Kumar, Antonio Lazcano, Leslie Orgel, John Reader, Julius Rebek, Manuel Santos, Paul Schimmel, Bill Waas, Malcolm White, and Xianglei Yang.

Finally, my warmest thanks to Landes Bioscience, Ron Landes, and Cynthia Conomos, for their efforts during the long process of publication of this volume.

*Lluís Ribas de Pouplana  
Barcelona 2004*

# CHAPTER 1

---

## The Early Earth

Oliver Botta and Jeffrey L. Bada

### Introduction

The Earth is so far the only place in the Universe where life is known to exist. Is the Earth special, or are there other places both in our own solar system and beyond where life may have originated and either became extinct or still exists today? Hopefully, in the not to distant future we may find out. During the coming decades, spacecraft will search for evidence of life on Mars and Jupiter's moon Europa, which are considered to be the most promising places for the existence of extant or extinct extraterrestrial life within our solar system. Using remote sensing techniques, we will also begin to look for signs of life's chemistry on the extrasolar planets, which seem to be omnipresent companions of many main sequence stars. If the conditions that resulted in the origin of life on Earth are common throughout the Universe, it seems almost certain that life must exist elsewhere. However, to evaluate whether the Earth is a unique place, or simply an average rocky planet around an average star, we must access what the Earth was like before life began and how these conditions contributed to the processes thought to be involved in the origin of life.

### Raw Materials

The 'Big Bang' produced all the hydrogen now in the Universe, as well as about one third of the helium. In the early Universe there were no elements heavier than these simple light elements. Thus any solar systems that might have formed during those early distant times must have been made up primarily of gas giant planets. This implies that life, at least with respect to carbon based life as we know it, would not have arisen during the Universe's early existence.

And yet, about 10 billion years (Gyr) after the "Big Bang", the nebula that produced our own solar system had a complete inventory of the chemical elements, ranging from the light to the very heavy. Rocky planets could form from these materials, and carbon, oxygen, nitrogen, iron, sulfur and other trace elements that make life possible were present in abundance. Where did these other elements come from?

The main reaction by which stars like our Sun obtain their energy is the fusion of four hydrogen atoms into a  ${}^4\text{He}$  nucleus, a process called 'hydrogen burning'. As stars age, their hydrogen is eventually consumed and exhausted and the star then swells into a bloated object called a Red Giant. It is in Red Giant stars that the next stage of energy and element production begins by a process called "helium burning" in which carbon is the principal product. Carbon production involves first the fusion of two  ${}^4\text{He}$  atoms to produce an unstable  ${}^8\text{Be}$  nucleus, which in the interiors of hot helium burning stars exists long enough to fuse with another  ${}^4\text{He}$  atom, yielding a  ${}^{12}\text{C}$  nucleus. Fusion of another  ${}^4\text{He}$  atom with a  ${}^{12}\text{C}$  atom yields  ${}^{16}\text{O}$ . All carbon and oxygen that are present in all the organic compounds on Earth and in the Universe were once part of Red Giant stars.

Essentially all of the remaining chemical elements, including the very heaviest ones, are synthesized during the death of a star. In fact, most of the heavy elements are created during

great cosmic explosions, called 'supernovae', that are the inevitable fate of aged giant stars that are at least 5-10 times more massive than our Sun. While much of this process used to be merely informed speculation on the part of astronomers, it has now actually been seen taking place. A supernova in a satellite galaxy of our own Milky Way was observed in great detail in 1987 using state-of-the-art astrophysical methods, and these observations confirmed that heavy elements had been made in abundance during the explosion.<sup>1</sup>

As stars age and die, their elemental waste products are strewn throughout the Universe and become incorporated into accumulating interstellar clouds. Parts of these clouds eventually collapse, and contract to form denser clumps of material, which contain dense cores, the localized sites of star formation within the cloud. Further infall of material leads to the formation of protostars, and finally young stars. These star forming regions are characterized by the presence of one or several young stars that have a very high flux in the ultraviolet part of the electromagnetic spectrum. Some new stars in the well-known Orion nebula observed with the Hubble Space Telescope showed dark disks of material around them that are thought to be the initial stages of forming planetary systems.<sup>2</sup>

In the recent years, many extrasolar planets and collapsing dust clouds have now been found that estimates of how many planets—and how many Earth-like planets—there might be in the Milky Way and in the Universe at large are continuously revised. It now seems that a large fraction of stars give rise to planetary systems some which probably have Earth-like planets. It thus appears that there are many potential places in the Universe where life could have arisen.

But, with respect to the origin of life on Earth, we still need to know some important things. How long ago did the Earth form? What was the juvenile Earth and solar system like? What does the early rock record on the Earth tell us about this ancient period of Earth history?

## **Formation of the Solar System**

The formation of solar systems such as ours is thought to take place in a surprisingly short period of time. Within 10 to 20 Myr after the initial collapse of the interstellar cloud, planetesimals started to form in the inner region of the accretionary disk of dust and gas that surrounded our young star. Accretionary disks around stars, similar to the one that surrounded our young Sun, have been observed recently with the Hubble Space Telescope.<sup>3</sup> Astronomical observations of stars in their early evolutionary stage have shown that the lifetime of the accretionary disks is  $\sim 10^7$  years,<sup>4,5</sup> in agreement with the estimates from our own solar system.

Generally, the current scenario for the formation of the planets and smaller objects in our solar system suggests that within the residual accretionary disk small bodies and dust particles began to stochastically accumulate to form larger and larger planetesimals. The orbits of these planetesimals, of which thousands had formed, were not circular, but eccentric, leading to gravitational interactions and collisions. As these interactions continued, planetary embryos were formed by low-velocity collisions, and due to their increased gravitational pull, these began to be the dominant bodies within the disk. Computer simulations of these dynamical processes show that in about 50% of the cases numerous small rocky planets formed near the star while larger gaseous planets formed further out.<sup>5</sup> Although there were originally many more rocky planets in the inner part of the solar system, most of them were either pulled into the Sun or ejected out of the solar system, leaving only the four present day terrestrial planets, Mercury, Venus, Earth and Mars, in the inner solar system.

A significant number of extrasolar systems have been detected in the last several years (for a review see ref. 7). However, these systems appear to be very unlike our solar system, with large Jupiter-sized planets in very close orbits around the parent stars, implying that there may be completely different solar system formation mechanisms. The validity of the formation hypothesis for our own system, and therefore the formation of terrestrial planets that can harbor life, is still based on only our own case. Future planned space missions such as the NASA Terrestrial Planet Finder (TPF) or the ESA Eddington and Darwin spacecraft will be designed to detect Earth-like planets around other stars and help us to learn more about the abundance and formation of extrasolar planets, and perhaps their habitability

## Formation of the Earth

We know today from the uranium/lead dating method of meteorites that the solar system, and by inference the Earth, is  $4.55 \pm 0.11$  Gyr old. Our planet probably formed from planetesimals of material that has a composition similar to that of meteorites known as ordinary chondrites. This primitive class of meteorites contains less volatiles than other solar system objects such as carbonaceous chondrites and comets and are probably representative of the material that formed in region around 1 Astronomical Unit (AU), the distance of the Earth to the Sun.

Other types of meteorites have provided additional information about the formation of our planet. The meteorite types achondrites, stony-iron and iron meteorites are products of melting and differentiation that took place on their parent bodies and these meteorites can therefore provide constraints on the timescales over which planetesimal accretion and subsequent differentiation took place.<sup>8</sup> Recent results indicate that these processes occurred fast on meteorite parent bodies, but the inference to the large planets such as the Earth or Mars is not straightforward, mainly due to the controversy over whether the core formation in these planets was a single global melting event or if the accreting planetesimals were already differentiated. Other work suggests that core formation in the Earth took place  $< 80$  Myr after differentiation of the planetesimals in which iron meteorites formed.<sup>9,10</sup> Based on Hf-W (hafnium-tungsten) systematics of SNC meteorites, which are widely believed to be samples of the Martian crust, it was concluded that core formation on Mars took place within the first  $\sim 30$  Myr of solar system history.<sup>11</sup>

The Earth is the only terrestrial planet with a large moon. The current favored scenario for the formation of the moon involves a collision, around 50 Myr after the formation of the inner planets, between the Earth and a planetesimal about the size of Mars.<sup>12,13</sup> Some of the debris from that collision went into a close orbit, probably around 25,000 km, around the Earth and eventually aggregated to form the moon. Due to this close distance, and to the much higher spin rate of the Earth at that time, tidal forces were three hundred times stronger than today, greatly deforming the freshly formed crust, and if global oceans existed, causing oceanic tides much higher than today. These tides might have played a significant role in the formation of life on Earth, as we will discuss later. If this collision theory for the origin of the moon is indeed correct, it also provides indirect support to the hypothesis that there were probably many smaller planetary bodies that formed in the inner solar system other than the four that are present today.

Even after the moon formed, the Earth was still bombarded by bolides 10s of km in diameter at a frequency of about one collision every 1000 years or so. These impacts, as well as the decay of radioactive elements in the Earth's interior, caused the planet to stay in a molten state for a few million years or less after its initial accretion.<sup>14</sup> During this so-called "Hadean" period, the temperatures of the Earth were so high that the heavier elements sank towards the center of the planetary body forming the core, leaving the lighter elements to form the mantle and crust. As already mentioned, the timing of this differentiation is absolutely crucial for the origin of life, because it defined the oxidation state of the crust and consequently the composition of the atmosphere. The formation of the core itself is important for other reasons as well. On one hand, it depleted the crust and the mantle of the heavy elements and made carbon one of the most abundant elements in these regions. On the other hand, a metallic core is a prerequisite to the formation of a strong magnetic field around the planet. A magnetic field acts like a protection shield against the harsh particle radiation that is present everywhere in the solar system and the galaxy.

The oldest rocks that have been dated on the Earth are those found in the Acasta Gneiss complex in Canada, which were formed over 4 Gyr ago.<sup>15</sup> However, zircon crystals that are up to  $\sim 4.4$  Gyr old have been extracted out of younger rocks from Australia.<sup>16</sup> The chemical and isotopic data obtained from one of the zircon crystals indicates the presence of large bodies of water, perhaps even oceans, as early as 4.3 to 4.4 Gyr, nearly a billion years earlier than the earliest evidence for cellular life on Earth.<sup>16,17</sup>



## The Early Atmosphere, Ocean and Climate

During the Hadean period, the volatile compounds that were trapped inside the accreting planetesimals were released from the molten rock to form a secondary atmosphere. Any primary atmosphere (if one existed at all) must have been lost, as evidenced by the depletion of rare gases in Earth's atmosphere compared to cosmic abundances.<sup>18</sup> As a consequence of the simultaneous formation of Earth's core with accretion, the metallic iron was removed from the upper mantle, which would allow the volcanic gases to remain relatively reduced and produce a very early atmosphere that contained species such as CH<sub>4</sub>, NH<sub>3</sub> and H<sub>2</sub>. Since the temperature at the surface was high enough to prevent any water from condensing, the atmosphere would have consisted mainly of superheated steam along with these other gases.<sup>19</sup> However, this secondary atmosphere may have been lost several times during large impact events such as the one that formed the moon, and would have been replaced by further outgassing from the interior and resupply from later impactors.

It is believed that the impactors during the latter stages of the accretion process have originated from further out in the solar system and would have been comparable in composition to comets (see Fig. 1). Comets, whose volatile compounds are the most pristine materials surviving from the formation of the solar system, may have supplied a substantial fraction of the volatiles on the terrestrial planets, perhaps including organic compounds that may played a role in the origin of life on earth (see below). It has been suggested that the water present currently on the Earth was provided entirely from this source. However, recent measurements of the deuterium enrichment of water in comets Halley, Hyakutake and Hale-Bopp indicate that only a fraction of it was delivered by comets, whereas the largest fraction was trapped during the earlier accretionary phase.<sup>20</sup> The volatiles on comets are more oxidized than the ones in asteroids due to their similarity to interstellar ices and their higher water/rock ratio.

Without going into details on other effects such as ingassing of volatiles and loss of H<sub>2</sub> to space, it can be assumed that the atmosphere that developed on the Earth over the period 4.4 to 3.8 Gyr ago (perhaps several times if was erased by large impact events) was basically a mix of volatiles delivered by volatile rich impactors such as comets and outgassing from the interior of an already differentiated planet. This atmosphere was probably dominated by water steam until the surface temperatures dropped to ~100°C (depending on the pressure), at which point water condensed out to form early oceans.<sup>16</sup> The reduced species, which were mainly supplied by volcanic outgassing, are very sensitive to UV radiation that penetrated through the atmosphere due to the lack of a protective ozone layer. These molecules were probably destroyed by photodissociation, although there might have been steady state equilibrium between these two processes that allowed a significant amount of these reduced species to be present in the atmosphere. Overall, however, the atmosphere was dominated by oxidized species such as CO<sub>2</sub>, CO and N<sub>2</sub>. A similar atmosphere is present on Venus today, although it is much more dense than the atmosphere of the early Earth.

The climate on the early Earth at this stage depended mainly on two factors: the luminosity of the Sun and the radiative properties of the atmosphere. Standard theoretical solar evolution models predict that the Sun was about 30 % less luminous than today.<sup>21</sup> If the atmosphere of the early Earth was the same as it is now, the entire surface of the planet would have been frozen. However, as discussed extensively by Kasting,<sup>18,19</sup> a CO<sub>2</sub> rich atmosphere may have been present throughout the Hadean and Early Archean period and this would have resulted in a significant Greenhouse effect that would have prevented the oceans on the early Earth from freezing. Since there were probably no major continents existing during that period of time, silicate weathering (the long-term loss process on for CO<sub>2</sub> today) would have been low, and CO<sub>2</sub> would have been primarily contained in the atmosphere and ocean. Even with the assumption of a 70% present solar luminosity, a steady-state atmosphere containing ~ 10 bars of CO<sub>2</sub> could have resulted in to a mean surface temperature approaching 100°C.

In summary, the current models for the early terrestrial atmosphere suggest that it consisted of a weakly reducing mixture of CO<sub>2</sub>, N<sub>2</sub>, CO, and H<sub>2</sub>O with lesser amounts of H<sub>2</sub>, SO<sub>2</sub>, and



Figure 1. Comets, such as this one photographed in 1892 by E. R. Barnard (taken from R. S. Ball, "The Story of the Heavens", Cassell and Company LTD, London, Paris, New York, Melbourne, 1900), may have supplied the early Earth with some of the reagents (HCN, aldehydes/ketones, etc.) needed for the abiotic syntheses as well as some of the organic compounds needed for the origin of life.

H<sub>2</sub>S. Reduced gases such as CH<sub>4</sub> and NH<sub>3</sub> are considered to be nearly absent or present only in localized regions near volcanoes or hydrothermal vents.

There is, however, the possibility that the CO<sub>2</sub> concentrations in the early atmosphere were not high enough to prevent the formation of an ice-covered ocean.<sup>22</sup> If this was the case, the thickness of the ice sheet has been estimated to be on the order of 300 m, which would have been thin enough to allow melting by an impactor of ~ 100 km in diameter. The frequency of impacts of such ice-melting bolides has been estimated to be one event every 10<sup>5</sup>-10<sup>7</sup> years between about 3.6 and 4.5 Gyr ago, suggesting that there could have been periodic thaw-freeze cycles associated with the ice-melting impacts. The precursor compounds imported by the impactor or synthesized during the impact, such as HCN, would have been washed into the ocean during the thaw period. In addition, CH<sub>4</sub>, H<sub>2</sub>, CO and NH<sub>3</sub> derived from hydrothermal vents would have been stored in the unfrozen ocean below the ice layer which protected these gases from the ultraviolet radiation<sup>23</sup>. Following a large impact, the trapped gases would have been expelled into the atmosphere where they could have persisted for some time before they were destroyed by photochemical reactions.

## Organic Compounds on the Early Earth

Today, organic compounds are so pervasive on the Earth's surface that it is hard to image the Earth devoid of organic material. However, during the period immediately after the Earth first formed some 4.5 Gyr ago, there would have been no organic compounds present on its surface. This was because soon after accretion, the decay of radioactive elements heated the interior of the young Earth to the melting point of rocks. Volcanic eruptions expelled molten rock and hot scorching gases out of the juvenile Earth's interior creating a global inferno. In addition, the early Earth was also being peppered by mountain-sized planetesimals, the debris left over after the accretion of the planets. Massive volcanic convulsions, coupled with the intense bombardment from space, generated surface temperatures so hot that the Earth at this point could very well have had an "ocean" of molten rock.

Although temperatures would have slowly decreased as the infall of objects from space and the intensity of volcanic eruptions declined, elevated temperatures likely persisted for a few hundred million years after the formation of the Earth.<sup>16</sup> During this period, temperatures would have probably been too hot for organic compounds to survive. Without organic compounds, life as we know could not exist. However, by roughly 4 Gyr ago, and perhaps even a lot earlier, the Earth's surface had cooled to the point that liquid water could exist and global oceans began to form.<sup>16</sup> It was during this period that organic compounds would have first started to accumulate on the Earth's surface, as long as there were natural pathways by which they could be synthesized, or sources from elsewhere that could supply them to the Earth.

The origin of life of life as we know it on Earth required the presence of liquid water and an inventory of prebiotic organic compounds from which could undergo further chemical processing so that life could emerge. Although it appears that liquid water in the form of large oceans was present on the surface of the early Earth,<sup>16,17</sup> the source of the required organic molecules on the primordial Earth is not clear.

## Sources of Prebiotic Organic Compounds

Based on our current knowledge of the last common ancestor to all life today, the first organisms were probably some sort of heterotroph, which means that not only water and energy sources had to be available, but also a minimal set of organic compounds. In contrast, it would have been possible that the first organisms were autotrophic, meaning that they could convert CO<sub>2</sub> directly into the reduced organic molecules they need to live. But heterotrophic organisms are much simpler than autotrophic organisms, which require an elaborate array of protein biosynthesis reactions and enzymes in order to fix CO<sub>2</sub>. Therefore, a heterotrophic origin of life has been widely accepted, mainly based on insights gained over the last fifty years.<sup>23</sup>

Nearly a century ago, the Russian biochemist Alexander Oparin<sup>24</sup> and the British biochemist J. B. S. Haldane<sup>25</sup> proposed a theory that organic compounds could have been synthesized on the early Earth when gases in the atmosphere were subjected to some type of energy. Inspired by these ideas, and with the assumption that the atmosphere of the early Earth was reducing as proposed by his mentor Harold Urey,<sup>26</sup> Stanley Miller was the first to experimentally demonstrate the possible synthesis of organic compounds under prebiotic conditions.<sup>27</sup> He constructed an apparatus in which he could simulate the interaction between an atmosphere and an ocean (Fig. 2). As an energy source, Miller chose a spark discharge, considered to be the second largest energy source, in the form of lightning and coronal discharges, on the early Earth after UV radiation. Miller filled the apparatus with various mixtures of methane, ammonia and hydrogen as well as water, which was then heated during the experiment. A spark discharge between the tungsten electrodes, which simulated lightning and corona discharges in the early atmosphere, was produced using a high frequency tesla coil with a voltage of 60,000 V. The reaction time was usually a week or so and the maximum pressure 1.5 bars. With this experimental setup, Miller was able to transform almost 50% of the original carbon (in the form of methane) into organic compounds. Although almost of the synthesized organic

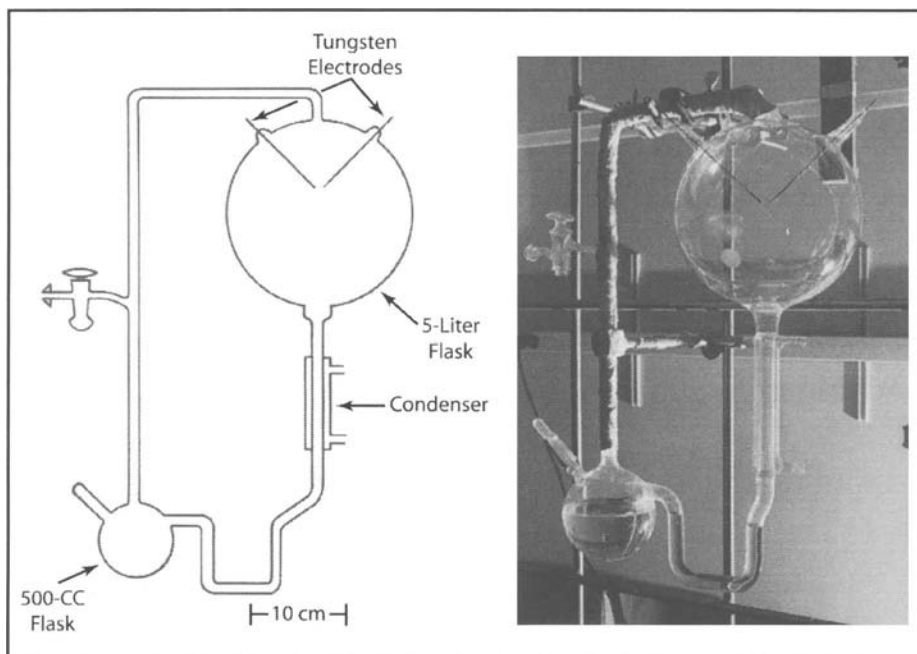


Figure 2. A diagram (right) of the original Miller-Urey apparatus showing the various components. A photograph of the actual apparatus is shown on the left (courtesy of Stanley L. Miller).

material was an insoluble tar-like solid, he was able to isolate amino acids and other simple organic compounds from the reaction mixture. Glycine, the simplest amino acid, was produced in 2% yield (based on the original amount of methane carbon), whereas alanine,<sup>28</sup> the simplest chiral amino acid, showed a yield of 1%. Miller was able to demonstrate that the alanine that was produced was a racemic mixture (equal amounts of D- and L-alanine). This provided convincing evidence that the amino acids were produced in the experiment and were not biological contaminants somehow introduced into the apparatus.

The other organic compounds that Miller was able to identify made it possible for him to propose a possible reaction pathway for the amino acids.<sup>29</sup> The proposed synthetic mechanism had actually been discovered in the mid-19th century by the German chemist A. Strecker.<sup>30</sup> It features the reaction of hydrogen cyanide, ammonia and carbonyl compounds (aldehydes or ketones) to form cyanohydrins, which would then undergo hydrolysis to form the  $\alpha$ -amino acids. Depending on the concentration of ammonia in the reaction mixture, varying amounts of  $\alpha$ -hydroxy acids are produced as well, which is what Miller found, with larger relative amounts of hydroxy acids being formed in a reaction mixture containing less ammonia.

It is important to note that neither purines nor pyrimidines, the nucleobases that are part of DNA and RNA, were investigated in the mixtures of the original Miller-Urey experiment. However, in experiments carried out soon after Miller's experiment by J. Oró and coworkers, the formation of adenine from ammonium cyanide solutions was demonstrated.<sup>31,32</sup> Later, it could be shown that the abiotic synthesis of purines and other heterocyclic compounds also works under the same conditions than the original Miller-Urey experiment, but in much smaller yields than the amino acids.<sup>33</sup> In addition, it has been found that guanine can be produced in a direct "one-pot" synthesis from the polymerization of ammonium cyanide.<sup>34</sup>

However, as should be clear from the earlier discussion, the atmospheric composition that formed the basis of the Miller-Urey experiment is not considered to be plausible by many

researchers. Instead, a weakly reducing or neutral atmosphere is to be more in agreement with the current model for the early Earth. Although Miller and Urey originally rejected the idea of nonreducing conditions for the primitive atmosphere, Miller later carried out experiments with CO and CO<sub>2</sub> model atmospheres.<sup>35,36</sup> He found that not only were the yields of the amino acids reduced, but that glycine was basically the only amino acid synthesized under these conditions. He found a trend, which showed that as the atmosphere became less reducing and more neutral, the yields of synthesized organic compounds decreased drastically. The presence of methane and ammonia appeared to be especially important for the formation of a diverse mixture of amino acids. The main problem in the synthesis of amino acids and other biologically relevant organic compounds with nonreducing atmospheres is the formation of hydrogen cyanide (HCN), which is an intermediate in the Strecker pathway and an important precursor compound for the synthesis of nucleobases.<sup>37</sup> However, as mentioned earlier, localized high concentrations of reduced gases may have existed around volcanic eruptions and in these localized environments reagents such as HCN, aldehydes and ketones may have been produced, which after washing into the oceans could have become involved in the prebiotic synthesis of organic molecules.

Because of supposed problems associated with the direct Miller-Urey type syntheses on the early Earth, a completely different hypothesis for the "home-grown" synthesis of organic compounds has been proposed. The hypothesis is based on one of the great oceanographic discoveries of the 20<sup>th</sup> century. In February 1977, using the submersible research vessel *Alvin*, scientists from the Woods Hole Oceanographic Institution (WHOI) approached one of the volcanically active regions deep off the Galápagos Islands, where water at temperatures of up to 350 °C was known to be spewing out of geothermal vents. Strong gradients of temperature, acidity and chemical composition were found to be present at the vent/seawater interface. The big surprise came when the crew of the *Alvin* detected rich and complex communities of animals around the dark, chimney-like openings of the vents. From the analysis of samples brought back from these vents, it has shown that bacteria which oxidize the H<sub>2</sub>S that spews from the vents form the low end of the food chain for these communities. These miniature deep-sea ecosystems are, however, not completely independent from solar radiation, since the chemosynthetic energy that the bacteria at these modern sites use to oxidize the sulfur-compounds comes from molecular oxygen (O<sub>2</sub>), which was produced by photosynthesis at the ocean surface.

A group of researchers, known colloquially as 'ventists', believe that the remarkable properties of the hydrothermal vent environments, particularly their protection from the harsh conditions caused by large impact events, might have played an important role in the origin of life.<sup>38</sup> Since it is thought by some that the last common ancestral organism of all extant life on Earth was a thermophile, several researchers have proposed the hypothesis that the organic compounds necessary for the origin of life were actually synthesized under vent conditions. For example, Shock and coworkers have calculated that thermodynamic-based equilibria favors the formation of compounds such as amino acids at hydrothermal vent temperatures,<sup>39</sup> especially in vents associated with off-axis systems.<sup>40</sup> However, at elevated temperatures associated with vent discharges, amino acids and other biomolecules have been found to rapidly decompose.<sup>41,42</sup> For example, amino acids are destroyed in time scales of minutes. The rate of hydrolysis for RNA at pH 7 extrapolated to elevated temperatures gives a half-life of 2 min at 250 °C for the hydrolysis of every phosphodiester bond; at 350 °C the half-life is 4 s. For DNA, the half-lives for depurination of each nucleotide at pH 7 are nearly the same as the hydrolysis rates for RNA.<sup>42</sup> It has been pointed out by Lazcano that if the origin of life was sufficiently long, all the complex organic compounds in the ocean, whether derived from home-grown synthesis or from exogenous delivery, would be destroyed by passage through the hydrothermal vents.<sup>43</sup> It thus appears that hydrothermal vents are much more effective in regulating the concentration of critical organic molecules in the oceans rather than playing a significant role in their direct synthesis.

## Prebiotic Organic Compounds from Space

Because of the difficulties discussed in the previous sections with the concept of “home-grown” synthesis of amino acids and nucleobases as a major source of these compounds for the origin of life, other hypotheses about the origin of these compounds were developed. In the early 1990s, Chyba and Sagan proposed that the exogenous delivery of organic matter by asteroids, comets and interplanetary dust particles (IDPs) could have played a significant role in seeding the early Earth with the compounds necessary for the origin of life.<sup>44</sup> They drew this conclusion from the knowledge about the organic composition of meteorites. It is important to note that, if this concept is valid, impacts on the early Earth not only created devastating conditions which made it difficult for life to originate, but at the same time perhaps delivered the raw material necessary for setting the stage for the origin of life. In an even wider view, this hypothesis could have profound implications on the abundance of life in the universe. The origin of the essential organic compounds needed for the origin of life is not constrained by the conditions on a particular planet. But rather organic compound synthesis is a ubiquitous process that takes place on primitive planetary bodies such as asteroids and comets. The possibility for the origin of life is thus considerably increased, provided the essential organic compounds are delivered intact to a planet that is suitable for further chemical evolution.

Carbonaceous chondrites, a class of stony meteorites, are considered to be the most primitive objects in the solar system in terms of their elemental composition, yet they feature a high abundance of carbon, more than 3 weight-% in some cases. The most extensively analyzed meteorites for organic compounds include the CMs Murchison (fell in 1969 in Victoria, Australia) and Murray (1950, Kentucky, USA) and the CI Orgueil (1864, France). The carbon phase is dominated by an insoluble fraction, with the rest being soluble compounds (Table 1). PAHs make up the majority (up to 80%) of the of the soluble organic matter, followed by the carboxylic acids, the fullerenes and amino acids, which are about an order of magnitude less abundant.<sup>45</sup> Other important compounds in context with the origin of life are the nucleobases. The purines adenine, guanine, xanthine and hypoxanthine have been detected as well as the pyrimidine uracil in concentrations of 200 to 500 parts per billion (ppb) in the CM chondrites Murchison and Murray and in the CI chondrite Orgueil.<sup>46-48</sup> In addition, a variety of other nitrogen-heterocyclic compounds including pyridines, quinolines and isoquinolines were also identified in the Murchison meteorite.<sup>49</sup>

More recently, it was found the CI type meteorites such as Orgueil contain a distinct amino acid composition in comparison to the CMs.<sup>50</sup> The simple amino acid mixture, consisting of just glycine and  $\beta$ -alanine, found in CI carbonaceous chondrites is interesting in the sense that it has been generally thought that a wide variety of amino acids were required for the origin of life. However, among the candidates for the first genetic material is peptide nucleic acid (PNA), a nucleic acid analogue in which the backbone does not contain sugar or phosphate moieties.<sup>51,52</sup> For the PNA backbone, achiral amino acids such as glycine and  $\beta$ -alanine, possibly delivered by CI type carbonaceous chondrites to the early Earth, may have been the only amino acids needed for the origin of life.

## The Prebiotic Soup and the First Living Entities

The organic material on the early Earth before life existed, regardless of its source, would have likely consisted of an wide array of different types of compounds, including amino acids, nucleobases, fatty acids, aromatic and heteroaromatic hydrocarbons among others. How these abiotic organic constituents on the prebiotic Earth were assembled into the first living entities is highly contentious. In modern biological systems, these compounds are part of oligomeric or polymeric molecular structures needed for catalysis and replication, so simple abiotic compounds were only the starting points for the chemical evolution that followed. For example, amino acids are the monomeric building blocks for proteins and enzymes, the structural and catalytic units without which life as we know it can not exist. Also, DNA and RNA, the molecules that encode and transcribe the genetic information in all terrestrial organisms, are com-

**Table 1. Abundances of soluble organic compounds found in meteorites. Amino acids concentrations have been determined for several CI and CM chondrites. All other data are for the CM chondrite Murchison (except those for the polycyclic aromatic hydrocarbons and the fullerenes, which are from Yamato-791198 and Allende, respectively). Taken from ref. 45.**

Compound Class		Concentration (ppm)
Amino Acids		
	CM meteorites	17-60
	CI meteorites	~ 5
Aliphatic hydrocarbons		> 35
Aromatic hydrocarbons		3300
Fullerenes		> 100
Carboxylic acids		> 300
Hydroxycarboxylic acids		15
Dicarboxylic acids & Hydroxydicarboxylic acids		14
Purines & Pyrimidines		1.3
Basic N-heterocycles		7
Amines		8
Amides	linear	> 70
	cyclic	> 2
Alcohols		11
Aldehydes & Ketones		27
Sulphonic acids		70
Phosponic acids		2

prised by mononucleotides, which contain nucleobases such as adenine, guanine, thymine, cytosine, and uracil, attached to a sugar-phosphate backbone. The most widely accepted scenario for the transition from abiotic to biotic chemistry is that the simple monomeric compounds present in the prebiotic soup somehow underwent polymerization, perhaps with the assistance of clays and minerals, and formed longer and longer chains or polymers which over time became increasingly more complex with respect to both their structures and properties.<sup>53</sup> Eventually, some of these polymers acquired the capacity to replicate, one of the fundamental and most important properties of living organisms.

What the first molecular self-replicating entity consisted of and how replication was accomplished is not known, but several suggestions have recently been made. The most likely candidates are nucleic acid analogs of DNA and RNA such as PNA. Stanley Miller and coworkers have found that the building blocks of PNA molecules, the nucleobase derivatives adenine- and guanine-N<sup>9</sup>-acetic acid and uracil- and cytosine-N<sup>1</sup>-acetic acid as well as and N-(2-aminoethyl)glycine (AEG), the molecule that makes up the PNA backbone, can be synthesized under likely prebiotic conditions and, under favorable conditions, could have been major constituents of the primitive milieu.<sup>52</sup> Still to be worked out are mechanisms for the polymerization of the monomers, but preliminary results indicate that AEG oligomerizes more efficiently at 100 °C than mixture of  $\alpha$ -amino acids at higher temperatures. However, even with PNA-like genetic informational molecules stability may be a problem. This could be overcome because the stability is highly sequence-dependent and the breakdown may be partly alleviated by blocking or acetylating the N-terminus.<sup>52</sup> However, other possibilities need to be considered because there may be other backbones and bases that were more abundant and more efficient for early biotic replication.

## Life as We Do Not Know It: Nonheterotrophic Hypotheses for the Origin of Life

There is a more radical hypothesis which has been proposed that discards the whole idea of a primordial soup. This ‘metabolic-life’ hypothesis, promoted primarily by Günter Wächtershäuser and coworkers in Germany, claims that life at the time of origin consisted of nothing more than a sequential series of chemical reactions that are catalyzed on mineral surfaces.<sup>54</sup> According to this theory, the first living systems on Earth were based a type of autotrophic metabolism of low-molecular weight constituents such as CO and CO<sub>2</sub> which are converted into biologically relevant compounds such as pyruvate at high temperature (100-250 °C)/high-pressure (0.2 - 200 MPa) vent-type conditions. The metabolism theory claims that life, at least in its beginnings, was nothing more than a continuous chain of mineral surface-associated self-sustaining chemical reactions with no requirement for genetic information. A primitive type of reductive citric acid cycle is often cited as a model. There is some experimental support for the hypothesis, although the conditions for the various individual reaction steps are very different,<sup>55,56</sup> and it remains to be established if the conditions used in these laboratory experiments are geophysically plausible and are therefore relevant to the origin of life. Of the various metabolic reaction schemes that have been proposed and investigated none have been demonstrated to be autocatalytic, nor are there any empirical indications that this indeed is even possible in a prebiotic context.

Whether a set of self-sustaining set of purely chemical reactions really constitutes a system that can be considered alive is debatable. Nevertheless, if self-sustaining reaction chains did arise on the early Earth, they could have played an important role in enriching the prebiotic soup in some molecules that were perhaps not readily synthesized by other abiotic reactions or derived from space. In this context, the metabolism theory can be viewed as simply a component of the prebiotic soup theory. But, regardless of its initial complexity, self-maintaining chemical-based metabolic life could not have evolved in the absence of a genetic replicating mechanism insuring the maintenance, stability, and diversification of its components. In the absence of any hereditary mechanisms, autotrophic reaction chains would have come and gone without leaving any direct descendants able to resurrect the process. Life as we know it consists of both chemistry and information. If metabolic life ever did exist on the early Earth, to convert it to life as we know it would have required the emergence of some type of information system under conditions that are favorable for the survival and maintenance of genetic informational molecules.<sup>57</sup>

Finally, in an even more radical theoretical hypothesis, it has been suggested that the origin of life did not occur on the surface of the Earth, but inside the crust. This “subterranean model” is based on the idea that there exists a biosphere that feeds off abiogenic hydrocarbons (petroleum) formed deep inside the Earth.<sup>58</sup> These supposed subterranean organisms are completely independent of photosynthesis. It is assumed that the habitability of the subsurface is enhanced in comparison to the surface, particularly during the early Archean. Therefore, it is concluded that there is a higher probability for the origin of life to occur in the deep subsurface and life then made its way to the surface when the surface became habitable after the end of the heavy bombardment. There are many fatal flaws in this hypothesis. For one, there is no evidence that abiotic hydrocarbons are formed deep within the Earth. In addition, according to the current view on the formation of the Earth which was discussed earlier, no primordial hydrocarbons would have survived the accretion process of the Earth (or any other planet) since the temperatures reached during accretion were high enough to decompose any reduced carbon compound in a short period of time.



## Conclusions

The early history of the Earth, the first 100 Myr or so, was dominated by the hot accretion of the planet followed by a relatively rapid cooling. There is evidence for the presence of liquid water on the surface at 4.4 Gyr ago. The proto-atmosphere, if it existed at all, was probably reduced, but it was removed from the planet early on by large impact events. The first “real” atmosphere was produced by outgassing from the crust, and was dominated by oxidized gases such as CO<sub>2</sub>, CO and N<sub>2</sub> with lesser amounts of H<sub>2</sub> and CH<sub>4</sub>. Only trace amounts of oxygen were present. In such an atmosphere, organic compounds, and amino acids in particular, with spark discharge as an energy source (“Miller-Urey-type”) are not produced. The influx of extraterrestrial organic compounds delivered by comets, asteroids, meteorites and IDPs may have been a major source of the compounds necessary for a heterotrophic origin of life on Earth. These compounds, perhaps in combination with other chemical products on the early Earth, may have been interacted on mineral surfaces in drying lagoons or other appropriate locations on a turbulent and chaotic surface to form more and more complex compounds. Eventually, through the process of chemical evolution, some of the compounds that developed the capability for catalytic activity and/or information storage (replication) became dominant. This scenario is a logical, but still a highly speculative pathway about the beginning of the first living entities, and there are still big gaps between the various stages that need to be filled. The following chapters address some of these remaining questions, such as the formation of the first nucleic acids and the role of replicating polypeptides in this scenario.

## References

1. For a review, see: McCray R. Supernova 1987A revisited. *Annu Rev Astron Astrophys* 1993; 31:175-216.
2. For pictures and press releases see: <http://opposite.stsci.edu/pubinfo/pictures.html>
3. Schneider G, Smith BA, Becklin EE et al. NICMOS imaging of the HR 4796A circumstellar disk. *Astrophys J* 1999; 513:L127-L130.
4. Storm SE. The early evolution of stars. *Rev Mex Astron Astrofis* 1994; 29:23-29.
5. Natta A, Grinin VP, Mannings V. Properties and evolution of disks around premain-sequence stars of intermediate mass. In: Mannings V, Boss AP, Russell SS, eds. *Protostars and Planets IV*. Tucson, University of Arizona Press, 2000:559-587.
6. Marcy GW. Personal communication.
7. Marcy GW, Cochran WD, Mayor M. Extrasolar planets around main-sequence stars. In: Mannings V, Boss AP, Russell SS, eds. *Protostars and Planets IV*. Tucson, University of Arizona Press, 2000:1285-1311.
8. Wadhwa M, Russell SS. Timescales of accretion and differentiation in the early solar system: The meteoritic record. In: Mannings V, Boss AP, Russell SS, eds. *Protostars and Planets IV*. Tucson, University of Arizona Press, 2000:995-1018.
9. Lee DC, Halliday AN. Hafnium-tungsten chronometry and the timing of terrestrial core formation. *Nature* 1995; 378:771-774.
10. Halliday AN, Rehkämper M, Lee DC et al. Early evolution of the Earth and the Moon: New constraints for Hf-W isotope geochemistry. *Earth Planet Sci Lett* 1996; 142:75-89.
11. Lee DC, Halliday AN. Core formation on Mars and differentiated asteroids. *Nature* 1997; 388:854-857.
12. Hartmann WK, Phillips RJ, Taylor GJ. *Origin of the Moon*. Houston, Lunar and Planetary Institute, 1986.
13. Lee DC, Halliday AN, Snyder GA et al. Age and origin of the Moon. *Science* 1997; 278:1098-1103.
14. Maher KA, Stevenson DJ. Impact frustration of the origin of life. *Nature* 1988;331:612-614.
15. Bowring SA, Williams IS. Priscoan (4.00-4.03 Ga) orthogneisses from northwestern Canada. *Contrib Mineral Petrol* 1999; 134:3-16.
16. Wilde SA, Valley JW, Peck WH et al. Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. *Nature* 2001; 409:175-178.
17. Mojzsis SJ, Arrhenius G, McKeegan KD et al. Evidence for life on Earth before 3,800 million years ago. *Nature* 1996; 384:55-59.
18. Kasting JF. Earth's early atmosphere. *Science* 1993; 259:920-926.

19. Kasting JF. Early evolution of the atmosphere and ocean. In: Greenberg JM, Mendoza-Gomez CX, Pirronello V, eds. *The Chemistry of Life's Origin*. The Netherlands: Kluwer Academic Publishers, 1993:149-176.
20. Meier R, Owen TC, Matthews HE et al. A determination of the HDO/H<sub>2</sub>O ratio in comet C/1995 O1 (Hale-Bopp). *Science* 1998; 279:842-844.
21. Gilliland RL. Solar evolution. *Global Planet Change* 1989; 1:35-55.
22. Bada JL, Bigham C, Miller SL. Impact melting of frozen oceans on the early Earth: Implications for the origin of life. *Proc Natl Acad Sci USA* 1994; 91:1248-1250.
23. Lazcano A, Miller SL. The origin and early evolution of life: Prebiotic chemistry, the preRNA world, and time. *Cell* 1996; 85:793-798.
24. Oparin AI. *Proiskhozhdienie zhiny*. Moscow, 1924, translated as "The Origin of Life". In: Bernal JD, ed. *The Origin of Life*. Cleveland and New York: 1957:Appendix I, 199-234.
25. Haldane JBS. *Rationalist annual* (1929). Reprinted. In: Bernal JB, ed. *The Origin of Life*. Cleveland and New York, 1957:243-249.
26. Urey HC. *The planets*. New Haven: Yale University Press, 1952.
27. Miller SL. A production of amino acids under possible primitive Earth conditions. *Science* 1952; 117:528-529.
28. Miller SL. Production of some organic compounds under possible primitive Earth conditions. *J Am Chem Soc* 1955; 77:2351-2361.
29. Miller SL. The mechanism of synthesis of amino acids by electric discharges. *Biochim Biophys Acta* 1957; 23:480-489.
30. Strecker A. Über die künstliche Bildung der Milchsäure und einem neuen dem Glycocoll homologen Körper. *Ann Chem* 1850; 75:27.
31. Oró J. Synthesis of adenine from ammonium cyanide. *Biochim Biophys Res Commun* 1960; 2:407-412.
32. Oró J, Kimball AP. Synthesis of purines under possible primitive Earth conditions. I. Adenine from hydrogen cyanide. *Arch Biochem Biophys* 1961; 94:217-227.
33. Yuasa S, Flory D, Basile B et al. Abiotic synthesis of purines and other heterocyclic compounds by the action of electrical discharges. *J Mol Evol* 1984; 21:76-80.
34. Levy M, Miller SL, Oró J. Production of guanine from NH<sub>4</sub>CN polymerizations. *J Mol Evol* 1999; 49:165-168.
35. Miller SL, Urey HC. Organic compound synthesis on the primitive earth. *Science* 1959; 130:245-251.
36. Schlesinger G, Miller SL. Prebiotic synthesis in atmospheres containing CH<sub>4</sub>, CO, and CO<sub>2</sub>. I. Amino Acids. *J Mol Evol* 1983; 19:376-382.
37. Ferris JP, Joshi PC, Edelson EH et al. HCN: A plausible source of purines, pyrimidines and amino acids on the primitive Earth. *J Mol Evol* 1978; 11:293-311.
38. Holm NG, Andersson EM. Abiotic synthesis of organic compounds under the conditions of submarine hydrothermal vents: A perspective. *Planet Space Sci* 1995; 43:153-159.
39. Shock EL. Geochemical constraints on the origin of organic compounds in hydrothermal systems. *Orig Life Evol Biosphere* 1990; 20:331-367.
40. Kelley DS, Karson JA, Blackman DK et al. An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30° N. *Nature* 2001; 241:145-149.
41. Bernhardt G, Lüdemann HD, Jaenicke R et al. Biomolecules are unstable under "black smoker" conditions. *Naturwissenschaften* 1984; 71:583-586.
42. Miller SL, Bada JL. Submarine hot springs and the origin of life. *Nature* 1988; 334:609-611.
43. Lazcano A. The tempo and mode(s) of prebiotic evolution. In: Cosmovici CB, Bowyer S, Wertheimer D, eds. *Astronomical and biochemical origins and the search for life in the universe*. Editrice Compositori, 1997: 70-80.
44. Chyba C, Sagan C. Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: An inventory for the origins of life. *Nature* 1992; 355:125-132.
45. Botta O, Bada JL. Extraterrestrial organic compounds in meteorites. *Surv Geophys* 2002; 23:411-467.
46. Van der Velden W, Schwartz AW. Search for purines and pyrimidines in the Murchison meteorite. *Geochim Cosmochim Acta* 1977; 41:961-968.
47. Stoks PG, Schwartz AW. Uracil in carbonaceous meteorites. *Nature* 1979; 282:709-710.
48. Stoks PG, Schwartz AW. Nitrogen-heterocyclic compounds in meteorites: Significance and mechanisms of formation. *Geochim Cosmochim Acta* 1981; 45:563-569.
49. Stoks PG, Schwartz AW. Basic nitrogen-heterocyclic compounds in the Murchison meteorite. *Geochim Cosmochim Acta* 1982; 46:309-315.

50. Ehrenfreund P, Glavin DP, Botta O et al. Extraterrestrial amino acids in Orgueil and Ivuna: Tracing the parent body of CI type carbonaceous chondrites. *Proc Natl Acad Sci USA* 2001; 98:2138-2141.
51. Egholm M, Buchardt O, Nielsen PE et al. Peptide Nucleic Acid (PNA). Oligonucleotide Analogues with an Achiral Peptide Backbone. *J Am Chem Soc* 1992; 114:1895-1897.
52. Nelson KE, Levy M, Miller SL. Peptide nucleic acids rather than RNA may have been the first genetic molecule. *Proc Natl Acad Sci USA* 2000; 97:3868-3871.
53. Ferris JP, Hill Jr AR, Liu R et al. Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* 1996; 381:59-61.
54. Wächtershäuser G. Life as we don't know it. *Science* 2000; 289:1307-1308.
55. Huber C, Wächtershäuser G. Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions. *Science* 1997; 245:245-247.
56. Cody GD, Boctor NZ, Filley TR et al. Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. *Science* 2000; 289:1337-1340.
57. Bada JL, Lazcano A. Some Like It Hot, but Not the First Biomolecules. *Science* 2002; 296:1982-1983.
58. Gold T. *The deep hot biosphere*. New York: Springer-Verlag, 1999.