

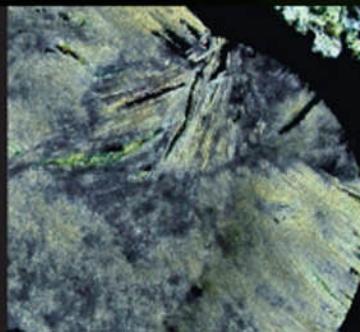


ANDREW M. SHAW

ASTROCHEMISTRY

FROM ASTRONOMY TO ASTROBIOLOGY

 WILEY



Astrochemistry

from Astronomy to Astrobiology

Andrew M. Shaw

University of Exeter



John Wiley & Sons, Ltd

Astrochemistry

from Astronomy to Astrobiology

Astrochemistry

from Astronomy to Astrobiology

Andrew M. Shaw

University of Exeter



John Wiley & Sons, Ltd

Copyright © 2006

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The Publisher is not associated with any product or vendor mentioned in this book.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, ONT, L5R 4J3, Canada

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data:

Shaw, Andrew M.

Astrochemistry : from astronomy to astrobiology / Andrew M. Shaw.

p. cm.

Includes bibliographical references and index.

ISBN-13: 978-0-470-09136-4 (HB) 978-0-470-09137-1 (PB)

ISBN-10: 0-470-09136-3 (HB) 0-470-09137-1 (PB)

1. Cosmochemistry – Textbooks. I. Title.

QB450.S53 2006

523'.02 – dc22

2006001883

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN-13: 978-0-470-09136-3 (HB) 978-0-470-09137-1 (PB)

ISBN-10: 0-470-09136-3 (HB) 0-470-09137-1 (PB)

Typeset in 10.5/12.5pt Times by Laserwords Private Limited, Chennai, India

Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire

This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Contents

Preface	ix
1 The molecular universe	1
Introduction	1
1.1 The Standard Model – Big Bang theory	1
1.2 Galaxies, stars and planets	3
1.3 Origins of life	4
1.4 Other intelligent life	9
1.5 Theories of the origin of life	10
Concepts and calculations	13
2 Starlight, galaxies and clusters	15
Introduction	15
2.1 Simple stellar models – black body radiation	15
2.2 2.725 K – cosmic microwave background radiation	20
2.3 Stellar classification	21
2.4 Constellations	26
2.5 Galaxies	31
2.6 Cosmology	36
Concepts and calculations	38
Problems	39
3 Atomic and molecular astronomy	41
Introduction	41
3.1 Spectroscopy and the structure of matter	41
3.2 Line shape	46
3.3 Telescopes	52
3.4 Atomic spectroscopy	56
3.5 Molecular astronomy	59
3.6 Molecular masers	77
3.7 Detection of hydrogen	79
3.8 Diffuse interstellar bands	80
3.9 Spectral mapping	81

Concepts and calculations	82
Problems	83
4 Stellar chemistry	85
Introduction	85
4.1 Classes of stars	86
4.2 Herzprung–Russell diagram	88
4.3 Stellar evolution	89
4.4 Stellar spectra	98
4.5 Exotic stars	102
4.6 Cycle of star formation	108
Concepts and calculations	110
Problems	111
5 The interstellar medium	113
Introduction	113
5.1 Mapping clouds of molecules	114
5.2 Molecules in the interstellar and circumstellar medium	117
5.3 Physical conditions in the interstellar medium	120
5.4 Rates of chemical reactions	123
5.5 Chemical reactions in the interstellar medium	130
5.6 Photochemistry	133
5.7 Charged particle chemistry	136
5.8 Polycyclic aromatic hydrocarbons	136
5.9 Dust grains	140
5.10 Chemical models of molecular clouds	145
5.11 Prebiotic molecules in the interstellar medium	151
Concepts and calculations	154
Problems	155
6 Meteorite and comet chemistry	157
Introduction	157
6.1 Formation of the solar system	158
6.2 Classification of meteorites	161
6.3 Meteorite mineralogy	162
6.4 Geological time	165
6.5 Chemical analysis of meteorites by μ L ² MS	168
6.6 The Murchison meteorite – kerogen	171
6.7 Meteorite ALH84001	173
6.8 Comet chemistry	180
6.9 Structure of a comet	180
6.10 Physicochemical conditions in a cometary coma	181

6.11	Chemical composition of comets	183
6.12	Cometary collisions	185
6.13	The Rosetta mission – origin of the solar system	187
	Concepts and Calculations	190
	Problems	191
7	Planetary chemistry	193
	Introduction	193
7.1	Structure of a star–planet system	194
7.2	Surface gravity	195
7.3	Formation of the Earth	197
7.4	Earth–Moon system	199
7.5	Geological periods	200
7.6	Radiative heating	202
7.7	The habitable zone	204
7.8	Extrasolar planets	206
7.9	Planetary atmospheres	209
7.10	Atmospheric photochemistry	215
7.11	Biomarkers in the atmosphere	219
	Concepts and calculations	221
	Problems	222
8	Prebiotic chemistry	225
	Introduction	225
8.1	Carbon- and water-based life forms	225
8.2	Spontaneous chemical reactions	227
8.3	Rates of chemical reactions	236
8.4	Endogenous production of organic molecules	237
8.5	Exogenous delivery of organic molecules	245
8.6	Homochirality	246
8.7	Surface Metabolism – ‘clay organisms’	249
8.8	Geothermal Vents – ‘black smokers’	251
8.9	RNA World hypothesis	253
	Concepts and calculations	256
	Problems	257
9	Primitive life forms	259
	Introduction	259
9.1	Self-assembly and encapsulation	261
9.2	Protocells	264
9.3	Universal tree of life	273
9.4	Astrobiology	274
9.5	Microbial Mars	281
	Concepts and calculations	283
	Problems	284

10 Titan	287
Introduction	287
10.1 Physical properties	289
10.2 The atmosphere	291
10.3 Temperature-dependent chemistry	294
10.4 Energy balance and the greenhouse effect	296
10.5 Atmospheric chemistry	297
10.6 Astrobiology on Titan	302
Concepts and calculations	305
Problems	306
Glossary of terms and abbreviations	307
Appendix A – constants and units	319
Appendix B – astronomical data	321
Appendix C – thermodynamic properties of selected compounds	323
Answers to problems	325
Bibliography	329
Index	335

Preface

Astrochemistry draws its inspiration, language, fascination, beauty, elegance and confusion from many different disciplines: starting with astronomy, passing through physical chemistry and ending with the new ideas of astrobiology. It is this breadth of fascination that I have attempted to capture in *Astrochemistry: from Astronomy to Astrobiology*. Choosing such a broad subject comes with the serious problem of how to limit the discussion of the details to allow an appreciation of the whole. I could have written an entire book on molecular astrophysics, looking at what molecules are doing in the various environments of space. I could have looked simply at the wonders of planetary chemistry, concentrating on the solar system or even just one planet such as Jupiter. Why does it have a giant red spot? Instead, I have chosen to apply a more general boundary condition for the book taking in all of the subjects but focused on the theme of “The Origin of Life”.

Astrochemistry starts with the origins of the Universe and the theory of the Big Bang, resulting in the formation of hydrogen, helium and a little lithium. Gravity pulls the matter together to form stars, galaxies and clusters of galaxies, all of which give off light in some form. The light tells the molecular story with information on the formation and evolution of stars and the role of atoms. At times these interesting subjects are buried in the disciplines of astronomy and astrophysics and I have tried to bring the pieces of the story together, concentrating on astrochemistry. The cycle of star formation ends with a supernova blowing huge quantities of material into the interstellar medium, now laden with all of the elements of the Periodic Table. Chemistry in the interstellar medium, with rather cold and tenuous conditions, is now possible and this controls the starting molecular inventory. To understand this fully, the subjects of quantum mechanics and kinetics need to be applied, through spectroscopy and chemical reaction networks, to the giant molecular clouds of the interstellar medium – the birthplace of stars and life?

Giant molecular clouds collapse to form stars and solar systems, with planets and debris left over such as comets and meteorites. Are comets and meteorites the delivery vehicles that enable life to start on many planets and move between the planets as the solar system forms, providing water and molecules to seed life? The planets have to be hospitable, however, and that seems to mean wet and

warm. Carbon-based life forms and liquid water seem to be the successful life-experiment on Earth from which we can draw some more general conclusions about the requirements for life in a view towards astrobiology. A look at prebiotic chemistry and primitive life forms on Earth poses interesting questions such as what is a cell and how big does it have to be? The guiding principles for prebiotic chemistry are the laws of thermodynamics that keep the origins of life and its understanding on the straight and narrow.

Finally, and tantalizingly for this book and astrochemistry, there is Titan. The Cassini–Huygens mission is now in orbit in the Saturnian system as the book is published. The Huygens probe has already made the descent to the surface of Titan and the data have been transmitted back successfully. Scientists, astronomers, astrochemists and astribiologists are trying to understand it. I have taken a brief look at Titan as a case study to apply all that has been learnt and to review the possibilities for astrochemistry in what is surely to be a very exciting revelation of the structure and chemistry of Titan.

Throughout the book I have tried to constrain the wonders of imagination inspired by the subject by using simple calculations. Can all of the water on the Earth have been delivered by comets: if so, how many comets? How do I use molecular spectroscopy to work out what is happening in a giant molecular cloud? Calculations form part of the big hard-sell for astrochemistry and they provide a powerful control against myth. I have aimed the book at second-year undergraduates who have had some exposure to quantum mechanics, kinetics, thermodynamics and mathematics but the book could easily be adapted as an introduction to all of these areas for a minor course in chemistry to stand alone.

Units and conventions

Astronomy is probably the oldest of the subjects that influence astrochemistry and contains many ancient classifications and unit systems that have been preserved in the scientific research of today. Distances are measured in light-years or parsecs, neither of which are the standard SI unit of length: the metre. This is not surprising when a light-year is 9.5×10^{15} m and is a relatively small astronomical unit of length! The correct SI convention for a light-year would be 9.5 petametres, written as 9.5 Pm. This is formally correct but would not help you in a conversation with anybody, as most scientists cannot remember the SI prefixes above 10^{12} . I have listed the SI prefixes in Appendix A and we shall use them where appropriate. However, I will use two units of length chosen from astronomy, namely the light-year and the astronomical unit. The light-year is the distance travelled by light in 1 year or 86 400 s and 1 ly is 9.5 Pm or 9.5×10^{15} m. Usefully, the distance to the nearest star is some 4 ly. The other length unit is the astronomical unit (AU), which is the average distance from the Earth to the Sun and is 1.49×10^{11} m, with the entire solar system being approximately 150 000 AU and the distance to the nearest star some 300 000 AU.

The unit of time is the second in the fundamental list of constants but it is convenient to use years when referring to the age of the Universe, Solar System or the Earth. I have chosen to use the SI prefixes in front of the symbol yr so that 10^9 years is 1 Gyr; the age of the Universe is 15 billion years or 15 Gyr, etc., and whenever this refers to a period of time in the past then 4.5 Gyr ago will be used explicitly.

The conventions of chemistry, particularly physical chemistry, are standard and appear in all physical chemistry textbooks and will be used here. The same courtesy has been extended to organic and inorganic chemistry and biology, so that the ideas of these subjects can be linked into the common theme.

Course material

I have put together a website for the book (www.wiley.co.uk/shawastrochemistry) where I have included the figures from the book to be downloaded into lectures. I have also included some links that I have found useful, corrections when required even some possible examination questions. I hope an adventurous professor will find these useful.

Acknowledgements

The book started as a survey of the literature to identify a research project, which in part it did, but during the work I discovered how interesting the subject can be and decided that it would make a good lecture course. The long-suffering students at the Department of Chemistry at Exeter University have enjoyed the course on two separate occasions and in two incarnations, most latterly as CHE2057 in 2005. The students saw the book at first draft and have contributed to removing the mistakes and suggested additions, pointing out where I said too much or too little. The refinements have helped and improved the text immeasurably. I have doubtlessly introduced more mistakes and for this I must take the full credit. The integrity of the book has been improved greatly by two very conscientious reviewers, to whom I owe a debt of gratitude. I must extend thanks to all scientists around the world who helped to put together the figures for the book. Busy people spent valuable time collecting the images that have added to the wonder of the subject. The reward for writing the book will be the spark of curiosity that may flicker in the mind of the reader.

Andrew M. Shaw

1

The molecular universe

Introduction

Chemistry without numbers is poetry: astrochemistry without numbers is myth. A molecule placed around a star, in a nebula, lost in the interstellar medium, on a planet or within a cell has the potential for very complex and beautiful chemistry but unless we can understand the local conditions and how the molecule interacts with them we have no idea what chemistry is really happening. To understand astrochemistry we need to understand the physical conditions that occur within many diverse molecular environments. The exploration of the molecular universe will take us on a long journey through the wonders of astronomy to the new ideas of astrobiology.

The origins of life provide the motivation and excuse to investigate astrochemistry in its broadest sense, looking at molecules and their local complex chemistry using all of the tools of physical chemistry to constrain the imagination of the astrobiologist in the field and to force a re-think of the rules of biology that are prejudiced by the experience of life on Earth. The complexity of the problem places demands on the theories of science, stretching the understanding of kinetics and thermodynamics into areas where large non-ideal systems are hard to understand, although, curiously, modelling the complex chemistry of a giant molecular cloud is not dissimilar to the models of biochemistry within a cell. The size of the chemical problem quickly grows, so that the chemistry of 120 molecules in a molecular cloud must be compared with the 4500 reactions thought to be required to make a cell work. The full understanding of the chemical reactions must be modelled as a network of coupled chemical equations, which for something as comparatively simple as a candle flame can contain 350 equations.

Our mission is to explore the molecular universe with an understanding of all of the local molecular environments and constrain possible chemical reactions using the concepts of physical chemistry. With such a wide brief we need a focus and I have chosen the origins of life on Earth and on all planets – astrobiology.

1.1 The Standard Model – Big Bang theory

About 15 billion years ago the Universe and time itself began in a Big Bang. Observations of the night sky show that stars and galaxies are moving away from

Table 1.1 The history of the Universe according to the Standard Model

Time since $t = 0$	Temperature	Comments
10^{-43} s	10^{32} K	Gravity is now distinct from the three other forces: strong, weak nuclear and electromagnetic
10^{-35} s	10^{27} K	Inflation of the Universe – the strong force separates
10^{-12} s	10^{15} K	Weak and electromagnetic forces separate. Neutrons and protons are formed by photon–photon collisions
10^{-2} s	10^{11} K	Electrons and positrons are formed through collisions of photons
1 s	10^{10} K	The Universe becomes transparent to neutrinos
180 s	10^9 K	Nucleosynthesis: hydrogen, deuterium, helium and some lithium
$3 - 7 \times 10^5$ s	3000 K	Light element atoms form, and the Universe is now transparent to radiation: cosmic background is emitted
10^9 yr	20 K	Galaxies form
Present	2.726 K	Stars and galaxies

us, telling us that the Universe is expanding. Extrapolating backwards in time leads to a point of common beginning, a singularity in space-time known as the Big Bang. Temperature is critical to the phases of evolution and subsequent cooling of the Universe, producing a number of critical times, detailed in Table 1.1. They are all predictions of the Big Bang Theory or the Standard Model of Cosmic Evolution.

Einstein's theory of relativity allows for the interconversion of energy and matter through the famously simple equation $E = mc^2$. Thus collisions between high-energy photons in the primordial fireball created particle–antiparticle pairs such as protons and antiprotons. After some 180 s and at a temperature of 10^9 K atomic nuclei such as hydrogen, deuterium, helium and some lithium were formed. The first three minutes of all time were chemically the dullest with no atoms or molecules. For a further 10^6 s the light atoms continue to be formed, marking a period where matter is created by *Big Bang nucleosynthesis*.

There are a number of astronomical pieces of evidence for the Big Bang Theory as we shall see, including the recent observation of the cosmic microwave background radiation, but it is far from a complete theory. However, predictions of the theory may be tested. One such prediction is the relative abundance by mass of He, which must be at least 25 per cent of the total mass. Helium is also made in stars and must contribute to the He density of the Universe and in all observations to date the observed abundance is greater than 25 per cent. There are problems associated with matter. Why is the Universe made from matter instead of antimatter? When was this decision made to stabilise matter from high-energy photons and particle–antiparticle pairs. Further, calculations of gravitational attractions of galaxies suggest the presence of large amounts of matter that cannot be seen, so-called dark matter. What is dark matter?

Table 1.2 Relative cosmic abundances of the elements

Element	Relative abundance	Element	Relative abundance
H	1	S	1.6×10^{-5}
He	0.085	P	3.2×10^{-7}
Li	1.5×10^{-9}	Mg	3.5×10^{-5}
C	3.7×10^{-3}	Na	1.7×10^{-6}
N	1.2×10^{-3}	K	1.1×10^{-7}
O	6.7×10^{-3}	Si	3.6×10^{-6}

The majority of the Universe is made from hydrogen and helium produced during the Big Bang, although some He has been made subsequently. The relative cosmic abundance of some of the elements relevant to the formation of life is given in Table 1.2, with all elements heavier than H, He and Li made as a result of fusion processes within stars, as we shall see later. The cosmic abundance is assumed to be the same as the composition of the Sun.

1.2 Galaxies, stars and planets

All matter formed within the Big Bang is attracted to itself by the force of gravity and after about 1 billion years massive proto-galaxies form. Gravitational contraction continues in more and more localised regions to form the galaxies we know today, including our galaxy, the Milky Way. The Milky Way is in a cluster of galaxies called the ‘local group’ that includes the Large Magellanic Cloud, the Small Magellanic Cloud and the Andromeda Galaxy (M31). Two of these, the Milky Way and the Andromeda Galaxy, are very luminous spiral galaxies.

The Milky Way was formed within 1 billion years of the Big Bang and has a mass of 10^9 solar masses. It formed from a large cloud of hydrogen and helium that was slowly rotating. As the cloud collapsed, conservation of angular momentum required matter near the axis to rotate very fast. As a result it spreads away from the axis and forms a flat spiralled disc some 120 000 ly in diameter and about 3300 ly thick. The Sun is approximately 30 000 ly from the centre. The nuclear bulge at the core of the galaxy contains old stars, and observations suggest that it must be hugely massive. Rapid rotation around the axis of the disc requires gravity and angular momentum, hence mass, to hold it together and this produced speculation about the existence of a black hole at the centre of the Milky Way.

The Sun formed some 4.5 Gyr ago (Gyr is a Gigayear or 10^9 years) from its own gas cloud called the solar nebula, which consisted of mainly hydrogen but also all of the heavier elements that are observed in the spectrum of the Sun. Similarly, the elemental abundance on the Earth and all of the planets was defined by the composition of the solar nebula and so was ultimately responsible for the molecular inventory necessary for life. The solar system formed from a slowly rotating nebula that contracted around the proto-sun, forming the system of planets called the solar system. Astronomers have recently discovered solar systems around

other stars and, in only the briefest of looks, this has revealed a large proportion of similar planetary systems; thus the formation of planets around stars is a common process. The distribution of mass in the solar system is primarily within the Sun but distributed rather differently among the planets. The inner planets, the so-called terrestrial planets of Mercury, Earth, Venus and Mars, are essentially rocky but Jupiter, Saturn, Uranus and Neptune are huge gas giants; this needs to be explained by the formation process. Most important, however, is the formation of a planet in a habitable zone, where liquid water can exist and have the potential for life – at least if you follow the terrestrial model.

1.3 Origins of life

The age of the Earth is established by radioisotope dating at 4.55 Gyr and for most of the first billion years it suffered major impact events capable of completely sterilising the Earth and removing any life forms. The geological fossil records reveal, however, that life existed some 3.5 Gyr ago and perhaps as early as 3.9 Gyr ago. The oldest known life forms were very simple by modern standards but already had hugely complicated structures involving membranes and genetic information. The rather surprising conclusion is that life may have developed in as little as one hundred million years and at most 0.5 billion years, to evolve from the primordial soup to a viable living organism that had adapted to its local environment.

Definitions of life

There are many problems with the definitions of life, although determining what is alive and what is not is intuitively easy. At the extremes of collections of matter are the human being and the atom, with all of the possibilities in-between. Classical definitions of life taken from biology, such as ingesting nutrients, excreting by-products, growth and reproduction, all serve as good markers of life although are almost certainly prejudiced by life on Earth. What about fire? A candle flame (Figure 1.1) clearly ingests nutrients from the air in the form of oxygen and fuel from the wax. It produces waste products; it can also grow to cover large areas and certainly looks as if it might reproduce itself by creating new fires through sparks. It is localised by both a temperature and concentration gradient and might indeed be alive. However, one flame does become a copy of itself in that it will burn whatever fuel and oxidant combination available to it. In a sense it evolves and lives for as long as it can adapt to its environment. The adaptation to the environment is seen on the right-hand side of Figure 1.1 where a candle flame is burning under conditions of zero gravity in the space shuttle. The shape of the flame in air is controlled by buoyancy: the hot air inside the candle flame air is less dense than the air around it and it rises. In zero gravity the hot air does not rise because its weight is zero and so the random thermal motion results in diffusion

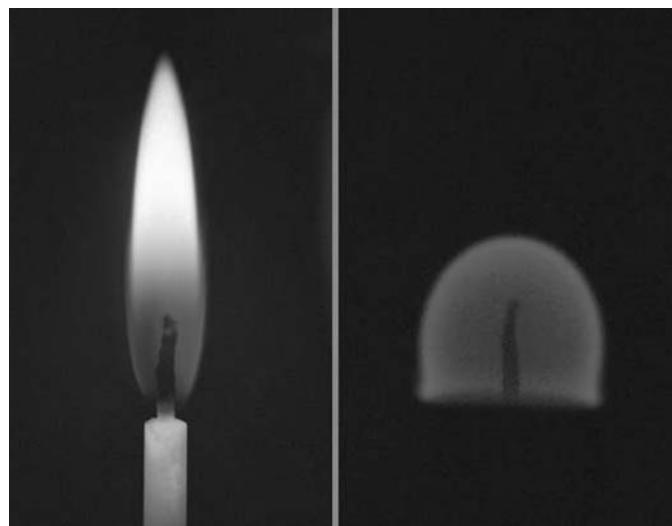


Figure 1.1 Two species of candle flame – dead or alive? The flame on the left is on Earth and the flame on the right is burning under zero gravity. (A colour reproduction of this figure can be seen in the colour section). (Reproduced from photos by courtesy of NASA)

of oxygen into the flame and combustions products away from the flame, hence the flame is now spherical. Even a complex set of chemical reactions, recognisable as a flame, has adapted to the environment. There is a consistent chemistry set within the 350 equations required to get the flame ‘metabolism’ chemistry to burn properly and as such it contains a recipe or a DNA. Other more impressively vague twilight life forms must include virus particles.

Viruses have no real metabolism and appear to exist in a dormant state until they find a suitable host. Then they hijack the metabolism and DNA replication apparatus of the host cell, switching the host into the production of huge numbers of copied virus particles, including some mutations for good measure. Finally, the cell bursts and the virus particles are released to infect a new host. The propagation of genetic information is important, as is the need for some form of randomisation process in the form of mutations, but it is not clear that there can be one definition for life itself. NASA has chosen the following definition:

‘Life is a self-sustained chemical system capable of undergoing Darwinian evolution.’

Alternatively:

‘A system that is capable of metabolism and propagation of information.’

Both are flaky, as even the simplest of thought will reveal.

Specialisation and adaptation

Cellular life may have arisen spontaneously, capturing whatever prebiotic debris that was present in the primordial soup. The encapsulation process provided the first specialisation within the environment, leading to compartmentalisation firstly from the external environment and then within the cell to provide areas of the protocell with dedicated adapted function. The external barrier in biology is provided by a cell membrane constructed from a bilayer of phospholipids with added sugars to make it rigid. The phospholipid molecules are amphiphilic, containing a long fatty acid chain of 10–20 carbon atoms at one end that are hydrophobic and a phosphate head group at the other end that is hydrophilic. It is the hydrophobic–hydrophilic characteristic at different ends of the molecule that make it amphiphilic. These molecules spontaneously form vesicles and membranes called liposomes in water when the concentration is above the critical micelle concentration. The network of chemical reactions trapped within a liposome could easily form a proto-metabolism but there is still the need for an information-bearing polymer.

Looking again at biology, genetic information is stored in all organisms as either DNA or RNA. These huge polymeric molecules contain the information for the replication of the building blocks of all organisms, the proteins. The four bases, G, A, T and C, pair together as A–T and G–C, the so-called Watson–Crick base pairs, which together with the deoxyribose sugar and phosphate backbone form the α -helix of the DNA molecule, shown in Figure 1.2.

The order of the bases is important along the length of the DNA and each sequence of three bases, called a triplet, represents the words in the genetic code. Each triplet codes for an amino acid so that AAA is lysine and UGU is cystine with signals for ‘stop’, such as UGA, and ‘start’ (no simple sequence but TATA is a reasonable example) to establish the beginning of a gene. More triplets are used to code for each of the 20 or so amino acids used in living organisms and the order in which they must be put together to form a protein (Figure 1.3). The information coded within the DNA is propagated from generation to generation nearly always correctly but sometimes with mistakes or mutations. Not all mistakes are bad; mistakes that provide an advantage in the local environment are good mistakes and allow evolution. The organism with the good mistake will evolve and adapt better to its surroundings, outgrowing less-well-adapted organisms.

Proteins are constructed from long chains of amino acids linked together by a peptide bond. There are 20 common amino acids that are coded within the genome and they are all of L-optical activity. Optical activity refers to the interaction of molecules with polarised light and divides molecules into three types: those that do not rotate the plane of polarisation of the light; those that rotate the plane of polarisation to the right; and those that rotate the plane of polarisation to the left. The two types of molecules that rotate light are called chiral molecules and those that do not are called achiral. The choice of one set of chiral molecules, called homochirality, over the other set is a marker for biological activity. Although amino acids may be produced in space on the ice mantles of interstellar dust grains, they are thought to be racemic mixtures, meaning that they have equal quantities of

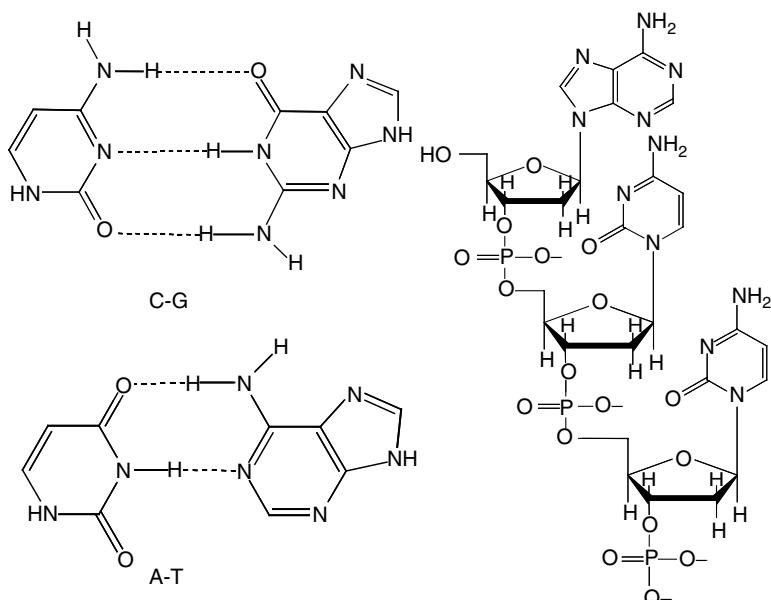


Figure 1.2 Watson–Crick DNA base pairs and the DNA backbone

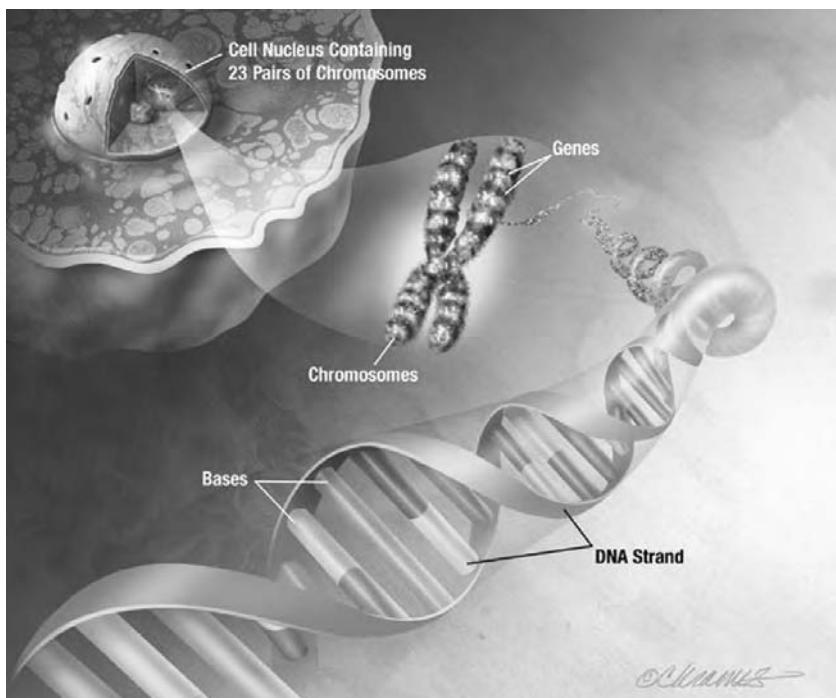


Figure 1.3 The genetic code. (Reproduced from Alzheimer's Disease Education by courtesy of the National Institute on Aging)

the L and D forms of the chiral molecules. Similar optical purity is seen in the bases of DNA and RNA and with biologically active sugars. Curiously, all sugars are D-enantiomers.

The origin of the *homochirality* is not known. There is a tiny energy difference between the optical isomers associated with a ‘parity violating energy difference’ of order 10^{-15} – 10^{-17} J, but in general homochirality will require biological amplification favouring one enantiomer over another, i.e. ‘enantiomeric amplification’. It has been suggested recently that organic synthesis in the circularly polarised light field around a star in the interstellar medium or due to chiral-specific surface reactions may also provide a mechanism for enantiomeric amplification and we shall discuss this later. Homochirality is, however, easily achieved by biological systems and may be considered as a *biomarker* – a marker for the existence of life.

It is the variety of life around the edges of the biosphere on Earth that is a testimony to its adaptation and ability to survive in harsh and extreme environments. The bacteria in the hot-water spring shown in Figure 1.4 have adapted to different temperatures and salinities.

Some bacteria require extreme temperatures, e.g. hyperthermophile organisms require hot water to live and will not survive below 90°C. The extremophile bacteria are from a general class of organisms that have adapted and thrive in extreme living conditions found in deep-sea marine environments and deep subsurface colonies. These bacteria may make up most of the collection of biological organisms on Earth that form the biota. The molecular classification of organisms based on the length of the genome suggests that the last common ancestor was a hyperthermophile



Figure 1.4 Hyperthermophile bacteria at Prismatic Lake in Yellowstone National Park. (A colour reproduction of this figure can be seen in the colour section). (Reproduced from a photo of Prismatic Lake by courtesy of National Park Service, Yellow Stone National Park)

bacterium. Far from being descended from the apes, we are actually descended from bugs. Measurements of the survival of bacteria in space suggest that, in the form of spores or dried cells, the survival in space is possible at least for the 6 years that the experiments have been taking place. The transfer of life from planet to planet is then a real possibility. The recent extensive analysis of meteorite ALH84001 suggests that there may be structures within the rock that look like fossil organisms. The meteorite was ejected from the surface of Mars probably by a collision and then made a rapid transit to Earth perhaps as quickly as 60 000 years. This meteorite express-delivery service suggests that not only are we descended from bugs but perhaps even from Martian bugs.

1.4 Other intelligent life

The prospect of intelligent life anywhere in the Universe has been puzzling astronomers and recently astrobiologists, and there have been some attempts to estimate probabilities. This led Drake to construct a now famous equation that collects the ideas together: the Drake equation. It is a mathematical representation of factors relating the probability of finding life and, in particular, an intelligent civilisation elsewhere in the Universe. This is an extreme example of ‘hypothesis multiplication’ and should be treated with caution. The equation is written:

$$N_c = R_s f_p n f_l f_I f_c L \quad (1.1)$$

where N_c is the number of intelligent civilisations in the Universe with whom we might communicate, R_s is the rate of formation of stars in the galaxy, f_p is the fraction of stars that have planetary systems, n is the average number of habitable planets within a star’s planetary system, f_l is the fraction of habitable planets upon which life arises, f_I is the fraction of habitable planets upon which there is intelligent life, f_c is the fraction of civilisations interested in communicating and L is the average lifetime of a civilisation.

The problem comes with assessing the values of the factors to place within the equation and this leads to some very optimistic or pessimistic estimates.

- R_s : There are approximately 10^{11} stars in our galaxy and given that the age of the galaxy is some 10 Gyr the rate of star formation is then approximately 10 stars per year.
- f_p : This could be as large as 0.5 but may be complicated by binary stars and other local factors: pessimistically it is 0.01 and optimistically 0.3–0.5.
- n : Our solar system may have had as many as three habitable planets (Earth, Mars and maybe Venus, at least for a while), however giant planet formation may have removed inner terrestrial-type planets by collision during the formation process: optimistically it is 3, pessimistically 0.01.

- f_l : Is life the product of a collection of simple chemical processes, in which case it should be everywhere, or is it more of a fluke: optimistically it is 1, pessimistically 10^{-6} .
- f_I : Intelligence on Earth took 4.5 Gyr to evolve and many stars do not live this long (dependent on their mass) so they choose some extremes: optimistically it is 0.5 and pessimistically 10^{-6} .
- f_c : If they are like human beings then everybody wants to talk so, lets say, 1.0.
- L : This has been a few tens of years in the case of our civilisation and we may yet destroy ourselves within 10 000 years: optimistically it is 1 billion years and pessimistically 100 years.

Performing the optimistic sum gives $N_c = 5 \times 10^9$ and the pessimistic sum gives $N_c = 10^{-13}$.

The Drake equation is just a mathematical way of saying ‘who knows’ but it does allow the factors that might control the origins of life to be identified; that said, it is probably the worst calculation in the book.

1.5 Theories of the origin of life

There is no one correct theory for the origin of life on Earth or any habitable planet, although many have been presented. The current set of ideas is summarised in Figure 1.5. Aside from the theory of creation, which seems particularly hard to test, the testable theories of the origins of life divide into two: *extraterrestrial* or panspermia, the theory that life was seeded everywhere somewhat randomly; and *terrestrial*, that life originated *de novo* on Earth or other habitable planets around other stars. The theories of terrestrial origin are more favoured but the recent discovery of habitable planets and life within any solar system suddenly makes panspermia more likely.

Terrestrial theories divide into an organic or inorganic origin for life. If life started with organic molecules, how were these molecules formed from prebiotic conditions, perhaps in what Darwin called a ‘little warm pool’ or a primordial soup theory? The endogenous production of organic material provides a continuous link from a prebiotic planet to a complete organism. However, conditions may have been habitable for organic material from outside the planet, exogenous origin, to land on Earth, perhaps delivered by meteorites or comets. The energy source for life could also be organic such as in photosynthesis or inorganic chemistry around hydrothermal vents. Indeed, perhaps inorganic material surfaces catalysed the formation of the first set of self-replicating molecules or a primitive organism called a ‘surface metabolite’?

Extraterrestrial theories suggest that life is formed wherever the chemistry will allow the formation of life, either randomly or perhaps directed by some guiding

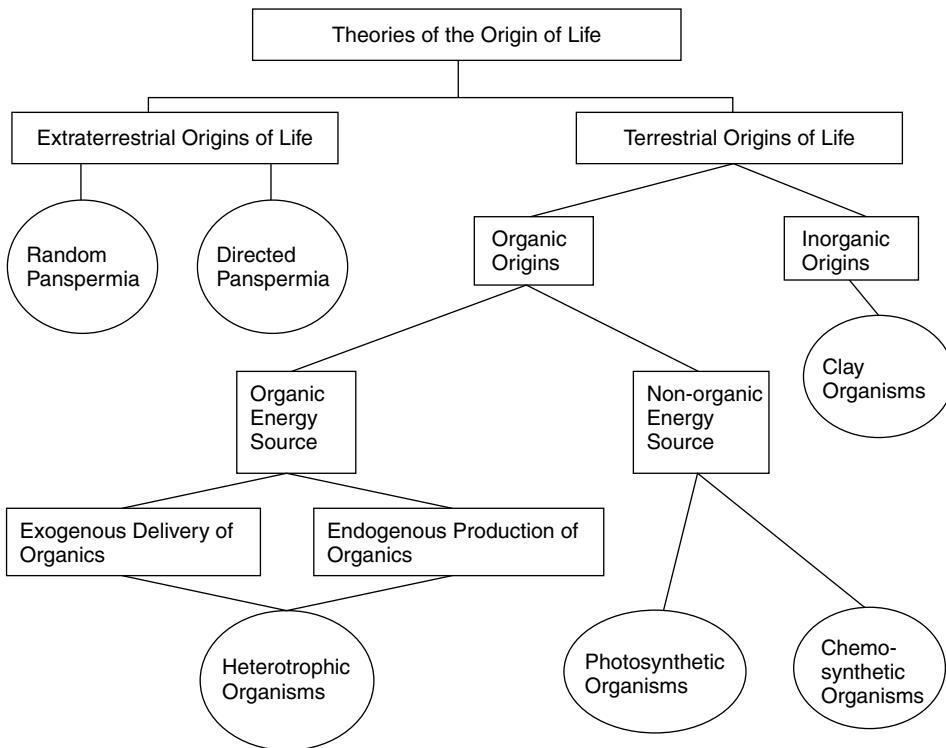


Figure 1.5 Theories of the origins of life. (Reproduced from Davis and McKay 1966 by courtesy of Kluwer Academic Publishers)

force. The discovery of extrasolar planets with a habitable zone that allows liquid water to exist suggests that conditions with the right energy balance and molecular inventory could produce life spontaneously. The idea of a prebiotic planet capable of supporting life must be ubiquitous. The temptation, however, is to assume that what we see today has a direct lineage with the prebiotic Earth or any planet and this cannot be the case. Mass extinction events abound in the fossil record on Earth, not least of which is the famous meteorite impact that killed the dinosaurs, as seen in Figure 1.6. Impacts in the early stages of the formation of the Earth were very common, as is witnessed by the cratering scars on the surface of the Moon. A similar stream of meteorite impacts must also have collided with the Earth and a large meteorite, several kilometres in diameter, certainly impacted on the surface of the Earth near the Yucatan peninsula in Mexico some 65 million years ago. The Chicxulub crater shows an impact large enough to cause global heating of the Earth's atmosphere and vaporisation of some if not all of the oceans, wiping out the dinosaurs and many more apparently evolutionarily fragile species on the planet at the time.

Evidence for early collisions is also present in the fossil record, suggesting that the diversity of species present on Earth has been reduced considerably on several occasions, perhaps removing some 90 per cent of species. Some organisms

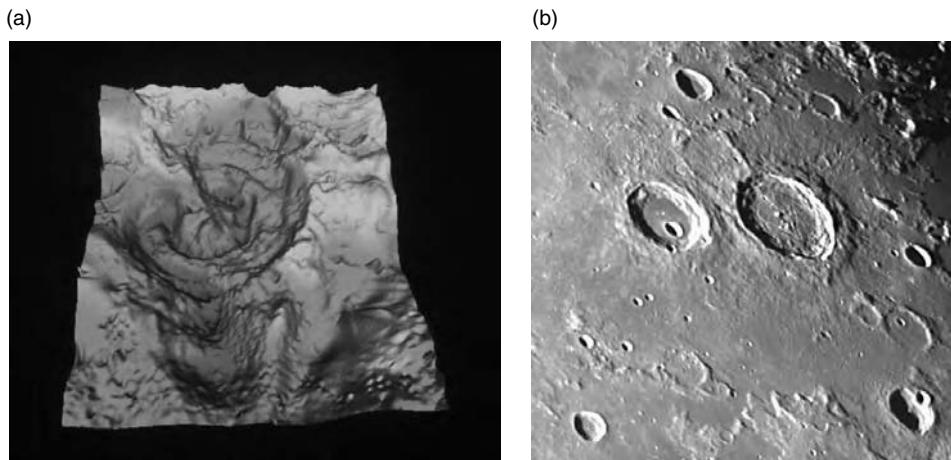


Figure 1.6 Impact frustration: (a) the Chicxulub crater, seen as a three-dimensional gravity map, thought to be responsible for the extinction of the dinosaurs; (b) the cratered surface of the Moon. (Reproduced from photos by courtesy of NASA)

may have survived by collecting biologically active molecules from the debris of other species. Replication and amplification result in the formation of the complex processes of translation and transcription of the genome, synthesis of proteins and the design of cellular metabolism. How modern life evolved may be a very different question to how life occurs spontaneously as a product of prebiotic chemistry and may look very different to the astrobiologist in the field. Astrobiology cannot be limited to the ideas of biology on Earth.

Concepts and calculations

Concepts

<i>Standard Model</i>	The theory of the origins of the Universe
<i>DNA base pairs</i>	The general structure of the information-bearing molecule used in biology on Earth
<i>Homochirality</i>	The general idea of biomarkers for the existence of life, in this case a preference for one optical isomer over another
<i>Extraterrestrial origins of life</i>	Life was delivered to the Earth (or any planet) by meteorites or cometary material, leading to the idea of panspermia
<i>Terrestrial origins of life</i>	The molecules of life were built on Earth, perhaps in the primordial soup or little warm pool
<i>Impact frustration</i>	A process typified by the extinction of the dinosaurs where the Earth's surface may have been sterilised of all life forms by the impact of a large meteorite or comet

Calculations

<i>Drake equation</i>	Use of this equation, including estimates of the optimistic and pessimistic calculations, for the existence of other life forms
-----------------------	---

2

Starlight, galaxies and clusters

Introduction

Look into the sky, day or night, and you will see starlight, light that when separated into its constituent wavelengths contains information about the stars. Whether it is our local star, the Sun, or many millions of not so local stars, there may be other worlds sharing the galactic starlight. Possibly the simplest form of light familiar even to early man was the light associated with heat. The red embers of a camp fire or the yellow colour of a candle flame are all due to the simplest model of starlight called black body radiation.

2.1 Simple stellar models – black body radiation

The simplest models of stars treat a mass of hydrogen and helium as a radiating black body that goes through a formation process, a mid-life and then a death in a sequence of events controlled by its initial mass. A black body is an object that absorbs and emits all wavelengths of radiation with equal efficiency. This produces a continuum electromagnetic emission spectrum over all wavelengths against which absorption features associated with the elements and molecules present in the stars can be seen. Properties that define a black body are:

1. A black body with $T > 0$ K emits radiation at all wavelengths;
2. A hotter black body emits more light at all wavelengths than a colder body;
3. A hotter black body emits more of its radiation at shorter wavelengths;
4. Black bodies emit and reflect radiation at all wavelengths with equal efficiency.

Black bodies have formed an important part in the development of the theory of quantum mechanics and were studied by the early quantum physicists, producing a number of laws relating the temperature of the black body to the photon flux, the

luminosity and the most intense wavelength, culminating in a complete description of the wavelength–intensity relation known as Planck’s Law from which we see for the first time Planck’s constant.

Stefan–Boltzmann Law

The first of the laws is the Stefan–Boltzmann Law relating the amount of energy emitted from a black body, F , to its temperature T :

$$F = \sigma T^4 \quad (2.1)$$

where F is the total flux or power per unit area, T is the temperature of the black body in Kelvin and σ is Stefan’s constant $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2}\text{K}^{-4}$.

Example 2.1

The part of the Sun that we can see is called the photosphere and has a surface temperature of 5780 K. The solar flux from every square metre of the surface is then given by Equation 2.1:

$$\begin{aligned} F_{\text{Sun}} &= 5.67 \times 10^{-8} (5780)^4 \quad (\text{W m}^{-2} \text{ K}^{-4})(\text{K}^4) \\ F_{\text{Sun}} &= 6.3 \times 10^7 \text{ W m}^{-2} \end{aligned}$$

Luminosity

The luminosity of a star, L , is the total rate at which energy is radiated by the black body over all wavelengths. Assuming that the star is a sphere, the total surface area is given by $A = 4\pi R^2$, where R is the radius of the star, then the luminosity of the star is given by:

$$L = 4\pi R^2 \sigma T^4 \quad (2.2)$$

This quantity is the total amount of radiation at all wavelengths radiating through the surface of the sphere and is simply the Stefan-Boltzmann Law multiplied by the surface area of the photosphere.

Example 2.2

The luminosity of the Sun may be calculated from Equation 2.2 knowing the radius of the Sun to be $R_{\text{Sun}} = 6.96 \times 10^8 \text{ m}$:

$$\begin{aligned} L_{\text{Sun}} &= 4\pi R^2 \times 5.67 \times 10^{-8} (5780)^4 \quad (\text{m}^2)(\text{W m}^{-2} \text{ K}^{-4})(\text{K}^4) \\ L_{\text{Sun}} &= 3.8 \times 10^{26} \text{ W} \end{aligned}$$

Wien's Law

Turning now to the wavelength distribution of the starlight. The emission from a black body must, by definition, produce radiation at all wavelengths, i.e. a wavelength distribution. It turns out for a black body that the wavelength at which the maximum radiation flux occurs is characteristic of the temperature and is given by Wien's Law:

$$\lambda_{\max} = \frac{w}{T} \quad (2.3)$$

where Wien's constant $w = 2.9 \times 10^{-3}$ m K and gives a value for λ_{\max} in metres.

Example 2.3

Taking the surface temperature of the Sun to be 5780 K and substituting into Equation 2.3 gives:

$$\begin{aligned}\lambda_{\max} &= \frac{2.9 \times 10^{-3}}{5780} \frac{\text{m K}}{\text{K}} \\ \lambda_{\max} &= 5.01 \times 10^{-7} \text{ m} = 501 \text{ nm}\end{aligned}$$

This is an interesting result, as with an absorption maximum in the blue-green part of the spectrum that is close to but not coincident with an absorption maximum in the spectrum for chlorophyll (see Figure 2.2).

Inverse Square Law

The amount of radiation reaching Earth from the Sun is fundamental to the energy balance of the planet and, for that matter, all planets around any star. Measuring the amount of energy arriving at the top of the atmosphere with a satellite-borne detector gives a flux of $f = 1370 \text{ W m}^{-2}$. But the radiation from the Sun is distributed equally in all directions and so this figure represents the amount of energy landing on every square metre of a sphere of radius d equal to the distance between the Earth and the Sun. This distance is called one astronomical unit, or 1 AU, and is $1.5 \times 10^{11} \text{ m}$ with an error of $\pm 1.5 \times 10^3 \text{ m}$!

The amount of radiation passing through each square metre depends on how far away you are from the star. Placing the star at the centre of a sphere radiating energy in all directions, at a distance of d from the star, the area of the sphere is given by $4\pi d^2$. As the sphere gets bigger the area of the sphere increases with d^2 , so the amount of radiation on each square metre falls by $1/4\pi d^2$ – the Inverse Square Law.

Example 2.4

The calculation for flux arriving at the Earth requires the Sun's luminosity and the distance from the Sun. The total solar flux ($F_{\text{Sun}} \times$ total area of the Sun) gives solar luminosity $L_{\text{Sun}} = 3.8 \times 10^{26}$ W and the flux at the Earth, f , is given by:

$$f_{\text{Earth}} = \frac{3.8 \times 10^{26}}{4\pi (1.5 \times 10^{11})^2} \frac{\text{W}}{(\text{m})^2} \quad (2.4)$$

$$f_{\text{Earth}} = 1362 \text{ W m}^{-2}$$

Note the potential confusion between luminosity of the Sun L and the flux at the Earth. The latter is quite naturally written as the amount of radiation arriving on every square metre of the Earth's surface and analogously the flux per square metre from the black body is also F . This calculation requires the total amount of radiation emitted by the Sun to be known, which is the luminosity of the Sun and not its flux.

Planck's Law for black body radiation

The Stefan–Boltzmann Law and Wien's Law for black body radiation have been unified into Planck's Law for black body radiation, from which Planck's constant was first introduced. Planck's analysis of the spectral distribution of black body radiation led him to an understanding of the quantisation of energy and radiation and the role of the photon in the theory of radiation. The precise law relates the intensity of the radiation at all wavelengths with the temperature and has the form:

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5 [\exp(hc/\lambda kT) - 1]} \quad (2.5)$$

where λ is the wavelength in metres, c is the speed of light and Planck's constant $h = 1.636 \times 10^{-31}$ J s. Plotting the curves for a number of temperatures yields a characteristic shape for the intensity of black body radiation, as shown in Figure 2.1.

The most intense curve in Figure 2.1 is for the Sun with a surface temperature of 5780 K, showing a maximum at 501 nm as calculated from Wien's Law. Interestingly, 48 per cent of the radiation emitted from the Sun is in the visible region of the spectrum from 400 to 800 nm. However, the energy flux (as opposed to the photon flux) from the Sun depends on the energy of each photon. The energy of the photon is related to the wavelength by the equation:

$$E = \frac{hc}{\lambda} \quad (2.6)$$

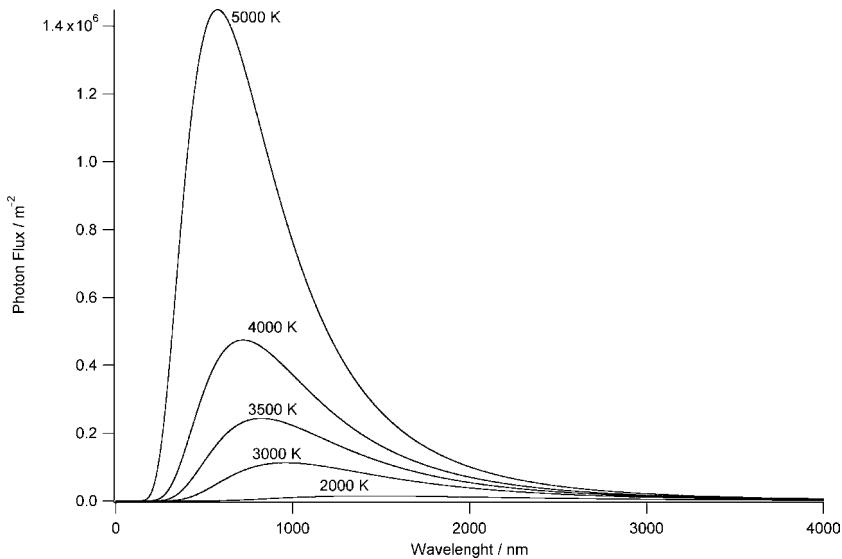


Figure 2.1 Planck curves for black bodies of different temperatures

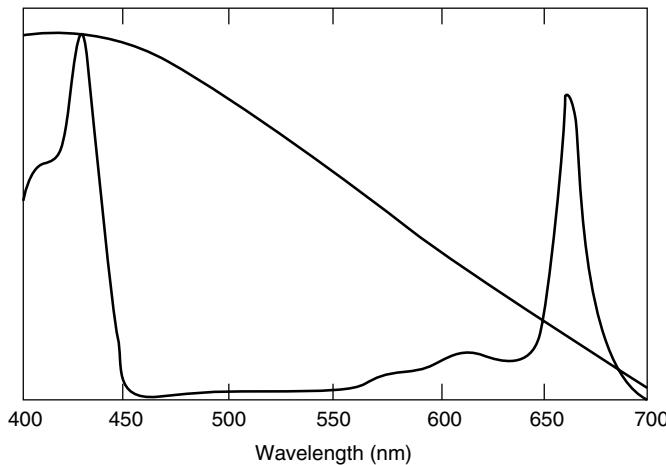


Figure 2.2 Absorption spectrum of chlorophyll overlaid with the energy flux from the Sun

where E is the energy, h is Planck's constant, c is the speed of light and λ is the wavelength of the radiation. The total energy flux from the Sun is the product of the photon flux and the photon energy, which has been overlaid on the absorption spectrum for chlorophyll in Figure 2.2.

The absorption maximum of chlorophyll at short wavelength is coincident with the maximum of the energy flux from the Sun, which might suggest that life around stars with similar temperature will also have evolved energy-trapping molecules such as chlorophyll. The longer wavelength absorption maximum is due to the

porphyrin ring structure of chlorophyll that has been optimised for the short wavelength maximum. This raises the interesting question that if a porphyrin ring is chosen as the light-harvesting structure, the ‘leaves’ on any planet around a Sun-like star will also appear green. The energy flux from the star will determine the colour of life on any orbiting planet. In short, life will look the same around stars of similar temperatures.

The black body radiation model for the continuum radiation from stars works well but it is not quite right. Careful consideration of the radiation profile shows deviations from the curves shown in Figure 2.1 due to the structure of the star itself. These deviations form the basis of a more detailed analysis including the effects of circulation within the star and will be left to others to explain; we shall use black body radiation as our model for stars.

2.2 2.725 K – cosmic microwave background radiation

One of the most impressive demonstrations of black body radiation is also a very powerful test of the Big Bang model of the origin of the Universe. The discussion of the Big Bang in Chapter 1 revealed an explosion of a ball of energy, converting the photon energy into matter and antimatter particle pairs and back again into photons. The expansion of the Universe would result in the cooling of this radiation. At around 4000 K hydrogen atoms formed and photon-matter scattering helped the cooling process, all of which suggest that there should be some of this remnant radiation left, albeit very cold. This fossil radiation should be from a pure black body and should show the Planck curve for its characteristic temperature. Various groups set about calculations of the temperature of the radiation, initially coming up with answers from 5 to 25 K; therein lies a detection problem. The black body temperature of the Earth is of order 275 K and there is hence a lot of background thermal radiation.

The development of new low-temperature detection technology and the launch of the Cosmic Background Explorer (COBE) satellite by NASA in 1989 helped to resolve this problem. The results from these observations were amazing – an almost perfect black body curve (Figure 2.3) with a black body temperature of 2.725 ± 0.002 K and a maximum wavelength of the radiation at $\lambda_{\max} = 1.05$ mm.

Now we have to be careful when making these measurements due to the relative motion of the observer with respect to a radiation source – the Doppler Effect. More of this in Chapter 3 but in general when you are moving towards the source of radiation the wavelength of the radiation looks shorter than it really is and is said to be blue-shifted, and receding from a light source makes the wavelength look longer or red-shifted. So you have to make allowances in the determination of λ_{\max} for the complete relative motion of the Earth relative to the radiation. The simplest correction is for the motion of the Earth around the Sun and then, more complicatedly, the Sun towards the centre of the Milky Way and the Milky Way towards the centre of the cluster of galaxies called the Local Group and, if that was not enough, how the Local Group is falling towards the Virgo cluster of

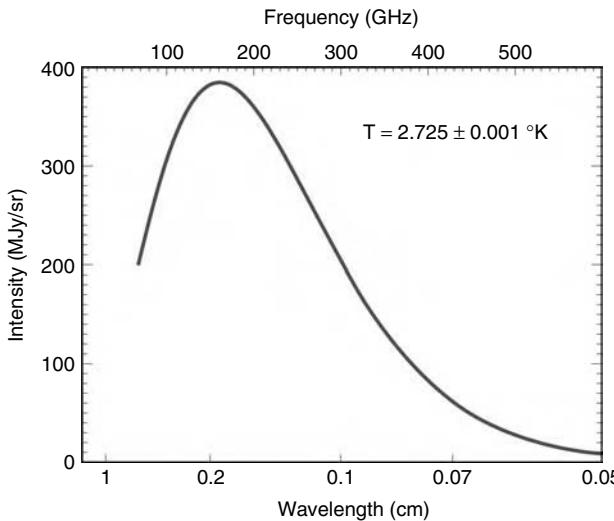


Figure 2.3 An almost-perfect black body spectrum for the cosmic background radiation. Figure courtesy of NASA/COBE Science Team

galaxies. It is curious how one scientific measurement produces estimates of lots of other quantities, especially when the first measurement has the accuracy of the 2.725 ± 0.002 K Planck curve. We now have an estimate for how fast the Local Group is falling towards the Virgo cluster: 570 km s^{-1} . We will explore more of the structure of the Milky Way and the galactic hierarchy later.

Returning to the origin of the Universe and the 2.725 K background, a further important observation is that the presence of the radiation is isotropic – it is seen equally intense in all directions, as the theory for a Big Bang predicts. This is strongly supporting evidence for the age of the Universe to be set at 15 Gyr. There are several interesting conclusions from the Planck curve: the beginning of the Universe was a quiet place and if there had been massive star formation it would have perturbed the black body curve; the Universe was hot at the start and has been cooling ever since; the radiation is *isotropic* – it is seen in all directions; and variations in the temperature are very small, of the order 1.6×10^{-5} K.

2.3 Stellar classification

Now that we have a simple model for the continuum spectrum of the stars based around the Planck curve, the temperature and the luminosity, we can make some observations and classifications of the stars. There are some constellations that dominate the night sky in both the northern and southern hemispheres and even a casual look should inspire wonder. Star hopping in the night sky should lead to the simplest observation: not all stars have the same colour. A high-quality photograph of the constellation of Orion (see page 2 of the colour plate section) shows stars

with very different temperatures and colours: these should be visible to the naked eye. The black body radiation laws tell us that this is due to the surface photosphere temperature and must contain information on the structure, age, composition and evolution of the star from which a complete classification has been derived.

Assuming that the radiation from a star follows a black body distribution, the ratio of the intensity of one colour against another is characteristic of the temperature of the star. This is achieved practically by placing a filter over the telescope and allowing only radiation from a part of the spectrum to be detected, and then blocking out another part of the spectrum and comparing the relative intensities. Astronomers choose the region from 400 to 500 nm, calling this the Blue or B region, and the band of radiation from 500 to 600 nm, calling this, for historical reasons, the Visible or V region. Hence it is possible to derive the B/V ratio as a measure of the temperature of the star. Of course the assumption underlying this temperature determination is that the star has a black body spectrum that follows the Planck curve. At this point a historical concept enters the discussion – magnitude.

Stellar magnitude

Astronomy is an ancient subject that has grown up with some curious conventions and standards, not least of which is the brightness of stars. This fundamental property is the first thing to be observed with the naked eye – clearly one star appears brighter than another. Based on the apparent relative brightness of stars to the naked eye, the astronomer Hipparchus constructed a catalogue of the stars, ranking their brightness into six categories called magnitudes. The brightest stars are of first magnitude and the faintest are of sixth magnitude. The magnitude of a star is still listed with the entries of stars in modern stellar catalogues following the tradition set by Hipparchus some 2000 years ago.

The naked eye classification is scientifically not very satisfying. Astronomers have now standardised the concept of magnitude and it remains part of the working vocabulary, however cumbersome it may seem, especially to a non-astronomer. The standardisation comes from quantitative observations showing that first magnitude stars are about 2.5 times brighter than second magnitude stars to the naked eye and that the ratio between the first and sixth magnitude stars is nearly 100. In 1850 this became the standard, so that a difference of one magnitude between two stars is $100^{1/5} = 2.512$ and some standard stars were chosen to calibrate the entire scale. The natural scientific measure of a star's relative brightness, however, is obviously the relative fluxes of radiation and the stellar magnitude difference is related to relative flux by:

$$m_1 - m_2 = 2.5 \log \left(\frac{f_2}{f_1} \right) \quad (2.7)$$

where f_1 and f_2 are the measured fluxes and m_1 and m_2 are the stellar magnitudes. Hence the two stars in Gemini, the twins called Castor and Pollux, have magnitudes

1.97 and 1.16, respectively. Plugging in the numbers into Equation 2.7 shows that Castor is emitting close to 50 per cent less light than Pollux – the brighter star has the smaller magnitude.

The introduction of the logarithm introduces a further complication: stars with very large fluxes can have a negative value for stellar magnitude. The brightest star in the sky is Sirius, which has a magnitude of -1.4 ; the full Moon is -12.73 and the Sun is -26.74 . Fainter objects have larger magnitudes so that the faintest naked-eye object in the sky is the Andromeda Galaxy at $+3.5$ but the Hubble space telescope can see down to 28th magnitude stars.

Absolute measurement of stellar flux is very hard work whereas relative measurements are easier. The flux of a number of standard stars has been measured with extreme care to act as ‘standard candles’ in a series called the North Polar Sequence. The choice of stars reflects the ancient requirement that the sixth magnitude star is just visible to the naked eye. Once the luminosity of the standard candles has been determined, all other objects in the heavens can be assigned a magnitude by using Equation 2.7 based on a relative measurement. Some examples of magnitude are listed in Table 2.1, which shows usefully when star hopping that the brightest objects in the night sky are planets and the brightest star is Sirius A: fairly easy to find from the belt of Orion.

The B/V intensity ratio is an excellent relative measure of magnitude and it is possible to derive a B/V magnitude and, using Equation 2.7, derive a calibration curve for the temperature of a star (Figure 2.4) so that the temperature of the star can be measured directly by telescopes. Now, with a measure of the luminosity of a star the radius can be determined, but there is a problem: the luminosity of a star as measured on Earth depends on how far away the star is – the Inverse Square Law – so the distance to the star must also be known to understand the absolute luminosity of the star.

Table 2.1 Magnitude of some heavenly objects (see also Appendix B)

Object	Magnitude	Absolute magnitude
<i>Solar system</i>		
Sun	-26.7	$+4.8$
Full Moon	-12.7	$+32$
Venus (maximum)	-4.3	$+29$
Jupiter	-2.6	$+26$
Mars	-2.02	$+24$
<i>Stars</i>		
Sirius A	-1.45	$+1.4$
Betelgeuse	-0.73	-6
Polaris	$+2.3$	-4.6
Brightest super giants		-8
Supernova		-19

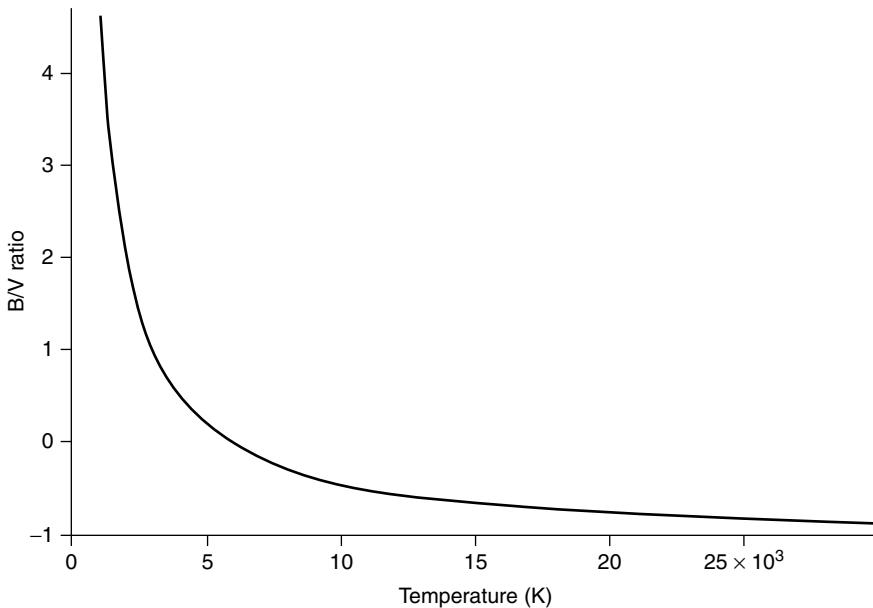


Figure 2.4 The B/V luminosity calibration curve

Distance to the stars

There is only one direct measure of the distance to the stars and this is based on the principles of parallax. Consider the observation of a star with respect to the background of stars. As the Earth moves around its orbit there is an apparent motion of the stars with respect to the background (Figure 2.5). Clearly, if the background stars are a very long way away they do not move and provide an apparently fixed background, hence nearer stars appear to move. Half the angle through which the star appears to move is called the parallax angle, given the symbol θ in Figure 2.5. Measuring the parallax angle and knowing the diameter of the Earth's orbit (and fully sophisticated calculations allow for the orbit to be an ellipse) gives the distance to the star. The sensitivity of the distance determination is limited by the accuracy of the measurement in the angle. The parallax of the star is defined as half the measured angle, as shown in Figure 2.5, and provides a unit of distance used by the astronomer called the *parsec* (pc), such that a star with a parallax of one arcsecond (arcsec) is, by definition, one parsec away. An arcsecond is a measure of the angle around a circle conventionally divided into 360° , each degree into 60 min and each minute into 60 arcsec. Hence, the distance to the star is given by:

$$d = \frac{1}{p} \quad (2.8)$$

So a star with a parallax angle of 0.1 arcsec is at a distance of 10 pc, which can be converted to SI units by knowing the Earth's orbit. One parsec is 3.26

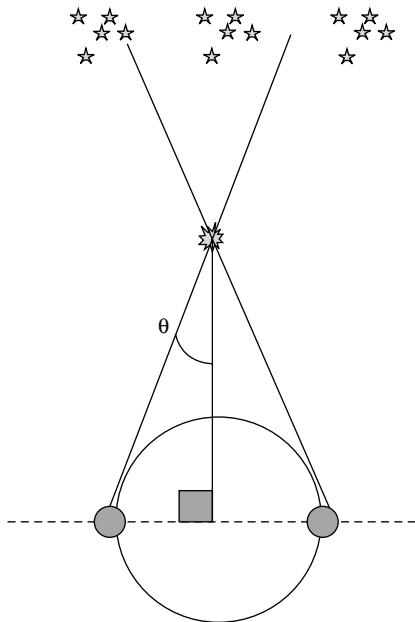


Figure 2.5 Parallax for nearby stars

light-years (ly) or 3.08×10^{16} m. The current limit for angle measurements with the telescopes is around 0.01° , which gives a length scale for this technique of $100 - 1$ pc or $326 - 4$ ly.

Example 2.5

The closest star to our Sun is one of the three stars in the α centauri system called *proxima centauri*. The distance to *proxima centauri* is 4.34 ly, so the measured angle can be calculated from Equation 2.8:

$$p = \frac{1}{4.34/3.26} = 0.751$$

$$\text{Parallax angle} = \frac{1}{0.751} = 1.331 \text{ arcsec}$$

The upper limit to the length scale is 326 ly and this seems superficially a long way, but putting this into astronomical context the Milky Way is greater than 100 000 ly in diameter. The distance measurement works well for astronomically short distances but all larger distances are inferred.

Combining measurements of the temperature and the luminosity enables the diameter of the star to be estimated. Returning to the concept of magnitude, the

distance to the star further complicates the magnitudes of stars and brings a new definition, misleadingly termed ‘absolute magnitude,’ suggesting a new measure of scale for the luminosity of a star. Rather than use the luminosity measurements directly, the magnitude scale is corrected to a standard distance of 10 pc so that all stars can be compared by their absolute magnitude at the standard distance (some absolute magnitudes are shown in Table 2.1) and of course the absolute magnitudes are a direct comparative measure of stellar luminosity. Comparison of absolute magnitudes with that of the Sun for example provides a route to the stellar luminosity.

2.4 Constellations

Constellations, their names and identifying them are not important directly to the scientific description of the Universe but they are important when making observations from Earth. The constellations are embedded in the history of astronomy and their names and conventions now form part of the astronomical terminology. To the non-astronomer this represents a significant barrier and we will spend some time de-mystifying the details. The problem derives from our very Earth-bound observations of the heavens and the striking beauty of the stars on a clear, light pollution-free night.

Celestial coordinates

Looking up from the surface of the Earth the heavens lie on the surface of a sphere and, just like the systems of latitude and longitude for defining the position of a point on the Earth, there are celestial coordinates for fixing the positions of stars. The equatorial coordinate system takes the angle from the celestial equator as the measure of *declination* of a star. Extend the line of the Earth’s equator onto the surface of the celestial sphere and this forms a line called the celestial equator, shown in Figure 2.6. The declination of the star is positive if the direction is towards the celestial North Pole, i.e. in the northern hemisphere, and negative if it is towards the celestial South Pole. The celestial poles are simply an extension of the axis of the Earth and the north and south poles on the Earth. Declination is the celestial equivalent of latitude and is straightforward – not so for the equivalent of longitude. The second angle is the *right ascension* and is the angular distance to the east of a fixed direction with the unit given in hours. The position of the star completes one full circle around the sky in 24 sidereal hours: the angle due west is 18 h 0 min 0 s and due south is 12 h 0 min 0 s. Sideral refers to the period of rotation of the Earth, which is 23 h 56 min 4.091 s. However, the fixed direction is not yet determined and this is the *first point of Aries*, which is the direction towards the Sun at the Vernal Equinox. Let us explain this concept further.

The Earth and other planets form a plane around the Sun called the ecliptic and the motion of all of the planets, including the Sun and the Moon, appears on this

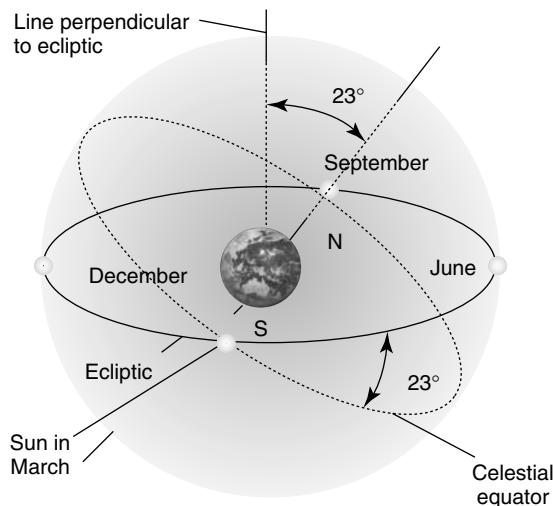


Figure 2.6 The celestial co-ordinates

line in the sky. However, the rotational axis of the Earth is not perpendicular to this plane and is inclined at 23° (Figure 2.6). During one half of the year the North Pole points towards the Sun and when the Earth has moved around to the other side of the orbit the North Pole points away from the Sun. Hence there is an apparent motion of the Sun towards the north in the sky in the northern hemisphere, reaching the summer solstice for the longest day in the summer. Conversely in winter there is a point at which the Sun reaches its lowest in the sky, called the winter solstice, and in-between there are two points when the Sun appears to be exactly on the equator and hence days and nights have equal length. In the spring on 22 March this is called the Vernal Equinox (two Latin words meaning ‘spring’ and ‘equal nights’) and on 23 September it is called the Autumnal Equinox. The direction in the sky towards the Sun at the Vernal Equinox is the first point of Aries.

The inclination of the Earth’s axis to the ecliptic has an important consideration for the differential heating of the hemispheres, producing the seasons. During the summer the longer days in the northern hemisphere provide more heat to the top half of the planet, increasing the convectional heating and the turbulence directly affecting the weather. The northerly progress of the Sun towards the solstice defines the Tropic of Cancer where the Sun will be directly overhead at 23° North latitude and will be seen in the constellation of Cancer. At the same time the Sun never sets over an area within 23° of the North Pole and this defines the Arctic Circle. Similar designations occur when the southern hemisphere has its summer, defining a region called the Tropic of Capricorn where the Sun is directly overhead 23° South of the equator and in the constellation of Capricorn. Similarly 23° south of the South Pole has a midnight sun and defines the Antarctic Circle. The Sun therefore makes a regular passage through the year against the background stars in the constellations but not through all constellations. The constellations visited by

the Sun are called the signs of the Zodiac, a word derived from the Greek for ‘zoo’ that refers to the constellations taking animal names and shapes in this region of the night sky.

Star signs

Star signs and the Zodiac are of little scientific significance but while we have our head in the heavens it is worth explaining it and completing a link. The Zodiac is shown in Figure 2.7 and the position of the Sun against the background constellations at the date of birth is the star sign, or perhaps more properly the Sun sign. A person born in November coincides with the Sun being in Scorpio, etc. – but enough of this non-science.

Ursa major

The constellations are products of mythical imagination of the ancient civilisations from around the world and we have adopted the Western names. Looking at the night sky, collections of stars are divided into constellations loosely based on patterns and images, as can be seen from a spectacular photograph of the constellation of Orion, the hunter, in Plate 3. Looking towards the north in the northern hemisphere is The Big Dipper or The Plough, all names for the constellation of Ursa Major or the Latin name for the Great Bear (Figure 2.8). Why the Great Bear?

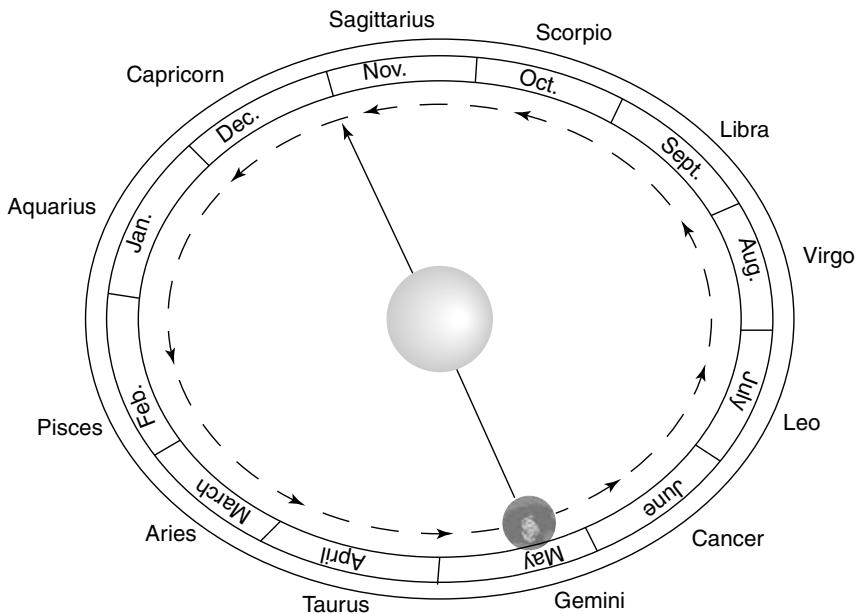


Figure 2.7 Signs of the Zodiac

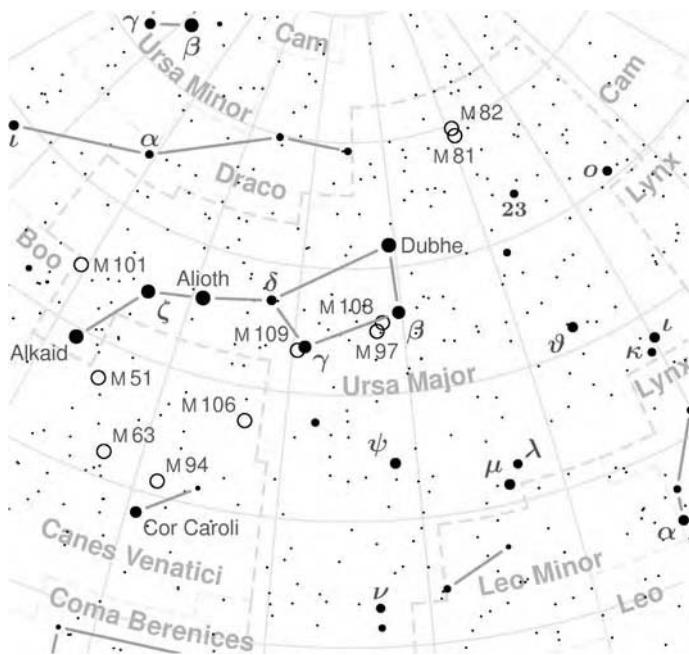


Figure 2.8 Constellation map for Ursa Major. (Reproduced by permission of Torsten Bronger)

Looking at the image of the constellation in Figure 2.9 this becomes immediately clear! The names of the constellations were translated from the Greek in the stories of Homer into Latin as the prominence of the Roman Empire grew. Stars within constellations have also been named, this time taking the Arabic translation of the original Greek. In the constellation of Orion the Hunter, the reddish star in the top left of the constellation, is called Betelgeuse, translated as ‘the armpit of the giant’.

The mythology of the constellations, such as why both Ursa Major and Ursa Minor (the great and little bears) have tails whereas terrestrial bears do not, is interesting and entertaining but there are some other important scientific labelling of stars derived from the constellations. The stars in Ursa Major are given Greek letter prefixes to denote the brightness of the stars in order: α -Ursae Majoris is the brightest in Ursa Major, taking the genitive of the Latin name of the constellations suggesting ‘the brightest star belonging to Ursa Major’. This is usually abbreviated to α -UMa (Figure 2.10).

The stars in the constellation are not next to each other in space and have no particular relation to one another, although they may be in the same galaxy. Turning to view Ursa Major side-on (Figure 2.11) reveals that the majority of the stars are between 60 and 70 ly away but η -UMa is 120 ly away. Within the constellations, however, there are other objects that are clearly not stars: the development of telescopes showed them to be clusters of stars and galaxies. These objects were catalogued by Charles Messier in the 18th century and are referred to as M1, M31,

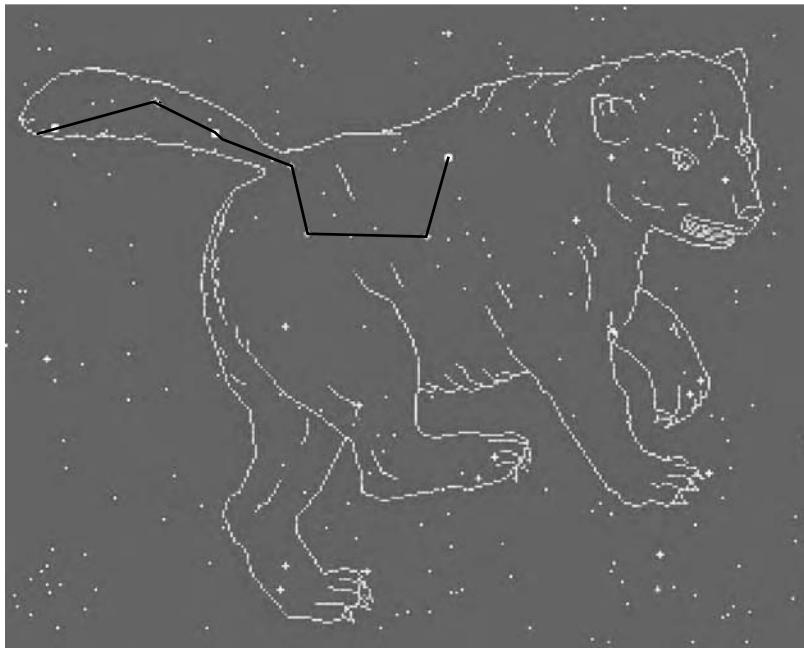


Figure 2.9 The Great Bear

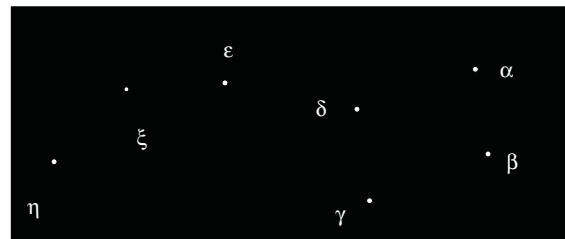


Figure 2.10 Labelling of the stars in Ursa Major

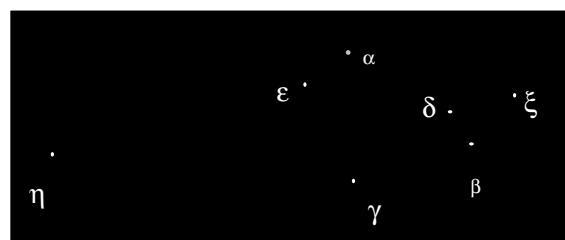


Figure 2.11 Ursa Major viewed side-on

etc. M31 also has the more familiar name of the Andromeda Galaxy, along with many of the other clusters within particular constellations. Updating the Messier catalogue in the 19th century produced the New General Catalogue (NGC) so that many objects are simply labelled with the catchy name NGC598 or M32. Two supplements to the NGC were published in 1895 and 1908, called Index Catalogues (IC), providing another string of memorable names such as IC10. All of them are rather dull names for such beautiful astronomical objects.

2.5 Galaxies

Some objects in the constellations are clearly not stars and even to the visible eye have a non-point-like structure and shape. These objects are galaxies and are divided into three broad classes following the pioneering work of Hubble: elliptical, spiral and irregular, all of which form part of the larger galactic hierarchy. Stars are collected into galaxies by the gravitational agglomeration of matter, and galaxies are collected into clusters. The image of empty space, however, remains true but with the idea that matter is clumped together in the galactic clusters. The classification has now been greatly extended and refined since Hubble's inception but we shall stick with the simpler classification. Galaxy classification essentially follows shape, with elliptical galaxies having an elliptical shape with little gas or dust associated with them. More elegant dusty, gassy structures are the beautiful spiral galaxies such as the Andromeda Galaxy (Figure 2.12) and our own galaxy the Milky Way (Figure 2.13). The final classification is a catch-all for the rest of the observed groups of stars.



Figure 2.12 Andromeda Galaxy. (A colour reproduction of this figure can be seen in the colour section). (Reproduced from photos by courtesy of NASA)

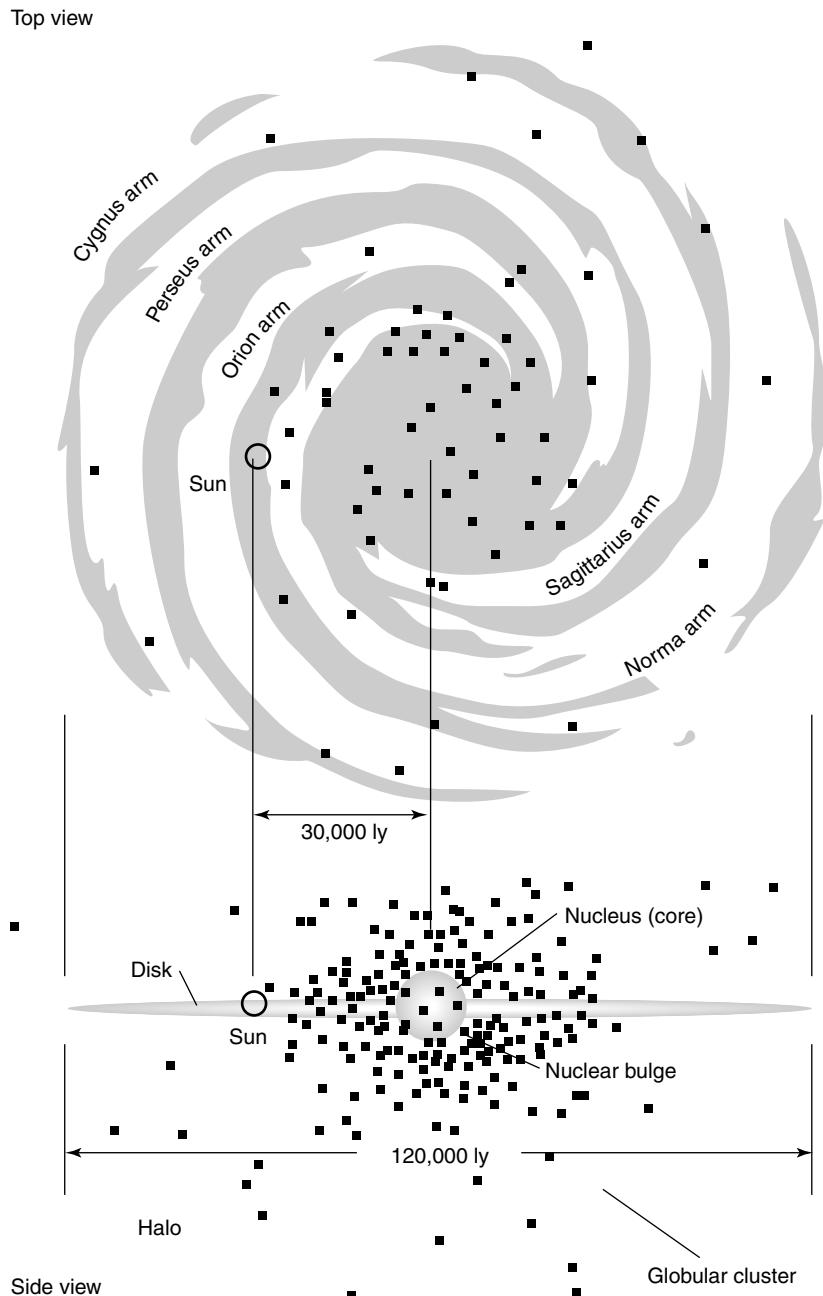


Figure 2.13 The Milky Way. (Reproduced from Figure 17.1 from “Astronomy, the Evolving Universe” 9th Edition, 2002, by permission of Boris Starosta and Michael Zeilik)

Table 2.2 Physical properties of the Milky Way

Quantity	Value
Mass	1.4×10^{11} solar masses (interior to the Sun's position)
Diameter	98 000 ly
Diameter of central bulge	33 000 ly
Diameter of the halo of stars and clusters around the centre bulge	300 000 ly (uncertain)
Thickness of the disc at the Sun's position	33 000 ly
Number of stars	4×10^{11}
Typical density of stars	1 star per ly^3 (local to the Sun)
Average density	$10^{-24} \text{ g cm}^{-3}$ or roughly one H atom cm^{-3}
Luminosity	$2 \times 10^{10} L_{\text{Sun}}$
Absolute magnitude	-20.5
Orbital period at the distance of the Sun	2.5×10^8 yr

The structure of the Milky Way is typical of spiral galaxies and we have discovered many of its properties. It is a large disc of luminous matter some 100 000 ly in diameter, about 3300 ly thick and the Sun is about 30 000 ly from the central spheroidal nuclear bulge: the physical properties of the Milky Way are summarised in Table 2.2. The galaxy is separated into six spiral arms with the Sun located in the Orion arm and the matter between the arms composed of essentially non-luminous interstellar gas. Hydrogen accounts for 96–99 per cent of the observed matter. The Sun has an orbital period around the central nuclear bulge of 2.5×10^8 years. Most interesting for the Drake equation is the number of stars, which is estimated at 4×10^{11} . The chemical composition of the Milky Way can be assessed from spectroscopy of the light coming from different regions and this will be considered in Chapter 3.

The Hubble constant, H

Having laid the ground rules for the classification of galaxies, Hubble set about measuring some of their properties, including their size, luminosity and a property called red shift. Red shift is a measure of the rate at which the galaxies are moving away from us and will be explained in Chapter 3. A plot of red shift versus distance was made initially for galaxies close by and then for galaxies further away and this plot was found to lie in a straight line. The resultant Hubble Law has the general form:

$$v = H \times d \quad (2.9)$$

where v is the radial velocity, H is the Hubble constant and d is the distance. The rate of recession of the galaxies can be determined from Hubble's Law and hence the rate of expansion of the Universe. Various estimates of the Hubble constant have been made by Hubble himself and others, including some very recent measurements

from the Hubble space telescope based on the distance of Cepheid stars (discussed in Chapter 4); the best estimate to date is $23 \text{ km s}^{-1} \text{ Mly}^{-1}$ (or the equivalent unit s^{-1}) with an error of around 10 per cent. Other data from supernovae measurements suggest a range of $19\text{--}21 \text{ km s}^{-1} \text{ Mly}^{-1}$. Let us take $20 \text{ km s}^{-1} \text{ Mly}^{-1}$ as a good number and from this calculate the age of the Universe.

With the assumption, and it is a big assumption, that the rate of expansion of the Universe has been constant, then the time of expansion so far is given by:

$$\begin{aligned} t &= \frac{d}{v} \\ t &= \frac{d}{Hd} \\ t &= \frac{1}{H} \end{aligned} \tag{2.10}$$

The age of the Universe is the inverse of the Hubble constant. The system of units must first be rationalised to convert all of the light-years into kilometres and then years, so that:

$$\begin{aligned} t &= \frac{1}{H} \\ t &= \frac{1}{20 \text{ km s}^{-1} \text{ Mly}^{-1}} \times 10^6 \text{ ly Mly}^{-1} \times 10^{13} \text{ km ly}^{-1} \\ t &= 4.5 \times 10^{17} \text{ s} \\ t &= \frac{4.5 \times 10^{17} \text{ s}}{3.3 \times 10^7 \text{ s yr}^{-1}} \\ t &= 1.5 \times 10^{10} \text{ yr} \end{aligned} \tag{2.11}$$

which gives an estimate for the age of the Universe of 15 billion years. The consequences of the assumptions need a little thought. At the beginning of the Big Bang the expansion rate would have been quicker than now and perhaps it has been slowing to the current rate, so the figure for the age of the Universe represents an upper limit. Of course this makes a big statement about the rate of expansion of the Universe: is it expanding or contracting, and is the rate of expansion constant or accelerating? These are questions at the frontier of cosmology.

The local group

The Milky Way belongs to a cluster of galaxies called the Local Group that occupies about 3 Gly in space and contains over 30 galaxies, although the precise number is hard to determine. Placing the Milky Way at the centre (Figure 2.14),

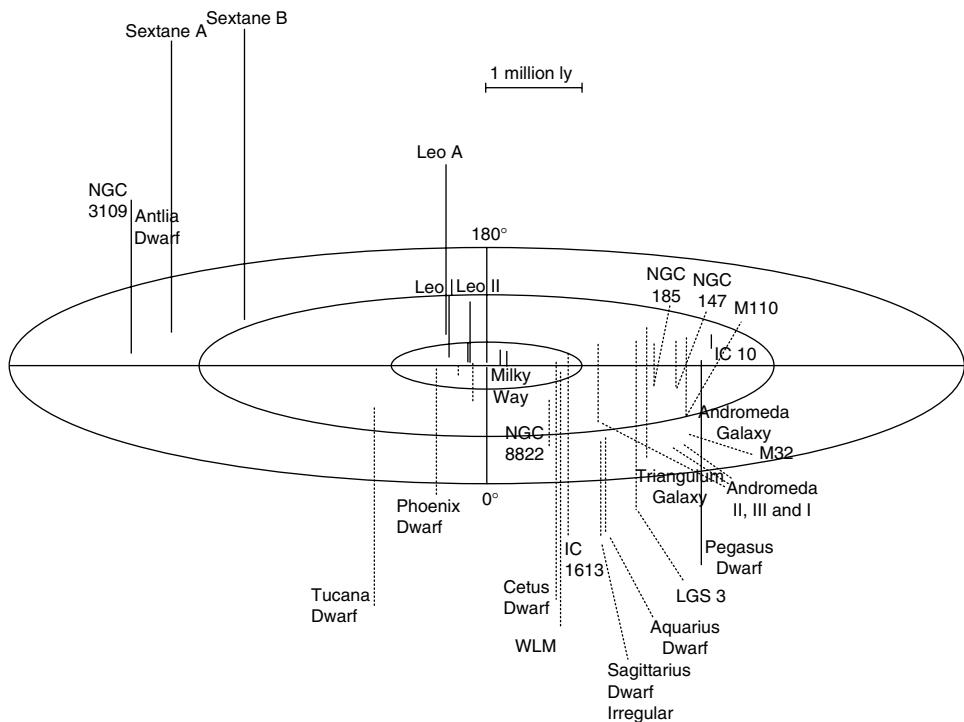


Figure 2.14 The Local Group

the Andromeda Galaxy is out on one extreme and these two galaxies orbit one another, trapping the other galaxies within this motion.

The Large and Small Magellanic Clouds (LMC and SMC) are prominent in the night sky of the southern hemisphere and are closest to the Milky Way at distances of 170 000 ly and 200 000 ly, respectively. A tenuous strip of hydrogen and a thread of stars connects the LMC and SMC, with the LMC containing about 20×10^9 solar mass of stars. The SMC is about a quarter of the mass but both have a similar population of young hot stars. The exciting observation of a supernova event called 1987A took place in the LMC and provided a rare look at the death of a star.

The Andromeda Galaxy some 2.2 Mly away and the halo of globular clusters make up a vast collection of stars within the galaxies of the Local Group. There is nothing to indicate that the Milky Way is unique in any particular form or has any properties that would confer on our Sun and the Earth anything so special as life. There are literally billions and billions of stars in our local neighbourhood and recent surveys of these suggest that most stars have some form of planetary system around them and are therefore potential homes for life. The size of the Universe and the numbers of galaxies and stars provide embarrassing flexibility for the Drake equation but it would be hard to believe that Earth is the only planet supporting life.

2.6 Cosmology

The overall study of the Universe is cosmology and addresses the concepts of the structure of space-time and the evolution of the Universe from the Big Bang, including the evolution of the four fundamental forces of nature: weak and strong nuclear forces, electromagnetism and gravity. Cosmology curiously but perhaps inevitably links all four in a demand for a grand unified theory or Theory of Everything (TOE). The study of galaxy clusters has led to observations of luminous arcs that are probably the largest objects observed in the Universe (Figure 2.15). The arc is as luminous as 10^{11} Suns and is some 300 000 ly long: three times the length of the Milky Way.

What causes the arc is not known precisely but it is common to a number of galaxy clusters and thought to be a lensing effect associated with the huge mass. Gravitational lensing occurs when light can be seen from behind a massive object because the light is bent around the mass due to the distortion of space-time. This was predicted by Einstein and has been observed directly several times. The luminosity of the arc is thought to derive from the light of stars and galaxies behind the galaxy cluster in the foreground that has been bent by the large gravitational field of the foreground cluster – like light dancing off the side of a cosmic wine glass. The problem comes with the calculations of the required mass for the lensing effect: it is simply much larger than the mass that can be seen in the galaxies. This large unseen mass has been termed dark matter and in some theories may account for some 99 per cent of the mass of the Universe. The amount of matter in the Universe determines whether it is contracting, expanding or remains constant in dimension over time. The descriptions of the Universe are intimately linked with the



Figure 2.15 Luminous arc located near the galaxy cluster 2242-02. (Reproduced from photos by courtesy of NASA)

descriptions of gravity and ultimately the TOE that will unite quantum mechanics and general relativity.

Planck, in an attempt to rationalise the problems of quantum mechanics and gravity, noted that a unit of length could be derived from the fundamental constants that appear throughout general relativity:

$$l_P = \sqrt{\frac{hG}{c^3}} \quad (2.12)$$

where h is Planck's constant, G is the gravitational constant and c is the speed of light. The unit of length is now called the Planck length, l_P , and is 1.6×10^{-35} m. It is this rather small length scale that defines the distance at which gravity becomes quantum mechanical. It also defines the Planck time of 10^{-43} s, which is the time it takes a photon to travel the Planck length and is the period of uncertainty at the beginning of the Universe. A number of Planck parameters have been noted, all connected with the scales over which gravitational attraction is quantised. The current best theory to unite quantum mechanics and gravity is *super string theory* and the Planck parameters are natural quantities in this theory.

Super string theory says that particles can be represented as vibrations in strings that are a Planck length long and under Planck tension, which is something like 10^{39} tons. Vibrations in these strings have different patterns, just like the allowed vibrations of a violin string, with each different resonance corresponding to different particles, whether an electron, photon or the particle thought to be responsible for the gravitational interaction, the graviton. The goal of string theory is to find a unifying description that will eliminate the 19 seemingly arbitrary numbers required to understand the Universe. The 19 numbers, including the constants of force interaction such as G , the charge and mass of the proton, Planck's constant and the speed of light, are in fact a rather limited set of numbers but nevertheless are apparently random significant numbers. The story of cosmology and the structure of matter will require observations both in astronomy and in particle physics to unlock the secrets of the Universe. Alas, we will not spend time discussing them further here: we must concentrate on the matter we can see and the life that has evolved from it.

Concepts and calculations

Concepts

<i>Black body radiation laws</i>	The ability of a body at a characteristic temperature to absorb and emit radiation at all wavelengths with equal and unit efficiency. The general application of all of the laws of black body radiation to the description of stars
<i>Planck's Law</i>	This law describes the complete wavelength–intensity profile of a black body characterised by a single temperature
<i>Light harvesting</i>	Light-harvesting molecules on a planet around a local star have developed an absorption spectrum to collect light at the maximum flux from the local star
<i>Cosmic microwave background</i>	The amazing result following measurement of the fossil radiation from the Big Bang
<i>Stellar magnitude</i>	The ideas of stellar brightness as observed in the night sky compared with the luminosity of the star
<i>Galaxies</i>	The three broad classifications (elliptical, spiral and irregular) of star clusters that also cluster together to form the Local Group that contains the Milky Way and the Andromeda Galaxy, along with the Small and Large Magellanic Clouds
<i>Cosmology</i>	The hierarchy of the Universe from planets and stars, through galaxies and clusters to the structure of the Universe itself. The search for a grand unifying theory to unite quantum mechanics and gravity

Calculations

<i>Black body laws</i>	Stefan–Boltzmann Law, luminosity, Wien's Law
<i>Stellar magnitude</i>	Use of the logarithmic scale relating radiation flux and magnitude difference
<i>Distance to the stars</i>	Parallax calculations are the only absolute measure of distance
<i>Hubble constant</i>	Calculations of red shift and the age of the Universe lead to a value of $H = 20 \text{ km s}^{-1} \text{ Mly}^{-1}$ and the age of the Universe at 15 billion years

Problems

Data for these problems can be found in Appendix B.

- 2.1** Assume that the stars Betelgeuse, Aldebaran and Rigel behave as black bodies:
 - (a) Calculate the maximum wavelength in the Planck curve for each star.
 - (b) Calculate the flux of radiation from each star.
 - (c) Calculate the luminosity of each star.
 - (d) Compare your results with the properties of the Sun.
- 2.2** The brightest star in the northern hemisphere is *Canis Major*, one of Orion's dogs, the other being *Canis Minor*. The luminosity of *Canis Major* is $26 L_{\text{Sun}}$.
 - (a) Calculate the surface temperature, assuming that it has the same radius as the Sun.
 - (b) Calculate λ_{max} .
 - (c) Calculate the energy of the photons at λ_{max} in joules and in kilojoules per mole.
 - (d) Calculate the stellar flux, f , at 1 AU from the star.
 - (e) Comment on the adaptation of life on a planet around *Canis Major* with this incident stellar spectrum.
- 2.3** Construct a spreadsheet (or write a program) to calculate the proportion of light leaving a star in the visible wavelengths. Compare the calculations for the Sun with those for Betelgeuse, Vega and Deneb.
- 2.4** Modify the spreadsheet to calculate the energy flux from each star and hence identify the wavelength most suited for energy harvesting in each case.
- 2.5** Black bodies emit radiation at all wavelengths. Using the spreadsheet, calculate the contribution to the visible spectrum from the cosmic microwave background in the visible region of the spectrum.
- 2.6** The distances to Betelgeuse, Aldebaran and Spica are given in Appendix B. Calculate the distance to these stars in parsecs and identify which distances can be measured on Earth using parallax.
- 2.7** Calculate the ratio of the stellar fluxes of Rigel and Spica on Earth and calculate the difference in magnitude of these stars. Which is the brighter? If Betelgeuse has an absolute magnitude of -6 , calculate the absolute magnitudes of Rigel and Spica.
- 2.8** Calculate the ratio of the photon flux from the Sun and the Full Moon on Earth and, using the measured solar flux on Earth, calculate the flux from the Full Moon on Earth.
- 2.9** If Betelgeuse has an absolute magnitude of -6 , using the information in Appendix B calculate the absolute magnitude of Capella and Deneb.
- 2.10** A spacecraft leaves Earth and the solar system; calculate the distance at which the Sun will no longer be visible to the naked eye from the spacecraft.

3

Atomic and molecular astronomy

Introduction

The colour of a star and the general emission spectrum is close to that of a black body with a temperature characteristic of the star. The black body approximation works well for young and middle-aged stars but not so well when the star is older and contains many more elements. The black body spectrum then shows many absorption features, as seen in Figure 4.2, with older stars containing more elements, as we shall see in Chapter 4. For the moment we will concentrate on what is causing the absorption features and what this tells us about the structure of matter and then move on to the understanding of the star. The emission and absorption of radiation by atoms or molecules is diagnostic both of the structure of matter and the interaction of radiation with matter. This general area of investigation is called spectroscopy.

The study of spectroscopy has provided all of the information required to make a positive identification of molecules in space. More interestingly, once the spectrum of a molecule or atom is understood accurately, the interaction of the molecule with its surroundings can be understood as well. Atoms and molecules, wherever they are, can report on their local conditions and be used as probes. We shall see many of these examples where knowledge of molecular properties provides insight into astrochemistry. For example, the understanding developed below will take us from the transition wavelength of H α to the radius of Jupiter.

3.1 Spectroscopy and the structure of matter

The observation of atomic spectra stimulated physicists in the early 19th century to develop the theory of quantum mechanics. This theory sets out to explain all physical phenomena at an atomic scale and atomic spectroscopy is an important validation. Quantum mechanics is flawed, however, notably in the description of gravity, but it is the best theory at present (although super string theory promises well) for the description of the structure of nuclei, atoms and molecules.

Quantum mechanics was formulated by some of the greatest scientists of all time (Einstein, Schrodinger, Heisenberg and Born) and famously they all disagreed about the interpretation of the theory, particularly its probabilistic view of the Universe. It centres on the complete description of matter by a mathematical function called a wavefunction. The theory postulates that all the measurable properties of a system that you would want to know can be extracted from the wavefunction by performing the right operation or mathematical manipulation on the wavefunction. This leads to a series of mathematical tools called operators that, when applied to the wavefunction, produce the required operation. So flippantly, but not incorrectly, the collar size of the reader would be determined by first working out the total wavefunction of the reader and then operating on this wavefunction with the collar-size operator. Unfortunately, quantum mechanics does not quite deliver what you would expect in that you cannot calculate the precise answer, only an expectation value.

Intrinsic to quantum mechanics is the concept of probability and all expectation values embody this idea. We can only know the expectation value of a bond length in a molecule and an expectation value of the energy of a system or an allowed energy level, never a precise number. Bringing these ideas together for the structure of matter produces, as a natural consequence of quantum mechanics, allowed quantum states within atoms and molecules. Interrogating the wavefunction with the Hamiltonian operator produces a series of allowed energy levels called eigenvalues, and the wavefunctions appear as eigenfunctions each described by a unique set of quantum numbers. The interaction of radiation with matter is an understanding of the quantum leaps between the quantum states within matter induced by the absorption of the photon. The general area of study is called spectroscopy and produces an extremely detailed and elegant understanding of the structure of matter.

The Beer–Lambert Law

The intensity of an absorption line in the spectra of atoms or molecules can be measured and it follows the empirical law called the Beer–Lambert law. Consider the amount of light entering a cylinder full of hydrogen gas, as shown in Figure 3.1. If the radiation is absorbed by the hydrogen gas, the amount of light emerging from the opposite end of the cylinder is reduced.

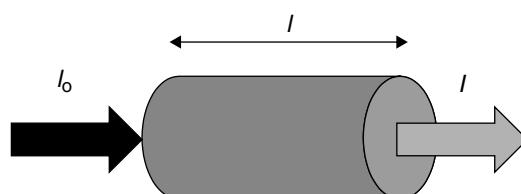


Figure 3.1 Absorption of radiation

The absorption of the radiation has been found empirically to increase so that a longer sample absorbs more light, as does a more concentrated sample, and it is found to vary exponentially:

$$I = I_0 10^{(-\varepsilon cl)} \quad (3.1)$$

where ε is the molar extinction coefficient ($M^{-1} \text{ cm}^{-1}$), c is the concentration (M) and l is the length (cm). The molar extinction coefficient is a measure of how strongly the radiation is removed from the beam and the dominant molecular mechanism is absorption by the molecule. The weaker way to remove light from the beam is for the molecule to scatter the radiation and this process does occur, however it can be four orders of magnitude weaker than the absorption event. The difference is that for absorption to occur the radiation energy has to be the same as a gap between the quantum states of the atom and is said to be resonant, whereas the scattering process occurs at all wavelengths with varying degrees of efficiency.

Some simple rearrangement of Equation 3.1 leads to the concepts of transmission $T = I_0/I$ and absorbance $A = -\log T$, with the quantity $\varepsilon c l$ called the optical density. The choice of units here for the extinction coefficient ($M^{-1} \text{ cm}^{-1}$) is appropriate for measurement of the absorbance of a solution in the laboratory but not so appropriate for a distance l of astronomical proportions. The two terms ε and c are contracted to form the absorption per centimetre, α , or, more conveniently (confusingly) in astronomy, per parsec. The intrinsic ability of a molecule or atom to absorb light is described by the extinction coefficient ε , and this can be calculated directly from the wavefunction using quantum mechanics, although the calculation is hard.

Example 3.1

The extinction towards a star at 450 nm is 0.24 and is attributed to an interstellar cloud containing dust particles with an extinction coefficient of 0.0032 pc^{-1} . Calculate the diameter of the intervening molecular cloud, expressing your answer in light-years.

$$\begin{aligned} \frac{I}{I_0} &= 10^{(-0.0032l)} \\ \log_{10}(0.24) &= -0.0032l \\ l &= \frac{-\log_{10}(0.24)}{0.032} \\ l &= 194 \text{ pc} = 631 \text{ ly} \end{aligned} \quad (3.2)$$

The empirical law of absorbance is all well and good but what are the atomic and molecular properties that control how much electromagnetic radiation is absorbed by a specific transition?

The interaction of radiation with matter

The interaction of radiation with matter is a huge and fascinating field of theoretical and experimental research and combines the ideas of quantum mechanics with the wonders of special relativity. These two great theories of modern science are required for a detailed description of the interaction of light with matter and the decoding of messages within the radiation. The simple picture is of a photon of light as a particle of both electric and magnetic field that oscillates in time but perpendicular to the direction of travel. There is, of course, the other simple picture of a photon, which is as a wave with a known wavelength. The wave–particle duality is an interesting problem that can be left to another time save to say that when making measurements on light if you ask a wave-type question you get a wave-type answer such as interference or diffraction, conversely, should you ask a particle-type question then you get a particle-type answer such as measuring the photoelectric effect. Interested readers should research Young’s double slit experiment with electrons, looking especially at what happens with low electron fluxes.

Either as a wave or a particle there are a number of possible interactions of photons with matter and for the moment we consider the atom, although any matter will do. The list of possible interactions is as follows (although it is not complete):

- elastic scatter
- non-elastic scatter
- stimulated absorption
- stimulated emission
- spontaneous emission

Elastic scatter involves the interaction of a photon with an atom that does not change the energy of the photon and conversely non-elastic scatter allows the photon to be absorbed by the atom and re-emitted at a different wavelength so that the atom absorbs some energy or gives additional energy to the photon. Absorption allows for the photon to be absorbed completely and destroyed, or formally annihilated, and the atom then absorbs both the energy and the momentum of the photon. Stimulated emission requires the atom to be in an excited state and when a photon of the right energy/wavelength passes by, the atom is stimulated into the emission of a photon of exactly the same colour as the passing photon. Finally, an atom may be in an excited state and spontaneously creates a photon, which is emitted, relaxing the atom to a lower energy state. All of these processes have to be considered when considering a spectroscopy problem in astronomy and we shall see examples of them all.

Einstein coefficients

An important quantity that needs to be known from an astrochemical observation is how much light is being removed by a particular species along the line of sight. If we know how many photons are being emitted by a star at a fixed wavelength resonant with a transition in the molecule, then we can calculate the number of molecules that must be present along the line of site. This quantity is called a column density and to know this number accurately we need to know the number of photons removed per molecule – the intrinsic strength of the transition, the oscillator strength, and the path length.

Einstein had an important role to play in the description of absorption with the development of the theory associated with the Einstein coefficients. Consider a transition from a low energy state to a higher energy state with a transition rate w given by:

$$w = B\rho \quad (3.3)$$

where B is the Einstein B coefficient ($\text{J}^{-1} \text{m}^3 \text{s}^{-2}$) and ρ is the density of the radiation at the frequency of the transition ν . Around a star, the Planck distribution given in Chapter 2 can be used in conjunction with the least-squares law to calculate the radiation density at a known distance from a star. The total amount of radiation removed by the intervening atoms is the transition rate for one atom multiplied by the number of atoms. Einstein also reasoned that if the atom was already in an excited state and radiation of the correct frequency was present then it could induce the downwards transition in a process called stimulated emission. The stimulated emission has to compete with any spontaneous emission: the natural rate at which the excited state will decay. So by analogy with Equation 3.3 Einstein wrote:

$$w = A + B\rho \quad (3.4)$$

where A is the Einstein coefficient of spontaneous emission. Einstein went on to show that the A and B coefficients are related by the expression:

$$A = \left(\frac{8\pi h\nu^3}{c^3} \right) B \quad (3.5)$$

The rate of spontaneous decay increases with ν^3 so that for higher frequency transitions, such as in the visible region of the spectrum, spontaneous decay is fast (of the order of nanoseconds) whereas rotational transitions or transitions between hyperfine levels within an atom are very slow.

Selection rules

Einstein coefficients provide good empirical relations between the rate of a transition and the density of radiation but quantum mechanics has something to say about

the rate or intensity of the transition from the fundamental structure of the atom or molecule. The atom or molecule interacts with the electric and magnetic field of the photon via the electron distribution, which must have, at least transiently, a positive and a negative end: in a general structure this is an electric dipole. The dipole associated with the transition is expressed quantum mechanically by the quantity called the transition moment, defined by the integral:

$$\mu_{fi} = \int \psi_i^* \hat{\mu} \psi_f d\tau \quad (3.6)$$

where ψ_i is the wavefunction for the initial state, ψ_f is the wavefunction for the final state and $\hat{\mu}$ is the dipole moment operator. This one integral defines spectroscopy. There are times when the integral has to be zero by symmetry and there are times when it is non-zero, producing the idea of allowed and forbidden transitions. The allowed transitions occur between allowed values of the quantum numbers and the allowed changes in the quantum numbers are called selection rules. A complete analysis then relates the rate of transition or intensity to the Einstein coefficient by:

$$B = \frac{|\mu_{fi}|^2}{6\epsilon_0 h^2} \quad (3.7)$$

Underneath all of the ideas of atomic and molecular detection, counting the number of molecules in a particular line of sight, requires the intensity of the transition to be calculated via the transition moment to the Einstein B coefficient. If the total photon flux through a sample is known and the transition moment is also known, then the absolute number of atoms or molecules present can be determined.

3.2 Line shape

The idea of a transition between two energy levels suggests that the transition will occur at only one precise frequency as a sharp spike in the absorption or emission spectrum. This is not the case and, in fact, the transitions have an intrinsic width and shape containing information about the local environment of the atoms. The line profile of an atomic transition has contributions from three effects:

1. Transitions have a natural linewidth associated with their lifetime (via the uncertainty principle) but this is usually small.
2. Pressure broadening allows for the presence of nearby molecules to perturb the positions of the energy levels, leading to a spreading of the transition frequency as a function of collisions and hence pressure.
3. Doppler broadening allows for observation that some molecules in a gas cloud may be moving towards the observer and some away from the observer in a line of sight.

Natural linewidth

The natural linewidth comes from the lifetime, τ , of the upper state of a spontaneous transition, which is related to the Einstein A coefficient so that $\tau = A^{-1}$: faster transitions have shorter lifetimes and vice versa, and similarly an allowed transition will have a short lifetime for the upper state whereas forbidden transitions will have a long lifetime. The lifetime consideration is very important in the laboratory where transitions have to occur on the timescale of the experiment, otherwise they are not observed. Hence in the laboratory allowed transitions are observed and in general (but not specifically) forbidden transitions are not seen. For astronomy this does not matter. So what if a forbidden transition has a lifetime of 30 million years – the Universe is 15 billion years old – if you wait long enough it will happen. The rules of spectroscopy need to be understood but in space anything goes!

Quantum mechanics enters here with a statement of uncertainty relating energy and time. If you know the lifetime of the excited state in a transition then you cannot know exactly the energy of the transition. This uncertainty principle is wrapped up in the following relation:

$$\Delta E \Delta \tau \geq \hbar \quad (3.8)$$

where ΔE is the uncertainty in the energy and $\Delta \tau$ is the uncertainty in the lifetime: the product of these two uncertainties is fixed at $h/2\pi = \hbar$. This is a very curious and counter-intuitive result and we do not have time to go into why this happens but all textbooks on quantum mechanics will have a discussion of Uncertainty Principles. The upshot is that all transitions are naturally broadened by the lifetime of the upper state, so-called lifetime broadened, and the uncertainty in the energy allows the distribution of photon energies to be absorbed by the atom or the molecule. The spread in transition wavenumber (the reciprocal of the wavelength in cm) is related to the lifetime of the excited state in picoseconds by

$$\delta \tilde{v} = \frac{5.3}{\tau} \frac{\text{cm}^{-1}}{\text{ps}} \quad (3.9)$$

The natural linewidth is the smallest contribution to the line profile of a transition and is only rarely seen as limiting within the laboratory. For an electronic transition with a lifetime of 10 000 ps the linewidth is of order 0.00053 cm^{-1} but for a rotational transition the lifetime linewidth $5.3 \times 10^{-15} \text{ cm}^{-1}$. The best microwave spectra recorded in the laboratory have a linewidth of a few Hz or 10^{-12} cm^{-1} , which is close (but not very) to the natural linewidth limit.

Pressure broadening

Atoms or Molecules in the gas phase are not alone and they collide with one another. The energy levels involved in a particular transition are no longer associated with one atom but transiently with those of the two colliding atoms. The

collision shifts the energy levels slightly and hence changes the width of the transition. Increasing pressure causes the molecules to collide more often, with the converse being true. Contributions to the linewidth of a transition from pressure broadening can also be calculated but we will not develop this further here. Consider a transition that is induced every time there is a collision. The time between collisions in a gaseous sample produces an uncertainty in time or a collision lifetime of the transition. The spread in frequency is given by:

$$\Delta\nu = (2 \pi \tau)^{-1} \quad (3.10)$$

from the uncertainty relation given in Equation 3.8. This simple analysis is the starting point for the more complete analysis of Lorentz and Debye. Astrochemically, pressure broadening can be irrelevant in the interstellar medium where molecules are essentially alone, but more important when molecules are in the atmosphere of a planet.

Doppler shift and profile

The natural linewidth would be limiting if the molecule were not moving but all molecules are in motion, either moving away from the observer or towards them. The apparent shift in frequency or wavelength of a transition that is seen in all stars is called the Doppler shift. This is the same shift that is observed when a police car is driving towards an observer with the sirens blaring. On approach the notes seem to get higher and when it has gone past the note seems to get lower. When talking about an optical transition in the visible range in, say, the spectrum of a star, the transition frequency of an atom can be shifted to longer wavelength when the star is receding or shorter wavelength when the star is approaching. These two effects are known as:

- **Red shifted** – Star is moving away from the Earth and the transition is shifted to longer wavelength.
- **Blue shifted** – Star is moving towards the Earth, and the transition is shifted to shorter wavelength.

Mathematically the shift in the wavelength due to the relative velocity is given by the expression:

$$\frac{\Delta\lambda}{\lambda} = \frac{V_{source}}{c} \quad (3.11)$$

where λ is the wavelength of the transition, v_{source} is the relative velocity of the source and c is the speed of light. This simple formula works well for velocities that are significantly less than the speed of light, however, as velocities approach $0.7c$ a relativistic correction needs to be made. In some circumstances in astronomy this may be important. Typically and more slowly, with a light source receding at

30 km s^{-1} and a transition wavelength of 500 nm there is a Doppler shift in the wavelength of the transition from the laboratory value of $\Delta\lambda = 0.05 \text{ nm}$, which is quite easily measured.

Even in the gas phase some atoms or molecules will be moving towards the source of the light and some will be moving away, each suffering the same Doppler shift in the wavelength of the radiation. The velocity of the molecules within the gas sample is a measure of their kinetic energy, which reflects the local temperature of the molecules. The effect of the Doppler shift in each direction, averaged over all the molecular velocities at a given temperature, is to create a Gaussian line shape with a linewidth given by:

$$\delta\lambda = \frac{2\lambda}{c} \left(\frac{2kT \ln 2}{m} \right)^{1/2} \quad (3.12)$$

where λ is the wavelength of the transition, k is the Boltzmann constant, T is the local temperature and m is the mass of the molecule. For a rotational transition of N_2 at room temperature the Doppler contribution to the linewidth of the transition is 70 kHz and is easily observed.

The Doppler shift has an important place at the heart of astronomy, as we saw with the relation between the red shift and the distance of a star related by the Hubble constant in Chapter 2. Leaving cosmological problems aside, a very powerful demonstration of the Doppler effect can be seen with Jupiter. Jupiter is a massive droplet of hydrogen and helium, a gas giant, which can be seen from the Earth with the naked eye but binoculars and telescopes reveal more beautiful structures, including the famous red spot. Simple astronomical telescopes can measure the reflection spectrum of the cloud tops of the planet, specifically including the spectrum of the hydrogen atom (Figure 3.2). Jupiter is not however stationary and it rotates about its axis much the same way as the Earth does. However, looking at the spectrum of the hydrogen atom, specifically the $\text{H}\alpha$ line (see Section 3.4), the Doppler shift from the rising and descending edges of the planet can be measured, from which we can deduce the speed and direction of rotation. The edge moving towards the Earth will have a blue-shifted $\text{H}\alpha$ line whereas the edge moving away from the Earth will have a red-shifted $\text{H}\alpha$ line. Hence from knowing the exact frequency of the transition the Doppler shift gives the rate of rotation of Jupiter.

Jupiter does not rotate at the same speed at all latitudes, similarly on Earth. Jupiter rotates fastest about the equator with a linear speed at the limb of some $43\,000 \text{ km h}^{-1}$ but much slower towards the poles. This suggests that the planet is fluid and has no obvious solid core – although strictly we should say the cloud tops have this differential rotation which leads to strong winds in the atmosphere running along permanent regions of high and low pressure creating ‘jet streams’ with wind speeds as fast as 100 m s^{-1} , some three times faster than those on Earth. The atmosphere is a turbulent place and the rapid mixing dynamics account for the red spot. The red spot is a giant high-pressure zone that rotates once every 7 days, leaving in its wake a series of eddy currents. But what causes its red colour?

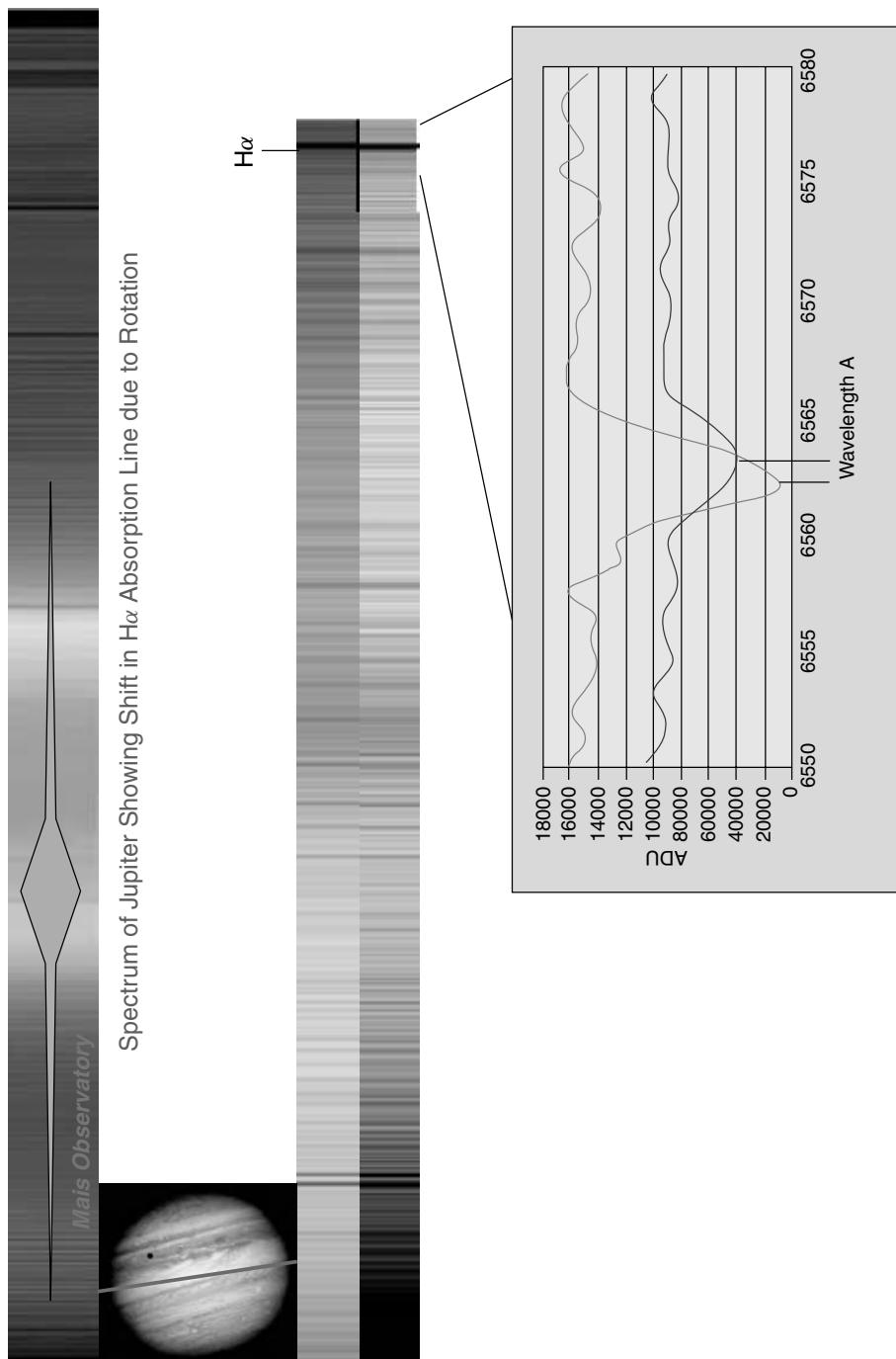


Figure 3.2 The reflection spectrum of Jupiter. (Reproduced by courtesy of Dale Mais, Mais Observatory)

Example 3.2

Calculate the Doppler shift that should be observed in the hydrogen atom transition at 656.300 nm for the ascending and descending limbs of Jupiter at the equator.

The linear speed of the limb of Jupiter (from the text) is $43\,000 \text{ km h}^{-1}$ and so the Doppler shift is given by:

$$\frac{\Delta\lambda}{\lambda} = \frac{V_{source}}{c}$$

$$\Delta\lambda = \frac{43 \times 10^6}{(60 \times 60)} \times \frac{656.3 \text{ nm}}{2.99 \times 10^8} \quad (3.13)$$

$$\Delta\lambda = 2.62 \times 10^{-2} \text{ nm}$$

For the ascending limb moving towards the Earth the wavelength is blue-shifted to 656.273 nm and for the descending limb moving away from the Earth the red shift is 656.326 nm. These Doppler shifts place some demands on the resolution of the spectrometer but, as can be seen from Figure 3.2, this is possible for modern telescopes.

Transition intensities and number densities

We have discussed the transition moment (the quantum mechanical control of the strength of a transition or the rate of transition) and the selection rules but there is a further factor to consider. The transition between two levels up or down requires either the lower or the upper level to be populated. If there are no atoms or molecules present in the two states then the transition cannot occur. The population of energy levels within atoms or molecules is controlled by the Boltzmann Law when in local thermal equilibrium:

$$I = N \frac{g_1}{g_2} e^{-\Delta E/kT} \times \text{Transition moment} \quad (3.14)$$

where N is the number of molecules, g_i is the degeneracy of the levels involved in the transition, ΔE is the separation between the levels and T is the temperature. The degeneracy of a level is the number of levels in an atom or molecule that have the same energy. The population of these levels is then equally probable and a set of degenerate levels will contribute more to the population of the molecule at that energy. There is one big catch when using the Boltzmann Law: the atom or molecule must be in thermal equilibrium. At room temperature and pressure this is certainly true, with molecules colliding and exchanging energy at the rate of 10^{19} s^{-1} . However, this may be unlikely if the density of molecules is very low, as in the interstellar medium where the density is 1 cm^{-3} and collisions between two atoms occur every 10^7 years.

Local thermal equilibrium (LTE) is an assumption that allows for the molecules to be in equilibrium with at least a limited region of space and remains an assumption when using the Boltzmann law for the relative populations of energy levels. The LTE assumption notwithstanding, observation of a series of transitions in the spectrum and measurement of their relative intensities allows the local temperature to be determined. We shall see an example of this in Section 4.4 where the Balmer temperature of a star is derived from the populations of different levels in the Balmer series.

3.3 Telescopes

Interrogating the light from stars and the absorption features of atoms and molecules in-between requires some fairly complicated optics in the form of a telescope. However, the telescopes are not restricted to the parts of the electromagnetic spectrum that you can see but use radiation from microwaves to gamma rays to observe the Universe. There is too much to learn about the optics or even adaptive optics of telescopes to be discussed here but there are some properties of telescopes that we must know because they are important for the identification of atoms and molecules. We shall discuss three telescope considerations: the atmosphere, the spatial resolution and the spectral resolution.

Atmospheric windows

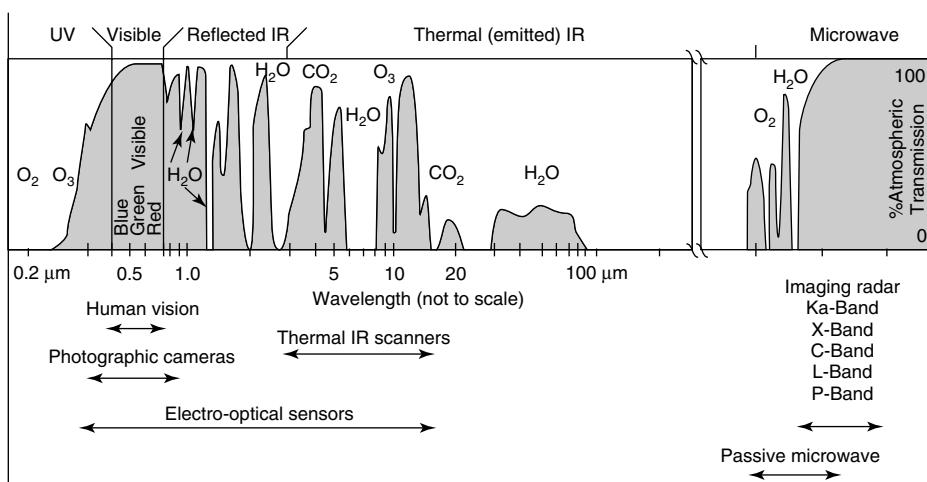
We are close to being able to consider atomic and molecular spectra in some detail but there is one intervening problem: the Earth's atmosphere. The atmosphere is a significant problem for atomic and molecular astronomy because absorption of radiation by all atmospheric components prevents the detection of molecules in space. In addition, the particulate or aerosol composition of the atmosphere causes light from the stars to be scattered, resulting in the apparent twinkling of stars. There are regions of the atmosphere in which the molecules of the atmosphere do not absorb and these provide astronomers with atmospheric windows (Table 3.1) for observations. In addition, observatories are built high on mountains such as in Hawaii (see Figure 3.13) to minimise the amount of cloud cover and atmospheric interference.

There are four atmospheric windows each revealing, broadly speaking, properties of atoms or molecules (Figure 3.3):

1. Visible UV spanning 300–900 nm contains information on electronic transitions in atoms and molecules.
2. Infrared (IR) 1–5 μm and 8–20 μm contains information on vibrational transitions in molecules and functional groups in solids.

Table 3.1 Atmospheric windows for astronomy

Wavelengths	Window
<300 nm	Absorbed by ozone
300–900 nm	UV-visible–near-IR window
1–5 μm	IR window between H_2O and CO_2 absorption features
8–20 μm	IR windows
1.8–1.1 mm	Millimetre window
0.8, 0.45, 0.35 mm	Submillimetre windows
1.3 cm–1.9 mm	Centimetre/millimetre windows
2 cm–10 m	Radio window
>10 m	Ionosphere absorption

**Figure 3.3** Atmospheric windows and molecular astronomy

3. Microwave and millimetre wave 1.3 – 0.35 mm contains information on rotational transitions.
4. Radio wave 2–10 m contains information on atomic hyperfine splitting such as the 21-cm line in H atoms.

Much of the electromagnetic spectrum has been used to investigate the structure of matter in the laboratory but the atmospheric windows restrict astronomical observations from Earth. Irritating as this is for astronomers on the ground, the chemical structure of the atmosphere and the radiation that it traps is important to the origins of life on Earth. The light that does get through the atmosphere, however, when analysed with all of the tools of spectroscopy, tells the molecular story of chemistry in distant places around the Universe.

Spatial resolution

The general drive to make telescopes bigger is driven by the need to catch as much light as possible so as to be able to see fainter objects or to look further into space. Improved spatial resolution will also allow two stars close to one another to be separated or resolved. It is important to know where the light entering the telescope came from as a diagnosis of the atomic and molecular environment. The minimum discernable separation between two stars is a measure of the resolving power of the telescope and is limited by diffraction of light. Light passing around two objects close together can interfere with itself and blur the image. So light passing close to the edge of the aperture, such as the eye, from two distant objects becomes blurred, hence the resolving power is related to the size of the eye and the wavelength of the light. The resolving power is given by:

$$\theta = 250\,000 \left(\frac{\lambda}{D} \right) \quad (3.15)$$

where θ is the resolution in arcseconds, λ is the wavelength of the radiation and D is the diameter of the collection optics – both have to be in the same units (metres for instance). The pupil of the human eye is typically 1 mm in diameter and for radiation in the visible at 500 nm or 500×10^{-9} m the resolving power is 125 arcsec. With this resolving power the eye is able to see features on the Moon that are around 200 km apart. Most telescopes in principle could have a much greater resolution than this, especially those with large 4 m collection mirrors. However, the atmospheric turbulence then begins to make the image shimmer and the theoretical resolution is not achieved on Earth. The best resolution on Earth in the visible region is around 1 arcsec. With this resolution modern optical telescopes can easily resolve objects on the surface of the Moon a few kilometres across (see Figure 1.6). More importantly from the point of astrochemistry is the spectral resolution of telescopes – the ability to tell the difference between one wavelength and another.

Spectroscopic resolution

Separate the light from the emission spectrum of the Sun and you will see the familiar rainbow colour spectrum but how small a wavelength difference can be detected? Is it possible to tell between 500 nm and 501 nm? The spectral resolution limits the ability of a telescope to tell the difference between two spectral lines and hence two different molecules. The smallest separation that allows two wavelengths to be distinguished is limited by the physics of dispersion and for sources of the same intensity, Lord Rayleigh determined that the dip between the two peaks should be $8/\pi^2$ or about 19 per cent.

Rayleigh chose the criterion as a rule of thumb from work performed on interference fringes of equal intensity (Figure 3.4). The bright straight-through fringe, $x\lambda_1$, is separated from a second fringe, $x\lambda_2$, when the maximum $I\lambda_1$ is over the

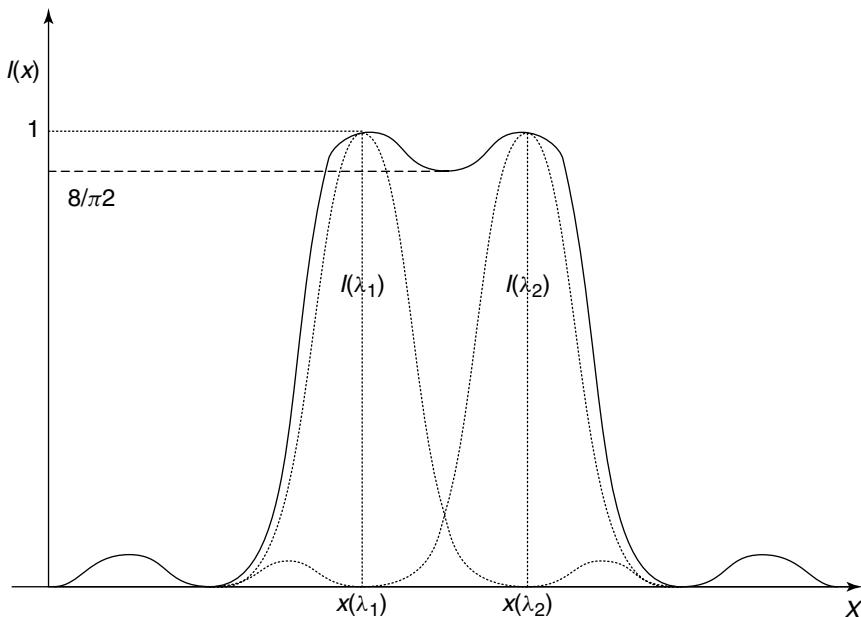


Figure 3.4 The Rayleigh criterion

first dark minimum of the second fringe pattern corresponding to the 19 per cent dip in the middle. It is only a rule of thumb and there are other conditions for the resolution of two equally intense signals but the Rayleigh criterion is useful for comparing the performance of different instruments. Astronomers use the definition of instrumental resolution as:

$$R = \frac{\lambda}{\Delta\lambda} \quad (3.16)$$

and in the laboratory this can reach some impressively large number but for telescopes the same is not true: this can be seen from the comparison in Table 3.2.

Measurements in the laboratory are much more sensitive to spectral separations and hence understanding the spectrum of a molecule uniquely for the purposes of identification. However, the astronomical observations are limited by the telescope resolution and the atmospheric shimmer but the intrinsic information content of microwave spectra provides the best method for molecular identification – the other regions of the spectrum have their uses, as we shall see.

Extraterrestrial astronomy

Satellite astronomical observations remove the problems of the atmospheric absorption windows but require more sophisticated orbital spacecraft. Telescopes mounted on aircraft in part can achieve better results but state-of-the-art results are obtained from orbiting satellites. Probably the most successful satellite telescope for public

Table 3.2 Comparison of spectral resolution

Instrument	Resolution
Microwave Laboratory Spectrometer	250 000
Infrared Vibration Laboratory Spectrometer	40 000
UV/Vis Laboratory Spectrometer	1000
High-resolution Laser Spectrometer (visible)	200 000
<i>Telescopes</i>	
UK Infrared Telescope (Figure 3.13)	3000–5000
James Clerk Maxwell Radio Telescope ($\lambda = 2\text{--}0.3\text{ mm}$) (see p. 3 of the colour plate section)	3000
Hubble Telescope Goddard High Resolution Spectrograph (visible)	20 000
Jodrell Bank Radio Telescope	3000
State-of-the-art resolution	$10^5\text{--}10^7$
Infrared Space Observatory Short-Wavelength Spectrometer	300

excitement in astronomy has been the launch of the Hubble Space Telescope (HST). In orbit some 600 km above the Earth, the HST no longer suffers from the effects of the atmosphere and has allowed us to resolve smaller objects, see further into the Universe and perform some sophisticated atmosphere-free experiments with four basic instruments:

1. Advance Camera for Surveys (ACS), observation range 200–1100 nm.
2. Near Infrared Multi Object Spectrometer (NICMOS), 0.8–2.5 μm .
3. Space Telescope Imagine Spectrograph (STIS), 115–1030 nm
4. Wide Field Planetary Camera (WFPC2), 115–1100 nm.

The HST remains in orbit at present although its future is now somewhat uncertain. A second orbiting telescope, which only had a limited lifetime due to the need for cooling of the detection technology, was the Infrared Space Observatory (ISO). A scientifically very exciting mission that orbited the Earth from May 1995 to November 1998 and had instruments on board covering the electromagnetic spectrum in the region 2.5–250 μm or $4000\text{--}40\text{ cm}^{-1}$. The ISO was a flying IR spectrometer and produced many important results that are still being analysed in the astrochemistry community.

Astronomical measurements have some limitations not present in laboratory investigations but a detailed knowledge of the spectra of atoms and molecules can be used to overcome the restrictions of resolution and atmospheric windows. The detailed knowledge is the key to this success, however, and confidence in the conclusions of astrochemical observations must come from the understanding in the laboratory.

3.4 Atomic spectroscopy

Dispersing the radiation from the star into its component wavelengths reveals that the spectrum of a star is not the continuous spectrum of a black body but there are

lines where the colour has been absorbed (see page 5 of the colour plate section). Here the important synergy between laboratory measurements of atomic spectra and the observation of stellar spectra leads to the identification and determination of the atomic composition of the stars and information on the temperature of the outer hot layer, the photosphere.

Structure of the hydrogen atom

Central to quantum mechanics and one of the reasons why it is most compelling as a theory, is the concept of quantisation. Without doing anything more than solve the equations, quantisation drops out as a natural consequence of the mathematics of the operators and the physical world. The allowed energy states of an atom or a molecule are labelled by integer quantum numbers; only certain values of the energy of an atom for example are allowed configurations of the electrons within the atom. The simplest atom is the hydrogen atom, with one electron and one proton, and it turns out to be the most important for astronomy.

The spectrum of H atoms is dominated by a series of lines, the strongest being $H\alpha$ at $\lambda = 656.3$ nm, then $H\beta$ at 486.1 nm and $H\gamma$ at 434.0 nm, converging at a wavelength of 365.0 nm (Figure 3.5). The $H\alpha$ line at 656.3 nm is in the red part of the visible spectrum and it is $H\alpha$ emission, the reverse process to absorption, that is responsible for the red colour of many gas clouds in space, such as the Orion nebula (see page 5 of the colour plate section).

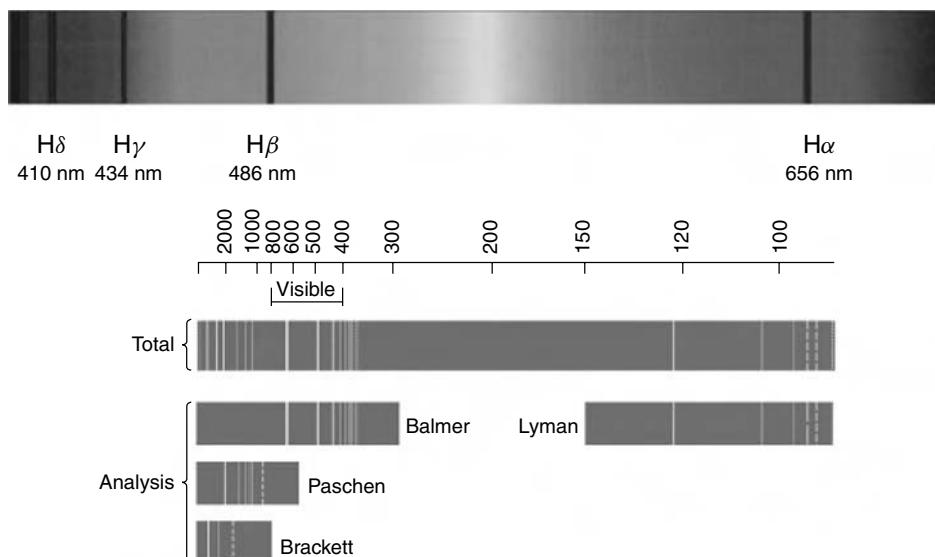


Figure 3.5 The hydrogen atom absorption spectrum and the named sequences. (Reproduced by permission of Pearson Publishing)

Each transition within a series is characteristic of a transition between two levels. The energy spacing between the levels determines the energy of the photon that is emitted or absorbed. The energy of the photon to be resonant with the transition of a fixed wavelength is given by the now familiar expression, $E = hc/\lambda$. The upper and lower energy levels are labelled n_2 and n_1 , respectively, and the transitions are only allowed if the quantum number changes by an integer amount. Hence, the series known as the Balmer series, named after the discoverer, has transitions observed at wavelengths given by:

$$\frac{1}{\lambda} = R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \quad (3.17)$$

where $n_1 = 2$, $n_2 = 3, 4, 5, 6 \dots$ and R_H is the Rydberg constant of $109\,677 \text{ cm}^{-1}$.

Example 3.3

Calculate the wavelengths of the first three transitions in the Balmer series. Using the value for R_H above with $n_1 = 2$ and $n_2 = 3, 4$ and 5 gives the following:

$$\begin{aligned} \frac{1}{\lambda} &= 109667 \left(\frac{1}{2^2} - \frac{1}{3^2} \right) \therefore \lambda = 656.532 \text{ nm} \\ \frac{1}{\lambda} &= 109667 \left(\frac{1}{2^2} - \frac{1}{4^2} \right) \therefore \lambda = 486.321 \text{ nm} \\ \frac{1}{\lambda} &= 109667 \left(\frac{1}{2^2} - \frac{1}{5^2} \right) \therefore \lambda = 434.215 \text{ nm} \end{aligned} \quad (3.18)$$

The first transition is in the red at 656 nm, with the other two transitions appearing in the blue at 486 and 434 nm. The separation between the transitions is also getting closer as the term in $1/n_2^2$ becomes very small, until finally the series converges on a limit with a wavelength of 364.740 nm.

The Lyman, Balmer, Paschen, Brackett and Pfund series

Other series are present in the spectrum of the H atom, starting from different initial levels and undergoing transitions to higher levels. There is formally no limit to the value of n in the upper state but eventually in the limit the electron is so far from the nucleus that it is no longer confined by the attraction of the proton and it leaves the atom: a process called ionisation. The ionisation event can be seen in any one of the series of the H atom when the separation between the lines of the series begins to converge towards zero – the ionisation limit at a fixed wavelength.

Lyman series: $n_1 = 1 \longrightarrow n_2 = 2, 3, 4, 5 \dots$ 121.6–92.1 nm in the deep UV
 Balmer series: $n_1 = 2 \longrightarrow n_2 = 3, 4, 5, 6 \dots$ 656.3–365.0 nm visible

Paschen series: $n_1 = 3 \longrightarrow n_2 = 4, 5, 6 \dots$ 4052.3–1458.8 nm infrared

Brackett series: $n_1 = 4 \longrightarrow n_2 = 5, 6, 7 \dots$ 7459.9–2279.4 nm infrared

Pfund series: $n_1 = 5 \longrightarrow n_2 = 6, 7, 8 \dots$ 12371.9–3282.3 nm microwave

There are in principle an infinite number of series beginning at higher quantum numbers with $n_1 = 6, 7, 8, 9 \dots$ but they become increasingly difficult to observe. For the higher n series to be seen in the spectrum the levels have to be populated, so some hydrogen atoms must be in the $n = 5$ level to see the Pfund series. We shall see that the presence of the Balmer series in the spectrum of a star is indicative of the stellar temperature, which is a direct consequence of the population of the energy levels. More of this in Chapter 4.

Atomic selection rules

One of the consequences of atomic energy levels, and the very foundation of all spectroscopy, is that energy within the atom can only take an allowed set of values controlled by the quantum numbers. It follows then that only energy changes corresponding to the allowed gaps between the energy levels can occur. The H atom spectrum seen in Figure 3.5 is seen in absorption, which means that photons of the right energy are removed from the incident light. The photon has energy related to its wavelength by $E = hc/\lambda$ and when this matches the energy gap between the levels of the atom it is absorbed. Hence, for the lowest transition between $n_2 = 2 - n_1 = 3$ in the Balmer series this is red at 656 nm. However, there is an implicit further condition that Δn , the change in the quantum number n , is allowed to be $\Delta n = 1, 2, 3, 4, \dots$, which is an example of a selection rule. There are three quantum numbers n, l , and m for a complete description of the energy levels of an atom and there are selection rules for transitions related to all quantum numbers. The selection rule comes from a consideration of conservation of energy and momentum. When the atom absorbs the photon, it is destroyed completely, which means that its energy and momentum are now taken up by the atom in the excited state. The l quantum number is a measure of the orbital angular momentum of the electron in a particular state and this must change when the photon is absorbed. The photon has one unit of angular momentum and so, when it is destroyed or created in either absorption or emission, one unit of angular momentum has to be destroyed or created (gained or lost) by the atom during the transition. This leads to the selection rule $\Delta l = \pm 1$. The concept of selection rules is common to both atoms and molecules, as we shall see later.

3.5 Molecular astronomy

The problem with molecular astronomy is knowing where to start and, more importantly, where to stop. The laboratory-based big brother of molecular astronomy is the field of high-resolution spectroscopy. The synergy between these two subjects

is crucial to the correct identification of molecules in space. Some spectroscopic techniques are better than others; the most important of these for identifying molecules is microwave or rotational spectroscopy. Even for microwave spectroscopy, the problems of interpreting the spectra and identifying the molecule become greater as the number of atoms in the molecule increases: life is easier and well understood for diatomic molecules and becomes complicated with triatomic molecules and polyatomics. It has been postulated that six atoms represents the upper limit for understanding the high-resolution spectroscopy of molecules: an interesting idea.

The next most useful is vibrational spectroscopy but identification of large molecules is still uncertain. In the laboratory, vibrational spectroscopy in the infrared (IR) is used routinely to identify the functional groups in organic molecules but although this is important information it is not sufficient to identify the molecule. Even in the fingerprint region where the low wavenumber floppy vibrational modes of big molecules are observed, this is hardly diagnostic of structure. On occasion, however, when the vibrational transition can be resolved rotationally then the analysis of the spectrum becomes more certain.

The least useful spectroscopy in the laboratory, as in astronomy, is the ultra-violet–visible (UV–Vis) spectrum. The high-energy photons associated with this branch of spectroscopy are indicative of electron rearrangements in molecules, with the obvious effect of changing the strength of the chemical bonds. Electronic excitation also adds energy to reactions: a reaction may not be possible in the ground electronic state but when the molecule is electronically excited it behaves essentially as a different and often highly reactive species. As a final rider, if rotational transitions can be resolved in the electronic transition between one vibration–rotation level in the ground state to another vibration–rotation level in the excited state then a fully resolved ro-vibronic spectrum is very useful in determining the identity and structure of a molecule, but only if you are lucky.

Microwave and millimetre wave astronomy

Without diving into the really exotic transitions in molecules just yet, the smallest separation between quantum energy levels in all isolated molecules is associated with rotational motion. Molecules cannot rotate arbitrarily fast but have to do so at a number of allowed frequencies and can only change the rate of rotation by jumping up or down to the next allowed rotational state. Molecules rotate about once every nanosecond and the spacing between the allowed energy levels corresponds to photons with frequencies of order GHz. These low-frequency photons correspond to long-wavelength radiation called microwave or millimetre waves. Absorbing radiation of this wavelength changes the rotational state of the molecule and hence the rotational energy. The microwave oven familiar to many produces radiation at 2.45 GHz that excites the rotational spectrum of liquid water. Collisions between the rotationally excited water molecules cause collisions with other molecules, removing some of the energy as translation. The temperature of a substance is a

measure of rotational, vibrational and translational energy of molecules, which is transferred to the atoms of mercury in a thermometer to give a temperature reading. It is perfectly reasonable, therefore, to talk of rotationally hot molecules and similarly vibrationally and translationally hot molecules and this is all part of the general concept called equipartition of energy, which is true at equilibrium at least.

The rotational energy levels for a simple diatomic molecule correspond to different end-over-end tumbling rates for the molecules and the energy levels are associated with the rotational quantum number N , which is allowed to take values of 0,1,2,3,4,5... but if there is any other form of angular momentum in the molecule, such as electronic angular momentum, then they can all couple together to produce a total angular momentum quantum number J , so the simplest case $J = N$ is just the end-over-end motion. Angular momentum in general is a hard concept and there is much interesting physics to be learnt in order to understand the rotation of large objects; this becomes particularly exciting when trying to understand how a gyroscope works for example. In general, however, angular momentum has some similarities with linear momentum, which will help with the understanding. Newton discovered that moving objects have a constant speed unless acted upon by a force and that you need a bigger force to move a heavier object. The property is called the inertia of a body; heavy bodies have a higher inertia and are thus harder to move. When it comes to making a body rotate, the resistance to rotational motion is called the moment of inertia; bodies with a large moment of inertia are harder to rotate. The moment of inertia depends on the structure of the body or molecule and tells you how the mass is distributed and how far apart the masses are – for diatomic molecules this gives the bond length.

Consider the simple diatomic molecule CO with a carbon atom at one end of the bond and an oxygen atom at the other. The moment of inertia is a measure of how heavy each atom is and the length of the bond between them. The moment of inertia carries important information regarding the structure of the molecule and, more importantly, is very useful in identifying a molecule. The energy separation between the allowed end-over-end rotations of a diatomic, E_J , is given by:

$$\begin{aligned} E_J &= \frac{\hbar}{8\pi^2 I} J(J+1) - D J^2(J+1)^2 \\ B &= \frac{\hbar}{8\pi^2 I} \end{aligned} \quad (3.19)$$

where J is the rotational quantum number 0, 1, 2, 3..., B is the rotational constant, D is the centrifugal distortion constant and I is the moment of inertia, $I = \mu < r^2 >$, where r is the bond length and μ is the reduced mass. The reduced mass for a diatomic molecule is given by

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad (3.20)$$

Reduced mass comes from the definition of moment of inertia, which will be discussed shortly.

The rotation of molecules can be treated within the rigid-rotor approximation, which assumes that the centrifugal distortion constant, D , is zero. The rigid-rotor approximation works well when the J quantum numbers involved in the transition are low, which is likely for molecules in the interstellar medium. The higher terms have to be added as the molecule begins to rotate more quickly. Taking Equation 3.19, the separation between the energy levels $J = 0 \rightarrow J = 1$, $J = 1 \rightarrow J = 2$, etc. is given by:

$$\nu = 2B(J + 1) \quad (3.21)$$

so that the separation between the levels is $2B$. Computing the rotational constant gives B of order 30 GHz for small molecules, which is in the middle of the microwave region. Absorption of a microwave photon by a molecule requires that the energy of the photon be absorbed, that the molecule rotates more quickly and that the angular momentum of the photon has to be absorbed by the molecule. The photon is then annihilated. Forcing the conservation of angular moment produces a selection rule so that J can only change by ± 1 unit of angular momentum (the angular moment of the annihilated photon), written as $\Delta J = \pm 1$. The separation between the transitions in the microwave spectrum is given by Equation 3.21, so if you know the spacing of the energy levels you can determine the rotational constant and hence the moment of inertia and the bond length. Gas-phase microwave spectroscopy is the most effective way to measure a bond length and also happens to fall happily in one of the atmospheric windows.

Example 3.4

The rotational spectrum of $^1\text{H}^{127}\text{I}$ consists of a series of lines with spacing 13.10 cm^{-1} . From this spacing it is possible to calculate the bond length of the molecule, assuming that it is a rigid rotor. The spacing is equal to $2B$ from Equation 3.21, which implies that the rotational constant $B = 6.55 \text{ cm}^{-1}$. Converting this to Hz gives $B = 1.96364 \text{ Hz}$. Rearranging the definition of the rotational constant from Equation 3.19 gives the moment of inertia, given by:

$$I = \frac{h}{8\pi^2 B} = \frac{6.626076 \times 10^{-34} \text{ J s}}{8\pi^2 1.96364 \text{ s}^{-1}} \quad (3.22)$$

$$I = 4.271371 \times 10^{-47} \text{ kg m}^2$$

The moment of inertia is $I = \mu < r^2 >$ and the bond length can be calculated by:

$$\mu = \frac{1 \times 127}{1 + 128} \text{ amu} = 0.99986 \times 1.66064 \times 10^{-27} \text{ kg}$$

$$r^2 = \frac{I}{\mu} = \frac{1.4271371 \times 10^{-47}}{1.660307 \times 10^{-27}} \text{ m}^2$$

$$r = 0.1604 \times 10^{-9} \text{ m} = 160.4 \text{ pm} \quad (3.23)$$

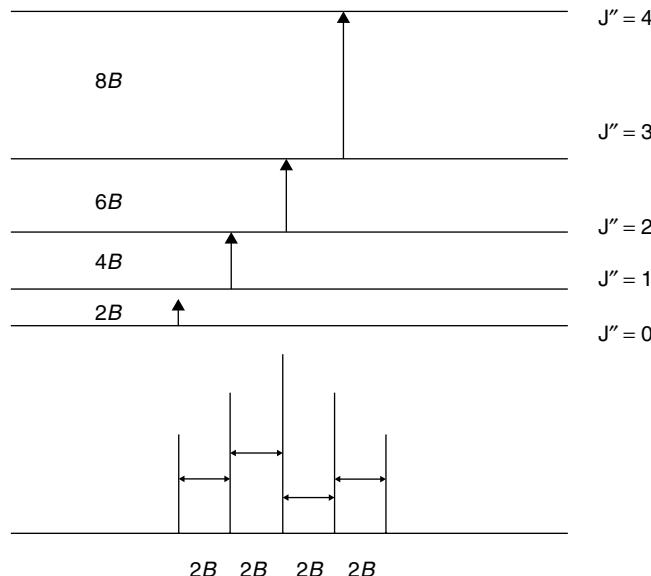


Figure 3.6 The energy-level diagram for a rotational spectrum

The microwave spectrum has a general appearance of that in Figure 3.6, where the separations and the transitions are identified. The transitions $J \rightarrow J + 1$ are labelled as the R-branch with the convention $R(0) = 0 \rightarrow 1$, $R(1) = 1 \rightarrow 2$, etc. and the $J \rightarrow J - 1$ transitions are labelled the P-branch with the convention $P(1) = 1 \rightarrow 0$, $P(2) = 2 \rightarrow 1$. It is common among astronomers and spectroscopists to talk about the P(1) transition in a molecule to identify precisely a transition. For the microwave spectrum the P and R transitions must be on top of one another so that P(1) and R(0) have the same frequency.

The spacing between the transitions observed in the spectrum is $2B$, which is accurate for low- J transitions but deviates as the centrifugal distortion of the molecule becomes important at high J . The simulated rotational spectrum for CO is shown in Figure 3.7 and shows regularly spaced lines at 115.271, 230.537, 345.795, 461.040... GHz, called a rotational progression. The separation between the energy levels is to a good approximation given by $2B$ and is of order 115 GHz for CO. The resolution required to see the separated transitions is of order 2 and so will easily be seen by all telescopes and laboratory spectrometers, which usually have a resolution of order 50 000. Identification of CO astronomically is then rather straightforward and resolution allows ^{13}CO to be identified positively. The spectrum also shows a characteristic intensity profile so that the R(0) transition is not as strong as R(1), with R(4) showing the greatest intensity in the rotational progression. This, as we shall see later on, is due to the rotational temperature of the molecule and the spectrum has been simulated with a rotational temperature of 40 K. The relative intensities of the transitions can then be used as a thermometer

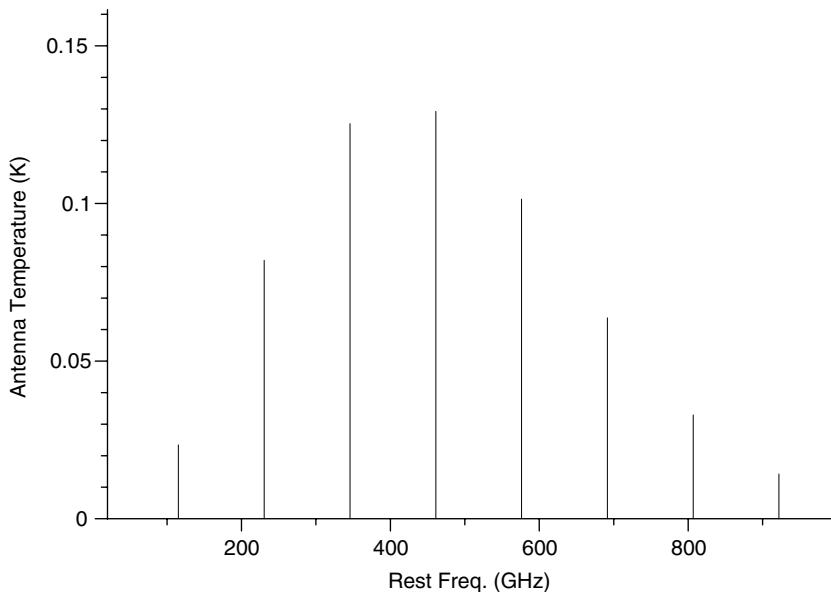


Figure 3.7 Simulated spectrum of CO with a rotational temperature of 40 K. Reproduced by with permission of Albert Nummeling, Chalmers University of Technology

reporting the temperature of the surroundings of the molecule, assuming that it is in local thermal equilibrium.

Example 3.5

The R(0) transition in ^{12}CO is at 115.271 GHz. Calculate the position of the R(0) transition in ^{13}CO in the rigid rotor approximation. The separation between the energy levels in the R-branch is between levels $E(J+1) - E(J)$ and in the rigid-rotor approximation is given by:

$$\Delta E_J = \frac{h}{8\pi^2 I} [(J+1)(J+2) - J(J+1)]$$

$$\Delta E = \frac{h}{8\pi^2 I} 2(J+1) \quad (3.24)$$

The mass dependence in this expression comes from the moment of inertia and ratio of the reduced masses for $^{12}\text{CO}/^{13}\text{CO}$. The isotope shift for the ^{13}CO transition is given by:

$$\frac{\text{R}(0) \ ^{12}\text{CO}}{\text{R}(0) \ ^{13}\text{CO}} = \frac{\frac{h}{8\pi^2 I(^{12}\text{CO})}}{\frac{h}{8\pi^2 I(^{13}\text{CO})}} = \frac{I(^{12}\text{CO})}{I(^{13}\text{CO})}$$

$$\begin{aligned} \frac{R(0) \text{ } ^{12}\text{CO}}{R(0) \text{ } ^{13}\text{CO}} &= \frac{\mu(^{12}\text{CO})}{\mu(^{13}\text{CO})} = \frac{\frac{12 \times 16}{12 + 16}}{\frac{13 \times 16}{13 + 16}} = \frac{6.8571}{7.1724} = 0.9560 \\ R(0) \text{ } ^{13}\text{CO} &= \frac{R(0) \text{ } ^{12}\text{CO}}{0.9560} = 120.576 \text{ GHz} \end{aligned} \quad (3.25)$$

The observation of the R(0) transitions in both isotopes of CO is very reassuring for a positive identification of the molecule but of course the relative amounts of each is also very interesting.

The moment of inertia is on our side in the case of CO but this is not so favourable for larger molecules where the components of the rotational progression become much more closely spaced and perhaps impossible to resolve. If we calculate the rotational constant for a linear molecule such as OCS we find that it is much smaller at around 6 GHz and now the resolution required to resolve an adjacent transition in a rotational progression is 115/6 or around 20. The limiting spacing resolvable by a ground – based telescope as a microwave transition is around 3000 so separations are very small, corresponding to rotational constants as small as 0.01 cm^{-1} or 30 MHz. Hence the spectra of all diatomic molecules and linear polyatomic molecules can be observed on Earth. However, the expression for the moment of inertia becomes more complicated and there are two other moments of inertia to be considered for larger molecules such as water and the problem becomes more complicated.

Moments of inertia

The moment of inertia for diatomic molecules is fairly easy to understand but for other linear molecules such as OCS or non-linear polyatomic molecules it is more complicated. The general expression for the moment of inertia is given by:

$$I = \sum_i m_i r_i^2 \quad (3.26)$$

which says that the moment of inertia is the sum of the mass of each atom multiplied by the perpendicular distance from the rotational axis running through the centre of mass of the molecule – this is easier to see in Figure 3.8 for the water molecule. The axes are labelled I_a , I_b and I_c by convention, with I_c as the biggest and so hardest to change the rotational angular momentum around this axis. Generally the order is $I_c > I_b > I_a$ for all molecules and you only need three axes to specify the complete rotation. The moment of inertia depends on the perpendicular

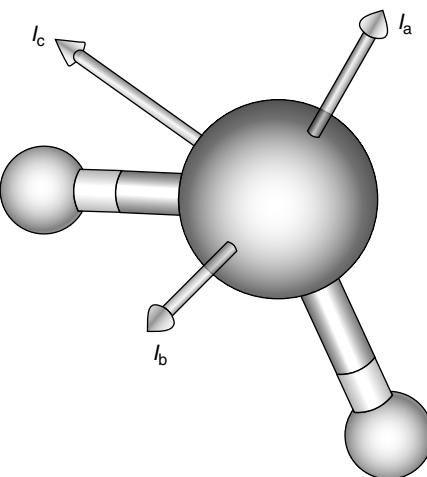


Figure 3.8 Rotation axes in the water molecule

distance from each of the axes so the bond angles and bond lengths control I , and thus the complete molecular geometry. However, as the molecule gets bigger I becomes less and less diagnostic of the molecular structure.

At first it does not look as if water rotates about these axes – surely it should just tumble around in all directions. Of course it does just tumble but you can describe this mathematically by the components along each axis. Tumbling around each axis should imply angular momentum around each axis and thus quantised energy levels for the spectrum. In a surprise turn, quantum mechanics prevents us from knowing the angular momentum around each axis because this is forbidden: the commutation relation between the operators; interesting as this is, you will have to research it elsewhere. We can, however, know the total rotational angular momentum J and its projection along any one axis at a time with quantum number K . The K quantum number can take values $-J \leq K \leq +J$, so for $J = 2$, $K = -2, -1, 0, 1, 2$. The selection rule involving K is $\Delta K = 0$, which adds any additional K -progression for each J value. This complicates matters somewhat and only careful measurement and simulation of the spectrum allows identification of the molecule. The spectrum of the worryingly simple molecule methanol is shown in Figure 3.9 but with the spectral resolution available in the laboratory the detail of this spectrum can be measured in full, revealing bond angles, bond lengths and how the bonds distort as the rotation becomes faster, an effect called centrifugal distortion. The rotational progression is very characteristic of the molecule and it can be identified confidently from a sequence of lines observed in the microwave spectrum in space. The problem becomes increasingly more complicated, however, with increasing numbers of atoms so that the rotational spectrum of glycine shown in Figure 3.10 looks too complicated and yet this is a molecule that contains only ten atoms. The synthesis of amino acids in the interstellar medium is an interesting

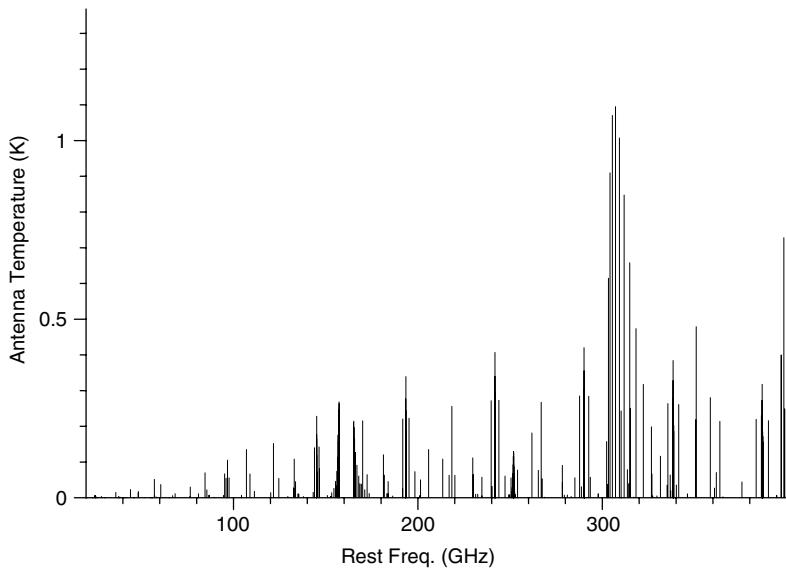


Figure 3.9 Simulated spectrum of CH_3OH at 40 K. Reproduced by permission of Albert Nummelin, Chalmers University of Technology

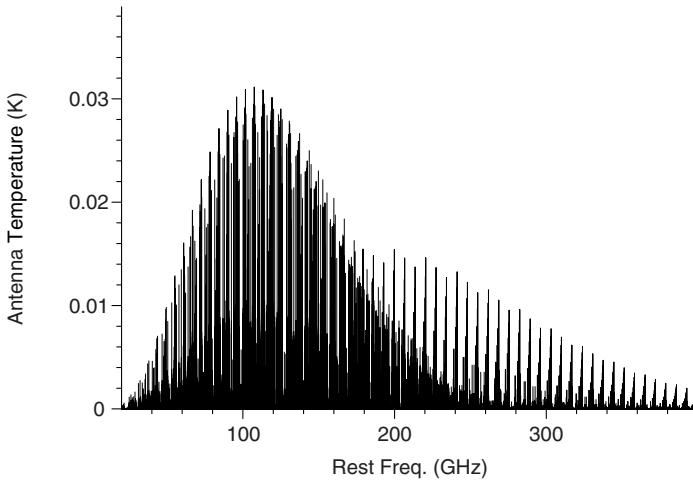


Figure 3.10 Simulated rotational spectrum of glycine at 40 K. Reproduced by permission of Albert Nummelin, Chalmers University of Technology

question for the origin of life and yet identification from the microwave spectrum is non-trivial.

The relative transition intensities of the spectrum are a good measure of the local temperature of the molecule. This raises the interesting question of whether the relative intensities of R- and P-branch transitions have the same rotational temperature and whether molecules in the same parts of space tell the same temperature.

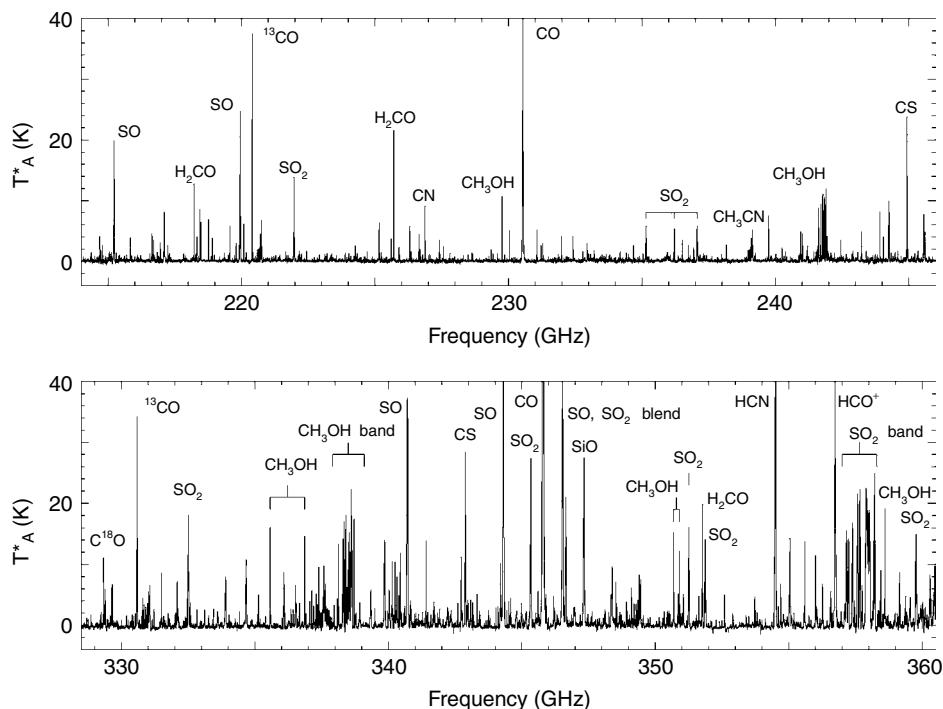


Figure 3.11 The microwave spectrum of the Orion nebula. Reproduced by permission of Geoffrey A. Blake, Owens Valley Radio Observatory

Microwave identification of molecules

Astronomical observations are complicated by the presence of many molecules in the field of view so that the measured spectrum is a complicated mixture of allowed rotational transitions such as in the Orion nebula (Figure 3.11). Identifying molecules is then a matter of experience based on knowledge of absolute transition frequencies and how the lines form rotational progressions. There are many things to consider in assigning the spectrum:

- Correct the line for the Doppler shift of the molecule
- Determine the lines within the rotational progression
- Molecules may be in different environments moving away or towards the telescope
- Isotope patterns must be present
- Selection rules

The Doppler shift correction is somewhat of a circular argument in that the best way to determine the Doppler shift is to measure the shift of a known spectral line.

Obvious species such as CO provide useful spectral standards and are plentifully abundant in space. The Doppler shift for CO can then be applied to other unidentified transitions to see if they are coincident with known transitions in the laboratory spectrum of a molecule. Different molecular environments may complicate matters, with some CO molecules along the line of site of the telescope having different Doppler shifts (Figure 3.12). The CO transition at 115 GHz may then appear to be split into several lines associated with a different Doppler shift in each cloud. The identification problem now also has to decide to which cloud the unknown transition belongs.

Atomic isotopes help detection so that the shift for ^{13}CO is perfectly predictable and isotope ratios are known from various sources, as we shall see later on. The change in the rotational constant with reduced mass is easily calculated and hence the change in the frequency of the $J = 1 \rightarrow J = 0$ can be calculated, as we saw in Section 3.3.

The complex problem of molecular identification is aided and constrained by the selection rules of quantum mechanics. A simple or so-called gross selection rule is that for a molecule to have a rotational spectrum the molecule must contain a permanent dipole moment. The dipole is a distribution of charge throughout the molecule along three perpendicular axes. For a diatomic such as CO the dipole moment must lie along the bond with a positive end of the molecule associated with the electronegativity of the species. Carbon will be partially positively charged and the oxygen atom will be partially negatively charged. The $\delta^+ \text{C}-\text{O} \delta^-$ electrostatic distribution along the bond enables the electric field of the photon to grab the molecule and cause it to turn: the positive and negative ends like to align with the field and this induces the rotational change. The permanent dipole moment selection rules poses a real problem for astrochemistry – hydrogen. The charge distribution of the H_2 molecule is completely symmetric – there is no dipole moment – and so one of the most important molecules in the astrochemical inventory is not detected by microwave astronomy.

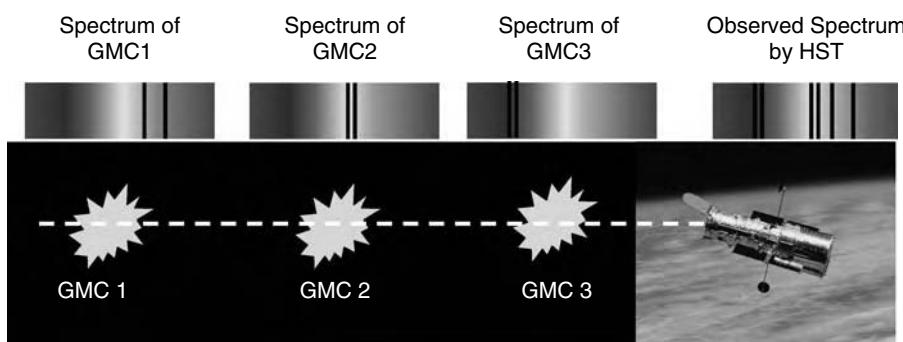


Figure 3.12 Line of sight through three Giant Molecular Clouds to the Hubble Space Telescope. (Reproduced by courtesy of Stsci and NASA)

Intensities of rotational transitions

Up to now I have hinted at the relative intensities of the rotational transitions providing a measure of temperature and we can now explore this further. In general, allowed transitions are strong and forbidden ones are weaker but quantum mechanics helps reveal more information about the environment of the molecule – local temperature. Rotational energy levels within a molecule are populated depending on the availability of energy from collision with other molecules, partitioning the energy into the levels based on how far apart they are – high energy levels are populated less than low energy levels. Collisions with molecules nearby is essentially the kinetic definition of temperature, so as long as there are collisions the molecules are said to be in *local thermal equilibrium* (LTE): the higher the temperature, the longer the rotational progression.

The rotational spectrum of a molecule involves transitions between energy levels, say the R(8) transition $J = 8$ and $J = 9$, but if there are no molecules rotating in the $J = 8$ level then there can be no R(8) transition. The local thermal collisions will populate some of the higher J levels in a general principle called equipartition. The general expression is the Boltzmann Law, given by:

$$\frac{n_2}{n_1} = \frac{g_2}{g_1} e^{-\Delta E/kT} \quad (3.27)$$

where n_i is the population of the i th level, g_i is the degeneracy of the level and ΔE is the energy separation between the levels. Degeneracy is the number of levels of the same energy and for a rotational level this is given by $(2J + 1)$ and this term controls the intensities early in the rotational progression. As the temperature increases, the number of J states populated increases, the populations of which are a balance of the degeneracy and the temperature, hence the rotational progression intensity shows a maximum associated with a specific value of J .

The temperature of a molecule within any astronomical environment may vary from the intense cold of the interstellar medium with a temperature of 10 K to the temperature within a sun spot 4000 K close to the temperature at which a molecule would fall apart. The relative intensity of transitions along the progression is given by a line strength factor:

$$I_R \propto (2J + 1) e^{-hB(J(J+1))/kT} \frac{(J + 1)}{(2J + 1)} \quad (3.28)$$

Hence the ratio of two transitions in the R-branch series may be used to derive the temperature of the molecule assuming that it is in LTE with its surroundings. Transitions in the P-branch have a different strength factor and this is given by:

$$I_P \propto (2J + 1) e^{-hB(J(J+1))/kT} \frac{J}{(2J + 1)} \quad (3.29)$$

A similar temperature analysis is possible for the P-branch transitions to determine the temperature, and usually the R-branch and P-branch temperatures agree. It is from analysis of the relative intensities of transitions that temperatures for the interstellar medium and the photospheres of stars may be determined.

Example 3.6

The relative intensities of the R(0) and R(1) transitions are given by the ratio:

$$\frac{IR(1)}{IR(0)} = \frac{[2(1) + 1]e^{-hB[(1)((1) + 1)]/kT} \frac{[(1) + 1]}{[2(1) + 1]}}{[2(0) + 1]e^{-hB[(0)((0) + 1)]/kT} \frac{[(0) + 1]}{[2(0) + 1]}} \quad (3.30)$$

$$\frac{IR(1)}{IR(0)} = \frac{3e^{-2B/kT} \frac{2}{3}}{1} = 2 e^{-2hB/kT}$$

Putting in the rotational constant for CO = 1.9313 cm⁻¹ but converting this to the SI units gives 5.719×10^{10} s⁻¹ and a temperature of 50 K gives a relative intensity of 1.803. The R(1) transition is nearly twice as intense as the R(0) transition. When the temperature drops to 10 K, the R(1)/R(0) ratio is 0.595 and it is this variation in relative intensity that enables the temperature of the interstellar medium to be determined: compare this with intensities in Figure 3.7.

Infrared astronomy

Rotational spectroscopy and microwave astronomy are the most accurate way to identify a molecule in space but there are two atmospheric windows for infrared astronomy: in the region 1–5 μm between the H₂O and CO₂ absorptions in the atmosphere and in the region 8–20 μm. Identification of small molecules is possible by IR but this places some requirements on the resolution of the telescope and the spacing of rotational and vibrational levels within the molecule. The best IR telescopes, such as the UK Infrared Telescope on Mauna Kea in Hawaii (Figure 3.13), are dedicated to the 1–30 μm region of the spectrum and have a spatial resolution very close to the diffraction limit at these wavelengths.

The resolution of the long-slit prism spectrometer is of order $R \sim 1500 - 3500$ and the long-slit grating spectrometer has a resolution of 400–40 000. These resolutions make the identification of some moderately large molecules possible, depending on the spacing of the energy levels within the molecule. The added advantage of being at the summit of Mauna Kea is the reduction in atmospheric absorbance. The spectrum of the atmosphere from the top of this extinct volcano is shown in Figure 3.14. The clear gaps between the water and CO₂ bands can be seen and if your target molecule is within one of the gaps then you book time on



Figure 3.13 UK Infrared Telescope situated close to the summit of Mauna Kea, Hawaii, at an altitude of 4092 m. (A colour reproduction of this figure can be seen in the colour section). (Reproduced with permission by UKIRT)

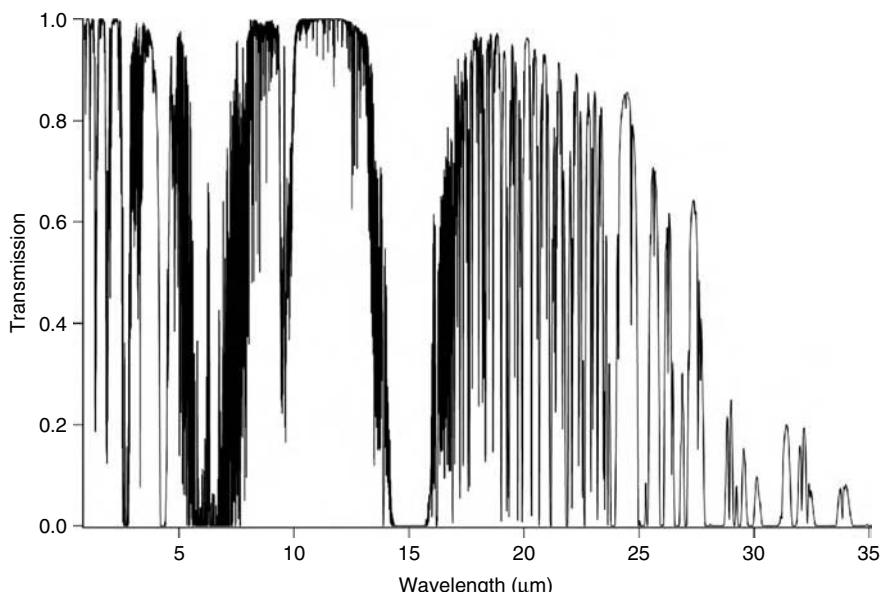


Figure 3.14 Infrared atmospheric absorption on Mauna Kea

the telescope and go and look for it. You will have to convince the telescope time manager that you are in with a good chance of making the observation and for this you need to know the vibrational structure of the molecule and make a prediction about the extent of any rotational progression. To do this we need to understand the origin of the ro-vibrational spectrum of a molecule.

The vibrational spectrum of a molecule is determined by the spacing of the vibrational energy levels of the molecule. Solving the Schrodinger equation for a simple vibrating-spring model of the chemical bond in a diatomic molecule gives the following expression for the vibrational energy levels:

$$E(v) = h\nu_e (v + 1/2) - \nu_e x_e (v + 1/2)^2 \quad (3.31)$$

The first term in the equation corresponds to the harmonic oscillator approximation with a vibrational quantum number v taking values 0,1,2,3,4. . . ., and ν_e is the fundamental frequency, multiplied by h , Planck's constant. The second term in the equation allows for the chemical bond to be broken and describes the anharmonic effects on the spring as it stretches. As the spring is stretched it must eventually break: the same is true of a chemical bond. There is one big difference between vibrational and rotational motion in a diatomic molecule – *zero point energy*. The vibrational quantum number can be zero but the energy corresponding to this state is non-zero, having the value $\frac{1}{2}h\nu_e$, so that at all temperatures, even absolute zero, the chemical bond is vibrating. The fundamental frequency of the vibration ν_e in the harmonic approximation is given by:

$$\nu_e = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \quad (3.32)$$

where k is the force constant for the chemical bond in question, typically 500 N m^{-1} , and μ is the reduced mass. The transitions between the vibrational energy levels of the molecules are in the infrared and, assuming they fall within the atmospheric windows, then they can be used astronomically. In general, if a molecule is to have a vibrational spectrum there must be a change in the dipole moment as the molecule vibrates, but there is no hard and fast rule for the change in the vibrational quantum number due to the anharmonicity terms. The selection rule then becomes a propensity rule with transitions involving $\Delta v = \pm 1, \pm 2, \pm 3 \dots$ observed routinely, although with decreasing intensity with bigger changes in quantum number.

As the molecule vibrates it can also rotate and each vibrational level has associated rotational levels, each of which can be populated. A well-resolved ro-vibrational spectrum can show transitions between the lower ro-vibrational to the upper vibrational level in the laboratory and this can be performed for small molecules astronomically. The problem occurs as the size of the molecule increases and the increasing moment of inertia allows more and more levels to be present within each vibrational band, $3N - 6$ vibrational bands in a non-linear molecule rapidly becomes a big number for even reasonable size molecules and the vibrational bands become only unresolved profiles. Consider the water molecule where $N = 3$ so that there are three modes of vibration: a rather modest number and superficially a tractable problem. Glycine, however, has 10 atoms and so 24 vibrational modes: an altogether more challenging problem. Analysis of vibrational spectra is then reduced to identifying functional groups associated

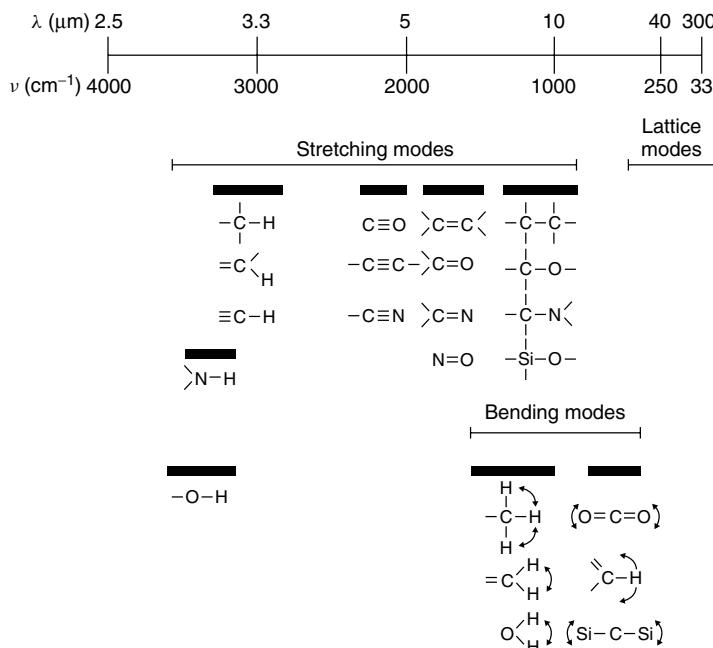


Figure 3.15 Infrared stretching frequencies (Adapted from Sandford 1996)

with the vibration of particular bonds at characteristic frequencies outlined in Figure 3.15. Thus in the case of glycine a group vibration associated with the $-\text{NH}_2$ and $-\text{COOH}$ groups is seen in the spectrum but this is not unique to glycine and could be any amino acid. The fundamental vibrational modes for the functional groups all occur in the infrared and are referred to by astronomers by their wavelength but by spectroscopists by their wavenumber. The wavenumber is simply the reciprocal of the wavelength, $1/\lambda$ (cm), and is chosen by spectroscopists as a transition with a bigger wavenumber has a higher energy. Hence stronger bonds have larger vibrational wavenumbers such as the C–H or O–H stretches, whereas the floppy modes of vibration associated with bending the bonds have lower wavenumbers.

Example 3.7

The ro-vibrational spectrum of $^{12}\text{C}^{16}\text{O}$ has four transitions in the infrared vibrational band $v = 1 \leftarrow 0$ observed at the following wavenumbers: 2135.55 cm^{-1} , 2139.43 cm^{-1} , 2147.08 cm^{-1} and 2150.86 cm^{-1} . Draw and label a diagram identifying the transitions involved in the spectrum and determine the harmonic vibrational wavenumber and bond length of $^{12}\text{C}^{16}\text{O}$. Assume that CO behaves as a harmonic oscillator and a rigid rotor.

The ro-vibrational spectrum is observed between the $v = 0$ and $v = 1$ vibrational levels of CO and between rotational levels. Looking at the spacing between the

transitions it is clear that one spacing corresponds to $2B$, twice the rotational constant, and the other corresponds to $4B$, four times the rotational constant, marking the gap between the P- and the R-branches. The four transitions can be assigned as $v = 1 \leftarrow 0$, P(2), P(1), R(0) and R(1), as shown in the energy level diagram for the spectrum in Figure 3.16. The harmonic oscillator and rigid rotor approximations allow the transition wavenumbers to be written according to:

$$\sigma(v, J) = \sigma_e(v + 1/2) + \overline{B}_v J(J + 1) \quad (3.33)$$

where the harmonic wavenumber and rotational constant are defined in terms of wavenumbers for the present with v and J as the vibrational and rotational quantum numbers, respectively.

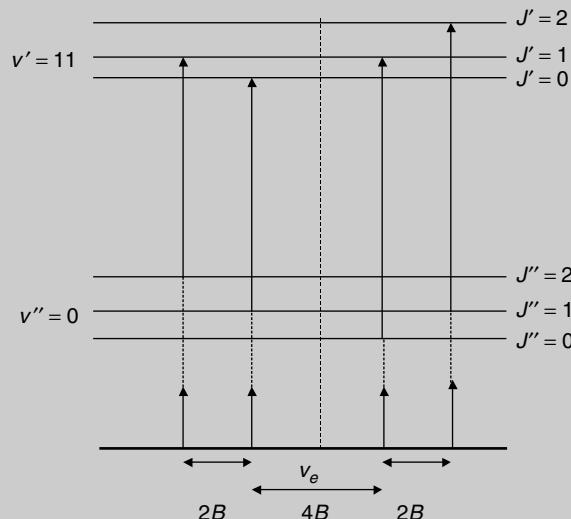


Figure 3.16 Energy level diagram for the CO ro-vibrational spectrum

The quantum number assignment of the spectrum and energy of the levels can then be made as follows:

1, 0	—→	0, 1	P(1)	$\sigma_e - 4 B$	2135.55
1, 1	—→	0, 2	P(2)	$\sigma_e - 2 B$	2139.43
1, 1	—→	0, 0	R(0)	$\sigma_e + 2 B$	2147.08
1, 2	—→	0, 1	R(1)	$\sigma_e + 4 B$	2150.86

By solving the equations for σ_e and B , average values for each can be determined as $\sigma_e = 2143.23 \text{ cm}^{-1}$ and $B = 1.913 \text{ cm}^{-1}$. The bond length can be estimated using the same ideas as in Example 3.4 to give a bond length of 113 pm.

Identification of molecules in space, even small molecules, by IR astronomy requires a rotational progression in the spectrum to be measured and resolved by the telescopes. For the transitions in the simpler molecules such as CO the telescope must be capable of a resolution of $2150/1.93 \sim 1114$, which is within the resolution limit of the UK Infrared Telescope (3000–5000). However, the rotational constant for CO is rather large and many molecules, especially polyatomic species, will have a rotation constant ten times smaller than this, placing the observation of a resolved rotational progression beyond the resolution of the telescopes. Confidence in the identification of the molecule is then severely dented. The problem is worse for visible astronomy.

Visible astronomy

Visible spectroscopy and astronomy involves transitions between electronic energy levels of molecules associated with different electron configurations within the molecule and hence a different bond structure. The chemistry of molecules in excited electronic states can also be very different and perhaps they are best considered to be different molecules. Each electronic level has a manifold of vibrational energy levels, which in turn has a manifold of rotational levels. The spacing between the rotational levels is now a measure of the difference in the rotational constants for the upper and the lower states and this can be a rather small number. Separations between electronic states of molecules fall in the visible region of the spectrum, typically with a wavelength of 500 nm, somewhere in the green. Converting the visible transition frequency to a transition wavenumber gives $20\,000 \text{ cm}^{-1}$. Taking the rotational constant of CO as a favourable rotational separation, a well-resolved rotational progression in a ro-vibronic spectrum requires a telescopic resolution of $20\,000/1.93 \sim 10\,360$. The Hubble Telescope Goddard High Resolution Spectrograph in the visible has a resolution of 20 000 and for small molecules a rotational progression can be resolved; this has been done for C_2 , which has a large rotational level spacing. However, rotational constants of 1 cm^{-1} quickly put the resolved rotational progression beyond the limit of the telescope.

Visible astronomy does, however, provide most of the atomic and black body spectra of stars and astronomical objects and is of course appealing to us because the human eye is uniquely adapted to detection in the wavelength range 300–800 nm. The appeal of colour pictures has lead to the development of false colour scales used routinely by astronomers to visualise the intensity of radiation at other wavelengths. The concepts of temperature and colour are linked by the black body radiation and it

is tempting to make images that are more intense look whiter so that lower intensities are shown in the red and higher intensities are shown bluer or more white. However, confusingly some of the cooler temperatures are shown blue, which is not what is expected from the black body temperature profile, where hotter objects will emit more blue light. Many astronomical pictures have been coloured in using the cool blue, red hot scale for temperature and hence intensity, producing some beautiful but nevertheless ‘false colour’ images. This is particularly prevalent with molecular or spectral mapping, as we shall see shortly.

3.6 Molecular masers

The ro-vibronic spectrum of molecules and the electronic transitions in atoms are only part of the whole story of transitions used in astronomy. Whenever there is a separation between energy levels within a particular target atom or molecule there is always a photon energy that corresponds to this energy separation and hence a probability of a transition. Astronomy has an additional advantage in that selection rules never completely forbid a transition, they just make it very unlikely. In the laboratory the transition has to occur during the timescale of the experiment, whereas in space the transition has to have occurred within the last 15 Gyr and as such can be almost forbidden. Astronomers have identified exotic transitions deep within molecules or atoms to assist in their identification and we are going to look at some of the important ones, the first of which is the maser.

Energy levels within an atom or molecule can be populated in several ways to produce more target species in the higher energy excited state than in the ground state. The population can occur by collisional processes such as between molecules in the interstellar medium and a balance can occur between the excitation process and a number of deactivation processes (Figure 3.17a). The population of level 2 can be subjected to:

1. *Radiative decay*. The excited state can spontaneously decay and emit a photon. The rate at which this happens determines a *radiative lifetime* determined by the Einstein A coefficient in Equation 3.4 where the radiative lifetime $\tau = A^{-1}$:
 - a. Electronic energy levels: the radiative lifetime is usually of order femtoseconds – picoseconds for an allowed transition between the levels.
 - b. Electronic energy levels: forbidden transitions involve transitions that break the selection rules. Radiative lifetimes are very long, even years, but on astronomical timescales (15 Gyr) they are still observed.

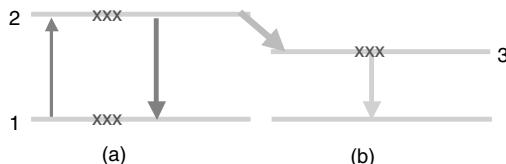


Figure 3.17 (a) Population excitation and (b) population inversion – maser transition

- c. Vibrational energy levels: radiative lifetimes are somewhat longer, of order picoseconds, although forbidden transition are also observed.
 - d. Microwave energy levels: radiative lifetimes are of order nanoseconds.
2. *Collisional deactivation.* Collisions with other molecules can stimulate the relaxation to the ground state. There are no selection rules for collisionally induced transitions.
3. *Non-radiative decay.* The population can switch away from level 2 to another level, level 3 in Figure 3.17b, producing a population inversion. More molecules are now in the excited state than the ground state.
4. *Stimulated emission.* The upper state can also decay by stimulated emission controlled by the Einstein B coefficient and the intensity of photons present of the same frequency.

Once a population inversion has occurred between an excited state and the ground state a photon of the same frequency as the transition can interact with the molecule containing the population inversion and stimulates the emission from the excited state. In a circumstellar environment the radiation field around the star contains all wavelengths and is hence likely to stimulate the transition. Transitions observed in the visible and IR part of the spectrum are called LASER transitions – Light Amplification by Stimulated Emission of Radiation – and result in the coherent production of light at the same frequency as the passing photon that is also in phase with that photon. These are very important properties of laser radiation and are exploited in spectroscopy. When stimulated emission from a population inversion involving rotational energy levels occurs in the microwave region of the spectrum the transition is said to be a MASER – Microwave Amplification by Stimulated Emission of Radiation. Rotational energy levels have longer radiative lifetimes and it is easier to populate them with the excess energy of a collision.

Maser transitions have been observed in many important molecules and have been used to carry out surveys of the entire sky. The 22.235 GHz water maser transition is the strongest transition in the radio universe and represents an interesting candidate for an interstellar broadcast frequency. If extraterrestrial intelligence is trying to communicate with us, the choice of the broadcast frequency is an important one and would be known to all intelligent life. Of course it would have a different label, 22.235 GHz being a distinctly Earthly label, but it is a fundamental transition frequency and is observed everywhere. Other maser transitions include the 6.7 and 12.2 GHz methanol maser, the SiO maser $v = 1, J = 7-6$, 301.8 GHz, which occurs between levels in the first vibration state of the SiO molecule, and finally the OH maser first discovered in 1963 and buried deep in the $^2\Pi_{3/2}$ electronic state of the hydroxyl radical near 18 cm. This is actually four transitions at 1612, 1665, 1667 and 1720 MHz, all of which must be seen as a group but not necessarily of the same intensity.

3.7 Detection of hydrogen

Detection of hydrogen is a particularly important problem for astrochemists because to a first approximation all visible matter is hydrogen. The hydrogen molecule is the most abundant molecule in the Universe but it presents considerable detection problems due to its structure and hence spectroscopy. Hydrogen does not possess a permanent dipole moment and so there is no allowed rotation or vibration spectrum and all electronic spectrum transitions are in the UV and blocked by the atmosphere. The launch of the far-UV telescope will allow the detection of H_2 directly but up to now its concentration has been inferred from other measurements. The problem of detecting the H atom, however, has been solved using a transition buried deep in the hyperfine structure of the atom.

The most famous H atom detection is the 21-cm line. The proton has a nuclear spin and the electron also has spin and there are two configurations for the atoms: either spins aligned or the lower energy spin-paired form. Collisions with other species in the local environment produce a population of the excited state (+,+ or $-,-$) but the transition back to the ground state (+, $-$) is forbidden so the excited state has a long lifetime of order 10^6 years. Transition to the ground state results in the emission of radio wave radiation at 1420.4 MHz with an equivalent wavelength of 21 cm (Figure 3.18). The emission of this line is either spontaneous, but only once every 10^6 years, or is induced by collision with other molecules leading to an actual lifetime in space of order 500 years. This apparently weak spectrum is compensated in weakness by the huge density of H atoms in space. The 21-cm line has been used to produce a map of H atoms in the radio sky and of course is yet another fundamental frequency that could be a broadcast frequency for communication with extraterrestrial intelligence.

A recent success in the detection of H species has been that of the molecular ion H_3^+ . All of the models of ion–molecule chemistry in hydrogen-dominated regions are controlled by reactions of H_3^+ but until recently the H_2^+ molecular ion had not been detected. However, the modes of vibration of H_3^+ provide for an allowed IR transition at 3.668 μm used for its detection. These ro-vibrational transitions have now been observed in a number of places, including the interstellar medium and in the aurorae of Jupiter. Not all astronomical detection and identification problems have been solved, however, and the most annoying and compelling of these is the problem of diffuse interstellar bands.

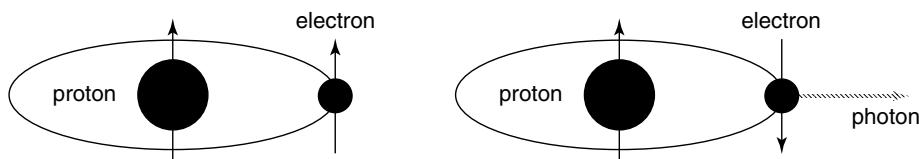


Figure 3.18 The 21-cm line in the H atom

3.8 Diffuse interstellar bands

One of the oldest problems in molecular astronomy concerns identification of the molecules responsible for *diffuse interstellar bands* (*DIBs*). Since their first observation in 1922 some 127 bands have been detected all over the electromagnetic spectrum, shown schematically in Figure 3.19, but the origins of the transitions, the so-called ‘carriers’ of DIBs, have not been determined.

The DIBs are seen along lines of sight in nearly all directions and in front of many stars and must be regarded as truly ubiquitous. They are particularly strong towards reddened stars where the blue light has been removed from the stellar spectrum along the line of sight by some blue-scattering species, possibly dust. The DIBs are in all regions of the spectrum and our best hope of understanding them must come from the microwave region where the instrumental resolution should resolve any rotational progressions. However, the ‘diffuse’ nature of the bands means that there are no nice progressions to resolve, perhaps suggesting that the molecule or molecules have large moments of inertia. Some bands appear to show an unresolved structure that might be consistent with a P- and R-branch structure but from a fairly large molecule. Some of the bands appear to be in families so that they are present together or not at all.

If you had to guess the identity of the carrier then you would have to guess H atoms because H atoms are 10 000 times more abundant than any other element. A theory for H₂ molecules as the carrier of the bands has been proposed but requires the molecule to be in a highly excited state before the spectral characteristics of the DIBs are observed. The match with the number of observed bands is then striking: more than half of the bands have coincidences with the excited H₂ spectrum. However, there seems no obvious mechanism of the production of highly excited H₂ and the astrochemical community has rejected the theory. There is almost certainly more than one carrier for the bands, which may include large carbon-containing species called polycyclic hydrocarbons and dust, as we shall see later. More exotic species such as bare carbon chains C₇⁻ and even C₆₀⁺ have been proposed based on small coincidences with the DIBs but as yet no generally accepted explanation is present. The biggest question, however, is how would you know when you have solved the DIBs problem?

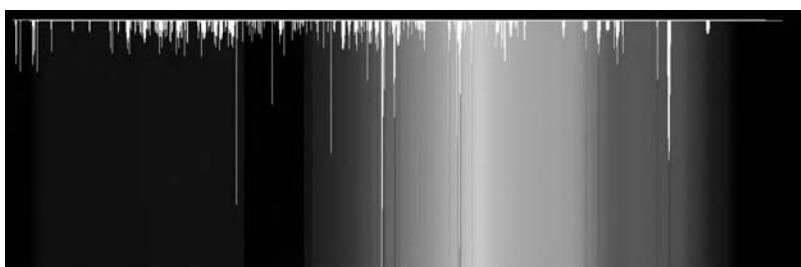


Figure 3.19 Diffuse interstellar bands. (A colour reproduction of this figure can be seen in the colour section). (Reproduced by permission of Peter Jeniskins, SETI Institute)

3.9 Spectral mapping

An obvious extension of measuring the spectrum and its successful identification in space is to look around and find out where else the molecule may be. This involves tuning the telescope to the frequency of say the 115 GHz R(0) transition in CO and looking around the entire night sky. In this way a spectral map of the heavens can be determined and the extent and distribution of matter visualised. Looking at the Milky Way and a number of wavelengths produces the beautiful spectral map shown in Figure 3.20. At the top of the figure is the 0.4 GHz and 2.4 GHz radiation in the radio continuum (a region of emission where there are no discrete transitions but instead a continuous emission profile over all frequencies), with the 21-cm line in the H atom below and the 115 GHz transition in CO showing that there is a correlation between the location of H atoms and CO molecules within the distribution of matter in the Milky Way. Three bands in the IR look at wavelengths from 100 to 3.5 μm that are all presented in false colour. These regions of the infrared correspond to absorption and scatter due to molecular vibrations and scatter by dust particles. The familiar visible region 400–700 nm is the only true colour image in Figure 3.20 and is the visible part of the emitted starlight within the galaxy. The higher energy photons in the X-ray and gamma ray regions are associated with star formation and bursts of stellar activity. A detailed look at our galaxy at all of these wavelengths reveals information about the stellar composition and age profile as well as the molecular inventory from which life has formed.

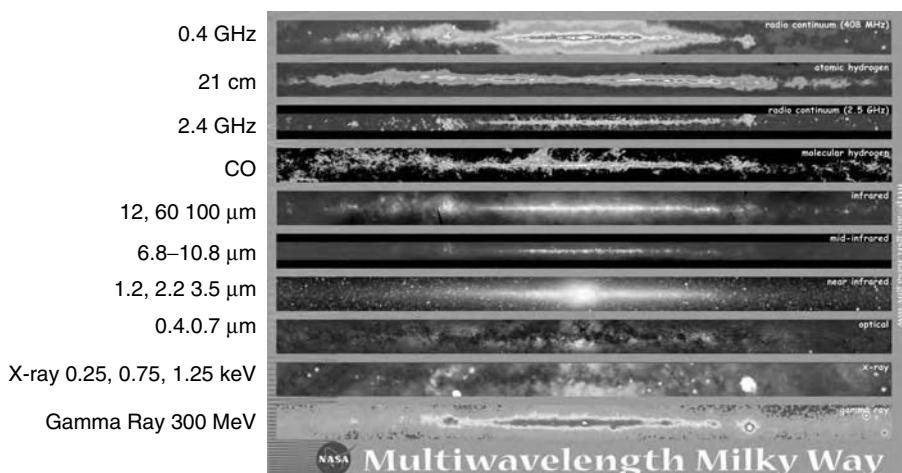


Figure 3.20 Multiwavelength view of the Milky Way. (A colour reproduction of this figure can be seen in the colour section). (Reproduced by permission of Dave Leisawitz, Observational Cosmology Laboratory, NASA Goddard Space Flight Center)

Concepts and calculations

Concepts

<i>Interaction of radiation with matter</i>	Light and matter interact by scatter and absorption. Absorption only occurs when the colour of the radiation is the same as an energy separation within the molecule, but scatter occurs at all wavelengths
<i>Doppler shift</i>	The wavelength of radiation detected depends on the relative motion of the detector and the emitter. If they are moving towards one another then the light is blue-shifted, shorter wavelength, and if they are moving apart then the light is red-shifted, longer wavelength
<i>Telescopes</i>	Instruments for collecting the light from the stars either outside the Earth's atmosphere in orbit or on the Earth collecting the light through the atmospheric windows. Each telescope has a finite spatial resolution
<i>Atomic structure</i>	The energy levels of the electron around the atom are controlled by quantum mechanics
<i>Spectroscopy</i>	The absorption of radiation between two energy levels within an atom or molecule
<i>Selection rules</i>	Some transitions in spectroscopy are allowed and some are forbidden
<i>Microwave spectroscopy</i>	The study of the rotational motion of molecules by the absorption of radiation – the technique of choice for the identification of molecules
<i>Infrared spectroscopy</i>	The study of the vibrational and rotational motion of molecules by the absorption of radiation
<i>Electronic spectroscopy</i>	The study of rotational, vibrational and electronic motion in molecules following absorption of radiation in the visible region of the spectrum
<i>Maser</i>	The microwave equivalence of a laser. A stimulated transition in a molecule that relaxes a population inversion in an excited state
<i>Spectral mapping</i>	Setting a spectrometer to a specific transition in an atom or molecule and mapping the sky

Calculations

<i>Beer–Lambert Law</i>	The loss of radiation from a line of site due to extinction
<i>Doppler shift</i>	Estimation of changes in wavelength with relative velocity

<i>H atomic spectrum</i>	Calculation of the allowed transition frequencies in the atomic spectrum of the H atom
<i>Microwave spectra</i>	Calculation of reduced mass, rotational constant, bond length, isotope shift, allowed transition frequencies and relative intensities
<i>Vibrational spectroscopy</i>	Calculation of diatomic energy level spacings, isotope shifts and allowed transitions in a vibronic spectrum

Problems

- 3.1** The extinction for dust particles is found to follow the empirical relation $\varepsilon_\lambda = 0.008 \lambda^{-4/3}$ mag pc⁻¹ (Equation 5.3 – a truly terrible astronomical unit for a scientist!). Calculate the extinction per metre at a wavelength of 501 nm. The light from a star at 501 nm is 0.25 of that expected for a star with that surface temperature and the extinction is attributed to dust in an interstellar cloud. Calculate the diameter of the cloud along the line of sight.
- 3.2** A Doppler shift for the H α transition is induced by the rotation of Jupiter.
- (a) Calculate the laboratory wavelength of the H α transition to three decimal places.
 - (b) Calculate the Doppler shift for the ascending and descending limb of the Jovian equator, given a rotational speed of 43 000 km h⁻¹.
- 3.3** The distance to the Andromeda galaxy is 4.3 ly.
- (a) Use Hubble's Law to calculate the red shift for the galaxy.
 - (b) The observed R(0) transition in CO in the Andromeda Galaxy is at 114.9 GHz. Calculate the Doppler shift for the transition.
 - (c) Account for the difference in these two calculations
- 3.4** The Einstein *A* coefficient for the electronic transition to the first excited electronic state at 65075.8 cm⁻¹ is 1.37 ns. Calculate the Einstein *B* coefficient for the induced transition rate.
- 3.5** The rotational constant for ¹²C¹⁶O is 1.9313 cm⁻¹. Calculate:
- (a) The moment of inertia of CO.
 - (b) The bond length of CO.
- 3.6** The observed transition intensity ratio for R(1)/R(0) of CO in a molecular cloud is 1.10. Calculate the temperature of the molecular cloud.
- 3.7** The R(0) transition is observed in ¹²CO and ¹³CO in two molecular clouds along the line of sight. Calculate the relative velocities of the two clouds if the R(0) transitions in each molecule are coincident. Comment on the possible reason for the relative velocities.
- 3.8** Calculate the transition wavelengths of the first four terms and the convergence wavelength for the Balmer series and identify the red transition responsible for the colour of the Orion nebula.

- 3.9** The ionisation energy of the H atom corresponds to removing the electron from the $n = 1$ energy level. Determine the convergence limit for the H Lyman series and hence estimate the ionisation energy of the H atom in kJ mol^{-1} .
- 3.10** Which of the named sequences in the spectrum of the H atom can be observed in atmospheric windows?
- 3.11** Calculate the reduced mass of ${}^1\text{H}{}^{35}\text{Cl}$ and ${}^1\text{H}{}^{37}\text{Cl}$.
- 3.12** The rotational constant for ${}^1\text{H}{}^{35}\text{Cl}$ is 10.937 cm^{-1} .
- Calculate the bond length of HCl.
 - Calculate the transition frequency of the R(0) transition.
 - Is the transition visible in an atmospheric window?
- 3.13** The rotational constants for ${}^1\text{H}{}^{35}\text{Cl}$ and ${}^2\text{H}{}^{35}\text{Cl}$ are 10.937 cm^{-1} and 5.449 cm^{-1} .
- Calculate the reduced masses of these species and rationalise the difference in the rotational constants.
 - If the force constant for all of the isotopes of HCl is 510 N m^{-1} , assume that the molecule is a harmonic oscillator and calculate transition wavenumbers for the $v = 0-1$, R(0) transitions for each of these species.
 - Would the transitions be resolvable using the UK Infrared Telescope spectrometer?
- 3.14** The IR spectrum of ${}^1\text{H}{}^{35}\text{Cl}$ contains two strong transitions at wavenumbers of 2886 cm^{-1} and 5668 cm^{-1} . Use these frequencies to determine v_e and x_e for the molecule.
- 3.15** The harmonic transition wavenumber for ${}^{12}\text{C} {}^{18}\text{O}$ is 2143.23 cm^{-1} .
- Use this value to calculate the force constant for the C≡O triple bond.
 - Calculate the harmonic transition wavenumber for ${}^{13}\text{C} {}^{16}\text{O}$.
 - Calculate the harmonic transition wavenumber for ${}^{12}\text{C} {}^{18}\text{O}$.

4

Stellar chemistry

Introduction

Stars seem to form in clusters that fall into two general types: globular clusters and open clusters. The globular clusters contain many thousands of stars but looking at their spectra suggests that they are deficient in heavy atoms such as iron. The absence of heavier atoms suggests that these stars are very old and formed from the materials in the primordial Big Bang. The open clusters contain 50–1000 stars and their spectra contain heavier elements suggesting that they are younger stars and formed from supernova remnants. To understand the statements about the formation of stars in globular and open clusters we need to know something of the process of star formation.

According to the current model of star formation, the first generation of stars must have formed from a cloud of interstellar matter composed of H, He and some Li formed during the Big Bang. A typical mass of a giant molecular cloud that we can see today is of order $10^6 M_{\text{Sun}}$ and some 100 ly in diameter with a density of one billion hydrogen molecules per cubic metre, $10^9 \text{ H}_2 \text{ m}^{-3}$. Consider what happens when a region of local high density of order $10^{11} \text{ H}_2 \text{ m}^{-3}$ and temperature of order 10 K begins to collapse under the influence of gravity. The mass of the cloud has to achieve a critical mass called the Jeans Mass if the gravitational collapse is to be spontaneous. Initially, gravity attracts the matter together, converting the gravitational potential energy into kinetic energy or temperature. The collapse near the centre is rapid and the temperature begins to increase to form a core of material with a temperature of a few hundred Kelvin. As the pressure increases, the gravitational attraction is balanced by the heat of expansion and the contraction stops. A protostar is said to have formed (Figure 4.1).

The temperature of the protostar results in the emission of IR radiation, radiating away some of the gravitational energy. This IR emission can be seen in the centre of big clouds where stars are beginning to form: regions called stellar nebulae. The envelope of material around the core continues to shower mass into the protostar and it begins to collapse again. This process takes about 10^6 years. As the contraction continues the core becomes optically thick, preventing the IR photons from escaping, and the temperature rises to 1500–2000 K.

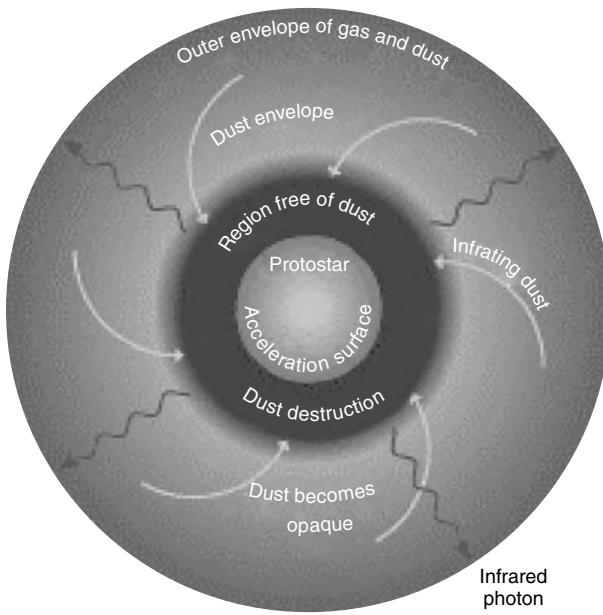


Figure 4.1 A protostar. (Reproduced from Figure 14.13 “Astronomy, the Evolving Universe” 9th Edition, 2002, by permission of Boris Starosta and Michael Zeilik)

At 2000 K there is sufficient energy to make the H₂ molecules dissociate, breaking the chemical bond; the core density is of order 10²⁶ m⁻³ and the total diameter of the star is of order 200 AU or about the size of the entire solar system. The temperature rise increases the molecular dissociation, promoting electrons within the hydrogen atoms until ionisation occurs. Finally, at 10⁶ K the bare protons are colliding with sufficient energy to induce nuclear fusion processes and the protostar develops a solar wind. The solar wind constitutes outbursts of material that shake off the dust jacket and the star begins to shine.

What is the importance of stars to our running theme, the origins of life? The nuclear fusion processes within the core of the star are the only source of the heavier nuclei and hence all elements with greater mass than H, He and Li. Secondly, and more strikingly on the surface of the Earth, the star is a possible source of heat for any planet on which life might evolve. Life on Earth certainly depends, at the surface, on the warming provided by our local star, the Sun: this is also the energy source driving photosynthesis and the conversion of CO₂ and H₂O into sugars. It is not, however, the only source of energy. Below the surface, volcanic activity provides thermal energy and a rich mineral diversity allowing exothermic chemistry to drive the reactions of life.

4.1 Classes of stars

The radiation coming from a black body is continuous in wavelength but dispersing the radiation from a star reveals absorption lines, which can be identified clearly

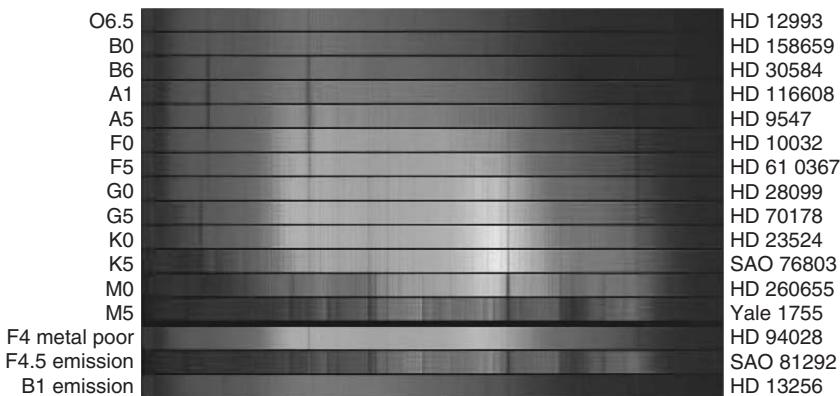


Figure 4.2 Spectra from different stellar classes OBAFGKM. (A colour reproduction of this figure can be seen in the colour section.) (Reproduced by permission of KPNO 0.9-m Telescope, Aura, NOAO, NSF)

from atomic spectroscopy. It was from an analysis of the spectrum of the Sun that helium was first identified. Direct comparison with laboratory spectra enables the chemical composition of the stars to be determined. This is a recurring theme – an understanding of the laboratory spectrum of an atom or molecule identifies it in astronomical environments and provides important information on the molecular environment. Investigation of the spectrum reveals a number of spectral features associated with the emissions of atomic and molecular species present within the visible disk or photosphere of the star (Figure 4.2).

The spectral features observed by astronomers have led to the classification of stars into seven broad classes outlined in Table 4.1, together with their surface temperatures. The highest-temperature class, class O, contains many ionised atoms in the spectrum whereas the older stars in class M have a much lower temperature and many more elements present in the spectrum of the star. Observation of a large number of the stars has lead to extensive stellar catalogues, recently extended by the increased sensitivity of the Hubble Space Telescope. Making sense of this vast quantity of information is difficult but in the early 19th century two astronomers

Table 4.1 Stellar classification

Spectral type	Surface temperature (K)	Spectral features
O	35 000	Ionised He and weak H absorption
B	20 000	Non-ionised He and stronger H
A	10 000	H absorption dominates
F	6 500	H absorption and heavier elements
G	5 500	H absorption weakening with increasing non-ionised heavy elements
K	4 500	No H lines and more non-ionised heavy elements
M	3 000	Many heavy element spectra and simple molecules

Herzprung and Russell independently analysed the spectral types, luminosities and temperatures to produce the Herzprung–Russell diagram.

The hottest stars have absorption features in the photosphere associated with lighter elements, some in highly ionised states, but the lower temperature stars have a more diverse atomic composition; the coolest stars show molecular emission spectra. This suggests an evolution of stars that involves the formation of heavier elements and ultimately molecules.

4.2 Herzprung–Russell diagram

The luminosity and surface temperature of a star are its two most important observable properties and are related to the mass at formation. A plot of luminosity against surface temperature produces the Herzprung–Russell diagram or HR diagram. There are alternative choices for axes, including class against colour magnitude, but when temperature is used the lower axis is plotted with increasing temperature from right to left as seen in Figure 4.3, here plotted with stellar luminosity relative to that of the Sun, L_{Sun} .

When all stars are plotted on the HR diagram there is a great swath of stars falling in the region from top left to bottom right that accounts for 92 per cent of all observed stars, with over half of them being of class M. This region is termed

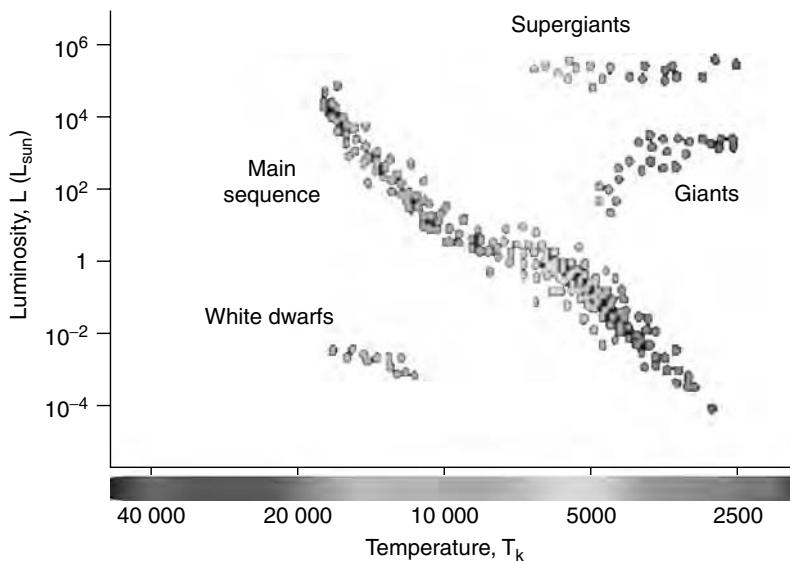


Figure 4.3 Herzprung–Russell diagram. (A colour reproduction of this figure can be seen in the colour section.) (Reproduced by permission of D.Padgett (IPAC/Caltech), W.Brandner (IPAC), K.Stapelfeldt (JPL) and NASA)

the *main sequence* and stars found this region are referred to as main-sequence stars, including our Sun. Most of the remaining stars, accounting for 7.5 per cent, appear at the bottom left-hand corner of the diagram and are termed *white dwarfs* because their low luminosities and high temperatures suggest that the stars are small in diameter.

A good example of a white dwarf is Sirius B, which has a black body spectrum with $\lambda_{\max} = 110$ nm corresponding to a surface temperature of 26 000 K, although its absolute magnitude and hence luminosity suggest that its radius is 4200 km, smaller than that of the Earth. The remaining 0.5 per cent of stars are found above the main sequence and are characterised by large radii and low surface temperatures. These stars are termed *giants* and *supergiants*.

The four general types of stars (main sequence, white dwarfs, giants and supergiants) provide a classification based on the fundamental observable properties but also suggest an evolution of stars. Astrochemically, the cooler giants and supergiants have many more atomic and molecular species that are the products of the nuclear fusion processes responsible for powering the stars. The nuclear fusion processes allow for the formation of more of the elements in the Periodic Table, especially the heavier elements that dominate life on Earth – principally carbon.

4.3 Stellar evolution

Big Bang nucleosynthesis produced only H and He atoms with a little Li, from which nuclei the first generation of stars must have formed. Large clouds of H and He when above the Jeans Mass condensed under the influence of gravitational attraction until they reached the temperatures and densities required for a protostar to form, as outlined. Nuclear fusion powers the luminosity of the star and also results in the formation of heavier atomic nuclei.

Observational evidence for the formation of protostars requires intense low-temperature objects to be seen within collapsing molecular clouds. There is an increasing number of observations of these pre-nuclear fission objects dominated by long-wavelength emission in the infrared and maser emissions (see Chapter 3). On the HR diagram this corresponds to an evolutionary line from the bottom right usually increasing above the main sequence, suggesting that they are large objects. A possible candidate for such an object is the Becklin–Neugebauer (B-N) object in the Orion nebula, which has an infrared luminosity of order $10^3 L_{\text{Sun}}$. The descent onto the main sequence follows a birth line that is determined exclusively by the mass of the protostar (Figure 4.4).

The B-N object may be considered to be a zero-age main sequence star that evolves with increasing surface temperature and luminosity at optical wavelengths. The descent from the right-hand upper quarter of the HR diagram, along with what has been called the Hayashi track or birth lines, precedes the entrance onto the

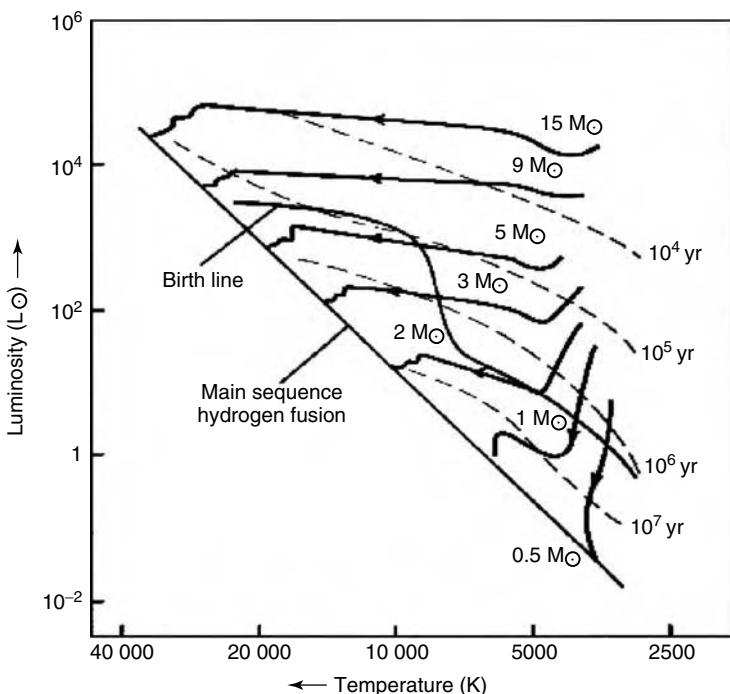


Figure 4.4 Birth lines for stars descending onto the main sequence, where M_{\odot} denotes solar mass

main sequence where nuclear fusion processes produce energy. The descent to the main sequence may take as long as 10^6 – 10^7 years and some stellar objects have been assigned to this phase. The best candidates are the T-Tauri stars (Figure 4.5) and their hotter relatives the Herbig Ae and Be stars. About 1000 T-Tauri stars have been observed so far and they occupy a key position in the cycle of star formation before the star joins the main sequence.

Energy and mass

Nuclear fusion processes derive energy from the formation of low-mass nuclei, which have a different binding energy. Fusion of two nuclear particles produces a new nucleus that is lighter in mass than the masses of the two fusing particles. This mass defect is then interchangeable in energy via Einstein's equation $E = mc^2$. Specifically, the formation of an He nucleus from two protons and two neutrons would be expected to have mass:

$$2m_p + 2m_n = 6.69688 \times 10^{-27} \text{ kg} \quad (4.1)$$

where m_p and m_n are the mass of the proton and neutron, respectively. However the observed mass of the He nucleus is $m_{\text{He}} = 6.644446 \times 10^{-27}$ kg, a mass

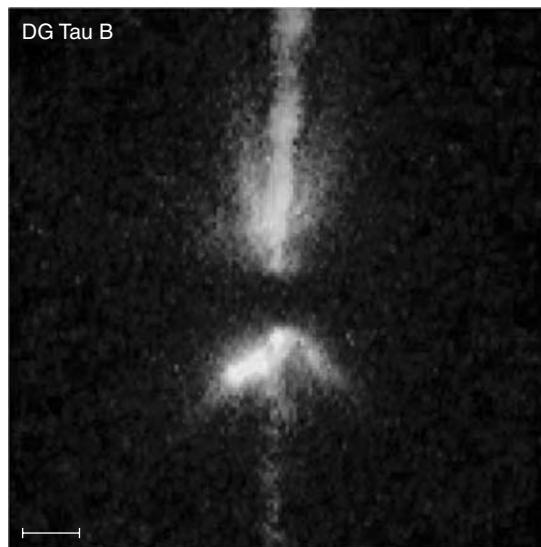


Figure 4.5 T-Tauri Young Stellar Object with characteristic bipolar jets. An excellent example of the complementary nature of Hubble's instruments may be found by comparing the infrared NICMOS image of D9 Tau B with the visible-light Wide Field and Planetary Camera 2 (WFPC2) image of the same object. The WFPC2 highlights the jet emerging from the system whereas NICMOS penetrates some of the dust near the star to outline more clearly the 50 billion-mile-long dust lane (the horizontal da/k band). The young star itself appears as a bright spot at the corner of the V-shaped nebula. Reproduced by permission of D. Padgett (IPAC/Caltech), W. Brandner (IPAC), K. Stapelfeldt (JPL) and NASA

defect of 5.242×10^{-29} kg that is the equivalent of 4.72×10^{-12} J released for the formation of each new He nucleus. This energy comes out in a variety of particles, including photons and thermal energy. Nuclear physicists choose not to work in joules and convert the energy into the unit MeV, where 1 eV is the potential energy acquired by an electron in moving through a potential difference of 1 V: 1 eV is 1.602192×10^{-19} J. Hence, the mass defect for the He nucleus formation process is 4.72×10^{-12} J/ 1.602192×10^{-19} J = 29.5 MeV. The nuclear fusion process is thus an exothermic process for the formation of nuclei with a mass defect. The binding energy per nucleon for atomic nuclei up to mass 220 amu is shown in Figure 4.6, featuring a broad maximum centred at the mass of Fe. Exothermic nuclear fusion processes will produce atomic nuclei up to Fe, releasing energy, but heavier elements cannot be formed in this way. For the heavier elements fission processes are exothermic, where splitting a heavy nucleus into its two daughter nuclei results in a mass defect and the release of energy. The fission/fusion line is shown in Figure 4.6 and marks the divide for the synthesis of nuclei in stars: nuclei with masses heavier than $^{26}\text{Fe}^{56}$ (the most abundant isotope on Earth) are not formed in stars so what is the formation process of the heavier nuclei that we find on Earth, such as $^{92}\text{U}^{235,238}$?

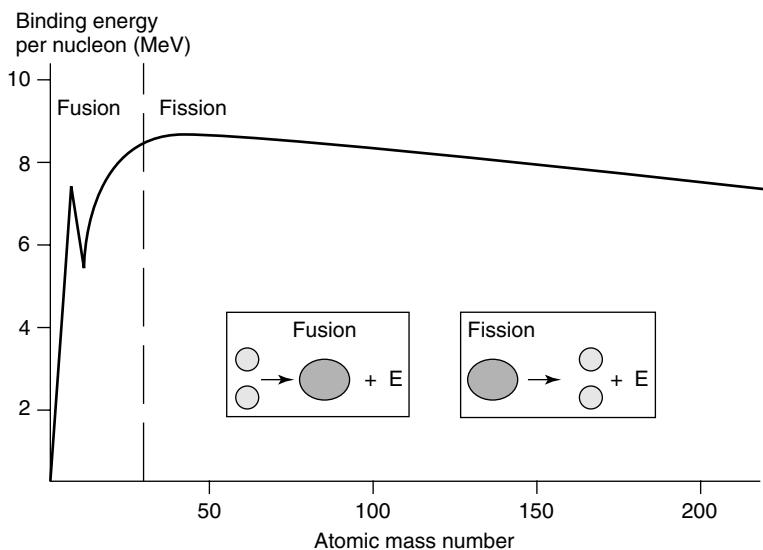
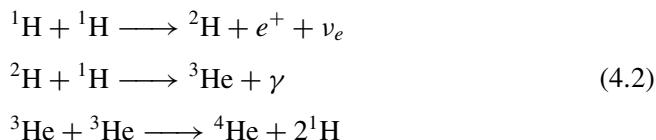


Figure 4.6 Binding energy of atomic nuclei

The nuclear generation processes in stars, however, do not follow simple fusion of nuclei one at a time. There are preferential networks of nuclear reactions called cycles that produce different atomic nuclei, but only some cycles are triggered in heavier stars.

The proton–proton cycle

The basic energy source in young stars is the fusion of protons into the ${}^4\text{He}$ nucleus in three steps. This process is occurring in our Sun and other main-sequence stars where it produces about 91 per cent of the total energy. The proton–proton cycle proceeds via three reaction steps:



The first reaction is a fusion of two protons to produce a ${}^2\text{H}$ nucleus, a positron (e^+) and a neutrino (ν_e). The second reaction is a proton capture with the formation of ${}^3\text{He}$ and a γ -ray. In the third reaction two ${}^3\text{He}$ nuclei fuse to give ${}^4\text{He}$ and two protons. The total energy released in one cycle is 26.8 MeV or 4.30×10^{-12} J. An important product of this process is the neutrino and it should provide a neutrino flux from the Sun that is measurable at the surface of the Earth. However, the measured flux is not as big as calculated for the Sun – the so-called neutrino deficit

problem. Allowing for the observed mass of the Sun, the total energy produced during one cycle is 6.4×10^{14} J and some 6×10^{11} kg of H is converted to He every second.

The triple-alpha process

At a temperature of 10^8 K, He starts to ‘burn’. The first stable product formed from a triple He nucleus collision is ^{12}C : the *triple-alpha process* (Figure 4.7). The process is less efficient than the proton–proton cycle, producing about 10 per cent of the energy per kilogram. Helium burning will convert the core of the Sun to carbon but the core will not be dense enough and not have sufficient pressure to allow the carbon nuclei to fuse. Thus the Sun will eventually run out of nuclear fuel and become a large carbon cinder. However, the triple-alpha process produces the most important atomic nucleus for the origins of life – carbon.

Evolution of the Sun (1-solar-mass star)

A Sun-like star spends 80 per cent of its life on the main sequence, called its *main-sequence lifetime*. The main-sequence phase ends when almost all of the H in the core has been converted to He. There is a greater energy flow to the surface during this phase and the luminosity of the star increases. The star is now poised to become a red giant. Hydrogen fusion reactions stop in the core but continue in the shell around the core. With no heating in the core it begins to contract under the force of gravity to produce more energy, an increase in the luminosity and an increase in the radius of the star. The Sun will then become a red giant. The core becomes a degenerate electron gas, which maintains a pressure and hence core diameter. As the temperature rises the temperature for the triple-alpha process is reached and He burning starts in a helium flash.

Eventually the He burning converts the core to carbon. Core thermonuclear shutdown occurs but fusion continues in the shell around the core, producing He followed by periods of helium flashes, causing wild variations in the luminosity. The Sun will then develop a super wind by convection that will blow off the overshell of the star, leaving a hot core behind. The expelled material forms a shell

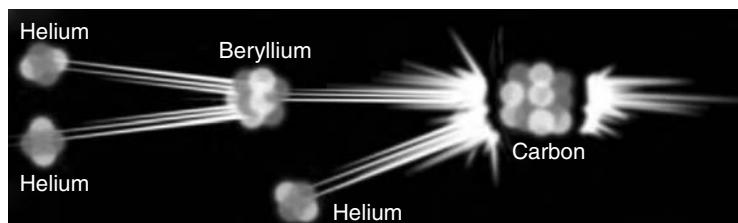


Figure 4.7 The triple-alpha process

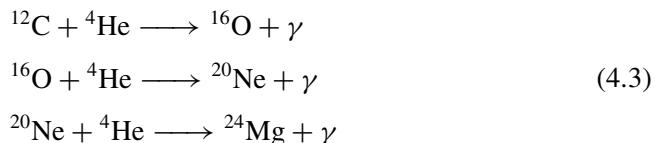
of heated gas called a planetary nebula (for historical reasons) and is imaged using the Hubble Space Telescope in a number of images, but notably NGC6369. The gas expands at speeds of 20 km s^{-1} . The nebula expands until it dissipates into the interstellar medium, carrying matter in the form of atomic nuclei, including carbon. The chemistry of the interstellar medium will be discussed in Chapter 5. The core does not have sufficient mass to ignite carbon burning so in about 75 000 years it becomes a white dwarf composed mostly of carbon. With no further energy sources the Sun will cool and become a black dwarf with a total stellar age of 10 Gyr: a carbon-cinder corpse of a star.

The Sun will expand to about 1.1 AU in diameter, engulfing the Earth. At the orbital speeds of the Earth, some 30 km s^{-1} , the atmosphere will be removed and the Earth will heat up. In somewhat less than 200 years the Earth will be vaporised, so in about 2 Gyr we will have to evacuate the Earth and look for a new planet on which to live.

Nuclear synthesis in heavy stars

The 1-solar-mass star is not capable of producing the heavier nuclei and the majority of the elements in the Periodic Table. With sufficient initial mass of H the He core produces ^{12}C , the point at which the star starts to die. The formation of an almost pure He core results in the end of the star's main-sequence lifetime. The core of the star begins to contract, raising the temperature to 10^8 K until the nuclear reactions involving He start.

At around $10^8\text{--}10^9 \text{ K}$ alpha capture also occurs, producing a number of nuclei:



Alpha capture can now continue in a similar way to give ^{28}Si , ^{32}S , ^{36}Ar , etc. Once ^{12}C is formed, however, a new cycle called the carbon–nitrogen–oxygen (CNO) cycle takes over and produces other nuclei (not all of which are stable) such as ^{13}C , ^{13}N , ^{14}N and ^{15}O , all of which can undergo alpha capture to produce other elements, as shown in Figure 4.8. The Sun has only a fraction of its energy coming from the CNO cycle but this will increase as it gets older. At higher temperatures the star can start burning carbon directly by fusion processes such as $^{12}\text{C} + ^{12}\text{C} = ^{24}\text{Mg}$.

Large-scale production of the heavier nuclei requires a heavier initial mass than that of the Sun. All nuclei up to iron were formed in stars with about 20 solar masses, as summarised in Table 4.2, which shows the important role of heavier stars in the cosmic nuclear synthesis of heavier elements.

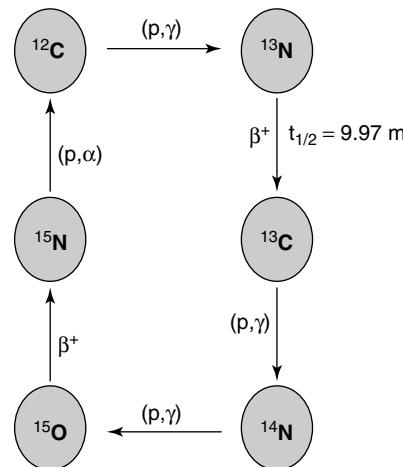


Figure 4.8 The CNO cycle $T > 1.6 \times 10^7 \text{ K}$ $M > 1.1 \text{ Solar Masses}$

Table 4.2 Stages of thermonuclear generation in stars

Process	Fuel	Major products	Approximate core temperature (K)	Minimum mass in solar masses
Hydrogen burning	Hydrogen	Helium	2×10^7	0.1
Helium burning	Helium	Carbon, oxygen	2×10^8	1
Carbon burning	Carbon	Oxygen, neon, sodium, magnesium	8×10^8	1.4
Neon burning	Neon	Oxygen, magnesium	1.5×10^9	5
Oxygen burning	Oxygen	Magnesium to sulphur	2×10^9	10
Silicon burning	Magnesium to sulphur	Elements near iron	3×10^9	20

Evolution of heavier stars

Red giant stars have very high luminosities but low surface temperatures (2000–6000 K) and to maintain these photon fluxes the stars have to be very large. Red giants are diffuse stars with a very tenuous photosphere surrounding a bright core and are generally hundreds of times bigger than the Sun. Betelgeuse is one example. The outflow of mass from all stars results in the formation of white dwarfs, of size typically that of the Earth and about the mass of the Sun. The fate of the white dwarf can be catastrophic.

As the density of matter increases, reaching $10^{15} \text{ kg m}^{-3}$, the atomic nuclei collapse resulting in a star consisting of neutrons – a neutron star. This occurs for stars greater than 1.4 solar masses. The stars are then thought to become unstable and blow off large amounts of matter from the outer regions of the stars. Such an event is called a *nova*, which may involve the core of the star resulting in

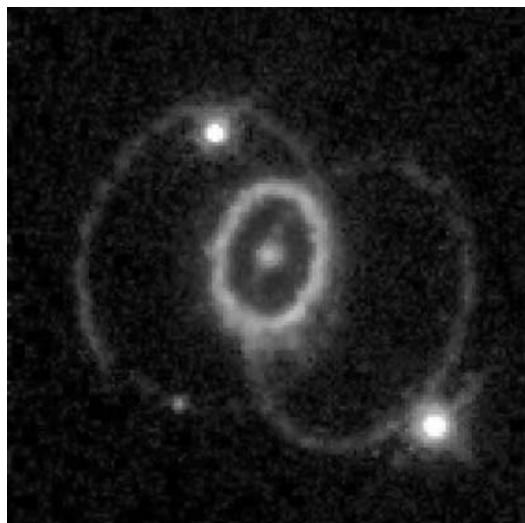


Figure 4.9 Supernova 1987A captured by the HST wide Field Planetary Camera 2. (Reproduced with permission from P.Challis, R.Kirshner (Harvard Smithsonian Centre for Astrophysics) and B. Sugarman (STScI) and NASA.). (A colour reproduction of this figure can be seen in the colour section.)

a *supernova* first captured in 1987, producing an amazing sequence of images (Figure 4.9) and rather casually named Supernova 1987A.

The instability of the core is thought to generate a huge shock wave that propagates throughout the neutron degenerate core and then, when it reaches the surface, theoretical calculations suggest that material is ejected from the surface with a kinetic energy of 10^{44} J and the energy released in the visible region of the year of the explosion is thought to be 10^{42} J. Shock waves of this ferocity will propagate through space, passing through interstellar clouds and making them unstable with respect to gravitational collapse. The shock wave will then trigger the collapse of the new cloud and start the formation of a new star.

Stars of mass greater than 1.4 solar masses have thermonuclear reactions that generate heavier elements (see Table 4.3) and ultimately stars of approximately 20 solar masses are capable of generating the most stable nucleus by fusion processes, Fe. The formation of Fe terminates all fusion processes within the star. Heavier elements must be formed in other processes, usually by neutron capture. The ejection of neutrons during a supernova allows neutron capture events to increase the number of neutrons in an atomic nucleus. Two variations on this process result in the production of all elements above Fe. A summary of nucleosynthesis processes is summarised in Table 4.4. Slow neutron capture – the *s-process* – occurs during the collapse of the Fe core of heavy stars and produces some higher mass elements, however fast or rapid neutron capture – the *r-process* – occurs during the supernova event and is responsible for the production of the majority of heavy nuclei.

Table 4.3 Stellar evolution

Mass	Evolutionary sequence
Low (<1 solar mass)	Protostar → pre-main sequence → main sequence → red giant → planetary nebula → white dwarf → black dwarf
Middle (5–10 solar masses)	Protostar → main sequence → red giant → planetary nebula or supernova → white dwarf or neutron star
High (>20 solar masses)	Protostar → main sequence → supergiant → supernova → neutron star

Table 4.4 Nucleosynthesis processes

Nucleosynthesis process	Elements formed
Big Bang nucleosynthesis (cosmic nucleosynthesis)	High-energy photon collisions produce antimatter–matter pairs. This leads to H,D, He and some Li nuclei
Proton–proton cycle	Hydrogen burning to produce He
Triple He collisions	^{12}C production
Alpha capture	Addition of ^4He to the nucleus
CNO cycle	Production of ^{13}C , ^{13}N , ^{14}N and ^{15}O
Neutron capture	Post-Fe nuclei

Origin of the elements

The discussion of the nuclear fusion energy sources within stars leads to the realisation that all atoms that make up life on Earth were formed in stars. The Big Bang nucleosynthesis produced large amounts of H and He with some Li but the remainder of the atomic nuclei were forged by fusion processes in stars. More interestingly, the elements that are heavier than Fe were formed by neutron capture events probably during supernova events. Looking at the Periodic Table there are large numbers of elements heavier than Fe that are important to the everyday understanding of chemistry on Earth and in particular biochemistry. Elements such as Cu, Zn, Br, I, Ag, Au, Pt, Pd and of course the very heavy lanthanides and actinides were all formed during a supernova.

The matter that made up the solar nebula from which the solar system was formed already was the product of stellar birth, aging and death, yet the Sun is 4.5 billion years old and will perhaps live to be 8 billion years but the Universe is thought to be 15 billion years old (15 Gyr) suggesting that perhaps we are only in the second cycle of star evolution. It is possible, however, that the massive clouds of H atoms, formed in the close proximity of the early Universe, rapidly formed super-heavy stars that had much shorter lifetimes and entered the supernova phase quickly. Too much speculation becomes worrying but the presence of different elements in stars and the subsequent understanding of stellar evolution is supported by the observations of atomic and molecular spectra within the light coming from the photosphere of stars.

4.4 Stellar spectra

Identifying stars in globular and open clusters as either old stars from the primordial explosion or new stars formed after a supernova event is based on the atomic composition of the stars. The primary way of identifying the elements in any excited state is to study the atomic spectroscopy of the stellar spectrum, such as in Figure 4.2 and identify the atoms by assigning the spectra. This becomes a complicated process for the heavier elements but is very informative even for the simple H-atom spectra.

Population of the Balmer series

Atoms and molecules have available to them a number of energy levels associated with the allowed values of the quantum numbers for the energy levels of the atom. As atoms are heated, some will gain sufficient energy either from the absorption of photons or by collisions to populate the levels above the ground state. The partitioning of energy between the levels depends on temperature and the atom is then said to be in local thermal equilibrium with the populations of the excited states and so the local temperature can be measured with this atomic thermometer.

The Balmer series is seen in many but not all stars because the first energy level in the series is an excited state with quantum number $n = 2$. There has to be a mechanism by which the excited state is populated and this is the local temperature of the star. Hence only if the star is sufficiently hot will the spectrum contain the Balmer series. The strength of the Balmer series within the stellar spectrum can be used to derive a temperature for the surface of the star to compare with black body temperature and the B/V ratio.

When the atom is in thermal equilibrium with its surroundings in the photosphere of the star, the population of the $n = 2$ level is given by the Boltzmann law:

$$\frac{n_2}{n_1} = \frac{g_2}{g_1} \exp(-\Delta E/kT) \quad (4.4)$$

where n_i are the relative populations of the levels, g_i are the level degeneracies, ΔE is the separation between the energy levels, k is the Boltzmann constant and T is the surface temperature of the star.

We have met the concept of degeneracy of energy levels with the $(2J + 1)$ degeneracy of the rotational levels of a rigid rotor. Here in the H atom the energy levels with energy determined by the n quantum number are also degenerate, associated with the orbital angular momentum of the electron in the orbital. The angular momentum quantum number is l and takes values from $n - 1 \dots 0$. When $l = 0, 1, 2, 3, 4, \dots$ these are referred to as s, p, d, f, g, ..., giving the names of the familiar s-orbitals, p-orbitals, etc. Each l -orbital is $(2l + 1)$ degenerate so there is one s-orbital, three p-orbitals and so on. The energies of the orbitals are different in an external magnetic field, as we shall see shortly.

Example 4.1

For the Sun the ratio $n_2/n_1 = 5.15 \times 10^{-9}$ and the Balmer series lines are not very strong. The difference in the energy levels for the H atom is given by the expression:

$$E(n=2) - E(n=1) = -R \left[\frac{1}{(n=1)^2} - \frac{1}{(n=2)^2} \right] \quad (4.5)$$

$$\Delta E = -R(1 - 1/4) = 82257.8 \text{ cm}^{-1} = 1.634 \times 10^{-18} \text{ J}$$

The degeneracies of each of the levels can be calculated from the value of n , the allowed values of l and the total degeneracy of the l -orbitals:

$$\begin{aligned} n = 1 \Rightarrow l = 0, \therefore g_1 &= 1 \quad 1s \text{ orbital} \\ n = 2 \Rightarrow l = 1 \quad (2l + 1) &= 3 \quad 2p \text{ orbitals} \\ \Rightarrow l = 0 \quad (2l + 1) &= 1 \quad 2s \text{ orbital} \\ \therefore g_2 &= 4 \end{aligned} \quad (4.6)$$

Substituting the value for the population ratio $n_2/n_1 = 5.15 \times 10^{-9}$ derived from the intensity of the transitions in the Balmer series into Equation 4.4 allows the Balmer temperature to be calculated:

$$\begin{aligned} \frac{n_2}{n_1} &= 5.15 \times 10^{-9} = \frac{4}{1} \exp(-1.634 \times 10^{-18}/kT) \\ \Rightarrow T &= 5780 \text{ K} \end{aligned} \quad (4.7)$$

This is a similar value to the temperature of the Sun derived optically or from λ_{\max} in the black body spectrum. In a colder star the Balmer series is weaker still, but in a hotter star the Balmer series lines are stronger. In very hot stars the Balmer series may become weaker again due to collisional ionisation of H atoms, removing the electrons from the atoms completely.

Heavy element atomic spectra in stars

The spectrum of the Sun contains the absorption lines associated with the atomic spectrum of heavier elements such as Fe (Figure 4.2), which indicates that the Sun is a second-generation star formed from a stellar nebula containing many heavy nuclei. The atomic spectra of heavier atoms are more complex. The simple expression for the H atom spectrum needs to be modified to include a quantum defect but this is beyond the scope of this book. Atomic spectra are visible for all other elements in the same way as for H, including transitions in ionised species such as Ca^{2+} and Fe^{2+} (Figure 4.2).

Unfortunately, the Schrodinger equation for multi-electron atoms and, for that matter, all molecules cannot be solved exactly and does not lead to an analogous expression to Equation 4.5 for the quantised energy levels. Even for simple atoms such as sodium the number of interactions between the particles increases rapidly. Sodium contains 11 electrons and so the correct quantum mechanical description of the atom has to include 11 nucleus–electron interactions, 55 electron–electron repulsion interactions and the correct description of the kinetic energy of the nucleus and the electrons – a further 12 terms in the Hamiltonian. The analysis of many-electron atomic spectra is complicated and beyond the scope of this book, but it was one such analysis performed by Sir Norman Lockyer that led to the discovery of helium on the Sun before it was discovered on the Earth.

The Zeeman effect

The spectra of stars is further complicated by the presence of a magnetic field. Rarely in the laboratory is the magnetic field strong enough to influence the spectra of atoms but the Zeeman effect, as it is called, does reveal some further information on the nature of the quantum numbers in the atom. The energy of the levels in the atom is controlled solely by the principal quantum number n but there are two further quantum numbers l and m that are required for the complete description of the electron within the atom. The l quantum number describes the orbital angular momentum of the electron and m describes how the angular momentum may be projected onto the axis of an external field, such as a magnetic field. The m projection quantum number is a difficult concept and it will suffice to say (unsatisfactorily) that it appears naturally in the solution of the Schrodinger equation.

Pictures of the orbital angular momentum of the electron in the hydrogen atom are familiar in the form of the shapes of atomic orbitals. The l number can take the values $l \leq n - 1$, so, for $n = 2$, l can take values 0 and 1. When $l = 0$ the orbital is called an s-orbital, which looks essentially like a sphere, whereas when $l = 1$ the orbitals are called p-orbitals, which have a more interesting shape (shown in Figure 4.10), and when $n = 2$ and $l = 2$ they are known as d-orbitals (Figure 4.11). The surface shown in each figure is the angular probability function for each orbital enclosing an area where there is a 90 per cent chance of finding the electron.

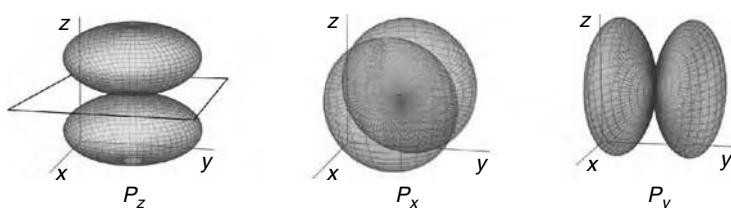


Figure 4.10 The p-orbital shapes

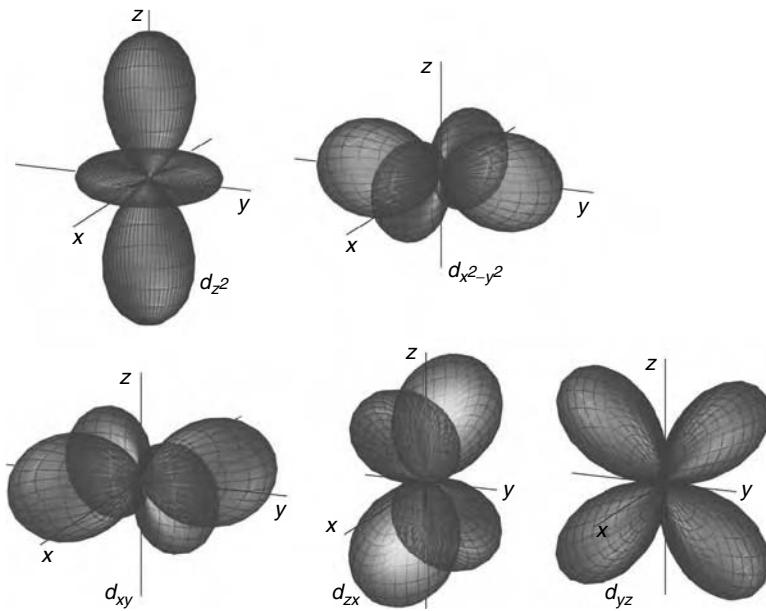


Figure 4.11 The d-orbital shapes

Notice that there are three p-orbitals and five d-orbitals all of the same energy with no external magnetic field: they are said to be *degenerate*. However, in the presence of the magnetic field, the value of the magnetic quantum number, m , can be seen due to the lifting of the degeneracy; each orbital now has an energy that depends on m . The magnetic quantum number can take values $-l \leq m \leq +l$ and there are $2l + 1$ such values: $2l + 1$ is said to be the degeneracy, as discussed in Chapter 3.

A small energy difference now arises between the normally degenerate levels of the atom in the presence of the external magnetic field, which gives an energy that depends on the magnetic quantum number M_J with energies given by $M_J g \mu_B B$, where B is the external magnetic field strength (in tesla), μ_B is the Bohr magneton ($\mu_B = 9.274 \times 10^{-24} \text{ J T}^{-1}$) and g is the Landé factor. The magnetic quantum number M_J is a measure of the total spin and angular momentum within the atom, and the strength of the interaction is determined by the Landé factor. Small but observable splittings appear in the spectra of atoms in stars, indicating the presence of a local magnetic field. Assuming that the effect of the spin is very small, the M quantum number takes the same values as the m_l quantum number for the p-orbitals and d-orbitals. The levels in the $n = 2$ state are changed in energy, with each m state taking a different energy revealing the three different p-orbitals, $m_l = 1, 0, -1$. In the $n = 3$ level the m quantum number is split into five levels associated with the d-orbitals, $m_l = 2, 1, 0, -1, -2$. The allowed transitions between the levels follow the selection rules $\Delta M = 0, \pm 1$, with the transition shown by the arrows in Figure 4.12. The result is a spectrum that is split into three components, with the

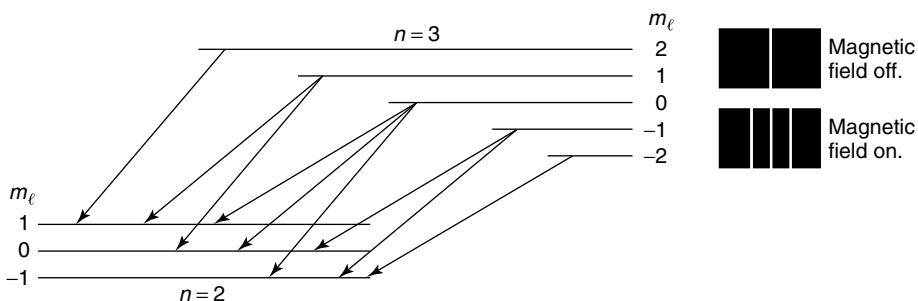


Figure 4.12 The origin of Zeeman splitting in the $n = 2-n = 3$ transition in the H atomic spectrum

Δm_l transitions at the zero-field wavelength and the $\Delta m_l = \pm 1$ transitions shifted to the $+\mu_B B g$ and $-\mu_B B g$. The separation between the two band extremes is $2\mu_B B g$, from which the magnetic field can be derived assuming that the g factor for the transition is known. This is a very simplified approach to the Zeeman effect in the hydrogen atom but it reveals two important things: (1) in a magnetic field the p- and d-orbitals are not degenerate and the need for all of the quantum numbers in the atom is immediately apparent; and (2) stars have magnetic fields.

In the visible region of the spectrum the shift is of order 0.01 nm T^{-1} and the Earth's magnetic field is 10^{-4} T , so the splitting is not seen in the laboratory without enhancing the magnetic field. The magnetic field on a fridge magnet is around 10^{-2} T , which is some hundred times stronger than the Earth's magnetic field. In some stars, however, such as white dwarfs or neutron stars, the magnetic field strength has been determined by the splittings in atomic spectra to be of order 10^3 T . The question is why do stars have magnetic fields and why are they so large? Not for the first time has the effect on observable properties of atoms been used to determine the local environmental conditions. The magnetic fields of stars vary extensively during its evolution so that the surface magnetic field of the Sun, a middle-aged main-sequence star, is $5 \times 10^{-4} \text{ T}$, whereas a pre-nova star could be 10^{-2} T . White dwarf stars may have magnetic fields as large as 10^2 T , with special classes of stars called magnetic stars, such as BD + 54 2846, having magnetic fields of 0.12 T. Neutron stars hold the record magnetic field with a surface magnetic field of 10^8 T and the spectra of atoms around this star would appear very different from the laboratory spectrum. All of these magnetic fields should be compared with that of an NMR spectrometer in the laboratory, which commonly has a magnetic field of 9.5 T – about the surface magnetic field of a white dwarf!

4.5 Exotic stars

During the discussion of the HR diagram in Section 4.2 it was noted that 92 per cent of all observed stars fall on the main sequence; 7.5 per cent are white dwarf

stars with high temperatures and low luminosities, whereas the remainder are giants or supergiants. A glance through the research literature, however, shows that more work is published on the minority 8 per cent of stars than the majority. We shall have a look at some of these stars to close this chapter.

Binary stars

Including binary stars in a collection of exotic stars is completely unjustified because nearly half of the stars in the Universe are binary star systems, which means that two or more stars orbit one another around a common centre or centre of mass. Some binary stars are so close to one another that they touch but other binary systems are separated by huge distances, taking hundreds of thousands of years to orbit each other once. The orbital period can be measured accurately by observing the spectra of the stars and combining these observations with the Doppler effect. Stars identified in this way are called spectroscopic binaries. Assuming that we view the binary pair of stars edge-on, then one of the pair will be receding while the other is approaching the line of sight. Spectra from the receding star are shifted, with absorption frequencies from the atomic spectra appearing at a longer wavelength than you would expect from laboratory observations – red-shifted. Yet on the approaching star the atomic transitions are shifted to shorter wavelengths than the laboratory measurements – blue-shifted.

The visible binary Doppler observation can be seen in Figure 4.13, illustrated for the sodium D lines at 586 nm in the yellow. When the two stars are travelling parallel to the line of site from the Earth there is no Doppler shift in the spectrum of the sodium lines. However, when the binary stars are moving away the sodium D lines are red-shifted on the right-hand limb and blue-shifted on the left-hand limb. This, of course, only works if the binary pair are end-on to the line of sight, otherwise there is no relative motion and no shift. When the two stars are again moving perpendicular to the line of sight there is only one line in the spectrum and the period of motion has been completed, once round. The orbital period of the binary can then be measured from the period of the spectral shifts.

Most spectroscopic binaries have periods ranging from days to months and are separated by distances of order 1 AU. A consequence of knowing the period of the star and separation, measured optically, is the determination of the mass of the stars. Assuming that the two stars are in circular orbit, for the sake of simplicity, then a centripetal force is required to keep them moving in orbit. Gravity provides this attraction and the two forces must be balanced. The complete solution of this problem is hard and only a combined mass can be derived without knowing some information other than the period of rotation.

All is not what it seems, however, because most binary stars have a very faint neighbour and so both the red- and blue-shifted lines are not observed and the spectroscopic motion of just the bright star is enough to make the period measurement. The faint binary partner will pass in front of the bright star during the period of rotation if the plane of the binary orbit is along the line of sight from the Earth.

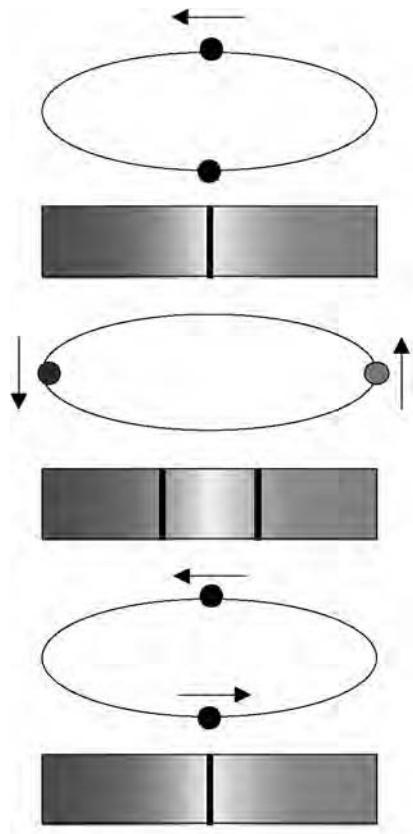


Figure 4.13 Doppler effect for a binary star pair

The result is an eclipsing phase when the dull star is in front of the bright star and an apparent variation in the intensity of the bright star.

In theory, planets can form around a binary pair and this would produce some interesting variations in the configuration of planets in a solar system. Planets might, for example, orbit both stars or orbit independently around each of the stars. Planets would be subject to the radiation from both stars, perhaps providing high levels of photon flux at damaging wavelengths. The atmosphere on Earth provides the protection from UV radiation, principally in the ozone layer. The presence of two stars would require thicker atmospheres to screen out the increased photon flux. A thick atmosphere and two local stars could make for some very interesting and beautiful sunsets.

Cepheid variable stars

Variable stars are so called because the luminosity of the star varies intrinsically and is not due to the passage of an eclipsing star. One such class of stars is the

Cepheid variables, named after the first observations of a variable star by the naked eye, the fourth brightest in the constellation Cepheus, δ -cephei, made in the early 1700s. In 1912 the astronomer Henrietta Leavitt made a series of observations of Cepheid stars in the Large and Small Magellanic Clouds. Although this was not known at the time, the two clouds are small satellite galaxies to our Milky Way and the stars within them are all at similar distances from the Earth. Empirically the observations showed that there was a relation between the period of light variation and the brightness or absolute magnitude of the star.

The emission of light from Cepheid stars has a characteristic light curve seen in Figure 4.14 for a Cepheid in the constellation of Perseus. The sawtooth pattern is characteristic of the class and enables the period of variation to be determined. The observation, however, that the luminosity and period are related has powerful consequences. The Cepheid variables fall into two classes: type I classical Cepheids have periods of 5–10 days and type II have periods of 12–20 days. The two types of Cepheids initially caused problems when determining the luminosity–period relation but the relation has now been determined. Type I Cepheids follow the expression

$$M = -1.9 - 2.8 \log_{10} P \quad (4.8)$$

where M is the absolute magnitude and P is the period in days. Similarly, type II Cepheids have the relation:

$$M = -2.8 \log_{10} P \quad (4.9)$$

These two types of stars form two distinct populations and have the very important property than we can use them as distance markers. By measuring the period of pulsation of the star and the magnitude, the apparent magnitude and the absolute

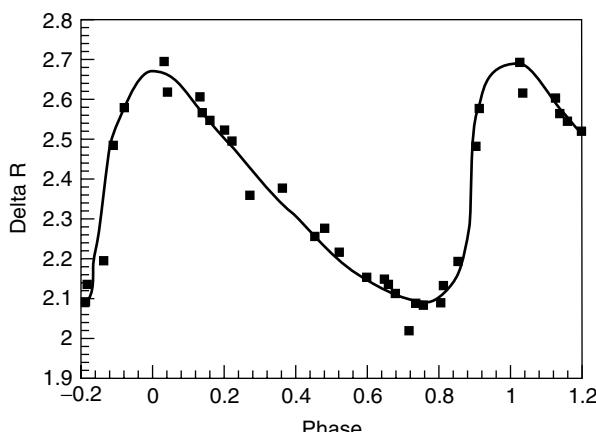


Figure 4.14 Light curve for the star SX Persei. (Reproduced by permission of David L. Dupuy, Virginia Military Institute)

magnitude can be compared and hence the distance to the star can be determined using the least-squares law. This powerful relationship has provided astronomers with an indirect way of measuring intergalactic distances beyond the limitations of parallax. Distances as far as 60 Mly have been measured and have enabled the position of the Sun in the Milky Way to be determined, not to mention distances to local galaxies, showing us that our galaxy is part of a local group of 30 or so galaxies including Large and Small Magellanic Clouds and the famous Andromeda Galaxy.

The cause of the variation in the Cepheid variables is interesting. The spectra show that the surface temperature of the outer layers of the star varies during the cycle and Doppler shifts in the spectra also indicate that the variation is associated with a change in the radius of the star, showing that the star swells and shrinks during the cycle. For the prototype Cepheid star, δ -cephei, the radial variation has been calculated to be 3 million km, which is a variation of 8 per cent in the diameter of the star. The maximum rate of expansion corresponds to the maximum luminosity and the maximum size occurs at the mid-point of the decline in brightness. The reason for this is associated with the formation of ionised H and He, which seem to act as a safety valve for the star. The ionisation of H and He when the star is at its smallest causes the radiation to be trapped in the star's outer layers and the star expands. As the expansion continues the He and H recombine to form neutral atoms again, allowing the radiation to escape and thereby cooling the outer layers.

Black holes

The transition from theory to observation is often a long and tortuous path and there is none more so than the conclusive detection of a black hole. The concept of a black hole is very quickly derived from an intuitively obvious argument founded in classical gravity, although the same result is derived from general relativity, assuming that the black hole is not charged or rotating – two properties you would not have guessed to be important for a black hole!

Assuming Newton's law of gravity given by the famous inverse squared relationship:

$$F = \frac{GM_sm}{r^2} \quad (4.10)$$

where G is the gravitational constant, M_s is the mass of a star and m is the mass of an object at distance r from the star moving away against the force of gravity, and assuming that the mass moves slowly (infinitesimally slowly), the amount of work done against gravity is $F\Delta r$ so that moving from r to an infinite distance from the star and thus escaping its gravitational field is given by:

$$\begin{aligned} W &= - \int_r^\infty \frac{GM_sm}{r^2} dr \\ &= \frac{GM_sm}{r} \end{aligned} \quad (4.11)$$

There is a sign convention here: the potential energy of the gravitational field is chosen to be zero at infinity, so the sign of the work is chosen to be negative. Letting r equal the radius of the star R , the energy required to escape from the star is equal to the kinetic energy of the mass m from which we can derive the escape velocity v_e :

$$\frac{1}{2}mv_e^2 = \frac{GM_sm}{R}$$

$$v_e = \left(\frac{2GM_s}{R}\right)^{\frac{1}{2}} \quad (4.12)$$

The mass of the object appears on both sides of the equation and thus cancels, hence the escape velocity is an intrinsic property of the star or planet from which you are trying to escape. So the concept of a black hole requires that the escape velocity should be set at the speed of light, the fastest possible velocity and a point at which not even light can escape, hence it must be black. Setting $v_e = c$ we can rearrange Equation 4.12 to calculate the radius at which a star or body of known mass will become a black hole:

$$R_S = \left(\frac{2G}{c^2}\right) M_s \quad (4.13)$$

where R_S is called the Schwarzschild radius, which for M_s in kg and R_S in km gives $R_S = 1.484 \times 10^{-27} M_s$. So taking the mass of the Sun as 1.99×10^{30} kg gives a Schwarzschild radius of 2.9 km for the Sun and 8.8 m for the Earth.

Example 4.2

What is the mass of a black hole if $R_S = 5.29 \times 10^{-11}$ m? Rearranging Equation 4.13 and substituting the value of R_S gives:

$$M = \left(\frac{c^2}{2G}\right) R_S$$

$$M = \left(\frac{(3 \times 10^8)^2}{2 \times 6.670 \times 10^{-11}}\right)$$

$$M = 3.6 \times 10^{16} \text{ kg} \quad (4.14)$$

The reason for calculating the mass required for such a small radius is that the radius is the Bohr radius of the hydrogen atom, so a black hole the size of an atom would be very heavy and somewhat elusive. Quickly many questions spring to mind about atomic sized black holes.

The point-of-no-return for a black hole is the surface at which a photon is trapped and from which it can escape. This is called the event horizon, which for the special case of a black hole, often called the Schwarzschild black hole, is non-charged, not rotating and spherical at a distance equal to its Schwarzschild radius. Rotation and charge cause a change in the shape of the event horizon but, aside from its mass, charge and angular momentum, no other physical property of matter falling into a black hole can be detected. Inside the event horizon, matter may form an infinitely dense centre, a singularity, warping space-time itself in particular leading to quantum gravity and perhaps links to other universes or parts of our Universe. Such is the speculation of science fiction.

Research into the properties of black holes is a very active area, including the idea that they may be detected. Detection, however, is far from certain at the time of writing. There are a number of suggestive observations showing mass flows from local regions of space to others, such as the core of the galaxy NGC 4261 seen by Hubble in 1992. However, one possible detection method has come from the ideas of Penrose and Hawking, leading to the idea of Hawking radiation: a fascinating idea to explore elsewhere.

4.6 Cycle of star formation

The description of star formation, evolution and then ultimately explosion in a supernova defines a cycle of star formation, as shown in Figure 4.15. The initial rather tenuous density of the interstellar medium forms into a cloud and immediately following the Big Bang this was a cloud made solely of H, He and Li. Collapse of this cloud leads through all of the phases of star formation and descent onto the main sequence, during which time the star begins to shine and radiate heat. The collapse of the cloud of material forms the stellar nebula around the young star, eventually accreting into a solar system of planets. The radiation flux from the star falls on the planets, heating and perhaps even supporting life. The star in the middle of the planetary system slowly burns up all of the nuclear fuel in its core and, depending on its mass, either burns out or undergoes a huge supernova explosion. The supernova sends material from the star and the surrounding planets back out into the interstellar medium, closing the cycle of star formation.

The supernova has performed a number of jobs: introducing heavier nuclei into the interstellar medium, such as oxygen and silica, and forming heavier elements such as uranium in the intense pressure of the explosion. Now the interstellar medium is enriched with different atoms ready to undergo chemistry, whether in the gas phase or on the surface of dust grains. Clouds of material now form under mutual gravitational attraction but can stay in a stable state for millions of years. The supernova also produces a shock wave in space, not just of high-speed molecular debris from the planetary system but perhaps ripples in space-time itself. The shock wave causes local variations in pressure and can trigger the collapse of a molecular cloud to form a star.

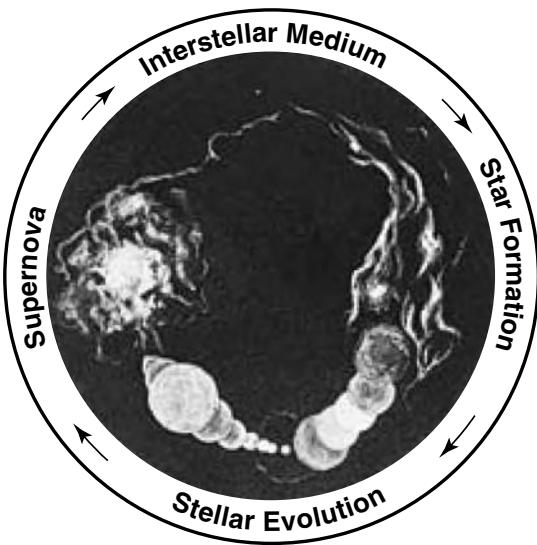


Figure 4.15 The cycle of star formation. Best resolution available. (Reproduced by courtesy of NASA)

The interstellar medium is thus a chemically diverse medium fed nearly all of the chemical elements by supernova explosions. Conditions in the interstellar medium produce a cocktail of molecules that ultimately find themselves back on the surface of planets during the formation of the new star and solar system. Does the interstellar medium seed life with molecules from space? The nature of interstellar medium chemistry might then add credibility to the formation of life in many places within the Universe and act as a panspermia model for the origins of life.

Concepts and calculations

Concepts

<i>Herzprung–Russell diagram</i>	The plot of luminosity versus temperature for all stars, resulting in the main sequence, red giants and white dwarfs. Stellar evolution leads to mass – dependent birth lines onto the main sequence
<i>Proton–proton cycle</i>	Nuclear fusion process leading to production of He
<i>Triple-alpha process</i>	Fusion of ^3He nuclei to form C nuclei
<i>CNO cycle</i>	Odd nuclei production forming ^{13}C , ^{13}N , ^{15}O and ^{15}N
<i>Supernova synthesis</i>	Production of nuclei above Fe
<i>Stellar evolution</i>	The birth of a protostar and its life as a pre-main-sequence star, its descent to the main sequence and death, starting with a red giant leading to planetary nebula and ending in white and black dwarfs. This sequence varies with mass
<i>Balmer temperature</i>	Observation of the Balmer series in the H atom implies population of the $n = 2$ level of the H atom and the temperature of the local environment
<i>Zeeman effect</i>	Observation of a splitting in the spectrum of an atom associated with the magnetic field of a star
<i>Black hole</i>	A gravitationally massive object (allegedly) from which light cannot escape. The Schwarzschild radius is the radius of the black hole that defines the event horizon: the distance at which light cannot escape
<i>Cycle of star formation</i>	The evolution of a star from collapse of an interstellar cloud, through formation of the star to the supernova and the ejection of matter back into the interstellar medium

Calculations

<i>Mass defect</i>	Energy associated with a mass defect in the fusion of low-mass nuclei and the fission of high-mass nuclei
<i>Balmer temperature</i>	Determination of the temperature of the star from the intensity of the Balmer series and the population of $n = 2$ in the hydrogen atom
<i>Zeeman splitting</i>	Splittings in the spectrum due to a magnetic field
<i>Cepheid variable</i>	Luminosity–period relationships
<i>Schwarzschild radius</i>	The event horizon of a black hole

Problems

- 4.1** During the temperature rise in protostar formation, molecular dissociation occurs followed by ionisation of atoms. Assume that molecules and atoms are in thermal equilibrium with the protostar:
- Calculate the dissociation temperature of CO (see Table 5.3).
 - Calculate the ionisation temperature of the hydrogen atom in its ground electronic state.
- 4.2** Water molecules have been observed in Sun spots on the surface of the Sun.
- Calculate the maximum possible temperature at which the water molecule can survive (see Table 5.3).
 - The photosphere temperature of the Sun is 5780 K. Can water molecules survive at this temperature?
 - The chromosphere of the Sun extends 2000 km from the surface with a temperature range of 4000–30 000 K. Will molecules survive in the chromosphere?
 - After a transition region the outer corona of the Sun has temperatures of 5×10^6 K. Will water molecules survive here?
 - Why are molecules not more prevalent in the spectrum of stars?
 - What will be the atomic character of the corona?
- 4.3** Sketch the HR diagram and identify the four major classes of stars. A star with radius 7×10^9 m has a surface temperature of 10 000 K. Plot its position on the HR diagram and indicate a birth line for this star.
- 4.4** A white dwarf star has a radius of 100 km:
- Estimate the luminosity of the star.
 - Assume that it has the same mass as the Sun and estimate its density.
- 4.5** The relative populations of levels in the H atom provide a measure of the temperature of the star. Calculate the relative populations in the first two levels of the atom that control the intensity of the Balmer series for Rigel with a surface temperature of 11 000 K.
- 4.6** The binary star system Sirius is separated by 30 AU. If the linear velocity at the extremes of motion is given by $v = r\omega$, where v is the velocity (m s^{-1}), r (m) is the radius of the circular orbital motion and ω is the orbital period (s):
- Calculate the Doppler shift on the $\text{H}\alpha$ transition if the orbital period is 10 years.
 - Is this shift detectable?
- 4.7** A type II Cepheid variable is observed in the elliptical galaxy M49 in the Virgo cluster of galaxies with a period of 30 days. Calculate the absolute magnitude of this star. If the total flux on the Earth from the star is $4.2 \times 10^{-22} \text{ W m}^{-2}$, estimate the distance to the Virgo cluster.
- 4.8** Calculate the Schwarzschild radius for the following masses:
- 1 kg;
 - a person;
 - a car.
- 4.9** Calculate the mass of a black hole if the radius is 1 fm – a subatomic black hole.

5

The interstellar medium

Introduction

Looking at the night sky it would appear that there is nothing between the stars but empty space and to a first approximation this has to be right. If you were to find an atom, then your best guess would be that it is 10 000 times more likely to be hydrogen than any other element. But simply looking into the sky also reveals stellar clusters, galaxies, our own Sun and its collection of planets that form the solar system, so the idea of empty space has to be only at one end of the spectrum. The density of matter in the Universe stretches from tenuous emptiness to massively dense stars and beyond to black holes. The interstellar medium (ISM) must have a complicated structure: if we define the edge of a star to be the top of its photosphere then the ISM is everywhere else.

From our discussion on the Big Bang it would appear that only H, He and some Li were formed in the beginning, from which first-generation stars formed. The evolution of stars results in stellar winds blowing off material into the ISM, now including heavy elements such as carbon. Violent supernovae add the higher mass elements, all being returned into the ISM. The cycle of star formation in the ISM repeats, allowing molecules to collect together and form clouds, from which stars reform, and new material, including new planetary systems and planets that may support life. The ISM cycle of molecular chemistry may be more important to the origins of life than the diffuse emptiness of space may at first suggest: the quiet processing of molecules in giant molecular clouds may be required to seed life.

Observations of molecules within the ISM started in 1904 with the detection in the visible spectrum of the singly ionised Ca⁺ towards the binary star system δ-Orionis. Later observations, again in the visible spectrum, resulted in the detection of CH⁺ and CN, the first molecules. Hydrogen atoms were detected using the 21-cm line in the radio spectrum in 1951 and the OH maser transitions at 1665 and 1667 GHz in 1963. It was, however, the observation of the radio spectrum at 115 GHz of the R(0) transition in ¹²CO emission that has changed the way we look at the structure of the Universe. Tuning the telescope to look at one transition frequency and scanning throughout the sky enables maps of CO number density

to be measured along all lines of site. The CO spectral mapping reveals structure in a way not seen with the naked eye, including regions of nearly no emission that are essentially diffuse interstellar space of very low density, regions where clouds of gas are forming and hugely dense giant molecular clouds. At the heart of the giant molecular clouds star formation begins to occur and remnants of supernovae or stellar winds: all are seen in the CO maps of the night sky, even the spiral structure of our own galaxy, the Milky Way.

To give the impression that all of the matter in the ISM is in the form of atoms or molecules would be wrong: one per cent of the matter comes in the form of dust grains typically $0.1\text{--}1\ \mu\text{m}$ in diameter and is responsible for a large part of the extinction or scattering of radiation. The chemistry occurring on the surface of dust grains is complicated, especially when covered with a mantle of water ice. Molecules can form on dust grains and desorb back into the ISM, adding surface chemistry to the rich diversity of physical and chemical conditions in the ISM.

5.1 Mapping clouds of molecules

The simplest molecules to map in the sky from Earth are the polar molecules, as they all possess dipole moments and hence rotational spectra. The obvious place to start is CO and the rotational emission from $J = 1\text{--}0$ at 115 GHz. Sky surveys at this frequency have revealed diffuse regions of the ISM and some very dense CO regions called giant molecular clouds (GMCs). The Orion nebula familiar in the northern hemisphere is only a small blister on the edge of two huge molecular clouds. The cloud has been shown on the constellation map in Figure 5.1, with the Orion nebula just below the belt of three stars. The GMC, however, extends over most of the constellation, occasionally breaking into flashes of nebula light such as the Horsehead and Orion nebulæ.

Looking at the CO frequency towards the Taurus molecular cloud (TMC), wonderful maps of CO turbulent density are revealed (Figure 5.2). The TMC is dense, perhaps collapsing to form stars, and filamentous, extending into the tenuous or diffuse ISM. The ^{12}CO R(0) emission for the TMC is shown in Figure 5.2 and in the first of many surprises in the chemistry of the ISM the emission for ^{13}CO $J = 1\text{--}0$ is shown in Figure 5.3 measured at 120 GHz – they are different.

The TMC CO isotope density is different in intensity, with regions of the TMC that are denser with ^{13}CO than with ^{12}CO , and different in mixing and degrees of turbulence. The problem with the chemistry of the ISM clearly involves mixing and turbulence, and shock waves and other processes in general that are physical transport processes – transporting matter without chemical change.

The TMC-1 is a cold ‘quiescent’ dark cloud some 45 ly away with a temperature of around 10 K. It is predicted to be only 100 000 years old and, up to 1998, 49 molecules have been detected within it.

The general extent of the molecular cloud can be mapped within the sky but the physical conditions and stellar activity lead to different chemical regimes, all of which must be considered if the chemistry of the ISM is to be understood. A

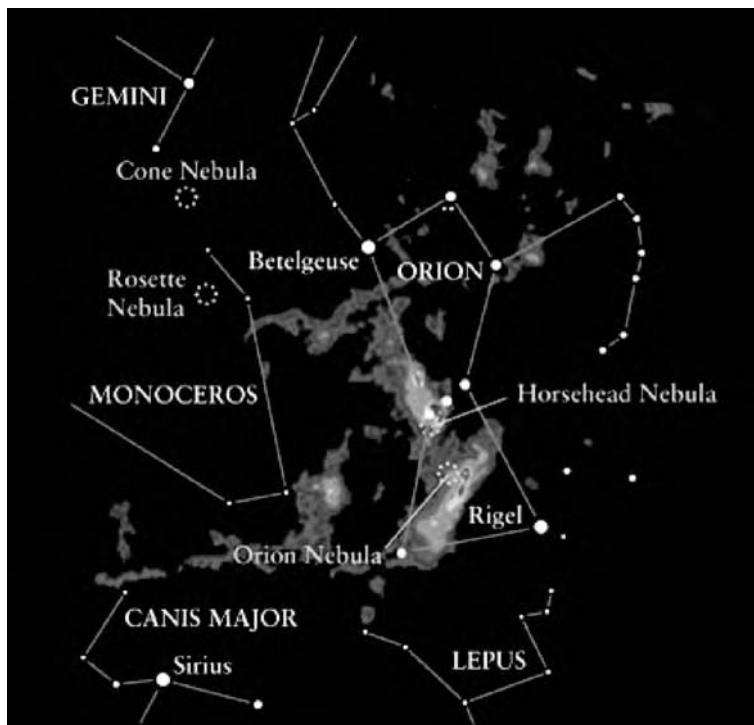


Figure 5.1 The Orion molecular cloud. (Reproduced by courtesy of Ronn Magdalena, NRAO)

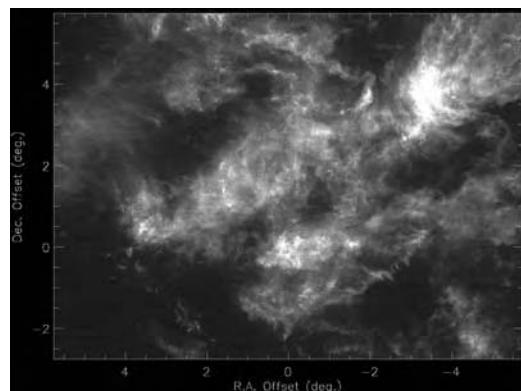


Figure 5.2 Taurus molecular cloud observed at 115 GHz ^{12}CO emission. (Reproduced by courtesy of Paul Goldsmith, JPL NASA)

schematic GMC is featured in Figure 5.4 and shows the transition from the ISM, where the density of molecules is very low and the temperature is 25 K, to the hot regions around a young stellar object (YSO). Hydrogen in the outer regions is in the HI state: simple neutral atoms. Moving towards the centre of a cloud,

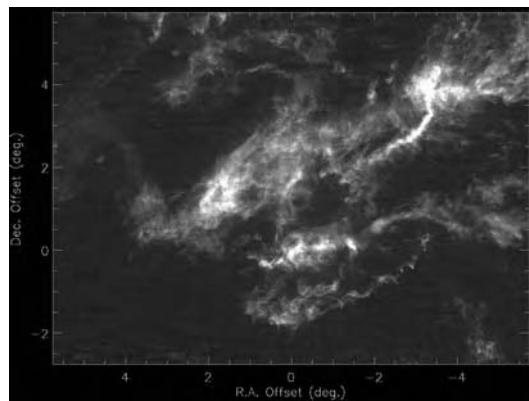


Figure 5.3 Taurus molecular cloud observed at 120 GHz ^{13}CO emission. (Reproduced by courtesy of Paul Goldsmith, JPL NASA)

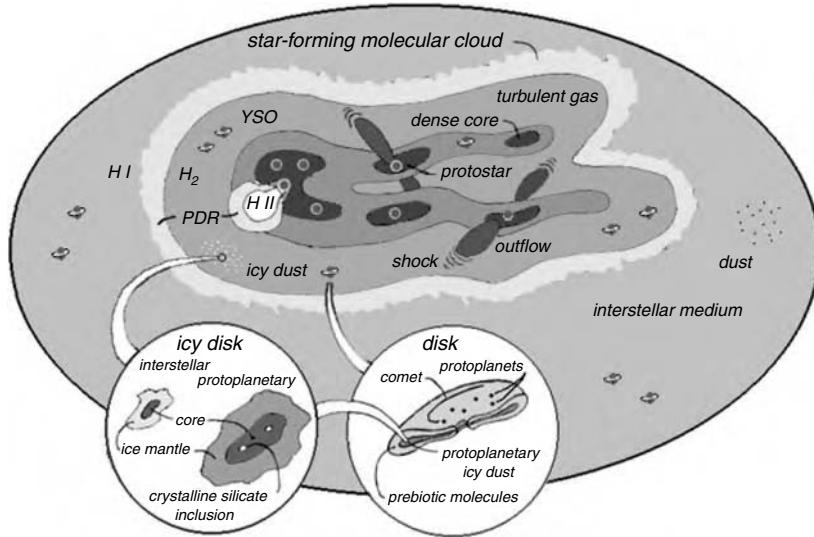


Figure 5.4 Structure of a giant molecular cloud

the molecular density and temperature rise and dust particle numbers also rise; some begin to form ice-layer surfaces that are bathed in radiation from the stars, processing the molecules on the icy surface. Hydrogen atoms can now combine on the surface of dust particles, producing H_2 at a much faster rate than the gas-phase recombination reaction. Young stellar objects pass through the initial stages of proto-star formation and begin to eject material into the GMC, having reached the T-Tauri phase of evolution.

Centred on hot young stars are the HII regions where the intense radiation field from the stars is sufficient to remove the electrons from H atoms and a new emission process is seen. An electron passing near to a proton momentarily looks

like a hydrogen atom but the electron is not captured: it slows down, with the emission of a photon. The HII emission is seen in partial electron-capture collisions. Photon-dominated regions (PDR) are hostile places in which to try to do chemistry. The physical conditions and hence the chemistry must be understood at all regions of the clouds and unfortunately all are seen along one line of sight through the cloud. It is hard to work out which chemistry is associated with which region from a simple line-of-sight observation, and different environments can be determined by the spectra of molecular probes in each region. But most of all it is in this giant cauldron of chemistry that prebiotic molecules may be forming.

5.2 Molecules in the interstellar and circumstellar medium

The first molecules detected in space were CH, CH⁺ and CN in the 1950s and since then the observation of over 100 molecules and molecular ions has been reported (Table 5.1), classified by number of atoms. Of all of the molecules discovered in the ISM only 12 elements are represented (H, C, N, O, S, Si, P, F, Cl, Al, Na, Mg), with H, C, N and O (sometimes referred to as ‘CHON’) being the most abundant. The usual rules of laboratory stability seem to be irrelevant: carbon for example does not need to have the requisite four chemical bonds because it is not

Table 5.1 Interstellar and circumstellar molecules* (Erhenfreund (2000))

		Number of Atoms									
2	3	4	5	6	7	8	9	10	11	13	
H ₂	C ₃	c-C ₃ H	C ₅	C ₅ H	C ₆ H	CH ₃ C ₃ N	CH ₃ C ₄ H	CH ₃ C ₅ N?	HC ₉ N	HC ₁₁ N	
AlF	C ₂ H	I-C ₃ H	C ₄ H	I-H ₂ C ₄	CH ₂ CHCN	HCOOCH ₃	CH ₃ CH ₂ CN	(CH ₃) ₂ CO			
AlCl	C ₂ O	C ₃ N	C ₄ Si	C ₂ H ₄	CH ₃ C ₂ H	CH ₃ COOH?	(CH ₃) ₂ O	NH ₂ CH ₂ COOH?			
C ₂	C ₂ S	C ₃ O	I-C ₃ H ₂	CH ₃ CN	HC ₃ N	C ₇ H	CH ₃ CH ₂ OH				
CH	CH ₂	C ₃ S	c-C ₃ H ₂	CH ₃ NC	HCOCH ₃	H ₂ C ₆	HC ₇ N				
CH ⁺	HCN	C ₂ H ₂	CH ₂ CN	CH ₃ OH	NH ₂ CH ₃		C ₈ H				
CN	HCO	CH ₂ D ⁺ ?	CH ₄	CH ₃ SH	c-C ₂ H ₄ O						
CO	HCO ⁺	HCCN	HC ₃ N	HC ₃ NH ⁺							
CO ⁺	HCS ⁺	HCNH ⁺	HC ₂ NC	HC ₂ CHO							
CP	HOC ⁺	HNCO	HCOOH	NH ₂ CHO							
CSi	H ₂ O	HNCS	H ₂ CHN	C ₅ N							
HCl	H ₂ S	HOCO ⁺	H ₂ C ₂ O								
KCl	HNC	H ₂ CO	H ₂ NCN								
NH	HNO	H ₂ CN	HNC ₃								
NO	MgCN	H ₂ CS	SiH ₄								
NS	MgNC	H ₃ O ⁺	H ₂ COH ⁺								
NaCl	N ₂ H ⁺	NH ₃									
OH	N ₂ O	SiC ₃									
PN	NaCN										
SO	OCS										
SO ⁺	SO ₂										
SiN	c-SiC ₂										
SiO	CO ₂										
SiS	NH ₂										
CS	H ₃ ⁺										
HF											

*A database of interstellar and circumstellar molecules is maintained at <http://www.cv.nrao.edu/~awootton/allmols.html> by A.L. Wootton with the current total at 129 (at time of going to press).

in a reactive chemical environment. There are many subvalent species, radicals and energetic isomers among the neutral species and then there are the molecular ions.

There is a high degree of unsaturation among the carbon-containing compounds, with many double and triple bonds, but few branched chains. There are few cyclic compounds but there are some three-membered rings. The species are not strongly oxidised and few have more than one non-carbon heavy atom. An eclectic mix of molecules with a chemistry only recently studied in the laboratory in the gas phase they certainly would not survive in solution. Does this molecular motley crew form the basis for building the molecules of life? Are the seeds of life sown in the vast, cold, giant molecular clouds in the ISM?

The chemistry of all of these molecules is fascinating but, concentrating on the origins of life, a detailed look at the organic species is appropriate to see what molecules are present and how they might have been formed. The only alkane detected directly in the ISM is methane but this is due to the problem of detection. All alkanes are non-polar and so do not have a pure rotation spectrum. However, there is one allowed vibration of methane that is infrared active and with the low moment of inertia of methane the vibration–rotation spectrum can be observed and a rotational progression identifies the molecule with confidence.

Of the alkenes (Figure 5.5) only ethene has been detected and of the aromatics only benzene has been seen unambiguously; surprisingly propene has not been seen despite its well-understood microwave spectrum. Of interest to the origins of life is the onset of polymerisation in HCN to produce cyanopolyyynes. These molecules could provide a backbone for the formation of information-propagating molecules required for self-replication. The survival of these species in a planetary atmosphere depends on the planet: oxidation would be rapid in the atmosphere of today's Earth but what of the early Earth or somewhere altogether more alkane-based such as Titan?

Oxidation of the alkanes introduces polarity into the molecule, making familiar molecules such as the alcohols, aldehydes and ketones (Figure 5.6). These molecules are the necessary building blocks for the majority of biological molecules such as sugars and acids. However, the presence of carboxylic acids and amino acids in the ISM is now somewhat contentious due to the uncertainty of the microwave observations of these species. The complexity of the microwave spectrum of glycine in Figure 3.10 is well known and reports of the detection of interstellar glycine have been made: there is also some consideration of the formation and observation of carboxylic acids but we shall return to this later.

There are some variations in the composition of chemical clouds as they evolve in time from dark clouds such as TMC-1 to giant molecular clouds such as Orion where the presence of light from young stars initiates photochemistry. The Orion molecular cloud chemical inventory contains several saturated species such as ethanol ($\text{CH}_3\text{CH}_2\text{OH}$) and its CN analogue ($\text{CH}_3\text{CH}_2\text{CN}$), the simplest carboxylic acid (CH_3COOH , acetic acid) and methylamine (CH_3NH_2).

It is highly unlikely that all organic molecules have gas-phase formation routes and many may be formed on the surface of dust grains, probably with ice mantles,

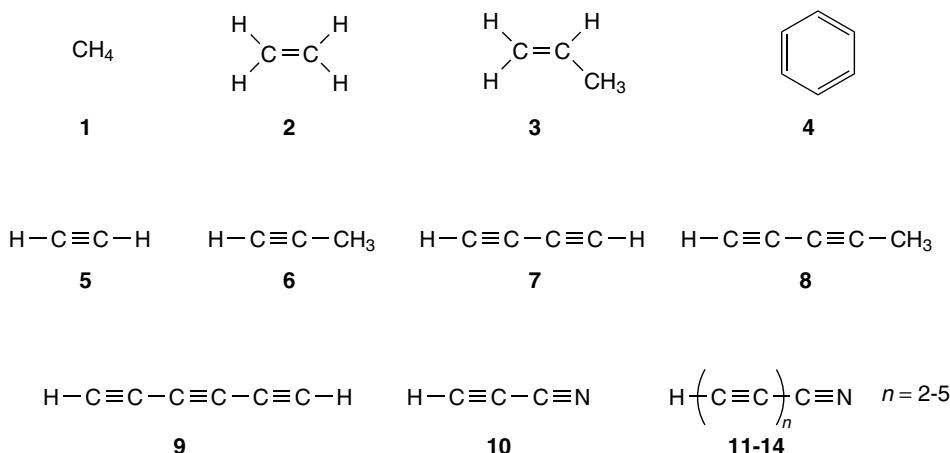


Figure 5.5 Alkanes, alkenes, aromatics and cyanopolynes: 1, methane; 2, ethene; 3, propene; 4, benzene; 5, acetylene; 6, propyne; 7, butadiyne; 8, pentadiyne; 9, hexatriyne; 10–14, cyanopolynes (the largest family with the longest molecule to date, 14). (Reproduced from Guillemin *et al.* 2004 by permission of Elsevier)

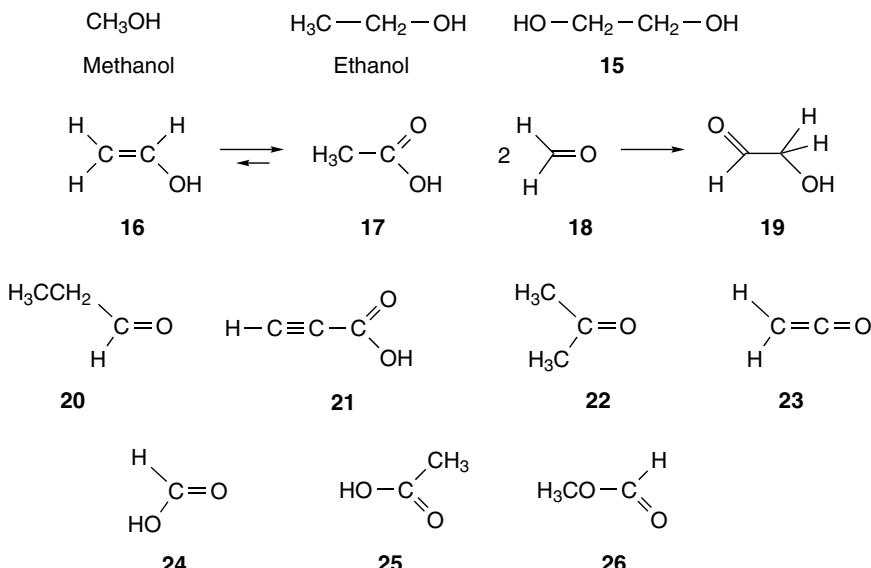


Figure 5.6 Alcohols, aldehydes, ketones and acids: 15, ethylene glycol; 16, vinyl alcohol; 17, acetaldehyde; 18, formaldehyde; 19, glyoxal; 20, propionaldehyde; 21, propionaldehyde; 22, acetone; 23, ketene; 24, formic acid; 25, acetic acid; 26, methyl formate. (Reproduced from Guillemin *et al.* 2004 by permission of Elsevier)

although the more exotic molecule may require photochemistry or photoprocessing to be synthesised. These ideas require an understanding of the chemistry of the ISM, how gas-phase reactions can result in organic synthesis and which molecules will survive and which will react. The development of ISM chemistry and its understanding follows a path:

- Detection of molecules within the cloud.
- Determination of molecular abundance within the cloud.
- Physical conditions within the cloud.
- Optical extinction and the role of dust particles.
- Constructing a chemical network of all possible chemical synthesis pathways.
- Solving the chemical kinetic equations and comparison with the observed molecular abundance.

This pathway sets the direction for the remainder of this chapter but has the same basic strategy for the understanding of meteorite chemistry, comet chemistry, planetary atmospheric chemistry, prebiotic chemistry and ultimately the chemistry of an organism. This pathway is the ‘molecule-up’ view of the Universe.

5.3 Physical conditions in the interstellar medium

The parameters that determine the chemistry in any region of the ISM are the average density of molecules and the temperature. These two parameters determine how often molecules collide with one another and with what energy. Whether a chemical reaction occurs depends on what has to happen to the reactants to become products. If a bond has to be broken, then there must be sufficient energy in the collision; too much energy, however, results in the bonds of the target reactant being broken. The available energy at the disposal of a chemical reaction is then, in a general sense, measured by the local temperature. In the discussion of the rotational and vibrational temperatures of molecules, the rate of re-partition of energy between molecules depends on the relative rates of collisions and the lifetimes of the excited states to establish *local thermal equilibrium*. In the ISM we have to make similar judgements.

Environments of the interstellar medium

Most modelling of ISM chemistry has concentrated on four broadly different environments for which we have good estimates of the local conditions:

- Diffuse interstellar medium: the atomic density can be of order $1\text{--}10^2$ atoms cm^{-3} and temperature is hard to define. The translational temperatures of

atoms and molecules may be as high as 100 K, especially in shock fronts, but with so few atoms is temperature a sensible concept?

- Giant molecular clouds: the GMCs have a lifetime of order 10^6 – 10^8 years and are the regions of new star formation. The Orion nebula (Orion molecular cloud, OMC) is some 50 ly in diameter and 1500 ly from Earth. The temperature within the cloud is of order 10 K and the atomic density is 10^6 cm^{-3} . The chemical composition is diverse and contains small diatomic molecules, large polyatomic molecules and dust particles covered with a thick ice mantle.
- Circumstellar medium: the region around a star will depend on the type and extent of its evolution. Young stars have high photon fluxes in the UV and all molecules are photodissociated and photoionised in these regions (called HII regions) i.e. they are photon-dominated regions (PDR). The star may also have ejected some dust during its collapse and there may be significant dust surface chemistry and scatter of the starlight.
- Photon-dominated regions: around stars where the intense UV and far-UV radiation fields are responsible for photolysis of all molecules, leaving atoms, some of which are ionised.

The state of matter within these regions needs to be determined before the balance of energy and chemistry can be understood. Extreme photon fluxes break all chemical bonds, prevent molecule formation and ionise atoms but as the density of species increases the UV and far-UV photons are absorbed and molecules begin to form. Chemical reactions are, however, slow in the gas phase due to the low temperature, and molecules condense out on the surface of dust particles, perhaps forming ice grains. Once on the surface, molecules continue to be photoprocessed by the starlight as well as by the continual bombardment of cosmic rays.

Visible extinction

Starlight from a distant star on its way to Earth can be scattered or absorbed by interstellar dust in a cloud. Scattering may be considered as the collision of a photon with a particle during which the photon changes its direction and is removed from the line of sight between the star and the Earth. Scattering is strongly dependent on particle size and wavelength and examination of the interstellar extinction curve indicates two sizes of particles: around 500 nm, responsible for short-wavelength scatter; and around 10 nm, responsible for short-wavelength scatter in the UV. Thus stars viewed through a dust cloud have short-wavelength radiation removed from their spectrum, leaving apparently more red light than the surface temperature would suggest. This is called *reddening* (Figure 5.7).

Dust particles appear to become oriented in interstellar magnetic fields, aligning themselves with the field lines to produce preferential absorption of light polarised

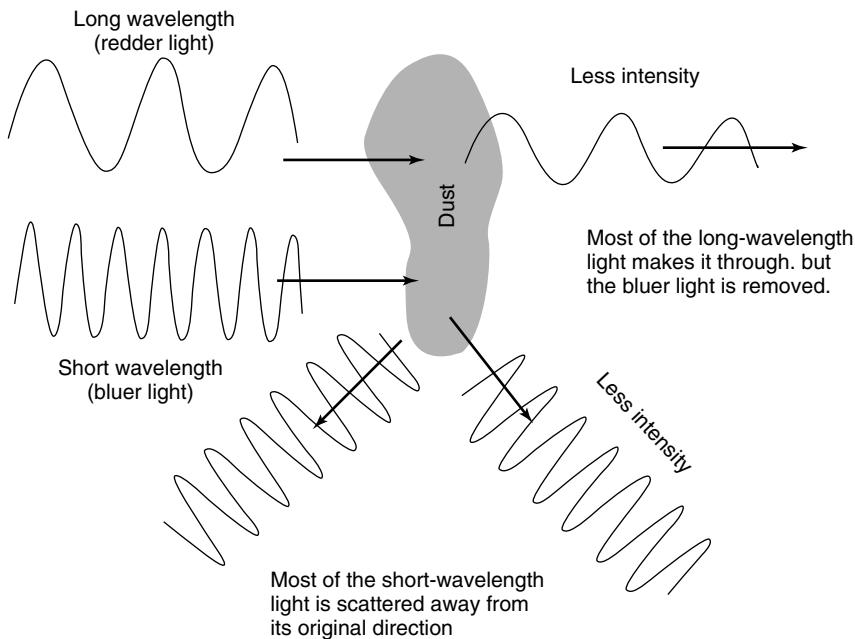


Figure 5.7 Visible extinction and the reddening of stars. (Adapted with permission from Fraser, 2002)

in one direction or another. Hence the passage of starlight through a dust cloud can result in an imbalance of light with one polarisation. Does this affect the synthesis of a molecule with one favoured chirality and hence is this a possible source of non-enantiomeric synthesis – the question of homochirality?

The interstellar extinction has a great effect on distance determination for stars. The B/V index derived in Chapter 2 will be distorted by the presence of interstellar dust, with an amount of radiation in the blue part of the spectrum removed. The difference between the observed colour index and the colour index on which it should have based its temperature is called the *colour excess*. We defined m to be the measured apparent magnitude, which must now be corrected by an amount A_v and added to the distance modulus equation:

$$(m - A_v) - M = 5 \log \left(\frac{d}{10} \right) \quad (5.1)$$

The estimate of the distance must now consider the estimate of the interstellar extinction A_v , best estimated by the reddening; A_v can take several values and in calculations of molecular cloud chemistry typical values are of order 1 but may be as much as 5. The distance calculation in Equation 5.1 can be significantly perturbed so that an A_v of 2.4 can reduce the apparent distance by a factor of 3.

The extinction is strongly size dependent, and hence wavelength dependent, and for particle sizes of around 100 nm the dependence on the extinction or scattering

cross-section is approximately inversely related to the wavelength:

$$E_\lambda \propto 1/\lambda \quad (5.2)$$

where E_λ is the fractional proportion of energy lost from the starlight at wavelength λ , and A_λ is:

$$A_\lambda \approx 0.008 \lambda^{-\frac{4}{3}} \quad (\text{mag pc}^{-1}) \quad (5.3)$$

where A_λ is the magnitude change per parsec when λ is expressed in microns. There are significant departures from this relation, notably at 0.2, 3, 10 and 30 μm , and these features are attributed to the chemical composition of the grains.

Polarisation can be introduced into this expression but practically the only process resulting in a strong polarisation dependence requires the dust particles to be non-uniform and aligned. The scatter would then depend on polarisation and become a differential scattering cross-section now dependent on the shape of the particles. If the particles are assumed to be needles, with the length three times the width (the aspect ratio), photons with their electric field parallel to the long axes are preferentially scattered but neighbouring particles will cancel out the effect unless they too are all aligned. The interstellar magnetic field is usually assumed to be 10^{-9} T based on the wavelength dependence of the Faraday effect and the observation of the Zeeman effect in atoms and molecules. Observations for our galaxy seem to suggest a magnetic field parallel to the plane of the galaxy with strength of 2×10^{-10} T and this is sufficient to align the dust grains.

5.4 Rates of chemical reactions

The chemistry of the ISM depends on the physical conditions and the molecules available within the local cloud, all of which is controlled by the rate of chemical reactions under the local conditions. Consider the standard alphabetical reaction between a moles of species A and b molecules of species B going to c moles of C and d moles of D given in Equation 5.4:



The rate of the reaction is related to the concentration of the species multiplied by a rate constant written generally as:

$$\nu = k[\text{A}]^a[\text{B}]^b \quad (5.5)$$

where ν is the rate of the chemical reaction, k is the rate constant and $[\text{A}]$ and $[\text{B}]$ are the concentrations of the reactants. The parameters a and b can be any number (including non-integer numbers) and for simple reactions represent the stoichiometry of the reaction. The total overall order of the reaction is given by $a + b$, so $1 + 1$ gives an overall second-order reaction but would be first order with respect to each of the reactants. The first-order rate equation for the

removal of a reactant may be written in terms of the rate of change of [A] with time as:

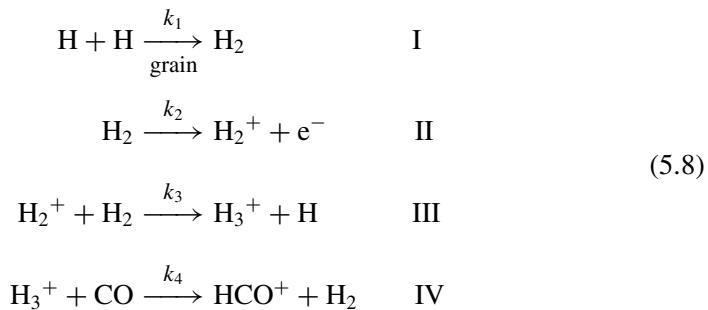
$$\frac{d[A]}{dt} = -k[A] \quad (5.6)$$

This equation can be generalised to any process within a complex set of chemical reaction pathways to allow for reactions generating A, written with positive rate constants, and reactions removing A, written with negative rate constants for example:

$$\frac{d[A]}{dt} = -k_1[A][B] + k_2[C][D] \quad (5.7)$$

There is a similar sequence of equations associated with the formation of C and D, all of which are coupled together to form a set of coupled differential equations. The set of differential equations can then be solved in time to determine the relative abundances of the species in the cloud. Let us make this clearer with an example.

Consider a kinetic model for the formation of H_2^+ towards the protostar GL2136. A simple model for the hydrogen chemistry of the cloud can be seen with the reaction scheme below:



This is our first example of a chemical reaction network so let us look at it carefully. Hydrogen molecules are assumed to form on the surface of a dust grain (more of this shortly), which then leave the surface to produce the gas-phase concentration. The overall rate of this process is given by the constant k_1 in Reaction I. Hydrogen is then ionised by photons or cosmic rays at an effectively constant rate k_2 to form H_2^+ . Rapid reaction with H_2 converts H_2^+ to H_3^+ and H atoms with a rate constant k_3 in step III and finally H_3^+ protonates CO to form HCO^+ in step IV, recovering H_2 present in steps I and II. HCO^+ is easily seen in the ISM, as is CO, and with difficulty we can obtain a density for H_2 but H_3^+ and H_2^+ are very hard to see. The chemical network enables the densities of each species in the cloud to be determined and estimates made of the detection sensitivities required to make the observations. The rate constants for each of the processes I–IV must be known, along with the initial densities of the species and the conditions within the cloud towards GL2136. The rate constants can be measured in the laboratory but there is a huge difference between chemistry at the temperature of the laboratory (298 K) and the temperature of the cloud (10 K).

Rate constant temperature dependence

The rates of a chemical reaction generally increase with temperature, an example of which is the hydrolysis of sucrose, which is 4.13 times faster at body temperature (35°C) than at room temperature (25°C). This represents a surprisingly large change in the rate of chemical reactions due to the simple rise to body temperature. It is then clear that cooling the same reactions down will slow the reactions and there is a big difference between body temperature and 10 K in an interstellar cloud.

Measurements of rate constants in the laboratory show that rate constants follow an empirical temperature dependence called the Arrhenius equation:

$$k(T) = A \exp\left[-\frac{E_a(T)}{RT}\right] \quad (5.9)$$

where A (s^{-1}) is the pre-exponential factor, T is temperature (K), E_a is the activation energy for the reaction (J mol^{-1}) and R is the gas constant ($\text{J mol}^{-1} \text{K}^{-1}$). For the simple bimolecular reaction:



A takes the value $8.0 \times 10^{10} \text{ s}^{-1}$, which is close to the collision rate at room temperature, and $E_a = 42 \text{ kJ mol}^{-1}$, which is the same order of magnitude as the energy required to break a weak chemical bond. Some consideration of the collision dynamics behind the Arrhenius equation throws light on the form of the equation, especially the temperature dependence.

The collision between two reactants to form products is the study of molecular reaction dynamics and as a recurring theme we need to understand some detailed quantum theory to predict astrochemistry fully. Collision between two reactants can result in the two molecules bouncing off one another and if no energy is exchanged this is called an elastic collision but if some kinetic, rotational or vibrational energy is exchanged it is called non-elastic. The third option is called reactive scatter and this leads to breaking of bonds and possibly the formation of new bonds. From kinetic theory the collision rate between two molecules A and B, Z_{AB} , assuming that they are hard spheres, is given by:

$$Z_{AB} = N_A N_B \sigma_{AB} \left[\frac{8kT}{\pi\mu} \right]^{1/2} \quad (5.11)$$

where N_i are the number density of molecules, σ_{AB} is the collision cross-section πr^2 (where r is the radius of the molecule), T is the temperature and μ is the reduced mass (as defined in Chapter 3, Equation 3.16). The terms in the equation relate how many molecules are present, N , to the size of the collision target, σ_{AB} , and the relative collision velocity at fixed temperature:

$$\bar{c}_{rel} = \left[\frac{8kT}{\pi\mu} \right]^{1/2} \quad (5.12)$$

For N₂ molecules in the air at room temperature \bar{c}_{rel} is of the order of the speed of sound, 370 m s⁻¹, σ is 0.43 nm² and $Z = 5 \times 10^{28}$ cm⁻³ s⁻¹. This is a very large number, which means that collisions between molecules occur very frequently and the energy can be averaged between them, ensuring the concept of *local thermal equilibrium*. Each molecule collides Z_{AB}/N_A times per second, which is about 5×10^9 s⁻¹ once every 0.2 ns. However, in the diffuse ISM where the molecule density is of order 10² cm³ the collision frequency is 5 × 10⁻⁸ s⁻¹ or a collision every 1.5 years.

Not all collisions result in a chemical reaction so there must be some consideration of the stereochemistry of the reaction. If the collision occurs with the non-reactive ends of the molecules pointing towards one another then the reaction will not occur. Only some collisions are favoured for the formation of a chemical bond if the three-dimensional geometry of the collision is correct. This eventuality is overcome by adding a steric factor so that molecules must be the right way round to react. Then there has to be sufficient energy to break a bond and form a new bond with the energy balance related to the thermodynamics of the reaction. The energy requirement is the exponential term involving the activation energy in Equation 5.9. Combining the collision number and the steric factor, P , with the activation energy gives an expression for the rate constant that has the correct temperature dependence:

$$k(T) = P\sigma_{AB} \left(\frac{8kT}{\pi\mu} \right)^{1/2} N_B N_A \exp \left[-\frac{E_A}{RT} \right] \text{ (cm}^3 \text{ s}^{-1}\text{)} \quad (5.13)$$

Collections of reaction rate constants exist for explosions, atmospheres and astrophysical reactions. One such collection is the UMIST Astrochemistry Database and here the form of Equation 5.13 is generalised to:

$$k(T) = \alpha \left(\frac{T}{300} \right)^\beta \exp \left[-\frac{\gamma}{T} \right] \text{ (cm}^3 \text{ s}^{-1}\text{)} \quad (5.14)$$

where α , β and γ are fitted experimental parameters, in some instances only derived from theory. Now we can return to the model of the proto-star hydrogen chemistry and look up the rate constants for reactions I–IV in the reaction scheme of Equation 5–8 and calculate the expected densities of the relative species. There are some different types of reaction in our reaction scheme and these are also stored in a standard form on the UMIST database. Direct cosmic ray ionisation is stored as:

$$k(T) = \alpha \text{ s}^{-1} \quad (5.15)$$

and for photoinduced reactions as:

$$k(T) = \alpha \exp[-\gamma A_v] \text{ s}^{-1} \quad (5.16)$$

where A_v is the extinction at visible wavelengths defined in Equation 5.3.

Unfortunately the rates of surface reactions are not stored in a common format because there is as yet no general consensus of the rates of these reactions and how they vary with temperature.

In principle, it is now possible to construct a complete network of interconnecting chemical reactions for a planetary atmosphere, a hot molecular core or the tail of a comet. Once the important reactions have been identified the rate constants can be looked up on the database and a kinetic model of the atmosphere or ISM molecular cloud can be constructed. Or can it? Most of the time the important reactions are hard to identify and if you are sure you have the right mechanisms then the rate constants will certainly not be known and sensible approximations will have to be made. However, estimates of ISM chemistry have been made with some success, as we shall see below.

Steady-state approximation

One useful trick in solving complex kinetic models is called the *steady-state approximation*. The differential equations for the chemical reaction networks have to be solved in time to understand the variation of the concentrations of the species with time, which is particularly important if the molecular cloud that you are investigating is beginning to collapse. Multiple, coupled differentials can be solved numerically in a fairly straightforward way limited really only by computer power. However, it is useful to consider a time after the reactions have started at which the concentrations of all of the species have settled down and are no longer changing rapidly. This happy equilibrium state of affairs may never happen during the collapse of the cloud but it is a simple approximation to implement and a place to start the analysis.

The steady-state approximation allows the concentrations of each species to be determined by assuming that nothing is changing significantly with time. Placing H_2^+ formed in the reaction network defined by Equation 5.8 into steady state requires that the processes leading to the formation of H_2^+ should have a zero effect on the rate of change of H_2^+ with respect to time: that is the derivative should be zero:

$$\begin{aligned} \frac{d[\text{H}_2^+]}{dt} &= k_2[\text{H}_2] - k_3[\text{H}_2][\text{H}_2^+] = 0 \\ [\text{H}_2^+] &= \frac{k_2}{k_3} \\ [\text{H}_2^+] &= \frac{10^{-17}}{10^{-9}} \frac{\text{s}^{-1}}{\text{cm}^3 \text{ s}^{-1}} = 10^{-8} \text{ cm}^{-3} \end{aligned} \quad (5.17)$$

Some simple algebra allows the concentration of H_2^+ to be established within the steady-state approximation. The simple model for the conditions in GL2136 has a temperature of 10 K determined from the rotational temperature of CO (which

is assumed to be in local thermal equilibrium) and a number density of $H_2 = 10^4 \text{ cm}^{-3}$. The UMIST database has a dark cloud cosmic ray ionisation rate of $k_2 = 10^{-17} \text{ s}^{-1}$ and a reaction rate of $k_3 = 10^{-9} \text{ cm}^3 \text{s}^{-1}$, which gives a steady-state concentration of $[H_2^+] = 10^{-8} \text{ cm}^{-3}$ – much less than one H_2^+ molecule per cm^3 !

Similarly, placing H_3^+ in steady state by setting the rate of change of $[H_3^+]$ from all processes creating and destroying H_3^+ gives:

$$\frac{d[H_3^+]}{dt} = k_3 [H_2^+] [H_2] - k_4 [\text{CO}] [H_3^+] = 0 \quad (5.18)$$

$$[H_3^+] = \frac{k_3}{k_4} \frac{[\text{H}_2]}{[\text{CO}]} [H_2^+]$$

Now the steady-state concentrations can be compared with the known observed concentrations. The H_2 concentration in GL2136 is determined from the concentration of CO, determined by the 115 GHz transition intensity and the assumption that the ratio $[H_2]/[\text{CO}]$ is of order 10^4 , the known relative abundance of H_2 and CO. Putting these numbers into Equation 5.18 gives a steady-state concentration of $[H_3^+] = 2 \times 10^{-4} \text{ cm}^{-3}$. The astronomical observation for the column density (the density of H_3^+ along the line of sight) is $H_3^+ = 3 \times 10^{14} \text{ cm}^{-2}$, which implies that the cloud is some 3 ly across in effective path length. Having discovered this important property of the cloud, we can go on to perform other steady-state calculations based on other species in the cloud and refine the estimate of the cloud diameter along the line of sight. A kinetic model of the cloud can then be established but there is the small matter of which reactions should be included in the model and the rate of reaction in the conditions of the cloud.

Activation energy

The activation energy is a measure of how much energy a reaction must acquire from its local environment in order for the reaction to start: it is an energetic barrier.

The energetic barrier is a measure of how many bonds have to be broken or formed from the available energy and is on the pathway of the chemical reaction (Figure 5.8) between the initial state of reactants and the final state of products. The activation energy can be changed by the introduction of a catalyst and, in an astrochemical context, this can be the surface of a dust grain. Figure 5.8 also defines the change in energy of the reaction, $\Delta_r G$, in going from reactants to products – more of this in Chapter 8.

The internal energy of a gas at a given temperature available as kinetic energy is given by the expression:

$$\varepsilon(T) = \frac{3}{2} RT \quad (5.19)$$

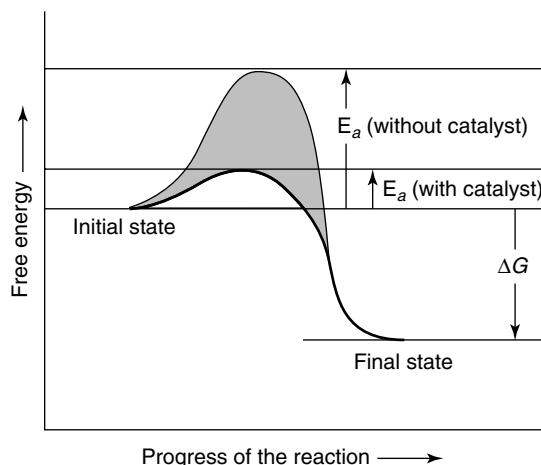


Figure 5.8 Definition of activation energy with and without a catalyst. (Adapted with permission from Fraser, 2002)

where R is the gas constant in kJ mol^{-1} and T is the temperature in K. The $3/2$ factor is usually dropped from discussion and the analysis of energy is usually related to RT . At a room temperature of 298 K RT is $2.479 \text{ kJ mol}^{-1}$, which may be compared with a typical bond energy of order 400 kJ mol^{-1} (see Table 5.2). The reality is that chemical bonds are not broken readily by collision with molecules in a gas and are generally chemically stable. Weaker bonds such as hydrogen bonds, which are about 40 kJ mol^{-1} , are obviously more vulnerable. When it comes to chemical reactions (the energy to start a chemical reaction) a bond must first be broken before new bonds can form, so the activation energy can be rather high.

The size of the activation energy gives some insight into the fundamental mechanism of a chemical reaction. Consider the rearrangement reaction



which has an activation energy of 160 kJ mol^{-1} . The formation of the transition state required to make the rearrangement occur clearly requires about a quarter of the bond energy to be provided to make the necessary changes in geometry. The exponential term in the Arrhenius equation takes the value

$$\exp\left(\frac{160}{2.479}\right) = 1.0 \times 10^{-28} \quad (5.21)$$

so at room temperature this reaction is highly unlikely. Matters can be improved by increasing the temperature to 1000 K, when the exponential factor becomes 4.4×10^{-9} . At temperatures of the photosphere of the Sun (5780 K) RT is 42.1 kJ mol^{-1} and chemical processes would be much faster but it turns out that there are only a few stable molecules on the Sun because all of the reactions

leading to the destruction of molecules are also very rapid. There is, however, some Sun spot chemistry where the temperature is about 1000 K cooler.

The improvement in the rate of chemical reactions is reversed when temperature is cooler and at temperatures as low as 30 K (a warm corner of TMC-1) the exponential term is of order 10^{-279} and nearly all reactions between neutral species are frozen out at 50 K. Two important classes of reactions survive: radical–radical chemistry and ion–molecule chemistry. The importance of these different reaction types will become apparent later with the construction of the models of molecular clouds. For the moment, however, laboratory measurements of reactions in radicals such as C₂H have shown that even with temperatures as low as 15 K the rate constant for reactions of the type:



can still be as high as $2.3 \pm 0.3 \times 10^{-10}$ molecule⁻² cm⁻⁶ s⁻¹ and must be included into the models of TMC-1.

The second class of reactions surviving in the cold TMC-1 conditions is ion–molecule reactions. These reactions involve the collision of a charged reactant with a neutral molecule and an understanding of the mechanism of reaction sheds some light on the reason for the rate of reactions at low temperatures. The charge on one reactant induces a dipole (charge-induced dipole) in the neutral species at long range with a $1/r^4$ interaction power law, with separation r . At long range the interaction is much stronger than for neutral–neutral reactions ($1/r^6$) and the target looks much bigger, there are larger reaction cross-sections and rate constants and crucially there is no activation barrier. In fact the charge-induced dipole attraction leads to a complicated temperature dependence and the activation energy may become negative! At the centre of all ion–molecule reactions in the ISM is H₂⁺, formed by ionisation of H₂ in the mechanism above (Equation 5.8). It then reacts with neutral hydrogen to form H₃⁺ via the reaction:



This is one of the fastest known reactions and occurs with nearly every collision to produce H₃⁺, which has been seen in diffuse and giant molecular clouds. However, because the rate of reaction is so fast the detection of H₂⁺ is going to be unlikely as it is quickly removed by the chemistry. Once H₃⁺ is formed it quickly protonates a number of species, particularly CO to form HCO⁺ and also O to form OH⁺, leading to the other ions containing protons seen in Table 5.1. The construction of networks of possible reactions requires a knowledge of the fundamental chemical physics of molecules and the possible chemistry in the local environment.

5.5 Chemical reactions in the interstellar medium

There are many possible chemical reactions that must be considered for all networks of chemical processes, whether within a cold dark cloud, a giant molecular

cloud, the atmosphere of a planet, the reactions of a hydrothermal vent or the metabolism of a cell. For the moment we restrict attention to the ISM gas-phase reactions and reactions mediated by the surfaces of dust grains.

The simple collision between two atoms may result in a new chemical bond but the new molecule still has the energy of the new bond and will dissociate immediately unless the energy is removed. The reverse of bond formation is called dissociation, recovering two atoms, but this requires sufficient energy to break the chemical bond, as shown in Figure 5.9, reaction (a). An obvious way to remove the energy of bond formation is to emit it as radiation, either as visible photons or a cascade of infrared and microwave photons passing down all of the vibrational and rotational energy levels of the molecule until it finally reaches the ground state. These processes may be rapid with large Einstein A coefficients for the optical transition but slower for the infrared and microwave transitions. The result is a radiatively stabilised molecule. The reverse process, absorbing radiation from anywhere within the electromagnetic spectrum, will result in radiatively heated molecules and there is a balance to be met between the radiative heating and cooling: equilibrium with the local radiation field. The absorption or emission of photons requires the molecule to have an allowed transition and hence a spectrum. A ro-vibronic spectrum exists for CO because it has a dipole moment and the rotational transitions are allowed in either direction. This cannot happen for H₂ because there is no allowed dipole microwave transitions and hence no way for two H atoms to collide and form H₂ and be radiatively stabilised. For this reason, the primary source of H₂ in the ISM, as indicated in Equation 5.8, is the association of two H atoms on the surface of a dust grain, with the dust grain absorbing the energy of the bond formation. The stabilised H₂ molecule once formed then desorbs from the surface into the ISM.

Radiative cooling processes for molecules compete with collisions between molecules for removing energy. Collision with another molecule, or in general

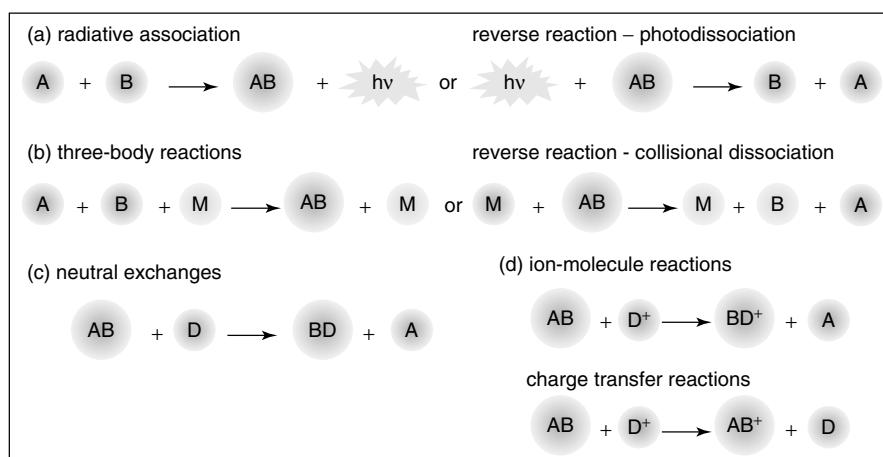


Figure 5.9 Reactions in the interstellar medium. (Adapted with permission from Fraser, 2002)

a third body (Figure 5.9, reaction (b)), can take away some of the excess energy of bond formation in the form of kinetic energy of the collision partners. Third-body stabilisation is most efficient when there are large numbers of molecules per cm³ and will be less important when the molecular densities are low. In the diffuse ISM the time between collisions can be rather long and so third-body cooling may be slow. Conversely, if the local temperature is high then a high-speed impact of a third body with a target molecule can break the chemical bond, causing the molecule to fall apart – collisional dissociation.

Neutral reactions of the type indicated in Figure 5.9, reaction (c), are subject to the restrictions of the activation energy and may proceed very slowly. This may not prevent the reactions from contributing, especially as the lifetime of molecular clouds may be 10⁶ years. Under cold conditions ion–molecule reactions (Figure 5.9, reaction (d)) dominate the chemical composition and the production of molecules. Charge transfer reactions present a further variation, all of which require the generation of ions and electrons either by photoionisation or high-energy collisions with cosmic rays.

The recombination of an ion with an electron (e^-) (Figure 5.10, reaction (a)) is the reverse of photoionisation. The ionisation energy is the energy required to remove the electron from the neutral species, as we shall see later, but the reverse process allows a charged species to capture a free electron. The resulting neutralised molecule or atom, however, now has the electron in an excited state and it must lose the ionisation energy to return to the ground state. The energy can be lost by analogous radiative processes seen in the neutral–neutral reactions discussed above or by losing the energy to a third body; a selection of variants is shown in Figure 5.10.

The last class of reactions is of importance for the formation of organic molecules and organic polymers. The carbon insertion reactions (Figure 5.11, reaction (i)) are ion–molecule reactions and so are favoured at low temperatures and result in the

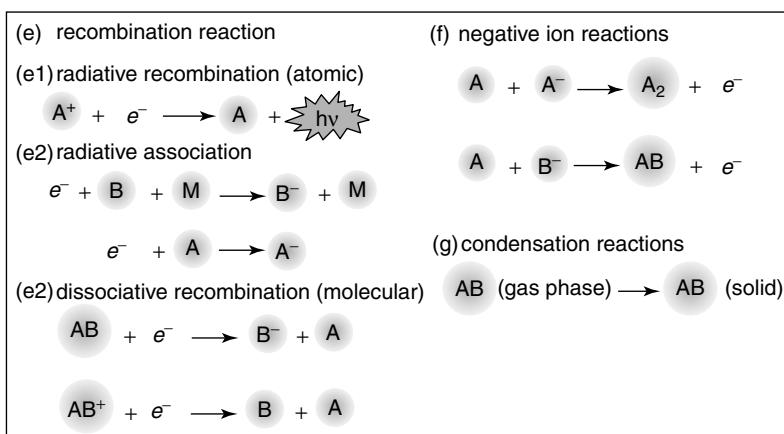


Figure 5.10 Electron–ion and electron–molecule reactions. (Adapted with permission from Fraser, 2002)

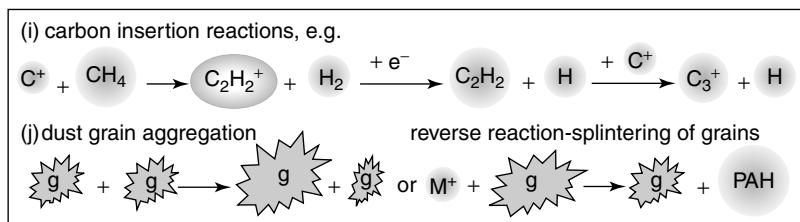


Figure 5.11 Hydrocarbon reactions leading to PAH synthesis. (Adapted with permission from Fraser, 2002)

insertion of C^+ into CH_4 . The resulting two-carbon species can then propagate to three-carbon and higher carbon chains. It is, however, unlikely that there is a gas-phase route for the synthesis of all carbon-containing species, an important class of which is the polycyclic aromatic hydrocarbons (PAH). The possibility of incorporating the low-mass hydrocarbons onto the surface of dust grains is important and new surface reactions occur here with lower activation energies, perhaps driven by cosmic ray bombardment or photochemistry (Figure 5.11, reaction (j)). Complex organic chemistry is now possible, returning more chemically diverse molecules to the ISM.

5.6 Photochemistry

Photochemical reactions are initiated by the interaction of a molecule with the local radiation field, perhaps from an embedded star. In cold dark clouds the radiation field may be rather small but as stars begin to form and emit radiation, initially at low energy but of every-increasing energy, photochemistry becomes important in generating a wide variety of reactive species to add to the diversity of species available for the chemical networks.

The breaking of chemical bonds initiated by photons may be written, for example



where the photon breaks the chemical bond and the excess energy from the reaction is shared among the dissociation product fragments. The energy, E , of the photon required to break the bond is given by Planck's Law $E = h\nu$, where h is Planck's constant and ν is the frequency of the radiation. This is usually scaled to allow for one mole of molecules and becomes $E = N_A h\nu$, where N_A is Avogadro's number, or is rewritten in terms of wavelength:

$$E = \frac{N_A h c}{\lambda} \quad E(\text{kJ mol}^{-1}) = \frac{119\,625}{\lambda} \frac{\text{kJ nm mol}^{-1}}{\text{nm}} \quad (5.25)$$

to give the energy in kJ mol^{-1} when λ is in nanometres. If the photon has a lower energy than the bond energy then it may be absorbed and held as internal energy of CO before relaxing to the ground state. The photon–molecule (or

atom) interaction opens a number of photoinitiated reaction pathways all with relative cross-sections. The cross-section is a measure of the size of the molecular target for the absorption of the photon and has a strong wavelength dependence. For high-energy photons around a young stellar object both photodissociation and photoionisation are important.

Photodissociation

A molecule within the radiation field of a local star may absorb a photon of energy sufficient to break the chemical bond, thus producing photolysis products and reactive species. Equation 5.25 relates the photon energy and the bond energy, and performing the calculation for CO (the strongest chemical bond) shows that a radiation of 105 nm is required to break the bond. The calculation has also been performed for some common chemical bond energies and the results are presented in Table 5.2. Laboratory measurements suggest that although the radiation at 105 nm is sufficient to break the CO chemical bond the absorption cross-section is rather small and the process is not efficient: the process is much faster when the wavelength is considerably shorter.

The data in Table 5.2 show that around a hot local star where the photon flux is high the photons will prevent all chemistry from occurring in so-called photon-dominated regions. Using Wein's Law, a star with a surface temperature of 28 000 K will have a λ_{max} of 104 nm and so all chemical bonds are susceptible to photodissociation and little or no chemistry is possible. The Sun has a photosphere temperature of 5780 K and λ_{max} is 501 nm in the green but there is a significant photon flux at wavelengths shorter than 501 nm, found by integrating Planck's radiation law. Removal of the short-wavelength radiation shorter than 323 nm is required to allow the C–C single bond to survive. This short-wavelength shield on Earth is the ozone layer, as we shall see, but there is no such shield in the ISM. However, molecules buried deep in ice-dust particles will protect the C–C bond and allow more organic chemistry to occur.

Table 5.2 Bond energies and photodissociation wavelengths

Chemical bond and bond order ^a	Bond energy (kJ mol ⁻¹)	Wavelength (nm)
C≡O (t)	1075	105
N≡N (t)	943	120
C≡N (t)	754	149
C≡C (t)	838	134
C=C (d)	612	184
C–C (s)	348	323
C–H (s)	415	271
H–H (s)	436	274

^a Bond order: (t) triple; (d) double; (s) single.

Photoionisation

Molecules in an intense radiation field can also undergo ionisation by removing an electron from a molecular orbital. The energy of the photon required depends on the ionisation energy (negative of the binding energy) of the electron in the molecule or atom. Ionisation energies for a number of species are listed in Table 5.3. Multiple electrons may be removed (usually one at a time), leading to second and third ionisation energies.

The ionisation process may be written:



and the energy balance requires the energy of the electron to be given by:

$$\text{KE } (e^-) = h\nu - \text{IE} \quad (5.27)$$

where $\text{KE}(e^-)$ is the kinetic energy of the electrons and IE is the ionisation energy. Astronomers use the convention $A = \text{AI}$, $A^+ = \text{AII}$, $A^{2+} = \text{AIII}$, but we shall use the conventions A , A^+ , $A^{2+} \dots$ associated with the atomic and molecular physics.

All of the wavelengths for ionisation are at very short, in the hard ultraviolet. Such wavelengths are present around hot stars where the flux of radiation above 90 nm is sufficient to ionise all molecules. The photon-dominated region (PDR) inventory is dominated by ions, many of which may multiply ionised species, and a population of electrons with different energies. The concept of an electron temperature is used to describe the energies of electrons in plasmas such as in a PDR or a flame. Temperatures as high as 10^6 K are calculated by using the degree

Table 5.3 Ionisation energies and wavelengths

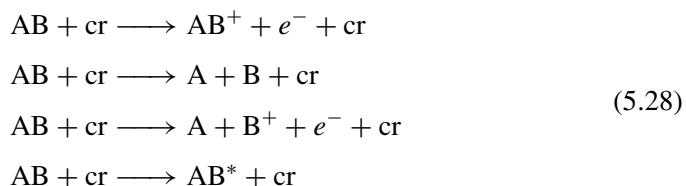
Atom or molecule	Ionisation energy (eV)	Ionisation wavelength (nm)
H I	13.6	90.73
He	24.59	50.48
He^+	54.51	22.8
Li	5.32	233.3
Li^+	75.63	16.4
C	11.26	110.2
C^+	24.38	50.91
O	13.62	93.19
O^+	35.11	35.53
H_2	15.43	80.45
CO	14.01	88.60
C_2	12.0	103.4
CN	13.8	89.9
CH_4	12.6	98.5

of ionisation, relating the temperature to energy via the Boltzmann constant, k_B , or gas constant R , for the molar quantity.

5.7 Charged particle chemistry

Primary cosmic rays (cr) emitted by stars bombard the ISM, including planets such as Earth. The particles are 84 per cent protons and 14 per cent helium nuclei stripped of both electrons called alpha particles; the remaining 2 per cent are electrons, heavier nuclei and more exotic particles.

The particles are usually of extremely high kinetic energy and hence high temperature and are capable of breaking chemical bonds or ionising species by direct impact: electron impact dissociation, ionisation and excitation (Equation 5.28):



The energy balance for the processes is conserved in the kinetic energy of the free electrons, as in Equation 5.27 for the photon energy balance. The cosmic rays in Equation 5.28 are more slowly moving and hence less energetic on the right-hand side of the equation compared with the left-hand side. Similarly, the molecules on the right-hand side may have internal energy in the form of rotation, vibration or electronic energy, represented as AB^* , to enforce the energy balance for the collision with the cosmic rays.

The last process results in the cooling of the cosmic rays because the particle leaves some energy with the colliding species. This is an example of non-elastic scatter of the cosmic rays and a lowering or ‘cooling’ of the electron/cosmic ray temperature. The excited molecule can then relax, emitting a combination of infrared or microwave radiation. The production of ionic species is by cosmic rays or photons, which is most important in dark molecular clouds, facilitating in particular the carbon insertion reactions leading to multi-carbon hydrocarbon species.

5.8 Polycyclic aromatic hydrocarbons

The carbon insertion process may lead to very large organic species and has stimulated the idea of the formation of polycyclic aromatic hydrocarbon (PAHs) species that have been attributed to broad features observed in the infrared spectra of strongly reddened stars. The PAHs rose to fame following the positive identification of the molecules in meteorites such as ALH84001, which we shall discuss in some detail in Chapter 6. The PAHs are large aromatic ring structures such as those seen in Figure 5.12, all of which look like the products of a lean-burn candle

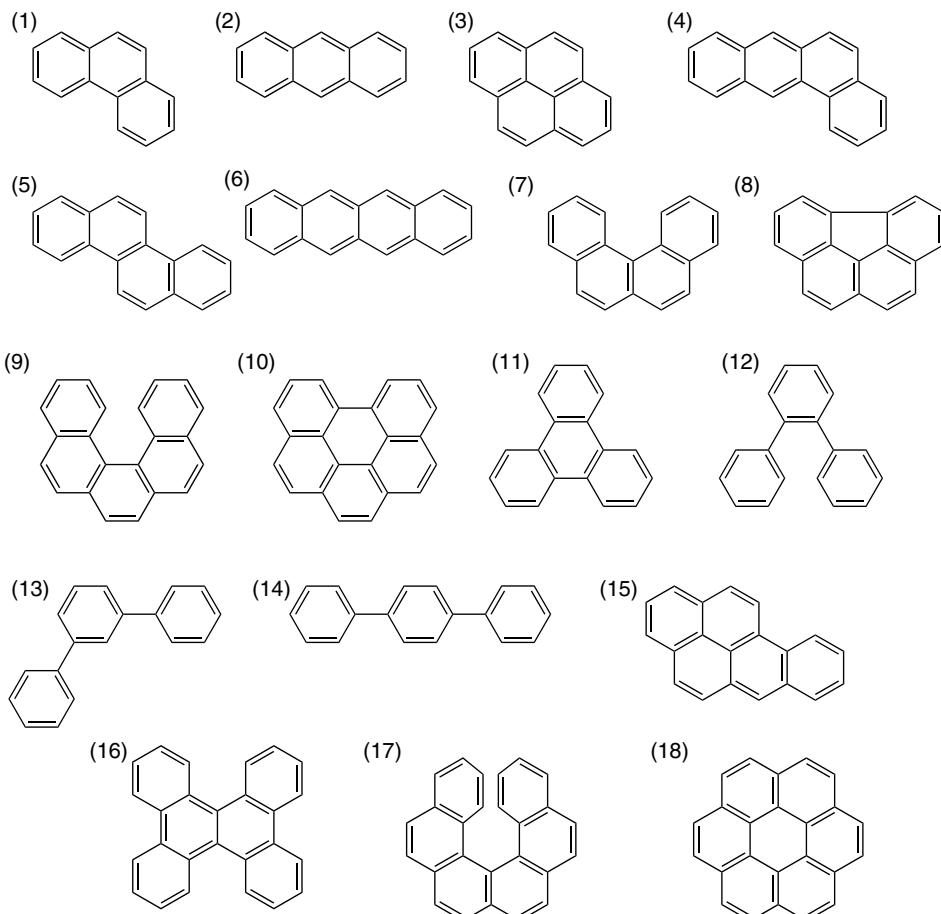
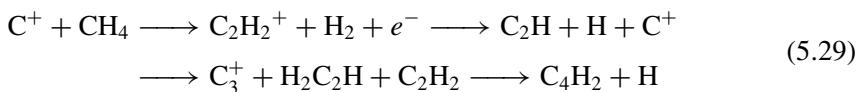


Figure 5.12 Polyaromatic hydrocarbon species: (1) phenanthrene, (2) anthracene, (3) pyrene, (4) benz[a]anthracene, (5) chrysene, (6) naphthalacene, (7) benzo[c]phenanthrene, (8) benzo[ghi]-fluoranthene, (9) dibenzo[c,g]phenanthrene, (10) benzo[ghi]perylene, (11) triphenylene, (12) *o*-terphenyl, (13) *m*-terphenyl, (14) *p*-terphenyl, (15) benzo[a]pyrene, (16) tetrabenzonaphthalene, (17) phenanthro[3,4-*c*]phenanthrene, (18) coronene

flame. More interestingly for the meteorite chemistry is that PAHs could also be the products of biological degradation such as found in oil deposits. They appear to be suggestive of complex chemical reactions, whether combustion, biological decay, cracking and perhaps even chemistry within the giant molecular clouds.

A complex chain of carbon insertion reactions is thought to initiate the formation of PAHs in the ISM and on the surface of dust grains by the formation of carbon chains:



The formation of carbon chains seems to be a natural consequence of the ion–molecule reactions and would be expected to have low activation energies. We have already mentioned the neutral–neutral radical chemistry that may also lead to carbon chain chemistry.

Laboratory measurements of the vibrational spectra of C₃, C₅ and C₇ show transitions that lie in the spectral region 2300–1700 cm⁻¹ or 4.35–5.88 μm, but again the usual caveat about the resolution of the IR instrument and the precise identification of molecules still applies. Infrared astronomy is still best at identifying families of molecules containing C–H or C–C stretch, whether aromatic or aliphatic. Laboratory measurements are, however, possible for these species both in the IR and in the visible, and the positive identification of C₂ emission in the Red Rectangle is without question, as in the identification of long chains up to HC₁₄⁻.

Propagation by positive carbon ions insertion leads to the formation of larger organic species but ring closure has to occur to produce the first aromatic species, benzene, and then further chain propagation and closure reactions to produce the PAHs. These processes are driven by polymerisation of acetylene via a sequence of stable radicals: two pathways are shown in Figure 5.13. The reaction starts with a three-body collision of acetylene followed by the formation of the propargyl radical (H₂C=C=CH), which goes on to react further to form the phenyl radical.

The second reaction pathway involves the isomerisation of acetylene to the vinylidene radical followed by further reaction with the acetylene to form benzyne and then the diphenyl radical, as shown in Figure 5.14. Addition of acetylene to the phenyl radical in a further four steps forms two fused benzene rings called naphthalene.

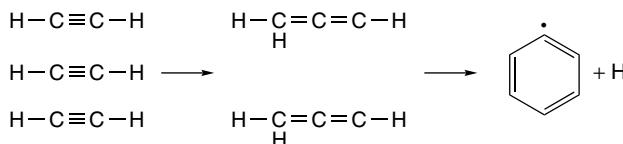


Figure 5.13 Acetylene polymerisation I

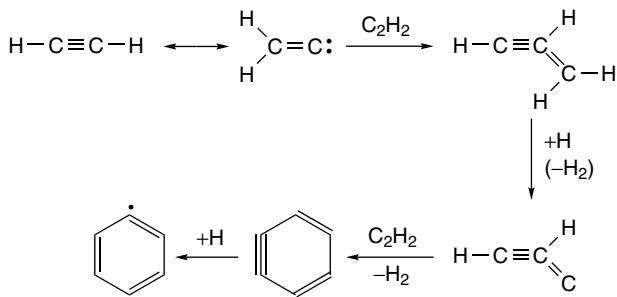


Figure 5.14 Acetylene polymerisation II (adapted from Tielens and Charnley 1996)

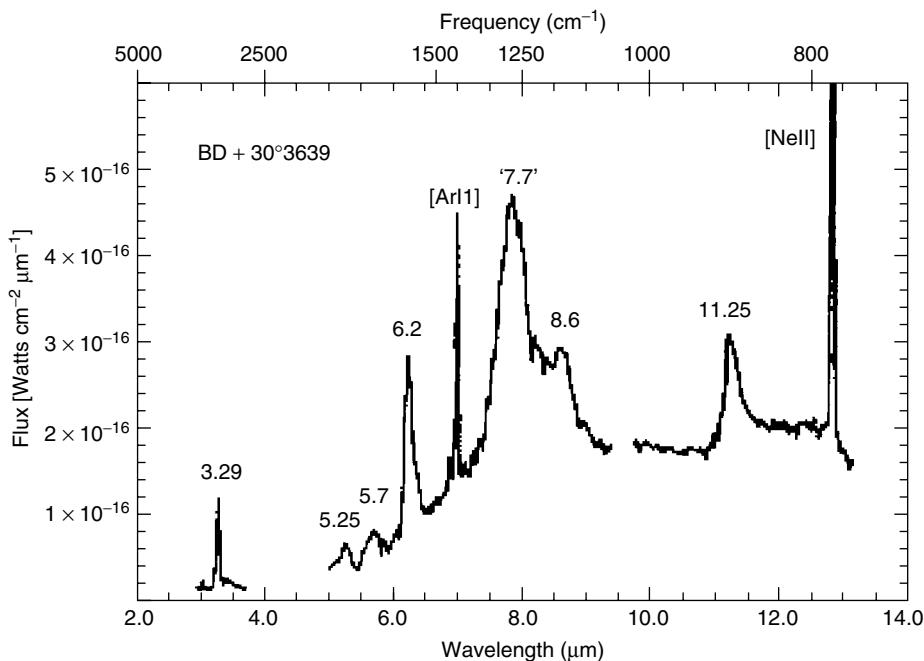


Figure 5.15 IR spectrum of the compact planetary nebula. (Reproduced by courtesy of Scott Sandford, NASA Ames Research Center)

There are some convincing observations of PAHs reported, such as the IR measurements towards the planetary nebula BD +30° 3639 shown in Figure 5.15. Identifying the Ar²⁺ and Ne²⁺ lines, the remainder of the features are attributed to the PAH spectra initially excited by UV radiation and now re-emitting in the IR. The spectrum has the following assignment: 3040 cm⁻¹ (3.29 μm), C–H stretch; 1905 (5.25 μm) is a combination overtone band of the C–H stretch; 1750 cm⁻¹ (5.7 μm) is assigned as a combination overtone C–H bend or C=C stretch; 1615 cm⁻¹ (6.2 μm) is the C=C stretch and other C=C, C–CH combination bands. Of course, these can only really belong to a family of C=C and C–H-containing molecules that could be all or part of the species shown in Figure 5.12. The observation and presentation of possible gas-phase mechanisms of production suggest that PAH chemistry and carbon chemistry are possible in the gas phase but also may occur on the surface of dust grains. Not all PAH molecules found in meteorites, however, need to come from biogenic processes such as the formation of oil in planets. The PAH synthesis in the ISM is well established.

It is tempting to take the carbon insertion mechanisms to the extreme and look for the completely unsaturated carbon allotropes of graphite and diamond. Graphite has been postulated for many years but there is at present no IR evidence for it in the ISM. This is partly due to the problems of detection. The gross selection rule for an IR spectrum requires a change in dipole moment during a vibration and the

symmetry of graphite means that nearly all of the vibrations are symmetric and so the spectrum is weak. However, there is a strong push to look for graphite to explain an observed extinction at 2175 Å, although the graphite assignment has yet to be confirmed. Graphite grains have been found in meteorites with the 2175 Å extinction but the grains have very different crystalline structure. Microdiamonds, however, have been found in meteorites and there is some suggestion of a C–H stretch associated with a tertiary carbon absorption at 2880 cm^{-1} ($3.47\text{ }\mu\text{m}$). What is conspicuous is the lack of other strong $-\text{CH}_3$ stretches associated with primary and secondary carbons that would be expected with molecular tertiary carbon stretches. The observation of microdiamonds in meteorites and the presence of the 7.46 and $3.47\text{ }\mu\text{m}$ features in the ISM suggest that microdiamonds may be present here.

5.9 Dust grains

About 1 per cent of the mass of the ISM is present in the form of dust grains formed in the outflows of dying stars. The main dust factories are the asymptotic Giant Branch stars, or red supergiants, referring to the position on the HR diagram. These stars have developed a large stellar wind once the stars have exhausted their hydrogen-burning phase. The reduction in the luminosity around the star protects the refractory molecules such as SiC and TiO, which form the aggregation nuclei for the particles. The dust aggregates into surprisingly crystalline structures, collecting oxygen from the ISM to become silicates. Condensation of volatile molecules on the surface results in the formation of increasingly large ice mantels. Recent IR measurements from the Infrared Space Observatory (ISO) have revealed some of these features in the spectrum of interstellar dust towards heavily reddened stars.

Chemical composition

Observation of the ISM using the ISO space observatory has produced an IR spectrum of interstellar dust that is characteristic (but not diagnostic) of its chemical composition, such as that seen in Figure 5.16. Comparison with studies in the laboratory suggests that the solid material of the particles is essentially silicate, similar to that found on Earth. The silicon may result from silicon-burning stars that also have high oxygen composition. These huge stars are likely to end in a supernova ejecting the silicates, particularly SiO, into the ISM. The diatomic SiO has been observed in the ISM. In addition, heavy-element-burning stars also produce carbon and the presence of mineralised carbon has been seen. The IR spectrum contains many compounds that depend on the type of molecular cloud in which the particle is formed and the nature of the surface covering, as shown in Figure 5.17.

Particles found in a diffuse cloud show a simpler chemistry consisting of silicate core, some carbonaceous particles and a thin ice mantel containing a small

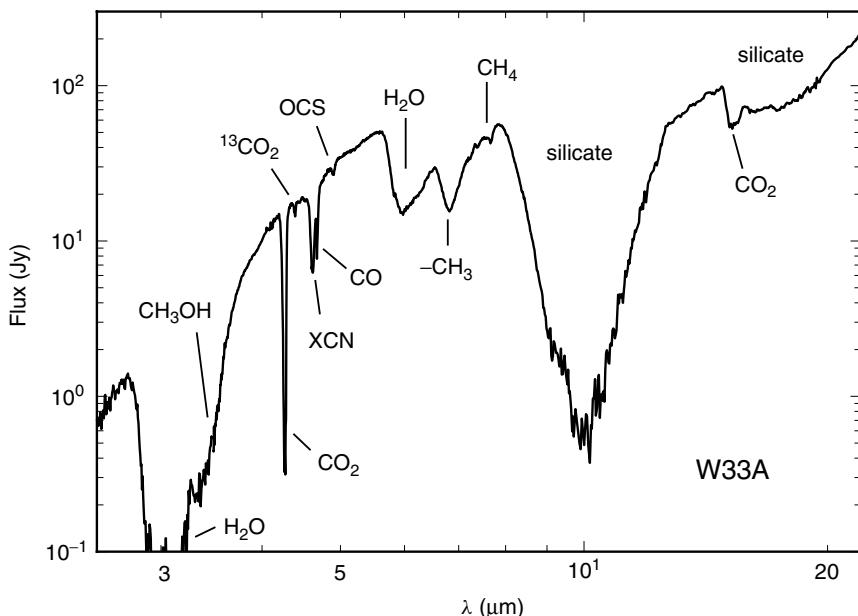


Figure 5.16 W33A Dust-embedded Massive Young star. (Reproduced from ISO data, Gibb, 2000 by permission of the American Astronomical Society)

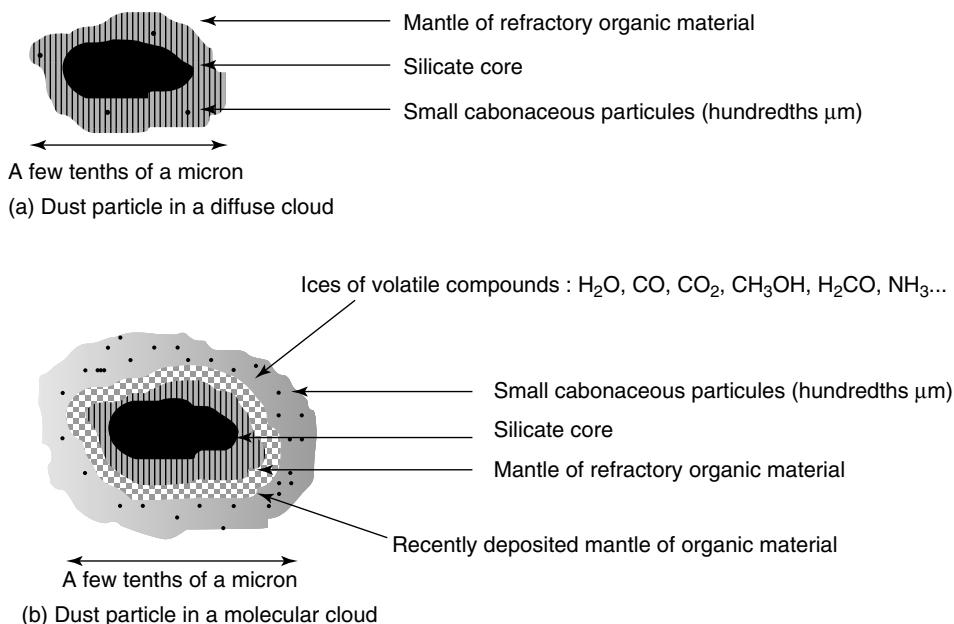


Figure 5.17 Structure of dust particles

inventory of volatile organic molecules. In a molecular cloud, the ice mantle is dominant and contains a rich variety of organic material that is processed by photolysis and cosmic ray bombardment. Our knowledge of surface reaction chemistry is, however, very poor by comparison to gas-phase chemistry, probably because of the difficulty in making laboratory measurements. Measurements of all surface reactions are at the forefront in ISM chemistry research and will change the understanding of these processes.

Surface reactions

Consideration of dust surface chemistry reads as a list of unknowns (Figure 5.18). Thanks to some remarkable observations made by the ISO we have some idea about the end product of several billion years of chemical processing on the surface of dust, such as the formation of ices, but the chemical mechanisms remain unclear. This gap in our knowledge has not stopped scientists from postulating a wide variety of roles for dust surface chemistry, many of which are possible but we do not yet know how probable.

The primary process initiating dust surface chemistry is the collision of a molecule from the ISM with the surface. The sticking probability is a measure of how often molecules will stick to the dust surface but this depends on the collision energy, the temperature of the grain surface and the nature of the chemical surface itself. The silicate surface is highly polar, at least for a grain of sand on Earth, and should attract polar molecules as well as atoms. The adsorption process can also be reversed, resulting in thermal desorption, both as the reverse of adsorption and by new molecules as the product of surface reactions.

Adsorption of molecules to the surface can result in dissociation of the molecule at the surface, producing reactive atomic species (see Figure 5.18). The dissociative adsorption process produces a selection of highly active species on the surface that are available for chemical reaction. However, the surface mobility of the species is not known and nor are any of the reaction rate constants. The role of a surface in reducing the activation energy of chemical reactions is well known and is the underlying principle of heterogeneous catalysis. The interaction energy with the surface overcomes the energy required to break the bonds to form the transition

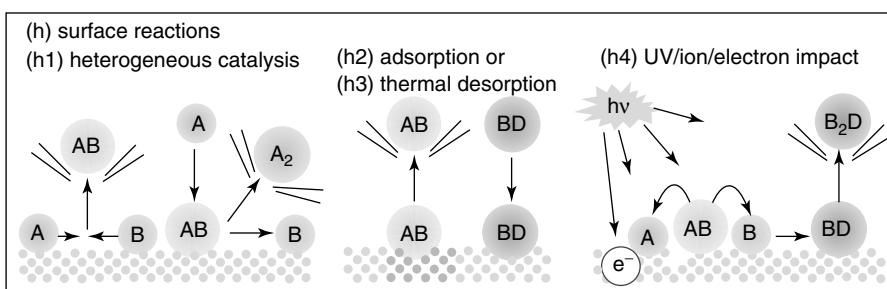
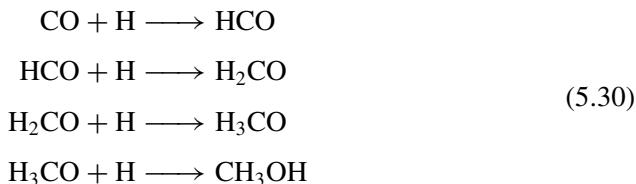


Figure 5.18 Surface reactions adapted. (Adapted with permission from Fraser, 2002)

state and so allows slow gas-phase reactions to proceed quickly. The most important heterogeneous reaction in the ISM is the formation of the neutral H₂ molecule in the gas phase. The reaction between two H atoms on a surface of a dust grain produces H₂, which then desorbs back into the gas phase. This is the primary mechanism for H₂ production and is assumed in all models of molecular clouds. Of course, the rate constants for these processes are not known and sensible estimates have to be made.

Organic synthesis in the interstellar ice

Heterogeneous catalysis is also proposed for the formation of the ice mantles around the particles. Co-adsorption of H, O and N atoms leads to the formation of water and ammonia–water ice on the surface, as deduced from ISO spectra. Adsorption of CO onto the ice surface provides a carbon source to initiate organic synthesis, for example, in the simple sequence of reactions:



The reaction scheme goes through the formation of HCO, formaldehyde and ultimately to the formation of methanol. It is hard to see how the formation of methanol previously seen in the ISM could result from a series of gas-phase reactions and so desorption of methanol from the grain surface may be responsible for the methanol concentration in the ISM. The simple scheme of reactions above shows the potential for hydrogenation of species on the grain surfaces. Many of the molecules in the ISM are unsaturated and subvalent (do not have all of its four chemical bonds) and adsorption onto the grain surface would result in the immediate hydrogenation of these species. Local thermal heating of the dust grain by a radiation field can increase the rate constant for surface reactions but we have no idea of the quantitative increase.

Surface chemistry increases the molecular diversity in the ISM, further enhanced by the presence of UV or far-UV radiation. Photolysis of molecules on the grain surfaces encourages radicals to be produced that undergo surface reactions but the rates of these fundamental processes are as yet unknown. Penetration of the radiation is controlled by the thickness and composition of the ice mantel and may protect larger molecules from photo-destruction. The dust surface is a fertile ground for organic synthesis.

Amino acid formation

The importance of organic synthesis within the ice mantel of dust grains remains poorly understood but there are some important laboratory reactions that point to

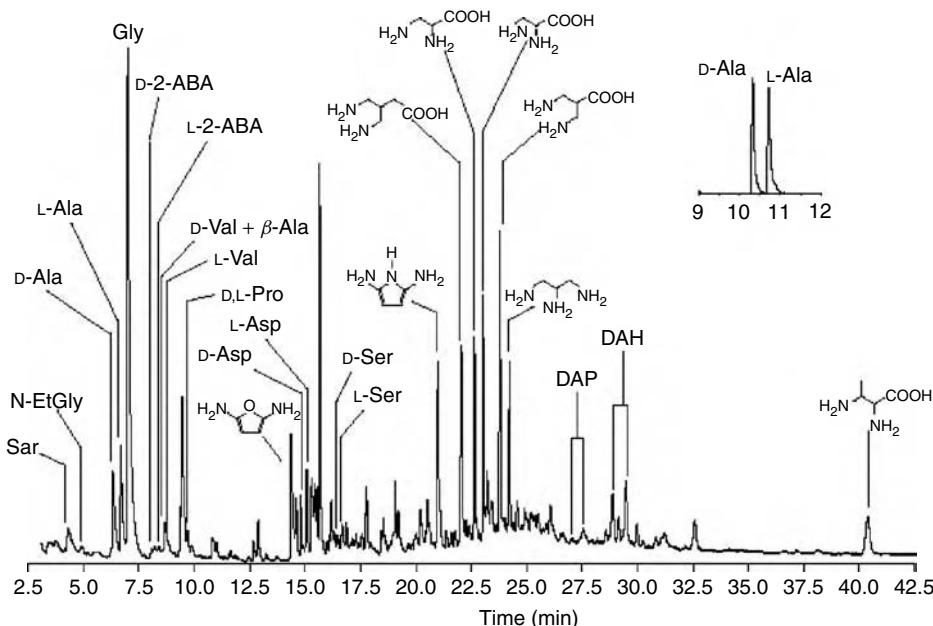


Figure 5.19 Formation of amino acids on ice surfaces irradiated in the laboratory (Nature Nature 416, 403–406 (28 March 2002) doi: 10.1038/416403a – permission granted). Data were obtained from analysis of the room temperature residue of photoprocessed interstellar medium ice analogue taken after 6 M HCl hydrolysis and derivatization (ECEE derivatives, Varian-Chrompack Chirasil-L-Val capillary column 12 m × 0.25 mm inner diameter, layer thickness 0.12 μm; splitless injection, 1.5 ml min⁻¹ constant flow of He carrier gas; oven temperature programmed for 3 min at 70°C, 5°C min⁻¹, and 17.5 min at 180°C; detection of total ion current with GC-MSD system Agilent 6890/5973). The inset shows the determination of alanine enantiomers in the above sample (Chirasil-L-Val 25 m, single ion monitoring for Ala-ECEE base peak at 116 a.m.u.). DAP, diaminopentanoic acid; DAH, diaminohexanoic acid; a.m.u., atomic mass units.

a rich diversity of biological molecules that should be considered. Irradiation of an interstellar ice analogue H₂O:CH₃OH:NH₃:CO:CO₂ (2:1:1:1:1) has been performed in the UV at 12 K. The ice is gently warmed through heating cycles of 12–40 K during the experiment and then analysed using gas chromatography after the exposure period. A typical gas chromatogram is shown in Figure 5.19, featuring a spectacular inventory of organic species, many of which are amino acids, and pointing to a possible pathway for their production. The gas chromatographic separation even resolves the enantiomers in the mixture and there would appear to be no specific enantiomeric enhancement. The evolution of amino acid chemistry in the origin of life must explain homochirality, the preference in life for one enantiomer, but the ice mantel synthesis in the laboratory does not show this selection.

The laboratory experiments do not expose the surface to polarised radiation and some authors have suggested that the polarisation of radiation due to the scatter

from dust particles must lead to organic synthesis within a polarised radiation field. Does this address the homochirality question by favouring the formation of one enantiomer over another? Possibly – more of this in Chapter 8.

5.10 Chemical models of molecular clouds

The complete chemical modelling of a cloud in the ISM looks tractable, if complicated, although there will always be assumptions, some of which will be rather gross, such as surface reaction rates, but they can all be refined in time. Molecular cloud modelling is at the frontier of astrochemistry and it would be hard to consider all of the results. The basic process of generating a kinetic model for a molecular cloud must follow the pathway:

- Chemical composition – determined from experimental observations providing reliable column densities for species in the cloud, including electron densities.
- Physical conditions within the cloud, including the temperature, the number density, the extinction (to give an estimate of dust composition) and electron temperature.
- Transport processes – diffusion, collisions and shock fronts transiting the cloud. Magnetic turbulence must be considered for mixing molecules.
- Photochemistry – an estimate of the radiation field from the formation of stars or existing stars within the cloud. A radiation field will require all photochemical-processing reactions to be included.
- Rates of reactions, where known, and their temperature dependence, otherwise estimates of the rate constants must be part of the assumptions.
- Decide on a network of chemical reactions: gas-phase reactions; surface of dust grains; photochemical processing.
- Identify target species that will represent the nature of the chemical populations within the cloud for comparison with observations.
- Set up the coupled differential equations and propagate the model in time, stopping to interrogate the concentrations of the target species.

Cloud properties

The lifetime of molecular clouds has been the subject of much recent research, both observational and theoretical. The surprising result is that molecular clouds

have shorter lifetimes than initially first thought. A typical GMC in which star formation is occurring is well defined:

- Mass of $\sim 5 \times 10^5 M_{\text{Sun}}$
- Radius of 120 ly
- Total hydrogen density $n(\text{H}_2) \sim 80 \text{ cm}^{-3}$
- Dynamic timescale $t_{cl} \sim 4 \times 10^6 \text{ yr}$

If we take the total observed mass in the Universe as $2 \times 10^9 M_{\text{Sun}}$ and divide this by the dynamic timescale for a GMC, this suggests that star formation is occurring at a rate of $\sim 500 M_{\text{Sun}} \text{ yr}^{-1}$, which is 100 times the currently observed formation rate. This calculation is riddled with assumptions and approximations, including the efficiency of star formation.

The lifetime of the molecular cloud is considered to be a time line running from cloud formation, star evolution and finally dispersion in a period that is several t_{cl} . The chemistry of the TMC and, to a good approximation, all molecular clouds must then be propagated over a timescale of at most 20 million years. The model must then investigate the chemistry as a function of the age of the cloud, opening the possibility of early-time chemistry and hence species present in the cloud being diagnostic of the age of the cloud. The model should then expect to produce an estimated lifetime and the appropriate column densities for the known species in the cloud. For TMC-1 the species list and concentrations are shown in Table 5.4.

Chemical networks

The choice of chemical networks is complicated and even for simple clouds such as TMC the species list is 218 species, with 2747 chemical reactions linking them. Network reduction mechanisms have been employed to reduce the number of reactions but preserve the chemical composition of at least the major species. All models must include simple ion–molecule chemistry with UV and cosmic ray ionisation initiation reactions, as shown in Figure 5.20.

Binary ion–molecule reactions are indicated by thin arrows (c.t. indicates charge transfer), the radiative association reaction of C^+ with H_2 is indicated by the thick arrow and the dissociative recombination reactions are indicated by dashed arrows leading to the neutral molecules inside the compound brackets (e^- indicates free electrons). The molecules indicated in bold are known (observed) interstellar molecules.

Gas-phase pathways to a number of bio-organically interesting molecules are possible, especially the carbon–nitrogen-containing species, as shown in Figures 5.21 and 5.22. The complexity of the network increases in denser GMCs such as in the Orion nebula and here the diversity of reactions leads to the formation of more organic compounds, many of which are more relevant to the origins of life.

Table 5.4 Relative abundances for species in TMC-1. (Reproduced by permission of Smith, 2004)

Species	Fractional abundance $N(x)/2N(\text{H}_2)$
H_2	0.5
CH	$\sim 10^{-8}$
CO	3×10^{-5}
CN	$\gtrapprox 4 \times 10^{-9}$
CS	—
OH	$\sim 3 \times 10^{-8}$
SO	2.5×10^{-9}
HCN	6×10^{-9}
HNC	4×10^{-9}
C_2H	4.2×10^{-9}
HCO^+	$\sim 8 \times 10^{-9}$
N_2H^+	—
NH_3	$\sim 5 \times 10^{-8}$
H_2CO	$\sim 6 \times 10^{-9}$
HNCO	$\sim 9 \times 10^{-11}$
C_3N	$\sim 4 \times 10^{-9}$
HC_3N	$(0.3-1) \times 10^{-8}$
HC_5N	5×10^{-9}
HC_7N	1.25×10^{-9}
HC_9N	1.5×10^{-10}
CH_3CCH	$(2-4) \times 10^{-9}$
C_4H	$(0.8-1.5) \times 10^{-8}$
$\text{CH}_3\text{C}_3\text{N}$	2.5×10^{-10}
$\text{CH}_3\text{C}_4\text{H}$	$\sim 10^{-8}$
HCS^+	—
CH_3CN	3.4×10^{-10}
CH_2CHCN	1.5×10^{-10}
SO_2	<

The molecular networks are formulated at a mechanism level using coupled differential equations such as those used in the demonstration of the steady-state approximation (Equations 5.17 and 5.18). The list of coupled differential equations is then propagated numerically in time using one of many standard numerical recipes for the representation of the derivative. The temperature of the cloud, the relative abundances and the extinction A_v for the cloud are assumed and the model is run. The output from the program is a series of relative concentrations for the cloud, which is then compared with the observed concentrations. The dependence of the composition on the initial starting conditions is also investigated, along with the lifetime of the cloud. There has been some success in predicting the concentrations but the fundamental properties of dust grain surface chemistry restrict the confidence in these models.

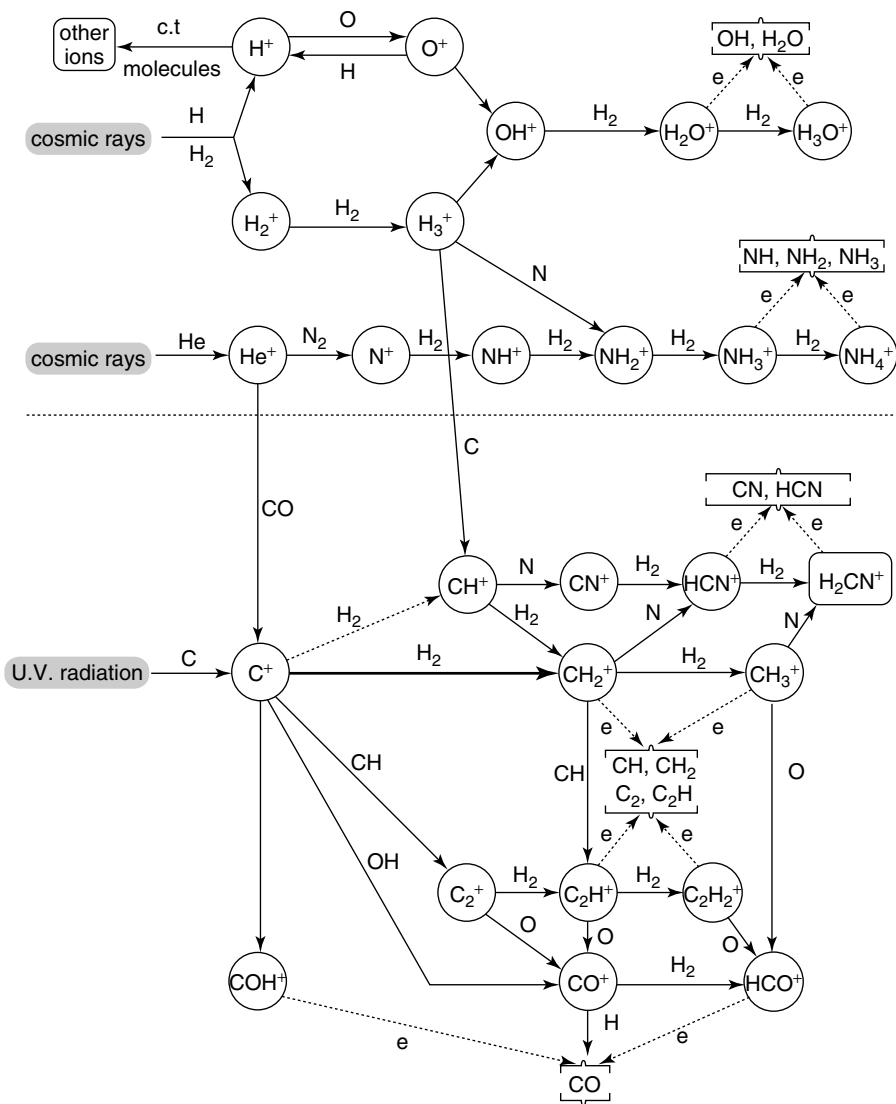


Figure 5.20 Simple molecule network for the beginning of TMC modelling. (Reproduced from Smith and Spanel, 1995, with permission from John Wiley Inc.)

There are some advantages of the temporal models of cloud chemistry associated with the concentrations of molecules at different times. Can we learn about the age of the cloud by its chemical composition or the age of an embedded star by the chemistry observed towards the object? Can the molecular environment be understood from the inventory of chemicals? Are there chemical diagnostics for planetary formation, star formation or even black holes? All of these questions are at the frontier of *Astrochemistry*.

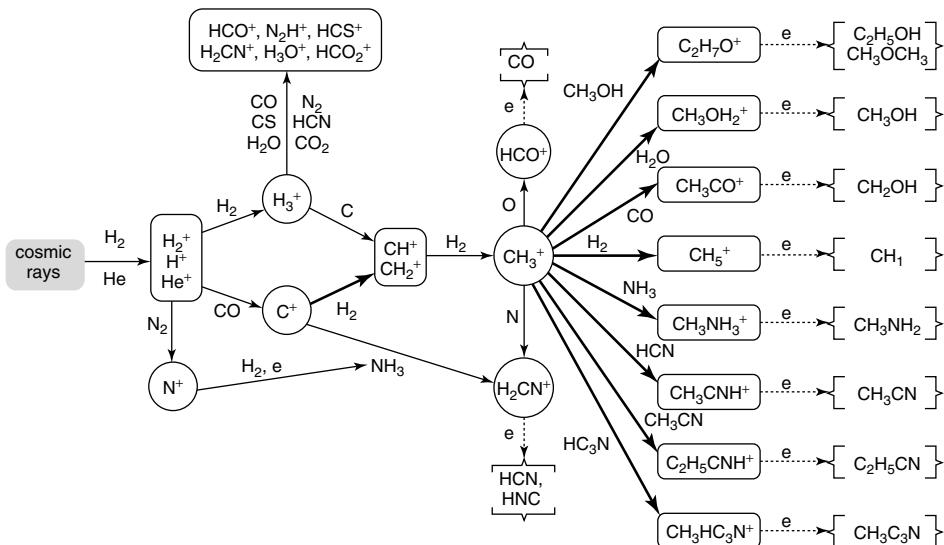


Figure 5.21 Initial reactions in dense molecular clouds. (Adapted with permission of John Wiley Inc.)

Deuterium enrichment

The quantum effect of zero-point energy has an interesting consequence in the ISM associated with the slight increase in bond strength of chemical bonds in deuterated species compared with non-deuterated species. The zero-point energy is associated with the fundamental frequency of vibration of a chemical bond and depends on the force constant and the reduced mass of the particles vibrating, as we saw in Equation 3.30. In the harmonic oscillator approximation, the energy of the lowest energy level, the zero-point energy, is given by $E(0) = \frac{1}{2}h\nu_e$, where $\nu_e = \sqrt{(k/\mu)}$, k is the force constant and μ is the reduced mass. The largest shifts in reduced mass occur between $H_2(\mu = 2)$ and $D_2(\mu = 1)$, hence the bond energies of deuterium-containing bonds are stronger: $D_0(H_2) = 432.1 \text{ kJ mol}^{-1}$ and $D_0(D_2) = 439.6 \text{ kJ mol}^{-1}$. This has important effects for the chemistry of deuterated species at a fixed temperature. In the body the enzyme chemistry is optimised for 37°C and the enzymes can break bonds at this temperature. Deuterium poisoning by $D_2\text{O}$ causes D atoms to be inserted into a number of important chemical reactions and the zero-point energy is reduced, increasing the bond strength. Chemical reactions breaking the bonds have a larger energy requirement and hence higher activation energy and a slower rate constant. The result is quantum poisoning of the enzymes. The same happens in the ISM.

The important reaction in deuterium chemistry in the ISM is the reaction between H_3^+ and HD , where HD is the primary source of D in the ISM:



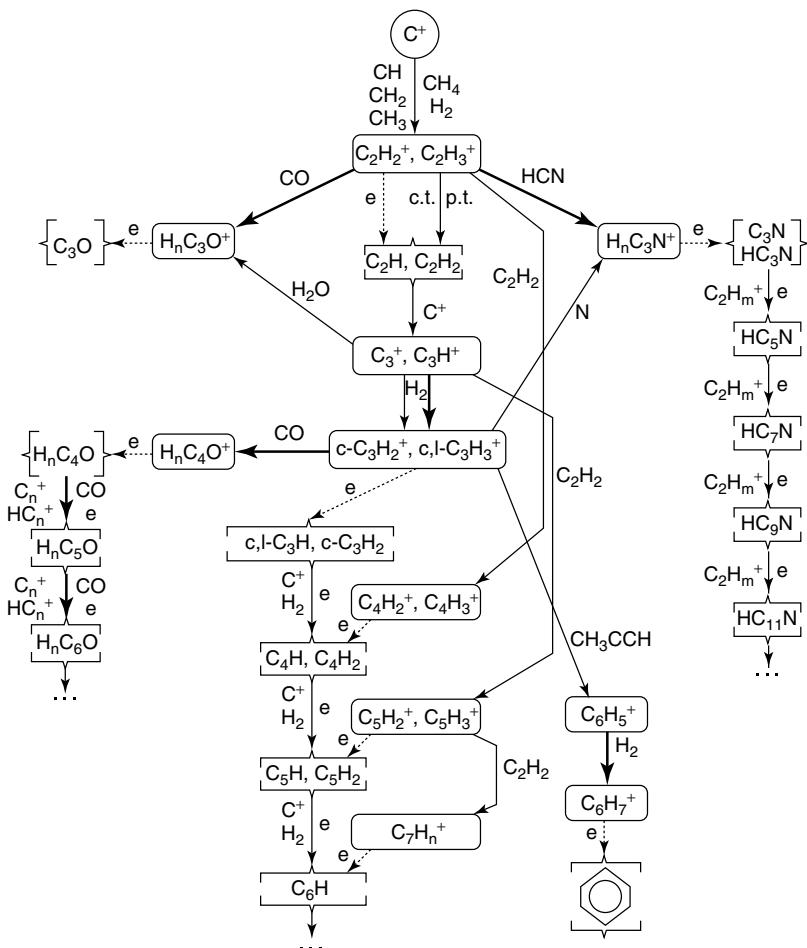
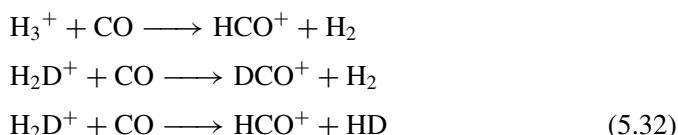


Figure 5.22 Organic synthesis in hot molecular cores. (Adapted with permission of John Wiley Inc.)

Once the H_2D^+ ionic species has been formed it is free to take part in fast ion–molecule reactions such as with CO in the following series of reactions:



The formation of HCO^+ and DCO^+ must be in the statistical ratio of 3:1, so the observed $\text{HCO}^+/\text{DCO}^+$ ratio in an astrochemical environment is a measure of the deuterium chemistry. All of the reverse reactions are possible in the reaction scheme but at 10 K, for example, the increased bond energy of DCO^+ prevents the reverse reaction from occurring and the DCO^+ becomes enriched. The H/D ratio is

routinely measured in species such as $\text{DCO}^+/\text{HCO}^+$ in places such as comets and GMCs and the ratios are important markers of the temperature and connectivity of the reaction networks.

5.11 Prebiotic molecules in the interstellar medium

One of the problems in looking for prebiotic molecules in the ISM is the instinctive prejudice towards their terrestrial counterparts. To draw this distinction we could usefully define prebiological molecules, biology being restricted to the study of organisms of Earth. Prebiotic could then be reserved for all pre-astrobiology. The search for prebiology has lead to the search and successful discovery of carboxylic acids and glycine in the ISM. Glycine has a number of isomers as shown in Figure 5.23, all of which may be synthesised in the ISM. The spectrum of glycine has been measured in the laboratory and simulated as in Figure 3.10, and there can be some confidence in the microwave spectrum of this species.

Searches in the 216 and 240 GHz regions of the spectrum should be rich in absorption features associated with each isomer, many of which will be overlapping, depending on the particular ISM molecular environment. Observations towards a number of objects have been made and are shown in Figure 5.24, where vertical lines show the positions of the glycine transitions. These observations and others convince astronomers that glycine has been observed towards these species. But is this important?

The list of molecules found in the ISM shown in Table 5.1 contains many carbon-containing species, which is not unreasonable given the nature of stellar evolution. The death of a low-mass star should produce C and O as products of the end of the nucleosynthesis process, the triple-alpha process. The death of the star also produces stellar winds ejecting C and O and it would appear that 12 of the other elements have been observed so far in the ISM. Among these elemental species only carbon shows significant polymerisation or chain chemistry that will survive in the ISM as either ion–molecule reactions or radical–radical reactions. This should lead to a subset of molecules that may be formed from the 12 elements in the ISM – CHON (carbon, hydrogen, oxygen, nitrogen) chemistry seems to be inevitable.

The search for other molecules such as pyrimidine and carboxylic acids is now underway to see how large possible prebiology molecules may become, but most interesting is the role of the dust grain surface. Our understanding develops apace for dust grain chemistry but it is clear that in the cold ISM, the chemistry on the surface of these particles is rich. The inventory of dust grain molecules is diverse and when stirred by UV photoprocessing CHON chemistry proliferates. The amino acid cocktail seen in laboratory experiments shown in Figure 5.19 suggests that organic chemistry becomes advanced within the grain ice mantles.

The cocktail of molecules, however, depends on the composition and condition of the cloud and the allowed chemistry on the surface or within the ice mantle. A complex network of chemical reactions exists, taking in energy from the

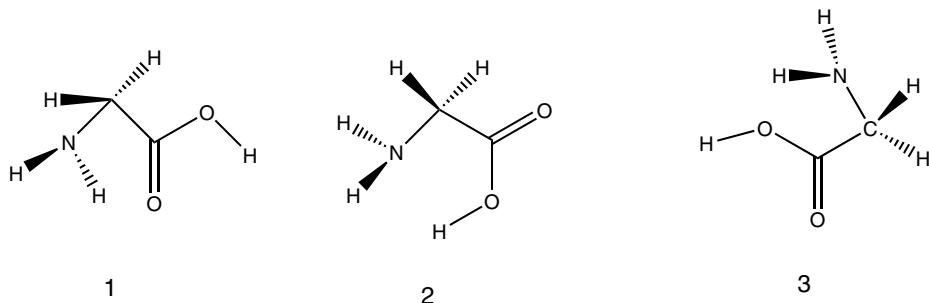


Figure 5.23 Isomeric structure of glycine

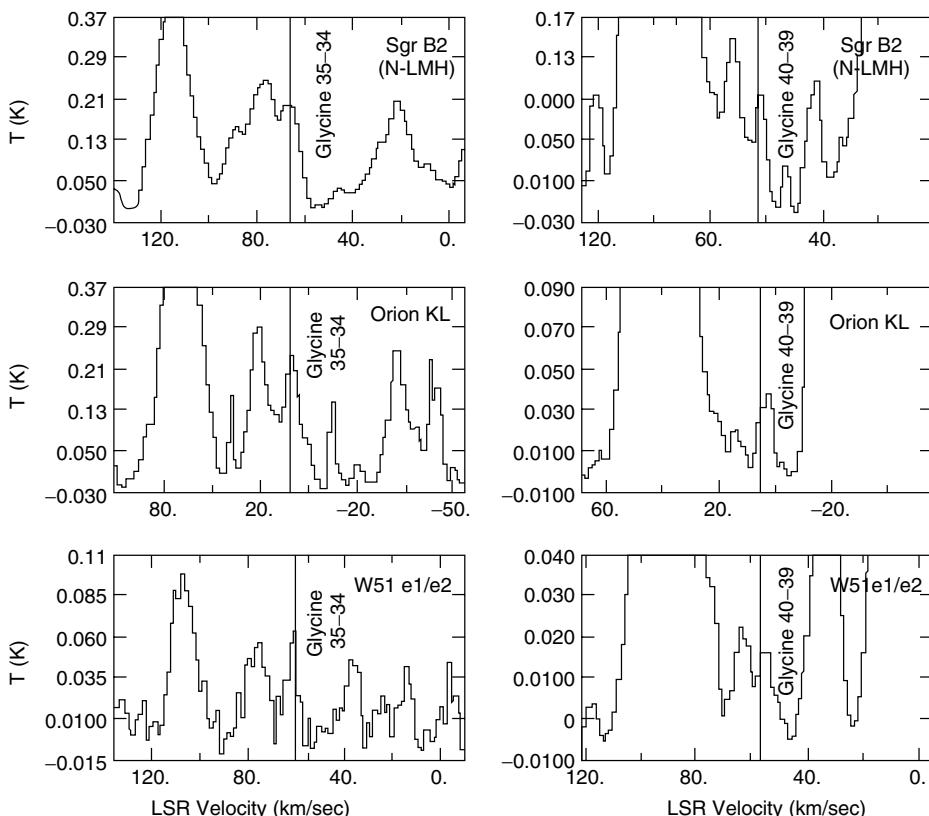


Figure 5.24 Glycine spectra (at 1 MHz resolution) observed at 206,468 MHz (line 21) in (a) Sgr B2(N-LMH), (b) Orion KL, and (c) W51 e1/e2 and at 240,899 MHz (line 26) in (d) Sgr B2(N-LMH), (e) Orion KL, and (f) W51 e1/e2. The abscissae give the rest frequencies based on the assumed LSR velocities: 64.0, 8.0, and 60.0 km s^{-1} for Sgr B2(N-LMH), Orion KL, and W51 e1/e2, respectively. The thick vertical lines mark the glycine lines at the frequencies corresponding to the measured velocities. The thin vertical lines denote the spectral features of other molecular species; question marks indicate uncertain identifications because of overly strong line intensities or inconsistent LSR velocities. (Reproduced from the *Astrophysical Journal* 593:848-867, 2003 by permission of the AAS)

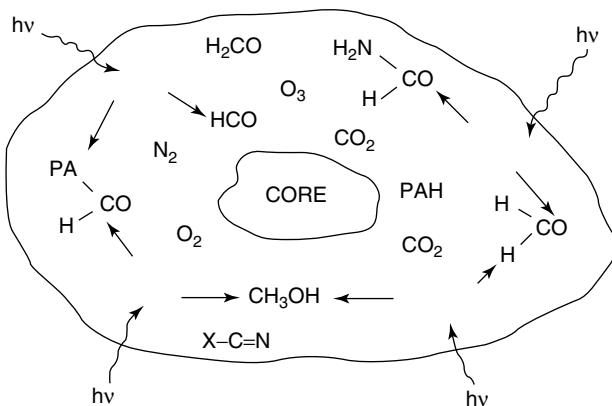


Figure 5.25 A CHON seed (Adapted with permission from Sandford, 1996)

cloud in the form of UV and thermal radiation, producing perhaps a ‘CHON seed’ (Figure 5.25) that could be dropped into the atmosphere or oceans of a proto-planet, seeding CHON chemistry in the planet. This addresses the proposal of exogenous organic synthesis and panspermia. The organic synthesis on the surface of the CHON seeds may be the seeds required for panspermia, which when planted on the surface of the host planet become life. Dropping the seed into water produces prebiological chemistry leading to life on Earth, whereas dropping the seed into an alkane such as on Titan would lead to a prebiotic chemistry associated with a different life form.

Concepts and calculations

Concepts

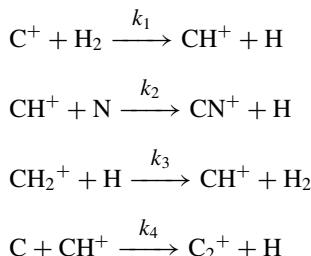
<i>Conditions in the ISM</i>	The variation of particle density from the tenuous ISM with 1 molecule cm ⁻³ to 10 ⁶ cm ⁻³ in GMCs, influenced by the radiation fields around stars
<i>Visible extinction</i>	The removal of light from a telescopic line of sight by either molecular absorption or scatter principally by dust
<i>Rates of reaction</i>	Rate constants and their temperature dependence leading to a differential rate equation
<i>Steady-state approximation</i>	The balance of chemical reactions leading to an apparent constant concentration of a species said to be in steady state
<i>Molecular processing</i>	The interaction between molecules in the ISM and the radiation and cosmic rays from a star leading to new ‘processed’ species
<i>Dust</i>	The role of dust grain surfaces in inducing different chemical processes enriching the ISM molecular inventory
<i>Chemical network</i>	A complex sequence of chemical reactions forming a network that controls the type and density of species that are formed and hence may be observed
<i>Biosynthesis</i>	Some of the molecular processing in the ISM leads to astrobiologically interesting molecules that might seed life

Calculations

<i>Visible extinction</i>	Estimate of the extinction along a line of sight due to dust, leading to the idea of reddened stars
<i>Rate of chemical reaction</i>	Calculation of the rate of chemical reaction when given the form of the constants
<i>Rate constant temperature dependence</i>	Calculation of rate constants at different temperatures, including collision numbers and concentrations of species in steady state
<i>Processing threshold</i>	Calculation of the rate of photodissociation and cosmic ray-induced molecular processing from photon and particle fluxes

Problems

- 5.1** Calculate the diameter of a cloud in the interstellar medium with a density of 10^2 atoms cm^{-3} if it is to collapse into a star with 1 solar mass. Assume that the cloud is spherical.
- 5.2** Repeat the calculation above for a giant molecular cloud with a density of 10^6 atoms cm^{-3} .
- 5.3** Calculate the collision number Z_{AB} ($\text{m}^{-3} \text{s}^{-1}$) for the collision between two hydrogen molecules and the collisions per molecule for:
- Diffuse ISM density 10^2 cm^{-3} and temperature 10 K.
 - GMC with a density of 10^6 cm^{-3} and a temperature 50 K.
- Assume that the collision cross-section for H_2 is 0.27 nm^{-2} .
- 5.4** The destruction of SiO molecules around a star is being monitored by the SiO maser transition. If the bond energy of SiO is $797.0 \text{ kJ mol}^{-1}$, calculate what surface temperature of a star that would be required to have a λ_{\max} sufficient to break the SiO bond.
- 5.5** Using the UMIST Astrochemistry Database format for rate constants, calculate the rate constants at 20 and 100 K for the following reactions, giving the units for the rate constants in each case:
- $\text{H}_3^+ + \text{CO} \rightarrow \text{HCO}^+ + \text{H}_2$, with $\alpha = 1.7 \times 10^{-9}$, $\beta = 0.0$ and $\gamma = 0.0$.
 - $\text{OH} + \text{CO} \rightarrow \text{CO}_2 + \text{H}$, with $\alpha = 1.17 \times 10^{-19}$, $\beta = 0.95$ and $\gamma = -74.0$.
 - Comment on the relative likelihood of neutral and ion–molecule chemistry.
- 5.6** The following is a simple reaction scheme involving CH^+ :



- (a) Derive an expression for the steady-state concentration of CH^+ in this network.
 (b) The reaction rates have the parameters in the table below. Calculate each of the rate constants at 20 and 100 K.

Rate constant	α	β	γ
k_1	1.0×10^{-10}	0	4640
k_2	1.9×10^{-10}	0	0
k_3	1.4×10^{-9}	0	0
k_4	1.2×10^{-9}	0	0

- (c) If the $[\text{H}]$ density is 10^6 cm^{-3} and $[\text{H}_2] = 0.5 [\text{H}]$, $[\text{CH}_2^+] = 10^{-3} [\text{H}]$, $[\text{C}^+] = 10^{-5} [\text{H}]$ and $[\text{C}] = 10^2 [\text{C}^+] = [\text{N}]$, calculate the steady-state concentration of CH^+ at both temperatures.

6

Meteorite and comet chemistry

Introduction

Hot cores in giant molecular clouds collapse to form young stellar objects with their own cloud or nebula around them. In the same way, our own solar system formed around the protostar that became our Sun and all of the planets from the solar nebula, which was nothing more than an average collection of dust and matter that had achieved critical mass. The collapse may have been triggered by a shock wave travelling through space, perhaps from a supernova, accelerating the rate of collapse so that above a critical mass the collapse was inevitable. The composition of the solar nebula was fundamentally responsible for the composition of the Sun and, with some processing, the chemistry of the planets. The physical conditions, especially temperature, during the formation of the early Sun determined the phase of matter on the planets and set the scene for the formation of the Earth and the evolution of life.

Meteorites and comets represent two types of astronomical bodies that are thought to have survived from the very beginnings of the solar nebula some 4.5 billion years ago. Meteorites land on the Earth and provide a unique chance to study the early solar system and are subject to the most intense laboratory examinations, whereas comets come close to the Sun, becoming very bright, and spectroscopy allows careful study of their chemical composition. Both objects carry secrets about the solar system and point to the conditions of planetary formation. Several missions, including the Giotto visit to Comet Halley, have flown close to comets to inspect the long tails, performing direct measurements and observations from within the tail. The ultimate comet exploration mission is pending in the Rosetta mission, which is planning to land on the surface of a comet and perform *in situ* measurements from the surface of the comet's nucleus. We will study these unique solar system species in some detail to unlock their secrets.

There are three definitions that pertain to meteorites:

- Meteor – an object entering the Earth's atmosphere that burns up completely during its passage through the upper atmosphere – a ‘shooting star’.

- Meteorite – an object entering the Earth’s atmosphere that survives the fall to the ground and is collected.
- Meteoroid – the name given to a meteorite or a meteor before it enters the Earth’s (or any other planetary) atmosphere.

Meteorites are essentially aggregated dust particles drawn from the inventory of matter in a giant molecular cloud, including initially the ice mantle of processed organic molecules: they are fossils of the solar nebula. Similarly, comets have an ancient origin in the debris of the solar nebula and are knocked into highly eccentric orbits around the Sun, returning at well-determined intervals. They are essentially composed of dust, rock and ice, which evaporate as it gets closer to the Sun, reaching maximum intensity at the point of closest approach, called perihelion. The origin, structure and chemistry of both meteorites and comets require an understanding of the formation of the solar system.

Meteors form when meteoroids enter the atmosphere. The mean orbital velocity of the Earth around the Sun is 29.7 km s^{-1} , so a particle stationary in space will enter the Earth’s atmosphere at this velocity. Gravity will further attract the particles up to a maximum of 11 km s^{-1} (the escape velocity) to make an entry velocity of order 40 km s^{-1} plus whatever additional velocity the particle may possess. This energy is converted into heat by the drag on the atmosphere, producing temperatures of order 2000 K – the glow seen from a shooting star. The final temperature depends on the properties of the particle, the atmospheric composition and the angle of incidence with the atmosphere.

Radar measurements of the Leonid meteor shower have measured entry velocities of order 70 km s^{-1} . Routinely, the Earth passes through the debris left in space by periodic comets to produce beautiful meteor showers such as the Leonid (Figure 6.1), with observation rates as high as 100 meteors per hour.

6.1 Formation of the solar system

The structure of the solar system is, theory suggests, not unique and is likely to form around any Sun-like star. Of course we have yet to observe the entire planetary structure around a local star, although we have now identified over 100 extra solar planets. A $1 M_{\text{Sun}}$ nebula is thought to undergo the sequence of events starting with a period quietly processing molecules in the molecular cloud, with chemistry controlled by the local temperature. The quiescent period may last 10^6 years but, once triggered, the cloud begins to collapse following the sequence shown in Figure 6.2.

At the centre of the cloud is the young stellar object destined to become the Sun. It accounts for approximately 99.9 per cent of the mass of the nebula and there are various examples of this in the heavens, including the classic pre-main sequence T-Tauri star. The star continues to evolve, blowing off bipolar jets (see Figure 4.5) and beginning a solar wind of particles. Of course, the star does not reach its full luminous intensity and the best theories suggest that the Sun was some 30 per cent less luminous when the Earth began to form.

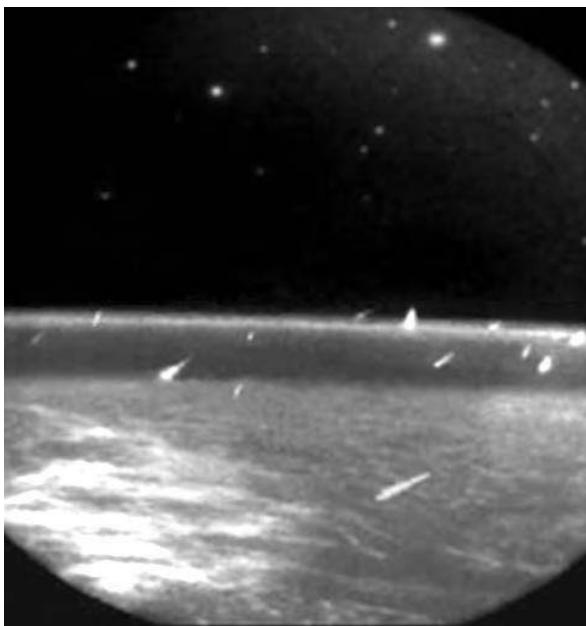


Figure 6.1 Leonid meteor shower. A composite of images taken by the Midcourse Space experiment satellite while in orbit around Earth. The satellite's perspective continuously changed relative to the path of the meteors, which caused their trajectories to be non-parallel. (Reproduced by permission of Jenniskens P., Nugent D., Tedesco E. and Murthy J. 2000; 1997 Leonid shower from space, *Earth, Moon and Planets* **82/83**: 305)

The collapsing nebula has some angular momentum that must be conserved during the collapse ((B) in Figure 6.2). The angular momentum is conserved by the formation of the disc and importantly there is a temperature profile, radially outwards from the proto-sun. The temperature at the Earth–Sun distance of 1 AU is thought to have been 200–800 K, falling to 100–400 K at 2.5 AU: comparatively warm out to the orbit of Jupiter. The heating comes from the dissipation of the gravitational energy, converting it to heat as the collapse proceeds. The disc temperature then depends on the rate of accretion and models allowing for a rapid collapse predict disc temperatures as high as 1200 K. At these temperatures all dust particles would begin to evaporate.

Tiny dust particles aggregate to form objects such as the asteroids and the planetesimals, and the accretion sequence is usually divided into three stages: aggregation of micron-sized dust particles into planetesimals that are around 1 m–10 km in diameter; growth of the largest planetesimals by gravitational attraction to form planets (planets form by accretion); and aggregation of the planetesimals by collision to form the terrestrial planets of Mercury, Venus, Earth and Mars. The rate of aggregation is critical in the early solar system: 1 km sized objects begin to feel the effects of gravity and will be attracted into the Sun, however, if the mass and energy are sufficient then the planetesimal goes into a stable orbit.

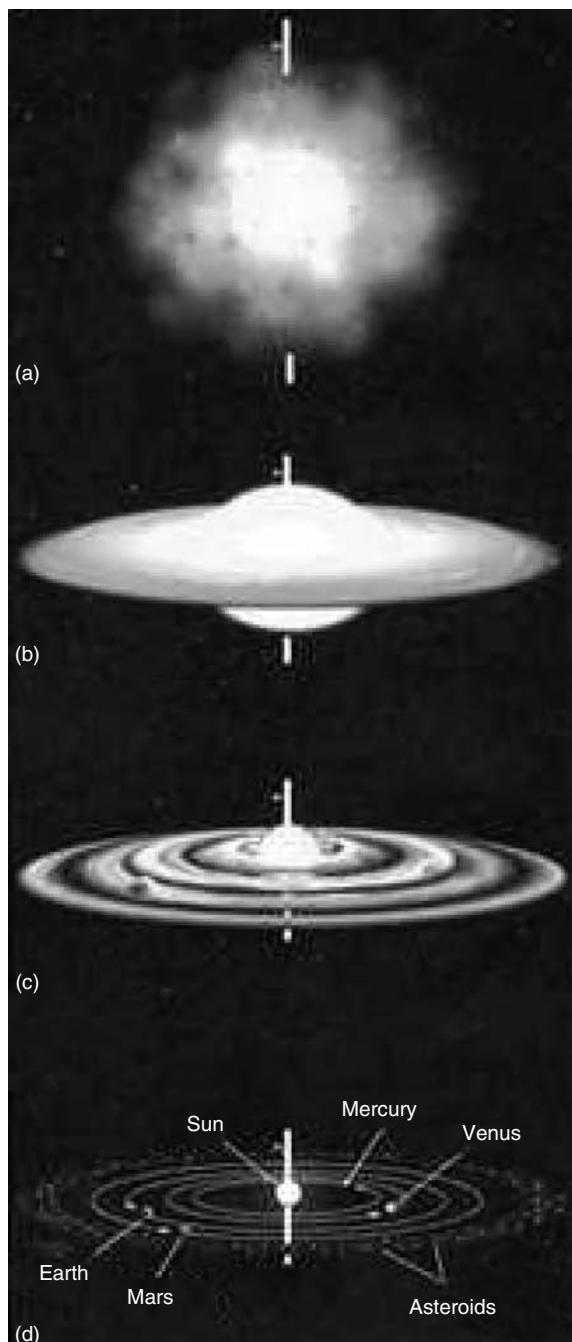


Figure 6.2 Formation of the solar system: (a) unstable molecular cloud possessing some angular momentum; (b) angular momentum conservation produces the disc shape under collapse; (c) matter accretion forms the planets; (d) the mature system of planets seen today evolves after 4 Myr

The volatile materials would have vaporised from the surface of the planetesimals once the temperature reached 160 K: below this temperature water sticks to silicate surfaces and condenses, ultimately freezing into ice. The new gaseous material is swept away from the planetesimals by the solar wind of particles, leaving bare planetesimals too small to acquire and maintain an atmosphere. The temperature gradient and location within the solar nebula are then important to the ultimate nature and composition of the planets themselves and interplanetary debris.

6.2 Classification of meteorites

Broadly speaking the classification of meteorites follows the geological mineral classification and with 275 mineral species reported so far this quickly becomes complex: some classes of meteorite have only one member. The mineral structure does convey essential information about the temperature at which the meteorite formed as well as the reduction–oxidation (redox) environment: was the environment in which it formed rich in oxygen? Meteorites have been classified into three broad classes:

1. Irons: composition principally of pure metallic nickel-iron. A large piece of unoxidised iron is very rare on Earth because it quickly oxidises to the iron ore.
2. Stony: principally silicates or rocky meteorites. It is harder to determine the extraterrestrial origin of these meteorites and it usually requires careful laboratory analysis.
3. Stony-iron: a mixture of the previous two classes.

The percentage fall for each class is shown in Table 6.1, where a ‘fall’ refers to an observed landing and a ‘find’ is a discovery after the landing. Curiously the largest find weight is of pure iron. Several meteorites that are positively huge in proportion are made simply of solid iron. The Willamette meteorite in the New York Museum of Science is a large chunk of what was once molten iron, practically without contamination. A surprising find from the Mars Rover mission is a pure iron meteorite, (see page 7 of the colour plate section) demonstrating that meteorite finds should be present on all planets. An obvious question, however, is why is it made of iron? As planets form they heat either by the accretion of gravitational energy or internally by radioactivity. The heating causes the rock of the growing planet

Table 6.1 Meteorite falls

Type	Fall (%)	Find (%)	Fall weight (kg)	Find weight (kg)
Stony	95.0	79.8	15 200	8300
Stony-iron	1.0	1.6	525	8600
Iron	4.0	18.6	27 000	435 000

to melt and the densest materials (Fe density 7874 kg m^{-3}) fall to the centre – the planet is said to be *differentiated*. Smallish planets become differentiated, including some large asteroids, and during a catastrophic early-solar system collision the small, differentiated planets are shattered, showering meteoroids of pure molten iron into space. The Martian meteorite must have once been part of a small planet that was differentiated and suffered a violent destructive collision. The collision debris landed on Mars and any other planet nearby.

The largest class of meteorite finds is stony meteorites, made principally of stone. The general stony classification is divided into three subclasses called chondrites, carbonaceous chondrites and achondrites, and it is at this level of distinction at which we will stop. Before looking at their mineral and isotopic structure in more detail, it is useful to hold the composition of the Earth's crust in mind here for comparison. The Earth's crust is 49 per cent oxygen, 26 per cent silicon, 7.5 per cent aluminium, 4.7 per cent iron, 3.4 per cent calcium, 2.6 per cent sodium, 2.4 per cent potassium and 1.9 per cent magnesium, which must have formed from the common origin of the solar system.

6.3 Meteorite mineralogy

Meteorite mineralogy (Table 6.2) is complicated and is diagnostic of the origin of the sample. Two subclasses of the stony meteorites are chosen of particular relevance to the study of the origins of life.

The chondrite subclass of stony meteorites is characterised by globules of once-molten material that quickly solidified from a temperature of 1600 K, called *chondrules*. These 1 mm particles are thought to result from the early proto-planetary disc and thus represent some of the oldest rock specimens from the origin for the Earth. They were probably molten before incorporation into a larger meteoroid and there is some evidence of re-melting. The composition is usually similar to that of the Sun and contains silicate minerals and various iron–silicate minerals, e.g. olivine ($(\text{Mg},\text{Fe})_2 \text{SiO}_4$). Chondrules (Figure 6.3) have been used as sources for the determination of isotope ratios that are thought to be indicative of the origin of the solar system. The chemical composition is assumed to be only weakly processed from that present in the proto-planetary disc. The chondrules are then samples of the material from the proto-planetary disc and would be good evidence for the composition of interstellar dust composition.

The chondrules could be a source for the interstellar molecules, processed in the ice layers of dust particles in the giant molecular clouds. The general problem of the precursors for chondrules is, however, not understood. Many theories abound and all of them have significant problems. However, the two basic types of chondrules (type I, essentially chondritic meteorite composition; type II, FeO-rich) suggest that solid carbon acted as a reducing agent for the FeO–silicate mixture. More interestingly for exogenous organic delivery to the Earth are the class of meteorites called the carbonaceous chondrites. These contain a layer of organic material,

Table 6.2 Meteorite mineralogy

Class	Subclass	Composition
Iron		Primarily iron and nickel and similar in composition to M-type asteroids: iron, 91%; nickel, 8.5%; cobalt, 0.6%. A recent find of an iron meteorite on the surface of Mars is shown on p. 7 of the colour plate section
Stony-iron	(a)	
Stony	(b)	
Chondrites	(c)	
Carbonaceous chondrites	(d)	Similar composition to chondrites but also contain organic compounds embedded in a <i>kerogen</i> -like substance

(continued overleaf)

Table 6.2 (*continued*)

Class	Subclass	Composition
Achondrites	(e)	Believed to derive from the Moon or Mars. Hard to identify

Source: Center for Meteorite Studies, Arizona State University

- (a) This meteorite was found in 1916 about 10 km north of Anaheim, discovered by Mr William Huiras who was mowing hay on his farm. It is believed to be related to a spectacular fireball observed some two years earlier. It is a crescent-shaped medium octahedrite, with a mass of nearly 12 kg and showing very clearly the dimpling or 'thumb-printing' of the surface that is often found on iron meteorites. The main mass is currently held in the Canadian National Meteorite Collection, Ottawa (Reproduced by permission of Johnston R.A.A. and Ellsworth H.V., 1921; *Transactions of the Royal Society of Canada* IV: 92.).
- (b) This L6 olivine-hypersthene chondrite fell in 1912 in Navajo County, Arizona, as part of a shower of stones estimated to be made up of as many as 14 000 individuals, ranging in mass from 6.6 kg downward and with a total mass of about 218 kg. It has a typical black fusion crust with well-marked flow lines. These furrows, which are formed by the atmospheric ablation of the meteorite, are deepest on the forward surface and radiate away from the apex. The incomplete nature of the fusion crust indicates that this sample is from a meteorite that broke apart as it struck the ground.
- (c) Mező-Madaras, Romania, ordinary chondrite. This close-up of a cut and polished face of Mező-Madaras measures ~7 cm from left to right. Chondrites are named for the nearly spherical, silicate-rich objects they contain called chondrules, which were among the first objects to have formed in our solar system. As is evident, Mező-Madaras has abundant large chondrules. (Reproduced from a photo by D. Ball, ASU).
- (d) Allende, Mexico, carbonaceous chondrite. This chondrite also contains chondrules (note the round cavity left by the removal of a large chondrule). In addition, Allende, like many carbonaceous chondrites, contains calcium–aluminium-rich inclusions (CAIs). Unlike chondrules, which are round and composed mostly of silicate minerals like olivine and pyroxene, CAIs are predominantly white to light grey in colour, irregularly shaped and rich in refractory (high-temperature) minerals like melilite and spinel. They are believed to pre-date chondrules by at least 2 million years. This specimen is ~11 cm from left to right. (Reproduced from a photo by D. Ball, ASU).
- (e) Sioux County, NE, eucrite. This achondrite specimen, a fall, has been broken open to reveal the stark distinction between its black, shiny and smooth fusion crust and its light-coloured interior. This specimen is ~7 cm from left to right. (Reproduced from a photo by J. Kurtzman).

generically called *kerogen*, between the chondrules. This is a rich source of organic compounds and will be discussed in detail later in the context of the Murchison meteorite. However, the temperature restrictions on the formation process, perhaps as high as 2100 K, suggests that if there were any organic species around the chondrule it must have turned to soot in the formation process. This does not explain the rich variety of organic species observed. Perhaps meteorites have a different role, namely transporting material between the planets once both the planet and the organic material have formed.

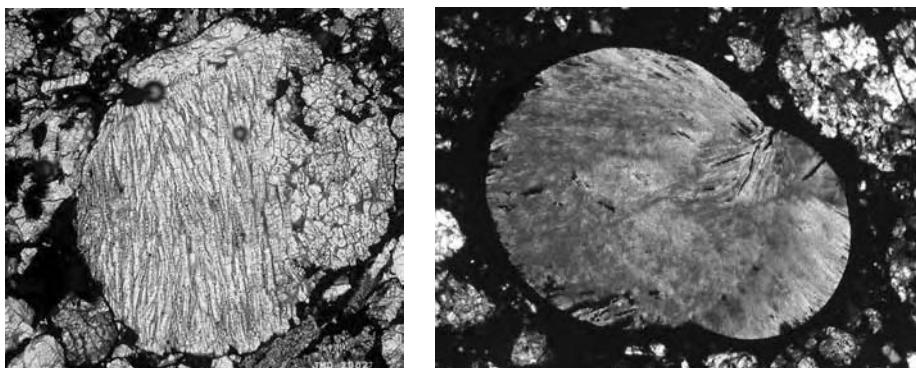


Figure 6.3 Chondrules. (A colour reproduction of this figure can be seen in the colour section). (Reproduced by permission of J. M. Derochette)

The idea of meteorite messengers between the planets has developed from careful observations of the chemical composition of several meteorites, the most extensive research being performed on ALH84001. There is compelling evidence that rocks can be ejected from the surface of asteroids or planets without complete melting, taking with them the chemicals present from the surface or local subsurface of the planet. Transit between the planets in this game of solar system billiards can take millions of years but some rocks come express-delivery in only 60 000 years. With a low-temperature ejection from a planet's surface containing bacterial spores able to survive the interplanetary voyage, seeding the solar system looks possible. Now, the analysis of carbon in meteorites may be the first look at ancient life on other planets.

Another feature of meteorites that proves to be important is the calcium–aluminium inclusions (CAIs), which, as the name suggests, show regions of enhanced Ca and Al. These micron- to centimetre-sized particles are some of the oldest objects known and have a similar temperature history. They probably formed at temperatures in the region 1700–2400 K and so are close to the centre line of the solar nebula. Although it is hard to be sure about the origin of these objects, there is agreement on their age based on radioisotope dating.

6.4 Geological time

Meteorites present an opportunity to look at geological time or the time told by radionucleotides within rocks. The oldest rocks found on Earth are not as old as the age of the Earth due to continual reprocessing of the Earth's surface. The oldest discovered rocks so far are the Acasta gneisses from Northwestern Canada, which are 4.03 Gyr, but these are young compared with the CAIs found in the Allende meteorite, which are 4.566 ± 0.002 Gyr or 4.556 billion years. The ages of these species are derived from the relative abundances of radioisotopes and their daughter species, as seen in Table 6.3.

Table 6.3 Isotopes and daughter products used in dating

Parent isotope	Stable daughter product	Currently accepted half-life values
Uranium-238	Lead-206	4.5 billion years
Uranium-235	Lead-207	704 million years
Thorium-232	Lead-208	14.0 billion years
Rubidium-87	Strontium-87	48.8 billion years
Potassium-40	Argon-40	1.25 billion years
Samarium-147	Neodymium-143	106 billion years

Radioactive decay is a nuclear process from an intrinsically unstable nucleus that emits alpha particles, beta particles and gamma rays. The loss of mass from the nucleus changes the element to one of a lower mass. Carbon dating uses the decay of the ^{14}C nucleus, a heavy and unstable isotope of carbon, to become the stable ^{14}N isotope. The overall process is written:



where the emission of the beta particle increases the number of protons in the nucleus by one. The stable ^{14}N nucleus is the daughter element to ^{14}C in this beta decay process. The rate at which this process occurs is constant and follows first-order kinetics so that the rate of decay of the isotope depends only on how many nuclei are present. This is written as:

$$\frac{dN}{dt} = -kN \quad (6.2)$$

where N is the number of isotope nuclei and k is the decay rate constant. The rate equation can be integrated to find the following expression for the number of nuclei at any particular time:

$$N = N_0 \exp(-kt) \quad (6.3)$$

where N_0 is the initial number of nuclei. The amount at any one time depends on how much there was to start with but the general decay process can be characterised by a half-life that is a measure of the time required for exactly half of the nuclei to decay. The half-life is derived from:

$$\begin{aligned} \frac{N_0}{2} &= N_0 \exp(-kt_{1/2}) \\ kt_{1/2} &= -\ln\left(\frac{1/2N_0}{N_0}\right) \\ t_{1/2} &= \frac{\ln 2}{k} \end{aligned} \quad (6.4)$$

The half-life is independent of the initial number of nuclei. For ^{14}C decay the half-life is 5717 years, whereas the ^{238}U decay half-life is 4.5 billion years. Carbon dating works well for timescales in the recent past and is used for dating objects such as the Turin shroud, but ^{238}U is better for timescales of the age of the solar system.

The decay process can be used as a radioactive clock by determining the ratio of parent nuclei to daughter nuclei within a sample and, assuming that there have been no other external sources of the daughter nuclei introduced into the geological sample, the age of the sample can be determined. There are two implicit assumptions hidden within this: the half-life is well known and for most nuclei these are good determinations, but how sure can we be that the half-life for ^{238}U is 4.5×10^9 years; and you have to be able to determine the elemental composition of the sample accurately. The latter requires some careful mass spectrometry analysis and this is in part responsible for the differences between the ages determined by the different radioisotopes – not all radioactive clocks tell the same time.

For the determination of events since the beginning of geological time the decay of ^{238}U is the most useful chronometer. The so-called lead–lead time comes from the measurement of the stable daughter isotopes of lead, both ^{210}Pb and ^{206}Pb . The full decay pathway with the method of decay for ^{238}U is shown in Figure 6.4. Similar decay of ^{235}U leads to the formation of ^{207}Pb . The determination of the

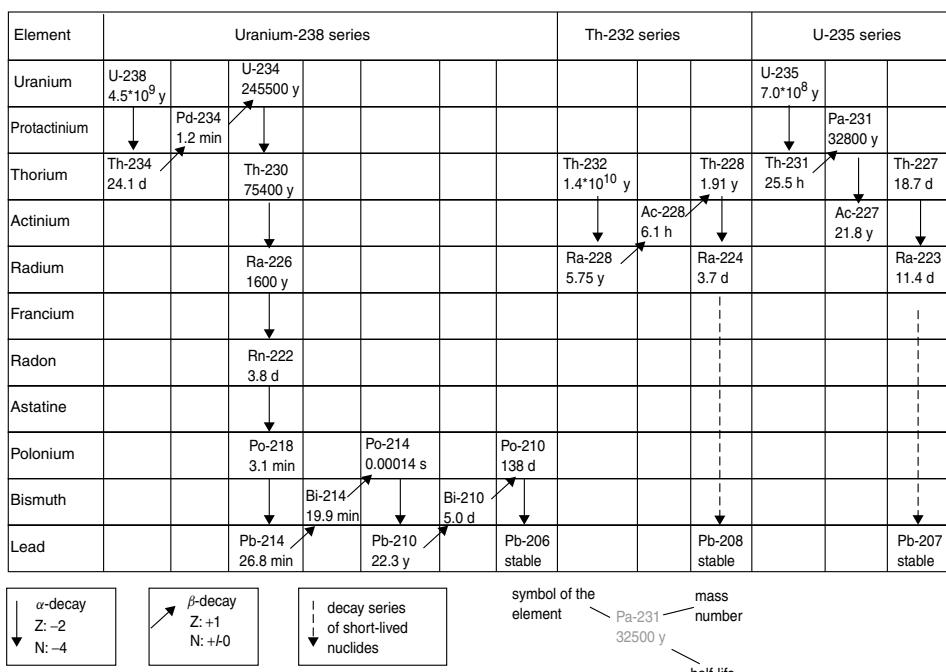


Figure 6.4 Decay pathways from ^{238}U to the two lead isotopes ^{210}Pb and ^{206}Pb

uranium/radon ratio can also be used and acts as a further check on the age of a particular sample.

Once the ratio of parent-to-daughter isotopes has been determined, the geological age of the sample is given by:

$$\begin{aligned} t &= \frac{1}{\lambda} \ln \left(1 + \frac{D}{P} \right) \\ \lambda &= \frac{\ln 2}{t_{1/2}} \end{aligned} \quad (6.5)$$

where t is the age of the rock, D is the number of atoms of the daughter isotope today, P is the number of atoms of the parent isotope today and λ is the decay constant, defined in Equation 6.5.

The Pb–Pb geological time determination from the two uranium isotopes has been used for some of the best radiological dating measurements of CAIs in the Allende meteorite, giving its age of formation as 4.566 ± 0.002 Ga – very close to the beginning of the collapse of the solar nebula. Other radioactive clocks can be used to determine when the CAIs were formed relative to the chondrules. The ^{27}Al isotope is radioactive and decays to the ^{26}Mg nucleus with a half-life of 0.73 million years. Comparison of the $^{26}\text{Mg}/^{27}\text{Mg}$ ratio in chondrules and CAIs leads to the conclusion that the CAIs are about 1.2 million years older than the chondrules.

For more recent dating processes carbon dating is used. With a half-life of 5717 years ^{14}C dating relies on the replenishment of the ^{14}C parent species, which occurs by neutron capture. The process occurs in the atmosphere as a result of bombardment of the Earth with neutrons:



This is the reverse of the process occurring in Equation 6.1, the decay of ^{14}C , and results in a nearly constant concentration of ^{14}C in the atmosphere. Any living organic synthesis on Earth, such as photosynthesis, will then capture the ^{14}C and produce a $^{12}\text{C}/^{14}\text{C}$ ratio in living things to be fixed. When a tree is used for wood in an object such as a museum artefact then the $^{12}\text{C}/^{14}\text{C}$ ratio changes and the age of the sample can be calculated using Equation 6.5 as before.

6.5 Chemical analysis of meteorites by $\mu\text{L}^2\text{MS}$

Chemical analysis of meteorites proceeds along classical analytical chemistry lines but with added precautions to prevent terrestrial contamination. Careful laboratory procedures have been developed, together with blank trials to enable the contamination in the laboratory to be eliminated. However, there is always the possibility of terrestrial contamination associated with the period of time on the ground before the ‘find’, in particular ice-melt water in the case of ALH84001. This meteorite

was found in the ice fields of Antarctica and had been there for some time allowing ice-melt water to soak into the meteorite and this has been proposed as the source for many of the prebiological molecules, as we shall see later.

Conventional techniques such as solvent extraction followed by gas chromatography mass spectrometry (GCMS), and acid digestions followed by GCMS have all been performed on meteorite samples but all information about the location within the sample is lost during a digestion processes. However, microprobe laser desorption studies allow the profiling of material within the meteorite sample.

Microprobe laser desorption laser ionisation mass spectrometry (μ L²MS) is used to provide spatial resolution and identification of organic molecules across a meteorite sample. Tracking the chemical composition across the surface of the meteorite requires a full mass spectrum to be measured every 10 μ m across the surface. The molecules must be desorbed from the surface with minimal disruption to their chemical structure to prevent fragmentation so that the mass spectrum consists principally of parent ions. Ideally, the conventional electron bombardment ionisation technique can be replaced with an ionisation that is selective to the carbonaceous species of interest to simplify the mass spectrum. Most information will be obtained if small samples are used so that sensitivity levels should be lower than attomolar (10^{-18} M); fewer than 1000 molecules can be detected and above all it must be certain that the molecules came from the sample and are not introduced by the instrument itself.

Designing a technique with these criteria in mind is not easy but is essential for a reliable analysis of the spatial distribution of carbonaceous material in meteorites. The devised technique is elegant and certainly worthy of study. Constituent neutral molecules of the sample are first desorbed with a pulsed infrared laser beam focused down to a spot, presently adjustable between 10 and 40 μ m. The laser power density is kept well below the plasma threshold to ensure desorption of neutral organic species with little or no fragmentation. The plasma threshold is the energy required to remove electrons from the neutral molecules, generating a population of electrons that can be accelerated and collide with other neutral molecules. These collisions produce further electrons, ultimately producing a mixture of species – ions, electrons and neutrals – that together are called a plasma. The plasma results in the fragmentation of desorbed molecules due to electron bombardment and hence fragmentation of the parent ions. This is avoided by careful choice of the laser power.

An ionisation laser is tuned to the selected class of molecules in the desorbed plume and preferentially ionises the selected molecules by a single-frequency pulsed ultraviolet (UV) laser beam that passes through the plume, as shown in Figure 6.5. The ionising laser is tuned at a specific electronic transition in the molecules of interest, namely the 280 nm transitions in aromatic molecules containing benzene rings, making the technique uniquely sensitive to the carbon-containing species. The resultant ionisation-selected ions are then extracted from the plume and pulled into a reflectron time-of-flight mass spectrometer. A reflectron is an electrostatic mirror that enables the length of the spectrometer to be doubled by sending

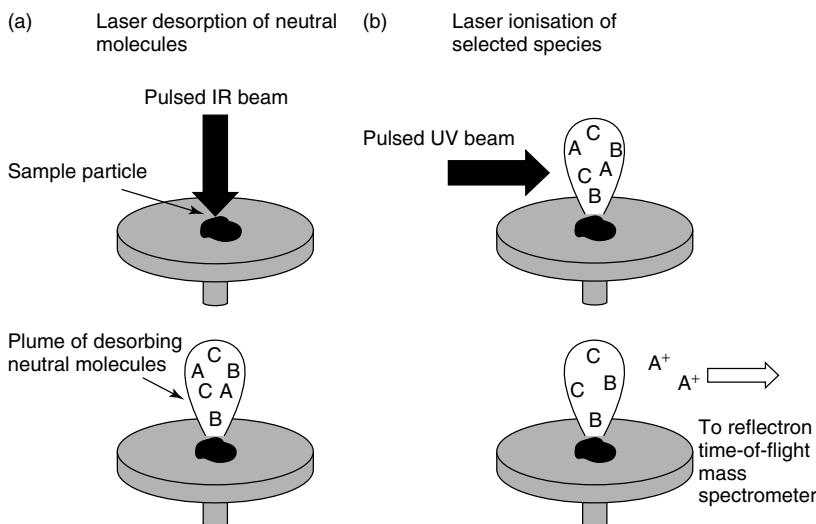


Figure 6.5 The $\mu\text{L}^2\text{MS}$ technique consists of (a) laser desorption followed by (b) laser ionisation to selectively produce ions, that are then extracted into a time-of-flight mass spectrometer for mass analysis. (Reproduced by permission of R.N. Zare, Stanford University)

the ions down a drift tube and then bouncing them back again. The increased length improves the mass resolution of the mass spectrometer.

Mass spectrometry is an electrostatic method of weighing ions. Bending charged species in a circle enables the mass-to-charge ratio of the species to be determined and not explicitly the mass. So ions with mass m and charge $1e$ arrive at the same time as ions of mass $2m$ and charge $2e$. The mass spectrometer in Figure 6.6 accelerates all charged species through a fixed potential difference from which they acquire a fixed kinetic energy of $\frac{1}{2} mv^2$, so that heavier ions move more slowly and lighter ions move faster. Measuring the arrival time of the ions at a detector enables the mass of the ion to be determined from the time-of-flight. The longer the path, the better the mass resolution, so in the design shown in Figure 6.6 a reflectron is used to send the ions down a 1-m flight path and back again. This flight time enables unit mass resolution to be achieved.

The desorption laser can be tracked across the surface of the meteorite so the plume of molecules can be associated directly with internal structures and morphology on the meteorite surface. The internal morphology of a meteorite is accessed by cleaving the sample in an ultraclean environment to prevent terrestrial contamination, followed by the laser desorption analysis. Nearly all meteorites have been studied in this way producing a mature field of research, however, the details of just two meteorites will be discussed in some detail. The Murchison and ALH84001 meteorites have provoked considerable interest, particularly the Antarctic meteorite ALH84001, which was the subject of a NASA announcement regarding life on Mars.

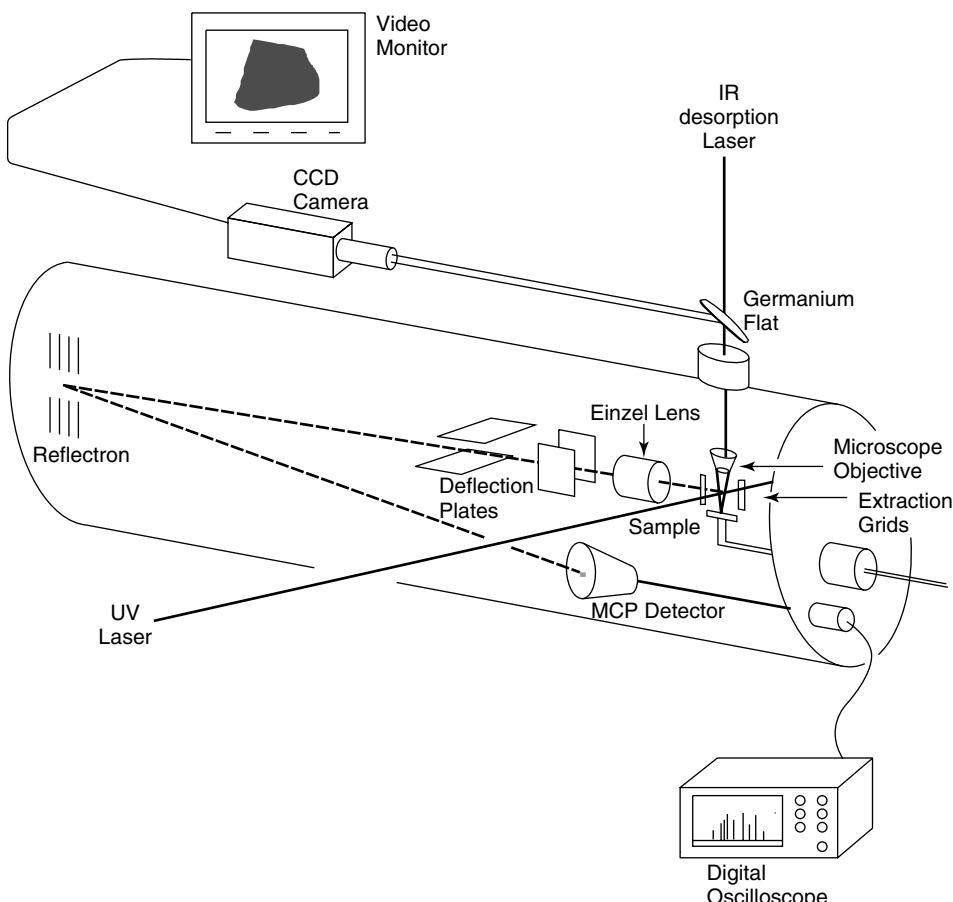


Figure 6.6 A time-of-flight mass spectrometer. The arrival time of the ions is proportional to the mass. (Reproduced by permission of R. N. Zare, Stanford University)

6.6 The Murchison meteorite – kerogen

The Murchison meteorite shown in Figure 6.7, like all meteorites, is named after the place from which it was recovered and in this case it is the town of Murchison, Victoria in Australia about 100 km north of Melbourne. The fall occurred in 1969 and was followed by an analysis of the chemical composition in some considerable detail. The Murchison meteorite is a carbonaceous chondrite containing about 2 per cent carbon, some as inorganic carbonates, and some as soluble compounds such as amino acids but the bulk as a macromolecular heterogeneous material referred to as *kerogen*.

The kerogen-like material from the Murchison meteorite is shown in Figure 6.8 and consists of a rich macromolecular carbonaceous material made from aromatic and aliphatic compounds observed in fluorescence following excitation at 280 nm.

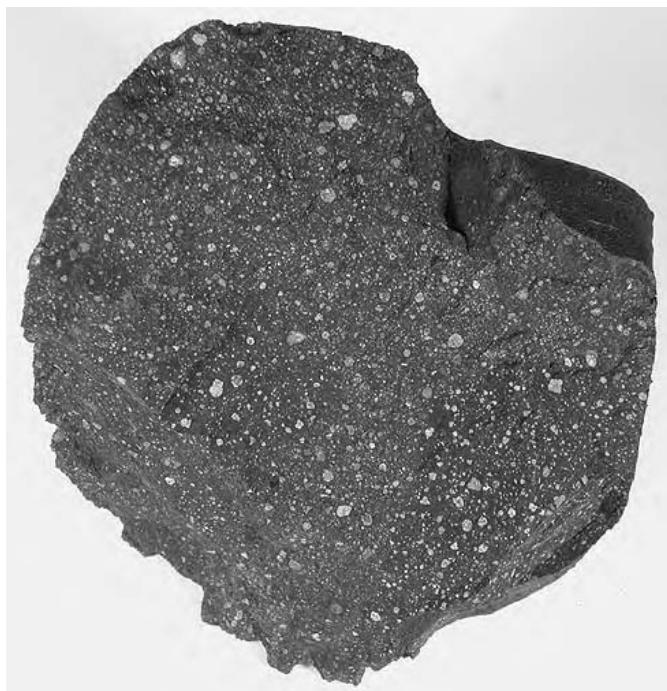


Figure 6.7 Murchison meteorite. (Reproduced by permission of Jim Strope)

Acid and base extractions from this material have been shown to form spontaneous structures in solution termed *coercevates* that could easily form the basis for prototypical membranes (more of this in Chapter 9). Hydrocarbons with chain lengths C15–C30 (both straight and branched chains) and of course PAHs, predominantly pyrene and fluoranthrene, polar hydrocarbons such as aromatic ketones, alkyl and aryl ketones, nitrogen and sulphur heterocycles and most intriguingly purine and pyrimidine analogues have all been observed from this rich carbonaceous cocktail of compounds. Why?

We must consider not just the composition of the meteorite but also what might happen to the chemicals once released onto the planetary surface. Hot aqueous extraction from the Murchison kerogen produces alcohols, ketones, mono- and dicarboxylic acids, amines and sulphonate and phosphonic acids – surely with such a suggestive mix of chemicals life is inevitable. Perhaps this is too good to be true. Reports from the early literature pre-1970 suggest that nitrogen heterocycles consistent with DNA and RNA were discovered. These discoveries have now been shown to be due to human contamination either after landing, during collection or during analysis. One counter here is the observation of an achiral mixture (1:1) of amino acids in the extraction, suggesting a non-biological origin with no preference for one isomer. This, however, does not ensure the extraterrestrial origin

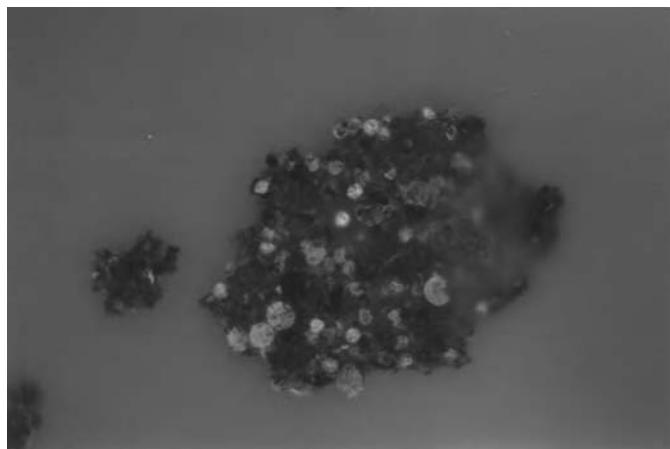


Figure 6.8 Kerogen. Moderately well preserved amorphous organic matter (AOM) with a rather dull matrix, but very common small (<10 µm in diameter), green-fluorescing micrhystridid acanthomorph acritarchs. The sample is from the deep sea, and it is probable that the AOM matrix was partially degraded within the water column but finally preserved in a dysoxic reducing environment. The sample lithology is a laminated black shale. The kerogen composition is Type III, in this case indicating partial oxidation and not terrestrial derivation. Sample RVT/DSDP534/125-4, 140–142 cm; Unit 7d, Callovian, DSDP Site 534, Blake Bahama Basin, western Central Atlantic. (Reproduced from slide prepared by Robertson Research, by courtesy of Les Riley from Tyson R. 1995, *Sedimentary Organic Matter: Organic Facies and Palynofacies*. Chapman & Hall: London)

of amino acids, as we shall see from the Strecker synthesis in prebiotic chemistry (Chapter 8).

The microprobe laser desorption studies reveal spectra such as that seen in Figure 6.9 showing a stegosaurus-like backbone of carbon compounds. The traces for the Murchison and Allende meteorites are compared with the biological deposit called fish clay found in Denmark. The fish clay is clearly a biological degradation deposit in the same way that oil is the product of biological photosynthetic products. If these deposits represent the detritus of life then should the same pattern of carbon molecules also be characteristic of life? Deep inside the meteorite sample there is no chance of decontamination from the period of time the meteorite was on the ground before the fall and hence the carbonaceous material must derive from the origin of the meteorite.

6.7 Meteorite ALH84001

Meteorite ALH84001 was found in the Allan Hills in Antarctica in 1984 and was the subject of an extraordinary NASA press announcement in 1996. It weighed 1.93 kg when it was collected (Figure 6.10) and is probably the most extensively studied lump of rock of all time. It is 95 per cent orthopyroxene, the volcanic rock mineral that accumulated in a molten-lava-reducing environment on Mars some

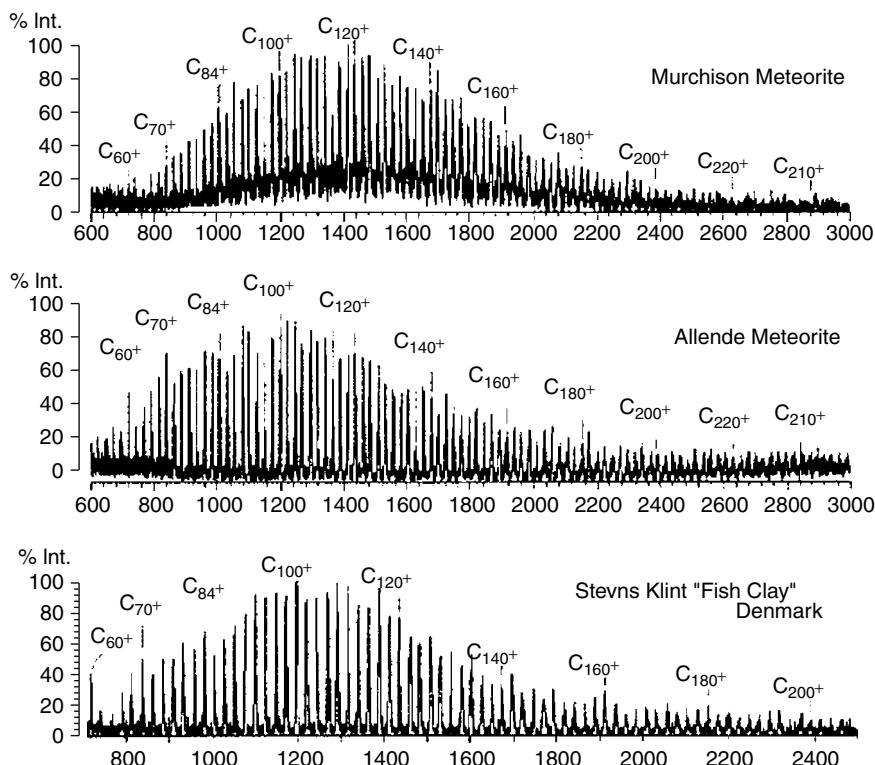


Figure 6.9 Laser desorption mass spectra of extracts of the Murchison and Allende carbonaceous chondrites and an extraction from the K/T boundary. (Adapted from Becker L, Poreda R. J. and Bunch T. E., 2000 by permission of the *Proceedings of the National Academy of Sciences* 97: 2979)

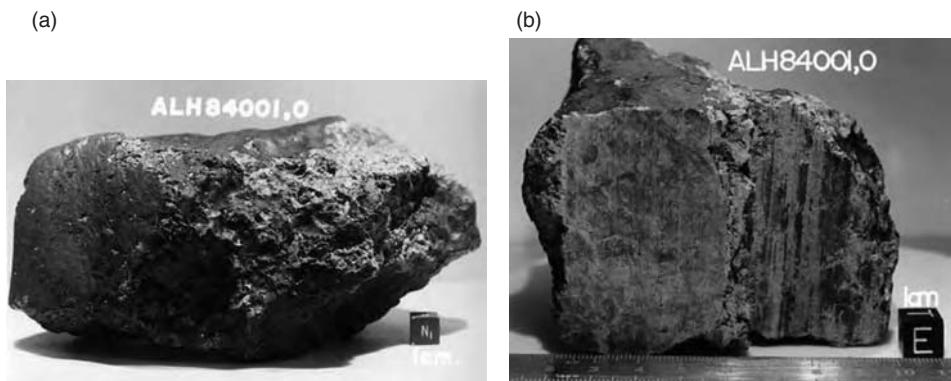


Figure 6.10 Meteorite ALH84001 – 1.93 kg of the most studied rock of all time: (a) as it was found with the outer edge showing a fusion crust; (b) after the sample was cut for analysis of the interior. (Reproduced from photos by courtesy of NASA)

4.5 Gyr ago. There is evidence that it was reheated as a result of a shock event about 4 Gyr ago and water flowed through it, depositing carbonate globules. It was ejected from the Martian surface by another impact event and orbited the Sun on its own for about 16 million years before intersecting the Earth's orbit and landing some 13 000 years ago, impacting into the ice of Antarctica where it was found.

A very definitive collection of information has been obtained to support the sequence of events just outlined but most striking is the level of confidence and precision in the details. Such was the enthusiasm and imagination regarding the information that during the press conference NASA announced that there was evidence for life on Mars based on data from ALH84001. We will look at some of the details of this extensive and controversial story.

Martian origin

The evidence for the Martian origin of ALH84001 is based primarily on comparison of gases absorbed within the meteorite and the composition of the Martian atmosphere determined by the Viking missions. The comparison with the composition from ALH84001 and the Viking lander is shown in Figure 6.11, both for the species and their isotopic composition. There is an impressive correlation between the two and this is particularly striking because the Martian atmosphere is unique within the solar system.

The anomalous micro-composition of the Martian atmosphere with regard to nitrogen, argon, neon, krypton and xenon has also been compared with trapped gases for the Martian meteorite collection (12 in total). The isotope ratios for

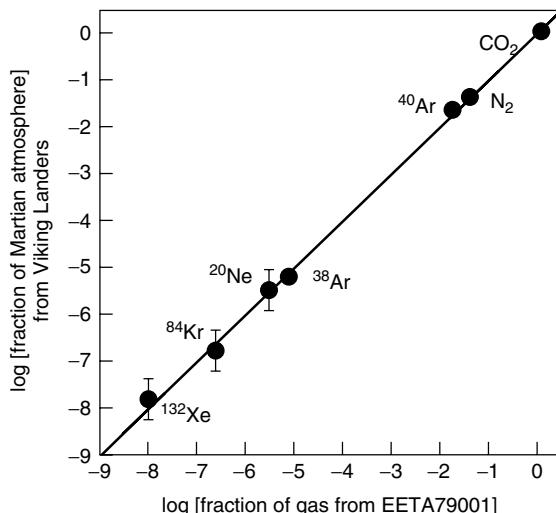


Figure 6.11 Comparison of composition and isotope data in released meteorite gases of EETA 790001 and the Martian atmosphere composition determined by Viking. (Reproduced by permission of Boggord, 1983)

$^{129}\text{Xe}/^{132}\text{Xe} \sim 2.6$, $^{36}\text{Ar}/^{132}\text{Xe} \sim 1$, $^{84}\text{Kr}/^{132}\text{Xe}$ and $^{20}\text{Ne}/^{22}\text{Ne}$ in the Martian atmosphere are plotted against one another for the meteorite. There are some problems with terrestrial contamination, especially if the meteorite is on the Earth for a long time prior to the find, but in general the evidence from the isotope data is convincing. All other evidence is circumstantial.

Organic material

The laser desorption measurements across the surface of ALH84001 have been used to determine the spatial distribution of PAH material. The tracking process is shown schematically in Figure 6.12 and shows an enhanced PAH concentration associated with the chondrules from the interior of the meteorite. The localisation to the interior reduces the likelihood of external contamination from the ice-melt water. The surface of the meteorite shows a fusion crust obtained during its high-velocity passage through the Earth's atmosphere. The entry velocity may have been as high as 70 km s^{-1} and resulted in melting of the surface, some of which would have evaporated during the fall. These temperatures would have converted any organic species to soot and sterilised any bacteriological spores.

Similar spatial resolution with the microprobe laser desorption studies detected significant PAH concentrations associated with the chondrules, also showing the characteristic stegosaurus backbone fragmentation pattern seen in Figure 6.13. The pattern is less well developed in the ALH84001 meteorite in Figure 6.13 compared with the other samples (Figure 6.9), due in part to the very small quantities of material probed in the experiment. The PAH signature is, however, unmistakable. Polyaromatic hydrocarbons are also thought to form at either high temperatures during combustion or in cold temperatures such as in molecular clouds. Some contend that PAHs are products of ancient bacterial fossilisation and decay but

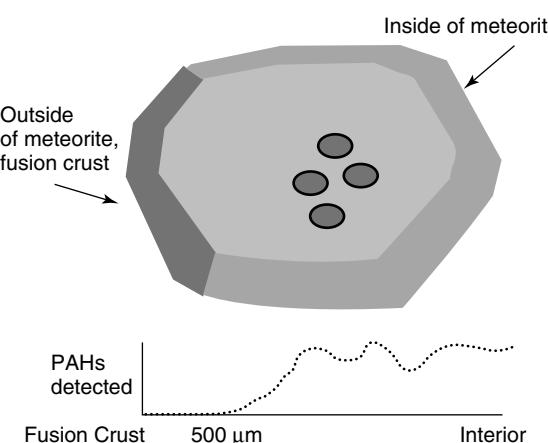


Figure 6.12 Profiling of PAH material. (Adapted with permission from Gibson, 2001)

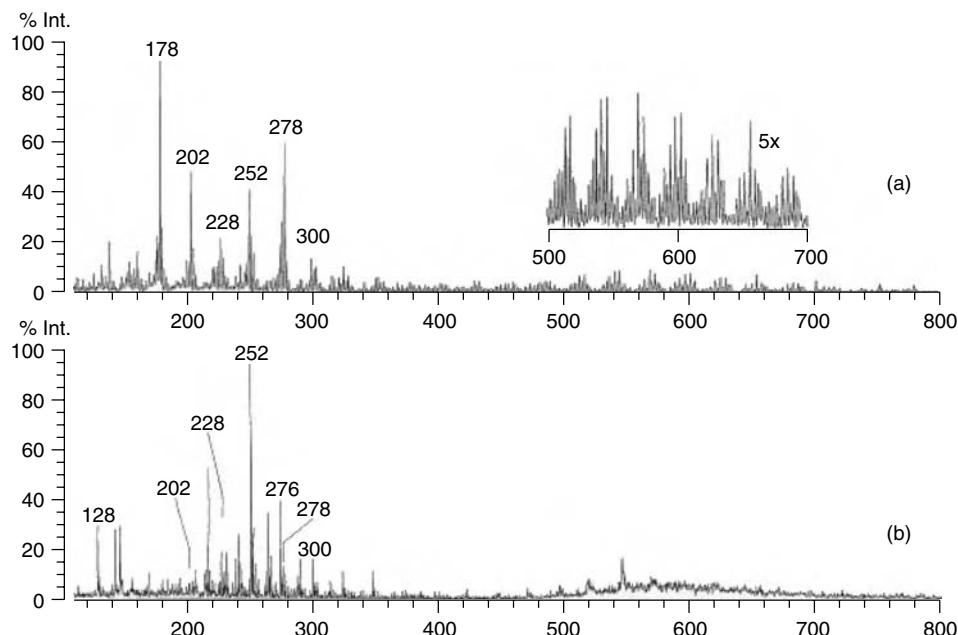


Figure 6.13 The $\mu\text{L}^2\text{MS}$ spectrum of ALH84001: (a) from the carbonate granules; (b) from the bulk meteorite. (Reproduced from Becher 1999 with permission from Elsevier)

this is hard to prove and the latter explanation was important in suggesting the hypothesis of ALH84001 evidence for life on Mars. Concerns were also raised about PAH contamination from ice-melt water in Antarctica but measurements from other chondrite meteorites with longer exposure times to the environment did not show increasing PAH concentrations with exposure time. The Martian origin of the PAHs is well established but the connection with a biological origin of this reduced carbon form is not well established.

Amino acid measurements in ALH84001 are almost certainly the result of Antarctic ice contamination. Amino acids are readily soluble in water but PAHs are practically insoluble. Isotopic measurements of ^{14}C show that terrestrial carbon is incorporated into the meteorite during extended stays in the Antarctic ice fields. In addition, microbial activity on the exposed surfaces provides an additional source of biogenic organic material that may be incorporated over time.

Biogenic structures

The debate regarding life on Mars was fuelled by the discovery of fossilised structures apparent in the electron microscope analysis of ALH84001, which produced the images seen in Figure 6.14. Structures suggestive of bacterial fossils were observed and hence the remnants of life on Mars. These structures, some 500 nm in length, were proposed as possible nanobacterium fossils and it raised questions

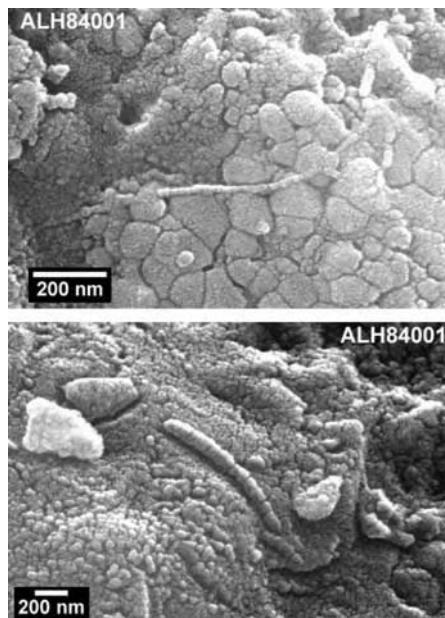


Figure 6.14 Fossilised structures in ALH84001. (Reproduced from Becher 1999, with permission from Elsevier)

regarding the lower size limit for a viable organism. We shall explore this idea further in Chapter 9 but it is an important point for astrobiology: what is required in the structure of a bacterium for it to be viable? Several studies have looked at minimal genomes for bacteria, including DNA/RNA and enzyme complement.

The structures seen in the ALH84001 meteorite could easily have formed from molten material solidifying rapidly or even by precipitation of minerals from saturated solutions. Neither explanation is as romantic as Martian nanobacteria. Similar sized features have been seen following electron microscopy analysis of basalt rock structures found in riverbeds, such as from the Columbia river (Figure 6.15).

The nanostructured fossils seem now to be less likely from Martian nanobacteria and probably are artefacts of the mineral formation process. Ironically, however, the discovery of the nanofossils may have been responsible for the massive interest in astrobiology, causing NASA to found the subject almost overnight.

Interplanetary organic material

Observations from meteorite falls suggest that organic compounds that may be prebiotic are available in a concentrated form within meteorite samples. Hot-water washes and other extraction processes show the production of many of the compounds required for life. Most importantly from the ALH84001 find is that material can be moved between planetary bodies and organic molecules can be transported

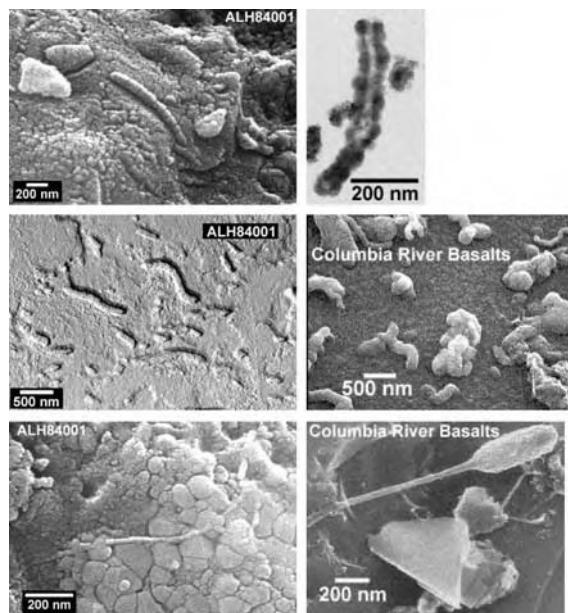


Figure 6.15 Fossil structures in ALH84001 and Columbian basalt. (Reproduced by permission of NASA, Johnson Space Center)

from one planetary environment to another. The evidence for this comes from magnetic domains within the meteorite structure. Some minerals act as collections of little bar magnets, which are ordered to produce a region within the meteorite that is magnetic. Laboratory measurements suggest that very low temperatures are required to scramble the magnetism to produce no domains with resulting magnetism. This temperature is 40°C and requires the interior of ALH84001 to have been below this temperature for extended periods of time. The low-temperature transfer of material between planets, without the rock getting too hot for life, allows for the possibility of bacterial spores or eucarya to survive the journey between planets. Almost one ton of material arrives on Earth from Mars each year, most of which has spent millions of years in space, but something like 1 meteorite in 10 million has made the trip in less than a year and it has already been shown that bacterial spores and microorganisms can survive in space for more than five years.

Experiments on board the NASA Long Duration Exposure facility have been performed with spores from the bacterium *Bacillus subtilis*, allowing them to be exposed to the extreme conditions of space. Low pressures and highly energetic particles are dominant in space and most importantly around the Sun, including an intense UV radiation field. It is the latter that is the most destructive in terms of viability of the spores, and under controlled conditions the extreme UV exposure is four orders of magnitude more likely to kill the cells than when screened. Crucially, however, not all spores were killed. Protection of the spores from the UV field for example within the interior of the meteorite suggests that the spores

can have a significant lifetime. Life in the solar system may have started on Mars and simply been transported to Earth via some express meteorite delivery service; this may have happened at any stage during geological time. The general concept of ‘Panspermia’ allows for life to be formed everywhere, surviving only where the planetary conditions are hospitable.

6.8 Comet chemistry

Interplanetary organic material is also present on the surface of comets, with a clear potential for cometary collisions with all planets during the formation of the solar system. Comets are regular visitors to the inner solar system in orbits that may be hugely elliptical and, like meteorites, would have made many collisions with the proto-Earth as it was forming. Comets originate from two places in the solar system: one is the Kuiper belt, a disc-shaped collection of icy debris at a distance of around 100 AU just outside the orbit of Neptune; and the second is the Oort Cloud at a distance as much as 50 000 AU that may contain a significant mass of material left over as a remnant of the formation of the solar system. No comet has been observed with an orbit consistent with an interstellar origin. The chemical composition of a present-day comet is then a good marker for the composition of the outer solar system and more interestingly the early solar nebula. Ground-based observations of some recent comets such as Hale-Bopp and Halley have revealed much about the structure and chemistry of comets but we have also begun to send spacecraft to intercept comets in orbit. The space probe Giotto made a very close fly-by of Halley, and the Rosetta mission launched in 2004 is ambitiously planning to land on the surface of a comet and perform *in situ* chemical analysis, returning the results to Earth.

6.9 Structure of a comet

The favoured theory for the structure of a comet is that of a dirty snowball and as such represents an environment in which surface chemistry can occur; in addition molecules can be processed by photons and cosmic rays. Cometary orbits take them close to the Sun, and the point of closest approach is called the *perihelion*. Passage close to the Sun produces two tails of material: a neutral molecule tail consisting of evaporated material from the nucleus and an ion tail produced by interaction of ionic species within the comet and then entrained in the magnetic field lines of the Sun (Figure 6.16). The core of the comet is some 1–10 km in diameter and is enveloped by the coma of the comet some 10^5 km in diameter, which becomes larger and more visible as the comet approaches the Sun. The dust tail and neutral molecule tail are generated by the thermal gradient and motion of the comet whereas the ion tail follows the magnetic field lines from the Sun. Both tail types are clearly visible in the images of the two comets Hale-Bopp and Halley in Figure 6.17. The two tails and the surface of the comet are three

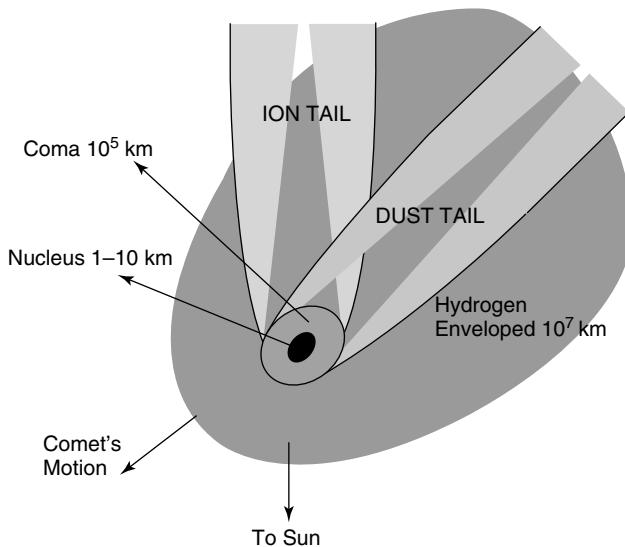


Figure 6.16 Structure of the comet showing the nucleus, the coma and two tails – an ion tail and a dust tail

different regions for the chemical processing of molecules. Ion–molecule reactions dominate the tail, with the colour deriving from ions in the plasma.

6.10 Physicochemical conditions in a cometary coma

The comet structure model proposed in Figure 6.16 shows clearly that the observation of molecules from Earth must be limited to those molecules present within the coma of the comet, and whilst they originate in part from the structure and composition of the nucleus the molecular observations are of the coma chemistry only. The coma observations will remain until we send a probe to land on the surface of a comet and report back the composition of the core. The Rosetta mission will do just this and we shall see the composition directly from the data it recovers, if successful.

The coma chemistry must be modelled as a function of temperature and composition and a mixture of the two. The nucleus is modelled as a collection of porous tubes on which there is an icy component dominated by water but also other volatile materials including CO, CH₄, CO₂, CH₃OH, HCN, NH₃, H₂S, C₂H₂, C₂H₆, and C₃H₄, depending on the particular choice of model. Heat from the Sun provides the required energy for vaporisation of the ice materials, energy that must be transported from the surface into the nucleus of the comet. The most volatile species will evaporate and some energy will be transported away from the coma. The transport of energy and molecules must be modelled simultaneously in a hydrodynamic–chemistry model for the comet. This mixing is responsible for the rich molecular inventory of the comet chemistry.

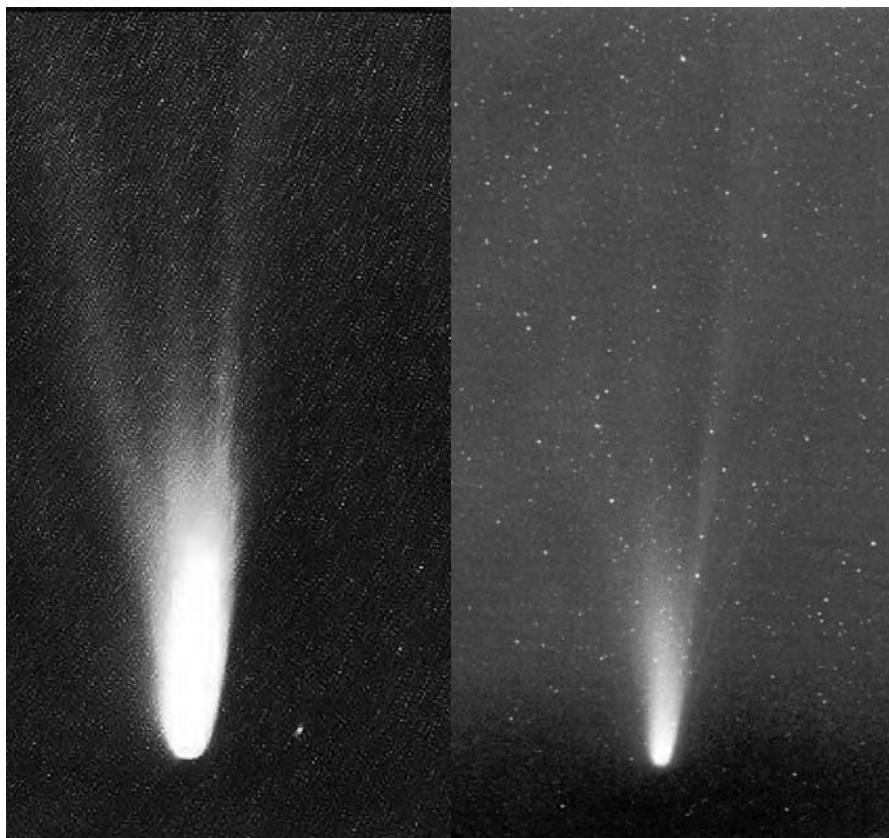


Figure 6.17 Tail structures of Hale-Bopp (left) and Halley (right) (1997/3/10, 4:38–5:00, Mt. Ho-Huan, Taiwan). (A colour reproduction of this figure can be seen in the colour section). (Reproduced by permission of Wei-Hao Wang, Institute for Astronomy, University of Hawaii)

Temperature plays an important role in the modelling of chemical pathways in the interstellar medium and must be modelled correctly for the comet. The energy balance is similar to that for all bodies orbiting a star and we will return to the ideas for the planets later. The energy balance considerations for an orbiting body are:

- Energy arrives from the solar flux heating the surface of the nucleus, the amount depending on the distance from the Sun, the inclination towards the Sun and the rate of rotation of the object.
- Some energy is reflected from the body, controlled by a property called the Albedo. Objects with larger albedos reflect more of the incoming light whereas lower albedos absorb more light. Clearly, albedo is a function of wavelength and this must be known for the surface composition of the comet.

- Energy is radiated from the surface of the comet simply as heat to the local interplanetary medium.
- Heat is needed to melt the icy surface, controlled by its heat of sublimation, which in turn is controlled by the chemical composition.
- Heat is transported through the layers of the ice into the nucleus of the comet. The temperature and rates of heat conduction are controlled by the coefficients of thermal conductivity.

The rate equations of transport and chemistry are solved simultaneously to give the predicted chemical composition of the coma as a function of time. At perihelion, the nucleus is producing 10^{28} molecules s⁻¹ into the coma, assuming a constant nucleus temperature of 10 K. The flux of molecules into the coma changes with distance from the Sun, with CO flux dominating at low temperatures because this compound has the lowest heat of vaporisation. At perihelion, water vaporisation dominates the coma composition. Photochemistry, initiated by the solar flux or radiation field as it is called, produces a population of fast-moving H atoms that initiate other radical chemistry, including the formation of species such as the OH radical. However, whether these molecules are observed or not depends on the lifetime of the molecules in the radiation field – are they photo-destroyed before they have time to contribute to the chemistry? The photoionisation cross-sections over the full-wavelength spectrum are not known even for the simplest molecules such as H₂O and CH₄ and are certainly not known for the larger organic species. The understanding of coma chemistry is limited by the laboratory data.

The cooling and heating of the comet as it passes by the Sun produces layers of ice that are rich in volatiles on the top but are more chemically complex further down. Surface chemistry clearly continues in the radiation field of the Sun and the snow line of the comet continually evaporates and then re-forms as the comet moves away from the Sun. The melting–freezing cycle allows volatile molecular fluxes such as that of CO to increase as the comet gets older. Considering the details for a comet in orbit around the Sun with a maximum distance (aphelion) from the Sun of 5 AU (somewhere near Jupiter) and a minimum distance (perihelion) 2 AU, a comet of initial radius 700 m loses about 4 m during every pass. The temperature of the nucleus depends on the rate of rotation of the body, producing its own day–night variation; typically the spin-period is about 6 h. The temperature varies from 50 K to 195 K during the day–night rotation at perihelion and 50–160 K at aphelion.

6.11 Chemical composition of comets

The recent close-encounter, especially with comet Hale-Bopp, focused the attention of the ground-based telescopes in all regions of the electromagnetic spectrum to produce the molecular inventory shown in Table 6.4. Chemical network models for

Table 6.4 Molecular inventory for Hale-Bopp observed at $r_h \approx 1$ AU

Molecule	Method ^a	[X]/[H ₂ O]	Remark
H ₂ O	IR	100	
HDO	Radio	0.06	
CO	Radio IR UV	23	Extended ^b
CO ₂	IR	20 ^c	
CH ₄	IR	0.6	
C ₂ H ₂	IR	0.2	
C ₂ H ₆	IR	0.6	
CH ₃ OH	Radio IR	2.4	
H ₂ CO	Radio IR	1.1	Extended ^d
HCOOH	Radio	0.09	
CH ₃ CHO	Radio	0.02	
HCOOCH ₃	Radio	0.08	
NH ₃	Radio IR	0.7	
HCN	Radio IR	0.25	
DCN	Radio	6 × 10 ⁻⁴	
HNCO	Radio	0.10	
HNC	Radio	0.04	Extended
CH ₃ CN	Radio	0.02	
HC ₃ N	Radio	0.02	
NH ₂ CHO	Radio	0.015	
H ₂ S	Radio	1.5	
OCS	Radio IR	0.4	Extended
SO	Radio	0.3	Extended ^e
CS	Radio UV	0.2	^e (from CS ₂)
SO ₂	Radio	0.2	
H ₂ CS	Radio	0.02	
NS	Radio	0.02	^e

Adapted with permission from Bockelee-Morvan *et al.* (2002).

^a Technique of observation.

^b Abundance given for nuclear + extended source.

^c Measured at $r_h = 2.9$ AU.

^d Abundance given for extended source.

^e Probably a daughter species.

the coma combined with the nucleus model allow good agreement to be achieved between the observations and the models. The molecules of interest to the origins of life and atmospheres are H₂O, CO, CO₂, NH₃, HCN and H₂S, which may deliver the important elements to the prebiotic Earth. Gas fluxes can be measured and compared with the models and also used to map the structure of the comet. Electron excitation of the O atom produces emission in the red at 630 nm and the O⁺ ion at 620 nm, which have been used to map the structure of the comet.

Clearly some molecular processing occurs in the coma to produce larger molecules such as methanol, formaldehyde and more complex still, but the detection and identification of bigger molecules are limited by the sensitivity and resolution of the instruments. The inventory, however, is probably a guide to an untold story

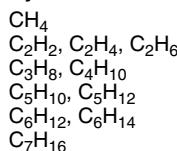
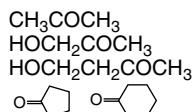
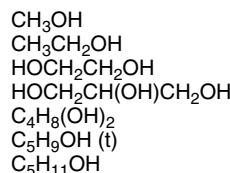
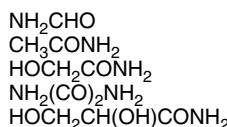
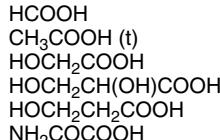
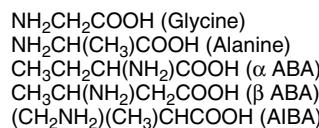
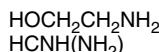
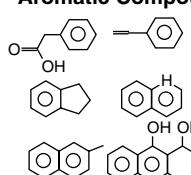
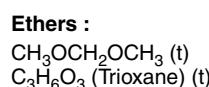
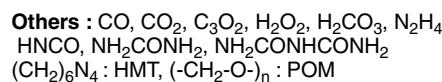
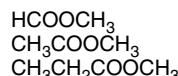
Hydrocarbons :**Ketones :****Alcohols :****Amides :****Carboxylic Acids :****Amino Acids :****Amines :****Aromatic Compounds :****Aldehydes :****Esters :**

Figure 6.18 Molecules detected in experimental simulations on cometary analogues

of a more diverse chemistry below the ice surface, similar to that on the surface of ice grains in the interstellar medium. Ultraviolet photon, cosmic ray and thermal processing of the comet's surface dominate at perihelion, producing a rich diversity of organic species. It is hard to say what species will be formed on the nucleus surface but simulations of cometary ice chemistry in the laboratory lead to the inventory of chemicals shown in Figure 6.18.

The analysis of the laboratory ice experiments reveals a diverse array of biological-looking molecules with the potential to seed life. Cometary collisions with planets could deliver these molecules to a primitive Earth or indeed to all planets within the solar system. The frequency of the collisions depends somewhat on the local star and the structure of the solar system forming around it.

6.12 Cometary collisions

The formation of the planets around the proto-sun initially started as a simple accretion process, aggregating small particles to form larger particles. This process was common to all planets, even the gas giants Jupiter and Saturn and to a lesser extent Neptune and Uranus. The planetesimals form at different rates and as soon as Jupiter and Saturn had reached a critical mass they were able to trap large amounts of hydrogen and helium from the solar nebula. The centres of Jupiter

and Saturn are thought to contain ten Earth-mass cores around which the giant gas planets formed. The planet-formation process is accompanied by some fairly pyrotechnic activity of the forming protostar, as we saw in Chapter 4. The T-Tauri phase of star formation results in the ejection of considerable quantities of material in the form of bipolar jets, as can be seen for the Tau B star in the constellation of Taurus (Figure 4.5). The ejection of the material accompanies the production of a significant radiation field around the star and the local temperature close to the star begins to rise. Increase in the temperature causes all water in the form of ice to melt and then ultimately vaporise. When the planetesimals are forming they will have insufficient mass to capture the water vapour by gravity and the vapour will be lost to the solar nebula. At some distance from the star, however, the temperature will be sufficiently low (~ 160 K) to allow water to stick to the mineral surfaces of meteorites, leading to the idea of a snow line in the solar system. For our Sun, the snow line forms at a distance at around 5 AU or around the orbit of Jupiter, as shown in Figure 6.19.

A new reservoir of comets may have formed at around 5 AU in a local orbit around Jupiter or at least perturbed by its gravitational attraction. A comet close to Jupiter would simply have been captured, delivering its chemical payload to the ever-increasing gas giant. Some comets would merely have been deflected towards the inner terrestrial planets, delivering a similar payload of water and processed molecules. Cometary impacts such as the spectacular collision of the comet Shoemaker-Levy 9 with Jupiter would have been common in the early formation phase of the solar system but with a much greater collision rate. Calculations of the expected collision rate between the Earth and potential small comets deflected from the snow line may have been sufficient to provide the Earth with its entire

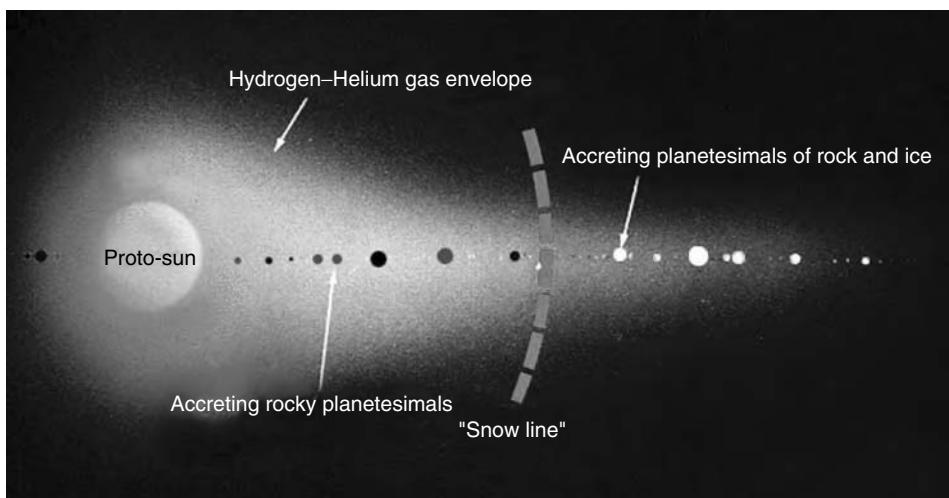


Figure 6.19 Snow line

water budget, forming the oceans and the atmosphere. The origins of the oceans and the atmosphere will be discussed in more detail in Chapters 7 and 8.

One primary observation from comets that have shown recent perihelion encounters has been the D/H ratio in water. From the three comets of Halley, Hyakutake and Hale-Bopp, observations of the D/H ratio seem to be 3×10^{-4} . Similar ratios have been measured for the molecules HCN/DCN, CH₃OH and NH₃, although we now have to be careful what is meant by the D/H ratio. The D/H ratio in water is 0.5, written as [HDO]/[H₂O], so it is not simply the ratio of the concentrations of the two species. The comparison with models at $T = 30$ K, dominated by ion–molecule reactions, reach similar D/H ratios, so specifically for water the ion–molecule reaction enrichment produces a ratio of 5×10^{-4} , in very close agreement with the cometary measurements. This suggests that the cometary material was formed within a cold interstellar cloud where ion–molecule chemistry dominates. The three comets in question are all Oort Cloud comets and would appear to have a D/H ratio that is approximately twice that found in water on Earth. Cometary origin of water on Earth therefore requires some D-enrichment process and one possibility comes from the activity of the Sun during the formation of the snow line. The T-Tauri phase is associated with increased X-ray emission, which could have led to significantly increased ion–molecule reaction chemistry in the early solar nebula. There would also be species-specific D-enrichment reactions controlled by the different ion–molecule reaction rate constants. Some evidence obtained from Earth-based observations of the coma suggests that different molecules do indeed have different D/H ratios but this must be modified by the ion–molecule chemistry induced by perihelion. To be sure about the composition of the cometary core directly it would be most useful to visit a comet.

6.13 The Rosetta mission – origin of the solar system

A direct measure of comet chemistry is planned with the landing of a probe on the surface of comet 67P/Churyumov-Gerasimenko, shown in Figure 6.20. The comet is in a short orbit around the Sun falling between the orbits of Jupiter and the Earth, with an orbital period of 6.6 years. Rosetta will rendezvous with the comet 67P while the coma activity is quite low at distances close to Jupiter and then follow the comet, at speeds of over 100 000 km h⁻¹, towards perihelion where the coma activity will significantly increase. The Rosetta mission will orbit the comet for over a year, mapping its surface and the changes in activity of the core, while making direct measurements from the lander. Comet approach will occur in the period January–May 2014 when a breaking manoeuvre will slow the relative velocities of the spacecraft and the comet to around 25 m s⁻¹ and finally 2 m s⁻¹ at the end of the rendezvous. Imaging at distances as close as 25 km will allow surface mapping and observation of significant surface features for the first time. The spacecraft will then enter into an orbit around the comet at a distance of around 25 km and a relative velocity of only a few cm s⁻¹.

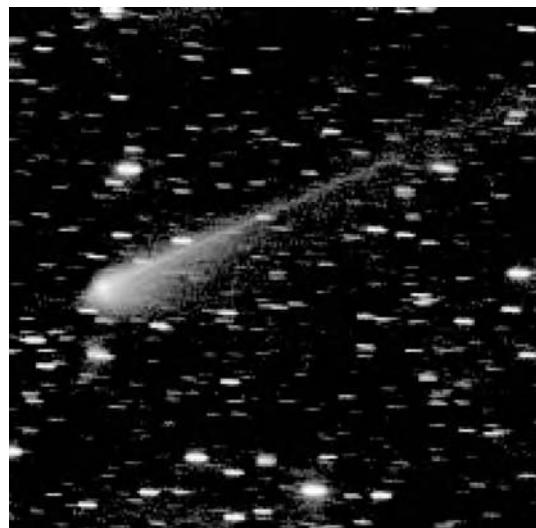


Figure 6.20 Comet 67P/Churyumov-Gerasimenko. (Reproduced by permission of DLR, Institut für Planetenforschung and Thüringer Landessternwarte Tautenburg)

The short-orbit comet 67P has been influenced by two close encounters with the gravitational field of Jupiter in the same way as comets in the snow line may have been deflected during the formation of the Earth. Prior to 1840 its perihelion distance was 4.0 AU when the Jovian encounter reduced it to 3.0 AU or 450 million kilometres. It decreased steadily until a further brush with the Jovian gravitational field in 1959 knocked its perihelion distance to 1.29 AU, which is its present-day value. The properties of the comet are summarised in Table 6.5.

The close-up images will allow the landing site to be chosen and the lander will be sent down to the surface in November 2014. The relative approach velocity for the landing will be less than 1 m s^{-1} , which is something like a slow walk. Once there, the scientific objectives will be met by the following series of measurements:

Table 6.5 Properties of comet 67P/Churyumov-Gerasimenko

Diameter of nucleus, estimated (km)	3×5
Rotation period (hours)	~ 12
Orbital period (years)	6.57
Perihelion distance from Sun (million km)	194 (1.29 AU)
Aphelion distance from Sun (million km)	858 (5.74 AU)
Orbital eccentricity	0.632
Orbital inclination (degrees)	7.12
Year of discovery	1969
Discoverers	Klim Churyumov and Svetlana Gerasimenko

Source: Reproduced by permission of the European Space Agency.

- Global characterisation of the nucleus and determination of dynamic properties, surface morphology and composition.
- Determination of the chemical, mineralogical and isotopic compositions of volatiles and refractories in the cometary nucleus.
- Determination of the physical properties and interrelation of volatiles and refractories in the cometary nucleus.
- Study of the development of cometary activity and the processes in the surface layer of the nucleus and the inner coma (dust/gas interaction).
- Global characterisation of asteroids, including determination of dynamic properties, surface morphology and composition.

Instruments on board will use X-rays to analyse the surface mineral structure and radio waves to obtain radar images of the internal structure of the comets. Most importantly for discerning the origins of the solar system are the two gas analysers dedicated to the determination of the composition of the volatile components from the surface as the coma develops and determining the D/H ratio. During the latter part of 2014 and towards the end of 2015 the lander and orbiter will follow the comet around the Sun, monitoring the increase in nucleus activity and the generation of the coma. In situ data will become available on the structure of comets and the generation of the coma, and real insight will be provided into the origin of the solar system and the formation of planetary atmospheres.

Concepts and Calculations

Concepts

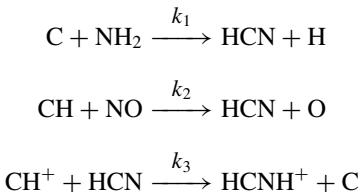
<i>Solar nebula</i>	The collapsing giant molecular cloud that leads to the formation of a star, specifically our Sun with its associated debris in the form of meteorites, meteors and comets
<i>Meteorites</i>	General classification into stony, stony-iron and iron, each with an interesting mineralogy, notably the carbonaceous chondrites
<i>Geological time</i>	Radiodating rock samples by the presence of radionuclei, notably the ^{210}Pb and ^{206}Pb daughters of uranium to give the best radiodate for the origin of the solar system
<i>Meteorite analysis</i>	The chemistry and morphology of meteorites, specifically ALH84001 and the Murchison meteorite
<i>Biogenic fossils</i>	Can some of the structures seen in meteorites, especially ALH84001, be attributed to astrobiology on other planets and what is the size of a cell?
<i>Comets</i>	Dirty snowballs with short and long orbits originating from the Kuiper Belt and Oort Cloud, respectively
<i>Molecular inventory</i>	An extension of the idea of molecular processing on dust grains in the interstellar medium to the surface of a comet
<i>Snow line</i>	The distance from the Sun at which water is stable on the surface of particles leading to comets. The presence of a large planetary mass such as Jupiter can then direct comets onto Earth, providing a source of cometary molecules to a prebiotic Earth

Calculations

<i>Black body temperature</i>	Temperatures of meteors and meteorites on entry
<i>Radiodating</i>	Estimate of the age of a rock from the radioisotopes
<i>Chemical networks on comets</i>	Extension of the idea of networks in the interstellar medium to molecular processing on the surface of comets

Problems

- 6.1** A meteor entering the atmosphere has a composition that is pure iron and may be treated as a perfect conductor.
- If the mass of the meteorite is 500 kg, the entry velocity is 70 km s^{-1} and initial temperature is 20 K, calculate the temperature of the meteor if all of the kinetic energy is converted to heat. The specific heat capacity of Fe is $S = 440 \text{ J kg}^{-1} \text{ K}^{-1}$
 - The latent heat of fusion for Fe is $L_{\text{fus}} = 28.93 \text{ kJ kg}^{-1}$. Calculate the energy required to melt the meteorite and re-calculate the temperature rise and hence the temperature of the meteor.
 - The latent heat of vaporisation of Fe is $L_{\text{vap}} = 6363 \text{ kJ kg}^{-1}$. If all of the kinetic energy is converted to melting and then vaporising iron, how much iron is lost during the descent?
 - What are the assumptions and sources of error in the calculations?
- 6.2** A meteorite was radiodated using two radionuclei:
- The $^{206}\text{Pb}/^{238}\text{U}$ isotope composition ratio of a meteorite is measured as 0.608. What is the age of the meteorite and when was it formed relative to the Sun?
 - The $^{208}\text{Pb}/^{232}\text{Th}$ ratio was measured as 0.350. Calculate the age of the meteorite using this radioisotope clock.
 - Account for the differences in the two calculations.
- 6.3** A biogenic structure in ALH84001 is 100 nm long and 25 nm in diameter. Assume that the structure is a cylinder and calculate the internal volume in zeptolitres (zL).
- 6.4** The volume of the world's oceans is estimated as $1.35 \times 10^{21} \text{ L}$. Calculate the number of cometary collisions of diameter 10^3 and 10^5 km required to deliver all of the ocean water. Assume that the comet is a sphere and the density of ice is 916.8 kg m^{-3} .
- 6.5** The following is a simple reaction scheme involving prebiotically interesting molecule HCN chemistry on a comet:



- Derive an expression for the steady-state concentration of HCN in this network.
- The reaction rates have the parameters in the table below. Calculate each of the rate constants at 50 K and 195 K (the perihelion diurnal temperature variation) and at 50–160 K (the aphelion temperature variation).

Rate constant	α	β	γ
k_1	3.26×10^{-11}	-0.1	-9.0
k_2	5.59×10^{-9}	0	10 814
k_3	1.8×10^{-9}	0	0

- (c) If $[C]=[CH]=[NH_2]$ density is 10^{12} cm^{-3} and $[NO] = 10^{-3} [C]$ and $[CH^+] = 10^{-9} [CH]$ at aphelion and $[CH^+] = 10^{-2} [CH]$ at perihelion, calculate
- (i) The diurnal variation in the steady-state concentration of $[HCN]_{ss}$ at perihelion.
 - (ii) The diurnal variation in $[HCN]_{ss}$ at aphelion.

7

Planetary chemistry

Introduction

The model for the formation of our solar system, or in general any solar system, by the collapse of a giant molecular cloud has to explain many different facts. The variation in the composition of the planets, for example, suggests something of the dynamics of the formation events. The gas giants Jupiter, Saturn, Neptune and Uranus are predominantly hydrogen and helium whereas the terrestrial planets of Mercury, Venus, Earth and Mars are essentially solid – rocky. Several planets also have moons that may have been formed at the same time as their parent planet accreted but may have been formed as a result of a collision between two objects, probably melting at least the surface of both. The latter giant-impact scenario is thought to be the origin of the Moon–Earth system.

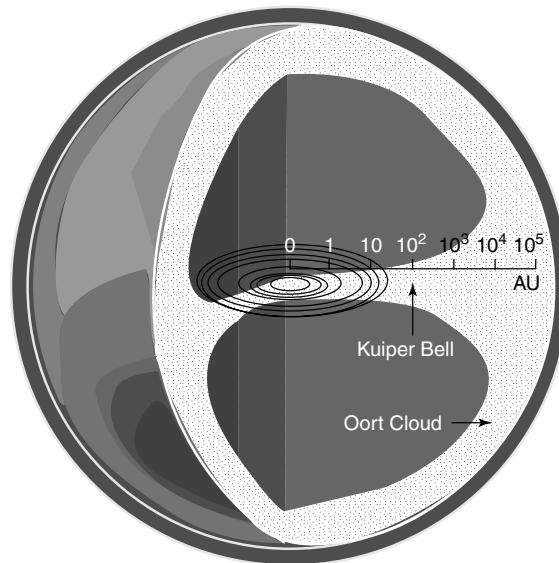


Figure 7.1 Structure of the solar system

The solar system is formed from the solar nebula and consists of the central star, the Sun, which accounts for some 99.9 per cent of the total mass: there are nine major planets and a collection of smaller bodies including the moons, asteroids and comets. The rocky terrestrial planets (Mercury, Venus, Earth and Mars) are nearest the Sun, with the gas giants or Jovian planets Jupiter, Saturn, Neptune and Uranus further out. At the edge of things is Pluto, with its huge moon Charon, the two objects perhaps making a double-planet system. Beyond this is the Kuiper Belt and then the Oort Cloud of comets (Figure 7.1). All of the planets orbit the Sun in the direction of the rotation of the Sun and the plane of rotation – the ecliptic. Mixed in with this collection of planets are the asteroids, which are thought to be the surviving debris of a set of planetesimals that occupied the solar system during its formation and have not formed into planets. Most are confined to the asteroid belt between Mars and Jupiter, which contains the biggest asteroid called Ceres some 1000 km in diameter and discovered over two centuries ago. Since then more than 6000 have been discovered. This collection of star, planets, asteroids and comets, all with similar angular momentum properties, make up the solar system. Simulations of collapsing clouds suggest that the observed structure is one of many possible solutions and is far from unique.

7.1 Structure of a star–planet system

The solar system is thought to date from 4.6 Gyr ago, as we saw from the CAI radiodating, and resulted from the collapse of a gravitationally unstable part of the Orion arm of the Milky Way. This region of space may have attained critical density as the result of a passing shock wave from a supernova nearby, with the compression of matter within the shock wave sufficient to trigger the gravitational collapse. The gravitational critical region of the local ISM inevitably has a small intrinsic angular momentum, which must be conserved during the collapse, and it is this angular momentum that is conserved within the rotation of the planets, firstly about their own axis and then around the central Sun. The Sun collapses to critical mass and begins to produce heat while the planets accrete from the remaining rocky and gaseous debris. Fortunately, large planetesimals form quickly and then attract under gravity, with a large subsolar nebula gas envelope forming the gas giants Jupiter, Saturn and Neptune. The subsolar nebula around the gas giants quickly attracts the majority of primordial H and He in the solar nebula, stripping it from the less-massive planets.

The evolution of the star has some dramatic effects on the nature of the surviving planets. Compression of the gas in the inner bulge of the solar nebula suddenly reaches the hydrogen fusion limit long before the proto-sun is properly formed and there is strong heating. The heating would have vaporised any condensed volatiles on the surface of planetesimals forming near the proto-star, notably at distances around 1AU and the forming Earth. The star starts its evolution track as we have seen, reaching the T-Tauri phase (Figure 4.5). The solar wind increases in a T-Tauri episode and the nebula in the nearer regions to the star is blown away. Hence, the

terrestrial planets are left without a large accreting gas cloud or inventory of low-mass molecules on the surface from which to form an atmosphere. The gas giants, however, are further away from the T-Tauri wind and continue to gather gas from the nebula apace as it is blown towards them.

The outline of the formation of the solar system, however, comes from theory and models based on the observational evidence of our own solar system. Unfortunately, even with the best spatial resolution from the Hubble Space Telescope of 0.046 pc or \sim 20 AU, all of the fine detail of the formation of planets in a stellar nebula is lost, even for the closest stars. The only system we have to study directly is the solar system shown in Figure 7.1. The planetary orbits extend 40 AU from the Sun to the orbits of Neptune and Pluto; these two planets rival for the furthest planet from the Sun due to the highly elliptical shape of Pluto's orbit. Beyond Pluto is the Kuiper Belt. This is a collection of comets that is believed to be the source of the short-orbit comets at a distance of 30–50 AU. Importantly, the plane of orbit of this subgroup of comets is in the plane of the solar system, which must be contrasted with comets that have longer orbital periods and orbital paths steeply inclined to the plane of the solar system. This second set of long-period orbit comets is thought to derive from the Oort Cloud at a distance of 100 000 AU from the Sun. The Oort Cloud is spherical in structure, representing the outer limit of the solar system but extending almost one-third of the way to the nearest star, which is 300 000 AU away. The composition and structure of the Oort Cloud is reminiscent of the giant molecular cloud that collapsed to form the solar nebula with a composition determined by the Orion arm of the Milky Way.

Understanding the atmosphere and structures of the planets themselves offers entire research careers for teams of scientists and, with the onset of planetary exploration, new data are continually being returned to Earth. We will concentrate solely on the Earth simply because it is the only planet we know to be inhabited. In Chapter 10 we will have a look at one of Saturn's moons, Titan, a body with a rich carbon chemistry that might support life. One other possible place of life in the solar system is below the ice caps of Jupiter's moon Europa. Europa is covered by ice that bears impact scars, and reflection spectroscopy shows the ice to be principally water. Below the ice could be large oceans of liquid water, heated by internal volcanism and the massive tidal rise and fall due to the presence of Jupiter. The ice layer is thought to be some 100 km thick and, although landing on Europa seems possible, drilling through the ice is a much harder problem.

It is tempting to take a tour of the solar system, stopping off at each planet to look at the chemistry from the origin of the red colour in Jupiter's great red spot to the volcanic activity of Io, but this would be another book for each planet. Instead, we will generalise the study to the formation of Earth-like planets in order to focus on the possibilities for life.

7.2 Surface gravity

The mass of a planet determines its surface gravity and hence the potential of the planet to capture and retain an atmosphere. Surface gravity also has implications for

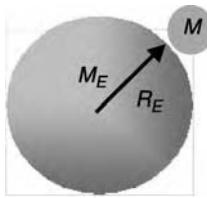


Figure 7.2 Gravitational weight

the nature of life on the planet, especially if it is to be motile. Simple measurements in the laboratory supported by astronomical observations allow Newton's Universal Law of Gravitation to be used to estimate the average density of the planets and hence a first look at their internal composition.

The surface gravity, g_o , is a measure of the weight of a person on the planet determined by the mass of the planet and the gravitational constant G (Figure 7.2) according to:

$$F_g = \frac{GM_e m}{R_e^2} = mg_0 \quad (7.1)$$

The gravitational constants has been known for a long time and quantifies the interaction between two masses. It is one of the constants that must be toppled if super string theory is to lead to a grand unifying theory. Measurements of the gravitational constant for the attraction between two masses produces a value of $G = 6.67 \times 10^{-11} \text{ kg}^{-1}\text{m}^3\text{s}^{-2}$. Local measurements of $g_o = 9.81 \text{ m s}^{-2}$ on the surface of the Earth combined with the radius of the Earth, $R_e = 6378 \text{ km}$, enables the mass of the Earth to be calculated using:

$$M_e = \frac{g_0 R_e^2}{G} \quad (7.2)$$

The mass of the Earth is $5.977 \times 10^{24} \text{ kg}$. Assuming that the Earth is a sphere of radius R_E , the mean density of the Earth can be calculated using:

$$\rho_e = \frac{M_e}{4/3\pi R_e^3} \quad (7.3)$$

Example 7.1

Using the numbers quoted above and the derived mass of the Earth gives $\rho_e = 5.52 \text{ g cm}^{-3}$, which, by comparison with the density of other materials measured in the laboratory, means that the Earth must be made of rock, and heavy rock at that. The mass of the other planets can be determined from their orbital periods and their radii can be measured, for example, from rates of transit in front of the Sun, and so the density of the other planets within the solar system can then be determined (Table 7.1).

Table 7.1 Planetary data (based on the Earth mass of 5.974×10^{24} kg and diameter 12 756 km)

Planet	Semi-major axis (AU)	Sideral period (Earth-year)	Mass (Earth-mass)	Diameter (Earth-diameter)	Density (g cm ⁻³)	Diurnal temperature variation (K)
Mercury	0.387	0.241	0.0558	0.381	5.50	100–700
Venus	0.723	0.615	0.815	0.951	5.3	730
Earth	1.000	1.000	1.000	1.000	5.518	200–300
Mars	1.524	1.881	0.107	0.531	3.96	145–300
Jupiter	5.203	11.86	317.9	10.86	1.33	165
Saturn	9.555	29.46	9.52	8.99	0.69	134
Uranus	19.22	84.01	14.54	3.97	1.27	76
Neptune	30.11	164.79	17.15	3.86	1.64	74
Pluto	39.44	248.5	0.0022	0.18	2.03	40

The density estimates in Table 7.1 show a distinction between the structures of the planets, with Mercury, Venus, Earth and Mars all having mean densities consistent with a rocky internal structure. The Earth-like nature of their composition, orbital periods and distance from the Sun enable these to be classified as the terrestrial planets. Jupiter, Saturn and Uranus have very low densities and are simple gas giants, perhaps with a very small rocky core. Neptune and Pluto clearly contain more dense materials, perhaps a mixture of gas, rock and ice.

7.3 Formation of the Earth

The Earth and the other terrestrial planets formed by the process of accretion produce planetesimals, which heat internally from the acquired gravitational energy and also from the decay of radioactive elements. The heating results in the *differentiation* of molten materials, separating by density; heavier elements sink into the core of the planet and give the Earth its molten core. All compounds present in the accreting materials will appear on the planet, including volatile materials such as the ices around interstellar grains and material resulting from the collisions of comets.

The early period of the Earth's history, known geologically as the *Hadean*, is associated with huge bombardment of the surface of the planet by meteors and comets. The sequence of events in the Earth's formation is shown in Figure 7.3, starting with the Hadean. Any volatile materials on the surface of dust grains or planetesimals deposited on the Earth will be removed and become part of the atmosphere, or more generally the volatile component inventory of the planet.

The bombardment continued for at least 0.5 Gyr until the oldest rocks on the planet, Acasta gneisses in northwestern Canada near Great Slave Lake at 4.03 Gyr, mark the beginning of a period geologically called the Precambrian. The first fossil record of life dates from 3.5 Gyr, leaving a window for evolution of perhaps 0.5 Gyr, although some say as short as 100 million years. The first life forms may not have survived for very long, having been made extinct by large-object collision with the newly fertile Earth – so-called *impact frustration* events.

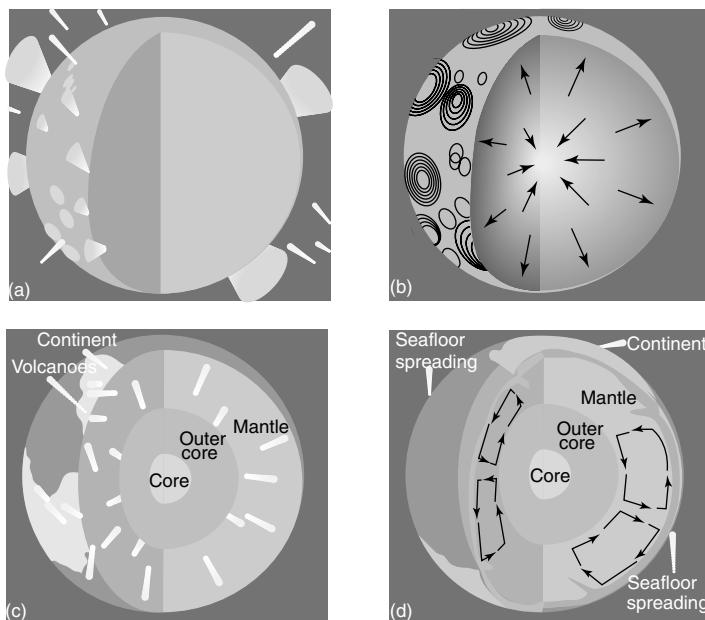


Figure 7.3 Formation of the Earth: (a) the Hadean; (b) melting of the planetesimals; (c) core fractionation; (d) mature continental structure. (Reproduced from Starosta and Zelik, 2002, with permission from Cambridge University Press)

Among the oldest rocks on Earth are those on Isua, an island off the coast of Greenland; they are 3.8 Gyr old, formed some 0.7 Gyr after accretion of the Earth. The rocks mark the beginning of the Archean period of geological time. The Isua rocks suggest that there was an extensive hydrosphere at this time, with erosion, transportation and deposition of minerals from water solution. The oldest lunar rocks, however, record an earlier high-temperature event – the Earth–Moon capture event.

There are two theories of the origin for the Earth: cold accretion and hot accretion. Cold accretion allows for the solar nebula debris to accrete into a cold Earth consisting of comparatively homogeneous mixtures of silicate and metal grains, as in chondritic meteorites. These particles simply stuck together by gravity to form a cold rock mass. The presence of ^{235}U , ^{238}U , ^{234}Th and ^{40}K provided radioactive heating to the core of the Earth, raising the temperature until the Fe–FeS melting occurred, a few hundred million years after formation (Figure 7.3). The molten iron then sank to the bottom of this melt, forming the core of the Earth. Gravitational energy is converted to heat and the entire planet melts at 1500 K in less than 10 Myr. This ‘Iron catastrophe’ liberates all of the condensed materials around the accreted matter, belching out a contribution to the primordial atmosphere in volcanic hysteria. The geological record forces the formation of the hydrosphere and life to be in the period around 3.8–3.5 Gyr ago, which means that the iron catastrophe occurred within 0.5 Gyr of the formation of the Earth.

The hot accretion theory is gaining prominence because of the observation that most of the asteroids are at least partially molten at a distance of 2 AU from the Sun. The hot accretion theory allows for already partially molten objects to accrete and the 1 AU catchment area of the forming Earth would also have contained partially molten asteroids. Hence, the Earth would have been built from rocks with soft centres and the impact would have released hot liquid rock, causing local heating and local volcanic activity. From the beginning the Earth would have been Hades itself. It is also thought that water-bearing objects such as comets and asteroids would have contributed to the early atmosphere and hydrosphere, increasing the surface temperature by the greenhouse effect – trapping the impact energy.

7.4 Earth–Moon system

Some time during the Hadean period a significant impact is thought to have occurred resulting in the capture of the Moon. The best theory for the origin of the Moon is an impact with a large Mars-sized object that was already pre-differentiated in the middle, so it already had a core. The co-orbiting body has been called Thea and must have been accreting matter within the same accretion catchment area as the Earth. The collision is thought to have occurred some 100 Myr after formation of the Earth. This model satisfies all of the boundary conditions associated with the orbital angular momentum of the Earth–Moon system and the similarity of the chemical composition. The giant impact would have vaporised the incoming body as well as the immediate surface of the Earth, throwing all of the water and molten rock into the atmosphere and into a very close belt, like one of the rings of Saturn. The rock vapour then would have condensed to reform the Moon, trapped by the Earth’s gravitational field (Figure 7.4). The hydrosphere and atmosphere of the early Earth are unlikely to have survived the impact and would have been lost to space.

The capture of the Moon has had some extreme effects on the conditions of early Earth. The rotational period of the Earth has been decaying since the Moon formation event, principally due to tidal forces. During the Devonian period some 400 Myr ago (comparatively recently) the length of the day was 21.6 h and there were 400 days per year. The rotational energy lost by the Earth is gained by the Moon, causing the Moon to move further and further away, so following the giant impact the Moon would have formed much closer to the Earth, with the rotation of the Earth perhaps as much as two or three times faster.

The much-closer Moon would have had a greater influence on the tidal rise and fall of the oceans. At the moment, the mid-ocean tidal rise and fall far removed from the land masses is of order 1 m but if the Moon had been formed at a distance of around 40 000 km the tidal variation would have been of the order 100–1000 m. Large regions of the Earth’s surface would have had a refreshed water supply every 4 h in the extreme cases of the model. Only well inland on the early land masses would there have been a dry environment, perhaps with fresh water replenishment. The early ideas of Darwin called for a ‘little warm pool’ to act

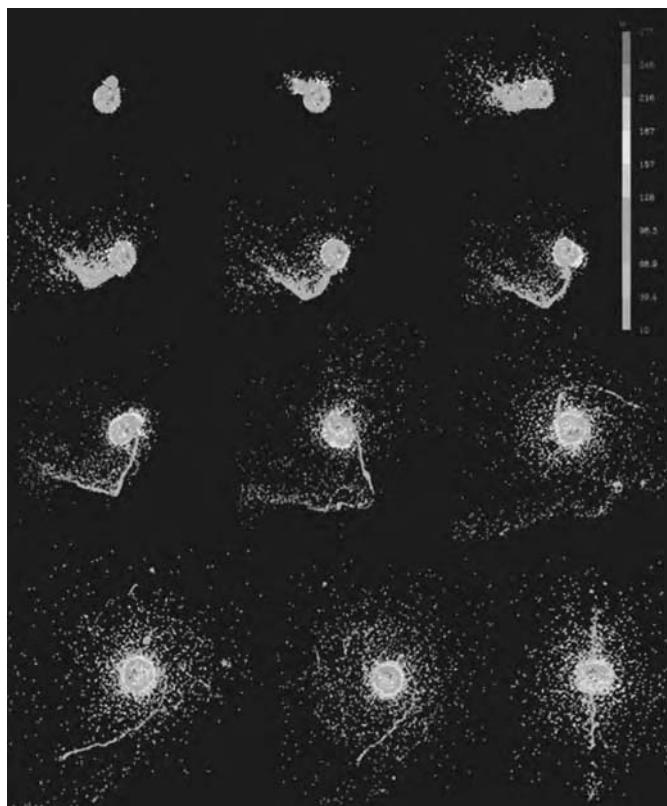


Figure 7.4 Modelling of the Moon-capture event. (Reproduced by permission of Southwest Research Institute)

as a chemical reactor for life. The primary role of the warm pool is to concentrate chemicals. The concentration of species in the oceans may have been rather low when averaged over the entire planet but because the water left a large rock pool at low tide the water would have evaporated and concentrated: the prebiotic soup. The Moon may have helped this.

7.5 Geological periods

Much of the information regarding the origin of the Earth, its atmosphere, biosphere and life forms comes from the geological record. Reading the rocks provides information on the composition of the paleoatmosphere by measuring the degree of oxidation of minerals. Information on the time at which organisms evolved with complete membrane structures is also seen in the fossil record and of course the regular mass-extinction events such as put paid to the dinosaurs. Geological time is divided into two eons, three eras and a number of periods and epochs as seen in the standard geological classification of time in Figure 7.5. The origins of life

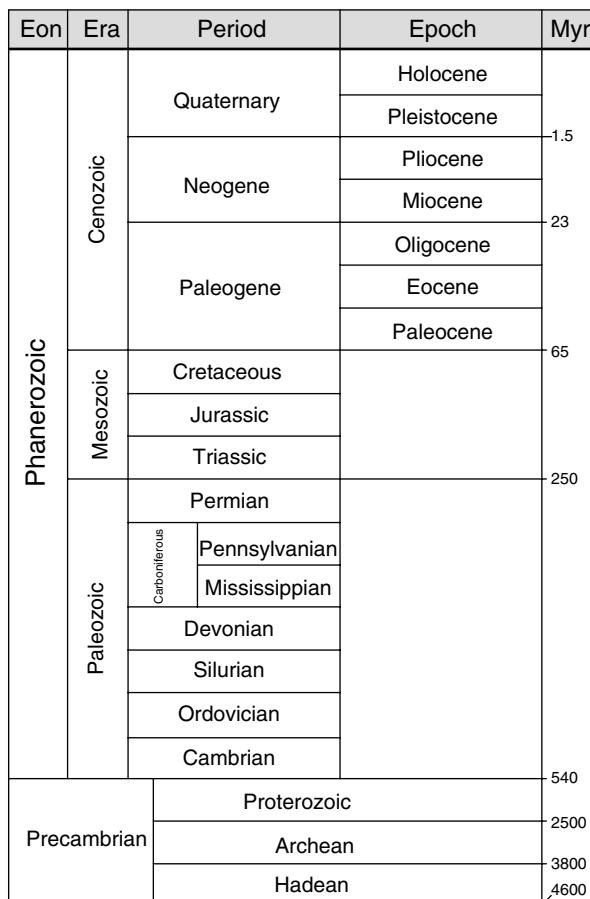


Figure 7.5 Geological time periods

must concentrate at the very bottom with the Hadean and Archean periods in the Precambrian. By the time of the Proterozoic period, life had solved all of the major problems of getting started.

The oldest minerals on the Earth (radiologically dated) are zircons from Australia, which are some 4.2 Gyr old, whereas the oldest rocks are the Acasta gneisses from Canada at 4.08 Gyr old. Western Greenland is home to approximately 50 per cent of the world's rocks over 3.6 Gyr old. The gneiss from the island of Akilia off the southwestern coast of Greenland shows an iron banded structure consistent with a water sedimentary deposition but the iron bands also show deformation, suggesting rocks having been forced underground. As in meteorites, the rock contains small carbon deposits, which have been proposed as remnants of early life forms, but this is questionable. The oldest undisputed life forms are stromatolites and associated carbonaceous material from rocks around 3.55 Gyr old in the Pilbara, Western Australia.

The water-sedimented rocks must have formed in the presence of significant amounts of water and CO₂ and have been deposited as carbonates. The removal of water and CO₂ from the Archean atmosphere leaves the surviving somewhat non-reactive N₂ to dominate the atmosphere. The chemical nature of the Archean atmosphere is very important for the processes of prebiotic chemistry, discussed in Chapter 8, where the contrast between oxidising or strongly reducing atmosphere is drawn. The geological record seems to suggest a mildly reducing atmosphere compared with the strongly reducing atmosphere of Jupiter and perhaps the solar nebula, dominated by the more chemically dynamic CH₄ and NH₃. Modelling of the oxygen concentration of the Archean atmosphere must be consistent with the presence of both oxidised iron deposits and reduced uranium in rocks that are 2.2 Gyr old. The data suggest that the O₂ concentration of the Archean atmosphere must have been more than 5×10^{-12} but less than 10⁻⁹ of the present atmospheric level (PAL). The geochemical evidence also points to a steady rise in the O₂ concentration throughout the period from 2.4–2 Gyr ago to levels of order 10⁻² PAL. This rise is associated with the beginning of photosynthetic contributions to the atmosphere. The Archean atmosphere of Earth is important to the understanding of prebiotic chemistry, leading to life. The available cocktail of chemical reactions, however, depends both on the composition of the prebiotic Earth as well as the available energy. The energy balance requirements of the Archean Earth are the same as those for any planet hoping to support life.

7.6 Radiative heating

The primary source of energy on the surface of the Earth or any planet is sunlight and the level of solar radiation arriving at the planet depends on the distance of the planet from the local star and the *albedo*. Albedo, A , is the measure of radiation reflected at all wavelengths from the surface, the cloud tops and aerosol particles without being absorbed by a body. A snow-covered field for example will reflect most of the radiation landing on it whereas a ploughed field will absorb more. The fraction of absorbed radiation is thus given by $(1 - A)$. An estimate of the effective equilibrium temperature of a planet can be obtained by assuming that the planet is a black body (Section 2.1) in the same way as a star and so will follow the Stefan–Boltzmann Law:

$$F = \sigma T^4 \quad (7.4)$$

where F is the total flux of the star, σ is the Stefan–Boltzmann constant ($5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$) and T is the temperature of the photosphere. For a planet of radius R_p the total emitting area is $4\pi R_p^2$. The absorbing area however is only the disc of the planet's surface presented towards the local star and has an area πR_p^2 . The energy balance is then given by:

$$4\pi R_p^2 \sigma T_e^4 = \pi R_p^2 (1 - A) F_p \quad (7.5)$$

where F_p is the stellar flux at the distance of the planet from the star. Rearranging this equation gives the effective surface temperature of the planet as:

$$T_e = \left\{ \frac{(1 - A)F_p}{4\sigma} \right\}^{\frac{1}{4}} \quad (7.6)$$

The solar flux can be calculated via Stefan's law from the observed surface temperature of the Sun, and the level of radiation at a known distance is calculated via the inverse square law (Figure 7.6).

Consider the amount of radiation arriving on the surface of the Earth at a distance of 1 AU or 1.5×10^{11} m. The total flux of the Sun is distributed evenly over a sphere of radius at the distance of the planet, d . From the luminosity calculation of the Sun, F , the solar flux at the surface of Earth, F_{Earth} , is $F/4\pi(1.5 \times 10^{11})^2 = 1370 \text{ W m}^{-2}$ from the least-square law of radiation discussed in Example 2.4 (Equation 2.4). Substituting this into Equation 7.6 with the estimate of the albedo listed in Table 7.2 gives a surface temperature for Earth of 256 K.

The calculated surface temperature of the Earth can be compared with a mean surface temperature of 288 K. The difference in these temperatures is attributed to global warming mechanisms in the atmosphere. Similar calculations can be performed for any object at any distance from a star of known temperature. The calculations for the other objects in the solar system are presented for the other planets in Table 7.2.

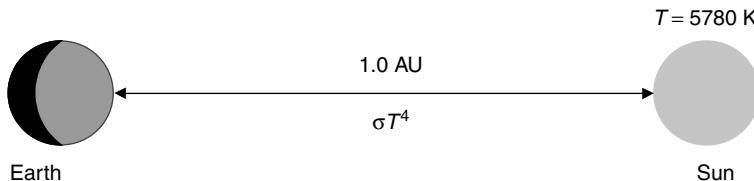


Figure 7.6 Surface temperature of the Earth

Table 7.2 Planetary information on surface temperature

Body	Radius (km)	Surface pressure (bar)	Albedo	Effective temperature (K)	Surface temperature (K)	Surface g (m s^{-2})
Venus	6052	92.1	0.77	227	735	8.60
Earth	6378	1.013	0.31	256	288	9.78
Mars	3393	0.063	0.15	217	223	3.72
Jupiter	71 490	Not defined	0.33	110	165	22.88
Saturn	60 270	Not defined	0.36	80	134	9.05
Uranus	25 500	Not defined	0.4	56	76	7.77
Neptune	24 765	Not defined	0.4	44	72	11.00
Titan	2560	1.5	0.2	85	92	1.25

The mean temperature of a planet depends both on the solar flux and the atmospheric composition and hence the radiation-trapping mechanisms of the atmosphere, notably the greenhouse effect. If liquid water is to be present on the Earth and act as the solvent for life, then the temperature of the Earth has to be maintained above 273 K. This requires two things: the solar flux must remain constant during the lifetime of the Earth; and the greenhouse or radiation-trapping mechanisms must remain constant or be able to accommodate the changes in solar flux. Current models suggest that the Sun's luminosity may have been some 20–30 per cent lower in the early stages of the solar system formation – the faint young Sun problem – and the Earth's surface temperature would have dropped below the freezing point of water.

7.7 The habitable zone

If liquid water is important for life, following the precedent set on Earth, then there is a distance from the local star that would receive sufficient radiation from the star to maintain the effective surface temperature of the planet above the melting point of water, 273 K. This condition defines a habitable zone around a star and is the range of distances from which an orbiting planet will have liquid water on its surface. As for the Earth, the habitable zone is determined by two factors: the effective surface temperature of a planet as determined by the flux arriving from the local star; and the radiation-trapping efficiency of the atmosphere around the planet. Inner and outer boundaries have been calculated for the habitable zone around the Sun based on the arriving solar energy and H₂O–CO₂ content of the atmosphere over geological time. A thick H₂O–CO₂ atmosphere would allow the Earth to maintain a habitable zone extending about 1.7 AU from the Sun, just outside the orbit of Mars, whereas the inner boundary is determined by the runaway-greenhouse effect as observed on Venus. If the surface temperature were too hot, above 373 K, this would vaporise all water on the surface of the planet. The inner boundary is around 0.85 AU so the habitable zone spans 0.85–1.7 AU for our Sun (Figure 7.7) but the current habitable zone spans 0.85 – 1.3 AU ($t = 0$) in Figure 7.7. The habitable zone was much larger when the Sun's luminosity was greater, and narrower when the luminosity was smaller.

The luminosity of the Sun has not remained constant throughout its lifetime and may have been as much as 30 per cent less than it is today during the early period of Earth formation: a faint young Sun. The flux of the Sun is decreasing at present so at some stage it must have gone through a maximum. The early Earth may have had a chilly beginning, if this was true and the greenhouse effect and the albedo of the Earth had remained constant (very unlikely) then T_e for the Earth would have been below 273 K for the first 2 Gyr of the Earth's evolution. As the Sun moved to maximal luminosity the habitable zone would have increased to a size encompassing the orbit of Mars before reducing again as the Sun reached old age.

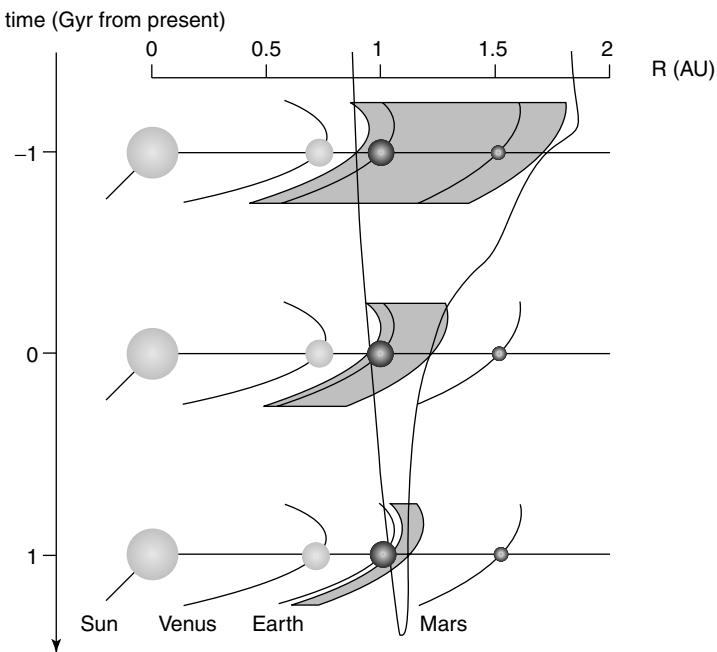


Figure 7.7 Continually habitable zone around the Sun

Example 7.2

Calculate the distance spanned by the habitable zone around the Sun at its current luminosity of 3.8×10^{26} W. Using Equation 7.5 for the flux at a distance corresponding to 273 K, the freezing point of water is given by:

$$\begin{aligned}
 4\pi R_p^2 \sigma T_e^4 &= \pi R_p^2 (1 - A) \frac{F}{4\pi d^2} \\
 d^2 &= \frac{(1 - A)F}{4\pi \sigma T_e^4} \\
 d^2 &= \frac{(1 - 0.31)3.8 \times 10^{26}}{4(3.1416)(5.67 \times 10^{-8})(273)^4} = 6.63 \times 10^{23} \text{ m}^2 \quad (7.7) \\
 d &= 2.57 \times 10^{11} \text{ m} = \frac{2.57 \times 10^{11}}{1.5 \times 10^{11}} = 1.72 \text{ AU}
 \end{aligned}$$

Repeating the calculation for the inner end of the habitable zone with $T_e = 373$ K, the boiling point of water gives $d = 0.92$ AU.

There are many sources of error with understanding the evolution of the Sun's luminosity but it raises the question of a 'continually habitable zone' around a

local star. Some estimates of the continually habitable zone suggest that the mean radius from the Sun should fall in the region 0.95–1.2 AU – a remarkably small region. However, the dependence on the atmospheric composition and the evolution of the atmosphere over time are critical. The position of the Earth around the Sun suddenly looks rather fortunate, with a large planetary object in the form of Jupiter diverting comets to deliver the hydrosphere, sufficient mass to maintain the atmosphere, within a continually habitable zone, greenhouse active atmosphere and perhaps importantly entwined in an Earth–Moon system from early in the planetary evolution.

7.8 Extrasolar planets

The radiation balance problem is not specific to our solar system and can be applied to all stellar planetary systems, although the number of planets outside our solar system – so-called extrasolar planets – is unknown. As noted earlier, the highest

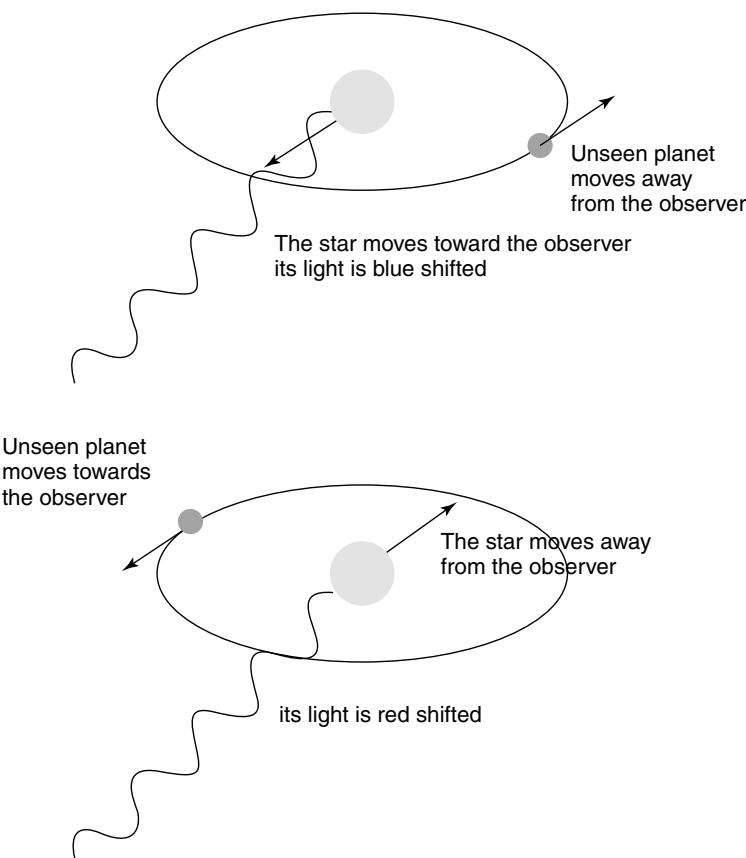


Figure 7.8 Doppler shift due to the presence of a planet around a star

spatial resolution of the Hubble Space Telescope is some 20 AU and so is unable to observe the detail of a solar system around a star. However, indirect observational methods have now been devised that have led to the discovery of extrasolar planets in orbit around their local stars. Now the description of a planet becomes important, especially as there are many binary star systems and a brown dwarf start might be detected and confused for a planet. The planet cannot be a source of light – pedantically of course it is a black body with $\lambda_{\max} = 10 \mu\text{m}$!

Current observations of extrasolar planets rely on indirect detection methods, usually by measuring the Doppler shift of stellar emission lines in response to the passage of a large planet through the line of sight. The local star and planet orbit with respect to a common centre of mass, which for most planets is very close to the star. However, it is possible to detect the wobble so that emission lines from the star are Doppler-shifted depending on the relative positions of the star and the planet (Figure 7.8). The wobble is very small so for Jupiter, the largest planet in our solar system, the effect on the Sun induces a Doppler shift of 13 m s^{-1} .

These detection techniques have found heavy planets with up to 400 Earth-mass but are unlikely to detect Earth-mass planets because the Doppler shift is too small. The first extrasolar planet to be discovered by the Doppler variation technique was 51-Pegasi, with the results shown in Figure 7.9. Precise radial

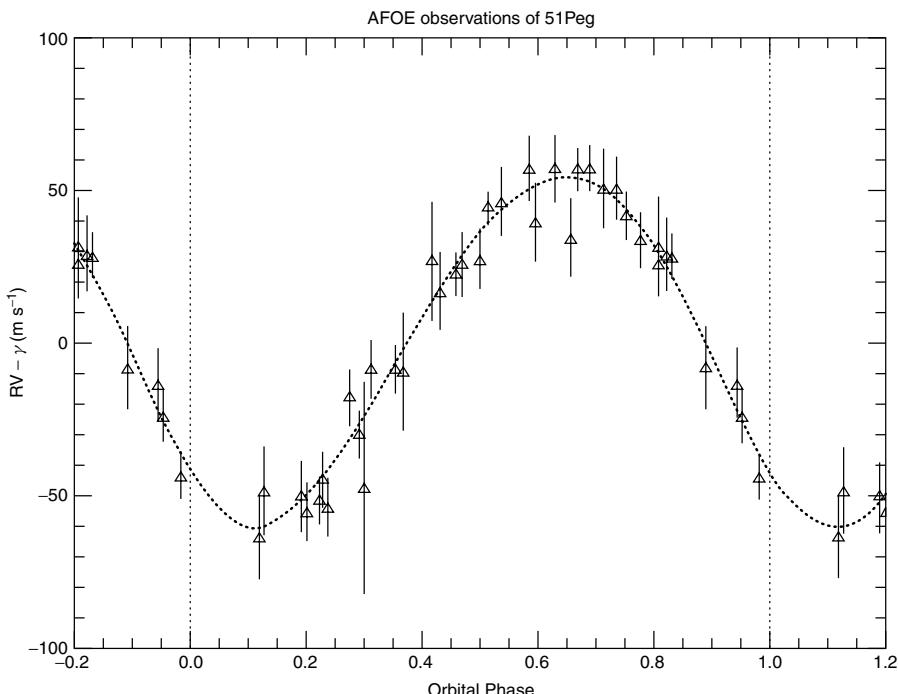


Figure 7.9 Doppler profile for 51-Pegasi. (Reproduced by permission of Sylvain Korzennik, Harvard-Smithsonian Center for Astrophysics)

velocity measurements of the star 51-Pegasi (HR 8729, HD 217014) with the AFOE (Advanced Fiber Optic Echelle Spectrometer) at the Whipple Observatory in Arizona indicate that the planet is about $0.45 M_{\text{Jupiter}}$ at a distance of 0.05 AU from the star. The large mass and the close proximity make the detection possible.

Similar radial measurements have been seen around the star τ -Bootes (Figure 7.10), consistent with the presence of a planet with a semi-major axis of 0.046 AU and a minimum mass of 3.9 Jupiter-mass.

Occultation experiments are another way to detect extrasolar planets and involve watching the transit of a planet across the disc of a star but these require precise photometry, especially as a Jupiter-sized planet removes only 1 per cent of the light from the stellar disc and an Earth-size object only 0.05 per cent of the light. Unfortunately, the luminosity of the Sun varies by more than 0.05 per cent and occultation is unlikely to observe Earth-size planets. Gravitational lensing measurements have potential but the best detection method for a planet, addressing the question of life as well, is a measurement of the infrared spectrum of the planet. This would reveal something of the composition of the atmosphere, especially biomarkers, and suggest a possible biosphere.

Direct detection of extrasolar planets requires a body with a luminosity some 10^9 times smaller than its parent star to be detected. The wavelength maximum

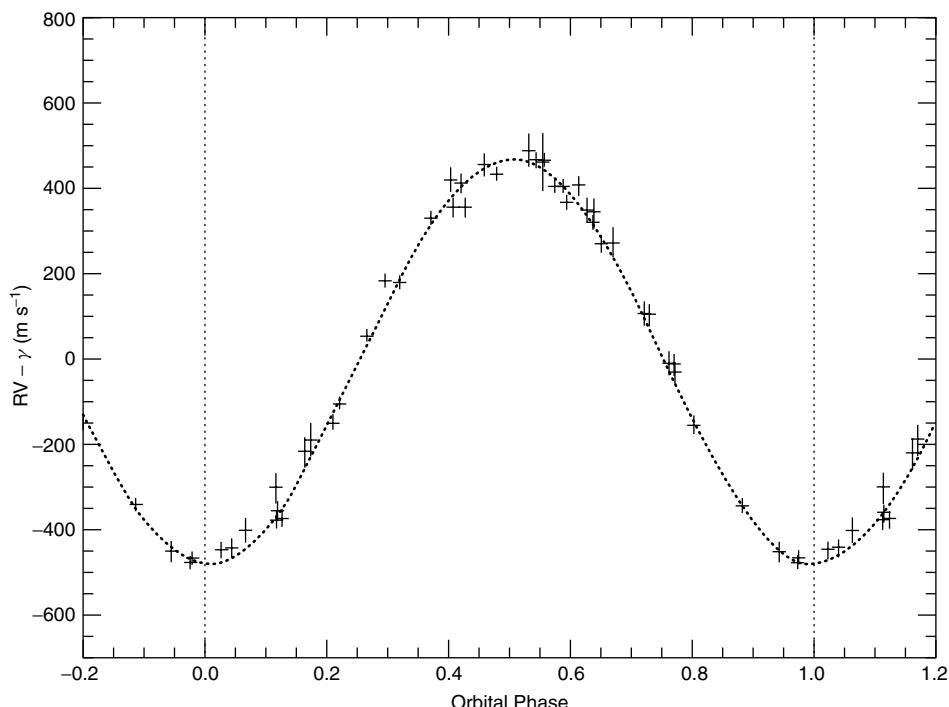


Figure 7.10 Radial velocity variations for τ -Bootes. (Reproduced by permission of Sylvain Korzennik, Harvard-Smithsonian Center for Astrophysics)

of a planet in a habitable zone with a surface temperature of 300 K is 9.7 μm in the near-infrared. This region of the spectrum is characterised by strong molecular absorptions such as CO₂ at 15 μm , water vapour at 6.3 and 16 μm , CH₄ at 7.65 μm , NO₂ at 7.8 μm and ozone at 9.7 μm .

Current extrasolar planets are all much larger than the Earth. The total count at present (9 September 2005) is 168 found in 144 planetary systems, of which 18 contain multiple planets. The first to be discovered was 51-Pegasi in the constellation of Pegasus by the radial velocity method. It is about 0.45 M_{Jupiter} and has an orbital period around the star of about 4.5 days. Of the 168 planets found so far only nine are present within a habitable zone around their star. The survey of the star catalogue for planets has only just started and we have found a large number of planets very quickly – solar systems, at least, are not special.

7.9 Planetary atmospheres

The preceding calculation of the thermal energy balance of a planet neglected any absorption of radiation by molecules within the atmosphere. Radiation trapping in the infrared by molecules such as CO₂ and H₂O provides an additional mechanism for raising the surface temperature – the greenhouse effect. The local temperature of a planet can then be enhanced over its black body temperature by the atmosphere.

Volatiles inventory

The starting chemical composition of the atmosphere is very important, especially if life is to evolve in perhaps as short as 0.5 Gyr. The inventory of volatile molecules present on the Earth comes from two sources: the internal reservoir and the external reservoir. The surfaces of the accreting particles would have brought to the Earth a collection of interstellar ice analogues containing molecules perhaps as varied as amino acids – somewhat of a head start when it comes to forming life. The alternative external reservoir from off the planet, perhaps from meteorite or cometary impacts, provides a late-accreting veneer of material delivered in a rather thermally traumatic event. The collision may well decompose all molecules thermally on impact, leaving a new cocktail of molecules to form from the cooling impact plasma. The late-accreting veneer model would also apply equally to Mars, Mercury and Venus, all of which should have the same chemical composition. The SNC meteorite collection seems to suggest, however, that Mars was different in at least several respects so perhaps the Earth did have a lucky hit with a particularly large comet.

The noble gas elements act as a record of the deposited material because they are essentially chemically inert and are also trapped within the ice of comets and meteorites. The late-heavy bombardment era must have affected both the Earth and the Moon similarly so an estimate of the collision frequency may be obtained by using the record of impacts on the Moon's surface. The collision rate calculated

should produce a delivery of material to the Earth of $1\text{--}4 \times 10^{25}$ g over the period of perhaps 0.6 Gyr. However, the origin of the volatiles has some problems with the noble gas record. The $^{132}\text{Xe}/^{84}\text{Kr}$ ratio is 10 times larger in carbonaceous chondrites than on Earth, which begs the question: where is all of the missing xenon? Similarly estimates of the cometary collision rates suggest that 6×10^{25} g of cometary material should have impacted during the same period, which represents something like 4–40 times the mass of the current oceans. Xenon has yet to be detected on comets but the Rosetta mission may provide the required information. The H/D and O isotope ratios found in comets such as Hale-Bopp and Hyakutake are essentially the same as found in the interstellar clouds but are much higher than the current sea water composition, suggesting that the comets made up perhaps 35 per cent of the oceans' water. The remaining 65 per cent must have derived directly from the solar nebula and has been processed by neutron-capture events in the increased solar wind activity during the T-Tauri star phase of the early Sun.

The delivery of volatiles to Earth and Mars must have been similar but where has the early Martian atmosphere gone? The atmosphere of the inner planets can be seen in Table 7.3. Cometary and meteorite impacts can deliver material to a planet but are also responsible for a process called impact erosion where the atmosphere could be lost due to an impact such as the Earth–Moon capture event. Current estimates suggest that impact erosion may be responsible for the loss of 100 times the current mass of the Martian atmosphere.

Atmospheric pressure

As the planet acquires a volatile molecule inventory it begins to develop an atmosphere, and in the case of the Earth this also includes the extensive circulation of water in the hydrosphere. The weight of the volatiles trapped in the atmosphere of a planet leads to a mean surface pressure, p_0 , given by:

$$p_0 = \frac{Mg_0}{4\pi R_p^2} \quad (7.8)$$

where M is the mass of the planet, R_p is its radius and g_0 is the surface gravity. The surface pressure for a number of planets has been calculated and reported in

Table 7.3 Atmospheres of the inner planets

Planet	Total (bar)	CO ₂ (%)	N ₂ (%)	O ₂ (%)	Ne (ppm)	⁴⁰ Ar (ppm)	³⁶ Ar (ppm)	Kr (ppm)	Xe (ppm)	H ₂ O (ppm)
Venus	92	96.5	3.5		7	70	35	0.05	<0.04	30–200
Earth	1.013	0.033	78		18.2	9340	31	1.14	0.087	<3%
Earth's surface inventory	70	98	1.5		0.4	190	0.6			
Mars	0.006	95.3	2.7		2.5	16 000	3	0.3	0.08	<100

Table 7.2. The mass of the Earth can be calculated from the surface pressure and gravity and this gives an estimate of 5.3×10^{18} kg.

The variation of atmospheric pressure with altitude, z , above the surface can be estimated by calculating the weight of a column of air of density, ρ , given by:

$$dp = -g \rho dz \quad (7.9)$$

Using the ideal gas equation, $pV = nRT$, the density ρ is given by:

$$\rho = \frac{M p}{RT} = \frac{mp}{kT} \quad (7.10)$$

where M and m are the relative molar mass and molecule mass, R is the gas constant, k is the Boltzmann constant and T is the temperature. Substituting Equation 7.9 into Equation 7.10 enables it to be integrated to give:

$$p = p_0 \exp \left\{ - \int_0^z \frac{dz}{kT/mg} \right\} \quad (7.11)$$

where p_0 is the surface pressure. This equation is known as the *Hydrostatic Equation* and can be simplified by assuming that g_0 is constant throughout the atmosphere (a reasonable assumption because the thickness of the atmosphere is small compared with the radius of the planet) and the temperature is constant. Temperature clearly varies with height but locally this is a reasonable place to start. Equation 7.11 reduces to:

$$p(z) = p_0 \exp\{-mgz/kT\} \quad (7.12)$$

As with all exponential expressions there is a characteristic scale over which the exponent falls to $1/e$ of its initial value and this is the *scale height*, H_s , which for the lower regions of the atmosphere is 6 km for $T \sim 210$ K, increasing to 8 km when $T \sim 290$ K. The expression for the scale height depends only on the atmospheric composition and, assuming that it is well mixed, this can be represented by a single mass. For the Earth the effective mass is 28.8, which is very close to the molecular weight of nitrogen. The pressure also controls the number density and, taken with the temperature, is important for constructing a kinetic model of the atmosphere.

Radiation trapping

The black body radiation profiles for a planet and the Sun have significantly different maximum temperatures with different spectral characteristics: planetary emission is principally in the infrared whereas stellar emission is dominated by the visible. Molecules present in the atmosphere may absorb the infrared radiation and re-radiate the radiation back to Earth. This was thought to be the role of glass

panes in greenhouses and is hence referred to as the greenhouse effect, although the prevention of convection within a glasshouse accounts for most of the temperature difference.

The dominant greenhouse gases on Earth are CO₂ and H₂O but a look at the absorption spectrum of water in the atmosphere seen in Figure 3.14 compared with CO₂ shows that water is the most important greenhouse gas. Other species such as CH₄, N₂O and O₃ together with particulates and clouds have to be included in all models of radiation trapping. The movement of heat around the atmosphere and hydrosphere is a complicated problem to describe physically and leads to very large global models. The input parameters of the global models are very important and must include accurate concentration profiles of all of the species, a region (longitude and latitude)-specific description of the land masses beneath the atmosphere (oceans and land have different effects on the atmosphere) and of course the solar flux of radiation. Models of global warming on Earth have to get the physics and chemistry right in all of the regions of the planet's atmosphere and then report with confidence an increasing or decreasing trend in average temperature of order 5 K against a background of 273 K. One important calculation to make is an estimate of the total heat absorbed by the atmosphere. The solar radiation flux arriving at the top of the atmosphere is 1370 W m⁻², which is then absorbed by the atmosphere. The result is an average ground-level solar flux that is 1342 W m⁻² but the global models report errors of 30 W m⁻² in this number – an error of nearly 10 per cent. The global warming calculations of temperature rise are reported against the background of this error, which only serves to highlight the difficulty in making the calculations and predictions.

Water vapour makes a sizeable contribution, and probably the largest, to radiation trapping and as the temperature increases the water vapour concentration increases. Temperature rises as a result of increased water vapour concentration and hence a mechanism for a positive feedback in the greenhouse effect that might lead to a *runaway greenhouse effect*. When the vapour pressure for water reaches saturation, condensation occurs and water rains out of the atmosphere: this is what happens on Earth and Mars. On Venus, however, the water vapour pressure never saturates and no precipitation occurs and the global warming continues to increase. Thus Venus suffers from extreme temperatures produced by both its proximity to the Sun and the presence of water vapour and carbon dioxide in its atmosphere.

Temperature profile

The temperature profile of a planetary atmosphere depends both on the composition and some simple thermodynamics. The temperature decreases with altitude at a rate called the *lapse rate*. As a parcel of air rises, the pressure falls as we have seen, which means that the volume will increase as a result of an adiabatic expansion. The change in enthalpy H coupled with the definition of the specific heat capacity

enables the change in temperature dT as a function of change in pressure dp to be determined and is given by:

$$\begin{aligned} dH &= V dp \\ \left. \frac{dH}{dT} \right|_p &= C_p \\ C_P dT &= V dp \end{aligned} \tag{7.13}$$

The variation of pressure with height derived in Equation 7.9 can be substituted to give the variation of temperature with height:

$$\begin{aligned} C_P dT &= -V \rho g dz \\ -\frac{dT}{dz} &= \frac{mg}{C_P} = \Gamma_d \end{aligned} \tag{7.14}$$

Dry air rising in the atmosphere has to expand as the pressure in the atmosphere decreases. This pV work decreases the temperature in a regular way, known as the *adiabatic lapse rate*, Γ_d , which for the Earth is of order 9.8 K km^{-1} . As the temperature decreases, condensable vapours begin to form and the work required for the expansion is used up in the latent heat of condensation of the vapour. In this case, the lapse rate for a condensable vapour, the *saturated adiabatic lapse rate*, is different. At a specific altitude the environmental lapse rate for a given parcel of air with a given humidity reaches a temperature that is the same as the saturated adiabatic lapse rate, when water condenses and clouds form! Clouds in turn affect the albedo and the effective temperature of the planet. Convection of hot, wet (containing condensable vapour) air produces weather and precipitation. This initiates the water cycle in the atmosphere. Similar calculations may be performed for all gases, and cloud layers may be predicted in all atmospheres.

The adiabatic lapse rate would suggest that the temperature of the Earth's atmosphere should fall at a constant rate with altitude and for the first 10 km or so this looks promising, shown schematically in Figure 7.11. This region defines the *troposphere* where all of the near-surface chemistry and weather occurs. However, at the *tropopause* a temperature inversion occurs due to the presence of local heating within the atmosphere. In the stratosphere both ozone and oxygen absorb radiation directly from the Sun and form a layer, trapping infrared, visible and UV radiation. This local heating as a direct result of the chemical composition causes the temperature to rise throughout the layer until the chemical composition changes and the molecules absorbing the radiation are no longer present.

The temperature inversion at the tropopause prevents mixing between the stratosphere and troposphere, with hotter air (less dense) sitting on top of cooler air (more dense). Pollutants (such as chlorofluorocarbons) present in the stratosphere have very long lifetimes (of order 30 years) and become persistent problems, especially

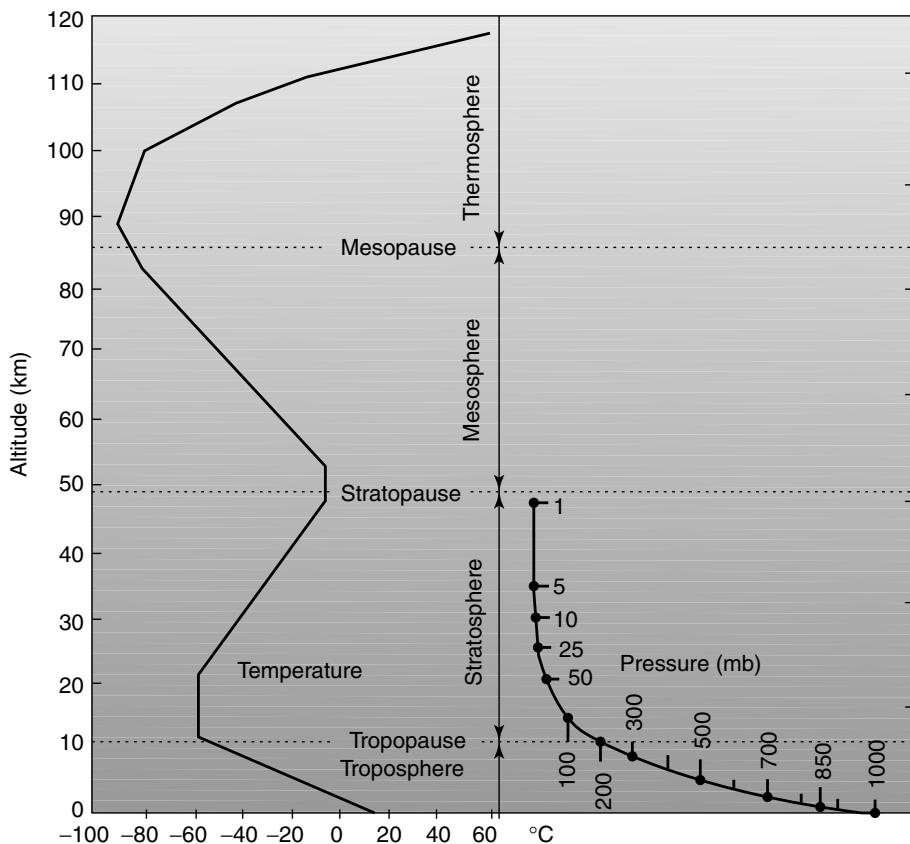


Figure 7.11 Temperature profile of the Earth's atmosphere. (Reproduced by permission of Oklahoma Geological Society)

with regard to the formation of ozone holes. Ozone also absorbs in the visible where the Sun's radiation is more intense but the heating effect is too small to compare with the cooling and there is a further temperature inversion called the stratopause. The atmosphere continues to cool throughout the mesosphere (lapse rate 3.75 K km^{-1}) where the lowest temperatures in the atmosphere are reached at 175 K with all of the consequences of low-temperature chemistry. Temperature in the thermosphere is a difficult concept due to the low concentrations of all molecules. Molecular collisions are rare and re-equilibration of thermal processes is not efficient. This can result in a very high temperature following a non-equilibrated absorption event.

The temperature of the planet within a habitable zone controls the rate of chemical reactions, the availability of chemicals and the phase in which the chemicals will be found. These have profound consequences for prebiotic chemistry and the origins of life.

Terraforming Mars

Current models of the Sun suggest that as it ages and moves off the main sequence of the HR diagram (Figure 4.3) it will become a red giant with a significant increase in size to a diameter of around 1.1 AU. This progress off the main sequence should occur in around 2 Gyr and will result in the total destruction of the Earth in about 200 years. This has prompted people to investigate the possibilities of making other planets habitable in the new habitable zone created by the enlarged Sun. Terraforming Mars is the first suggestion.

Mars has a similar rotation rate to the Earth, a surface gravity of around $0.38 g_0$ and a mean orbital radius of around 1.5 AU and so receives a solar flux something like 49 per cent of that on Earth at the present solar luminosity. The average surface temperature on Mars, however, is 213 K (-60°C) and would thus require 75 K of global warming to bring into a habitable zone. It is possible to assess the required levels of O₂, CO₂ and N₂ to provide a habitable atmosphere and then use the resources available on the planet, including the CO₂ polar ice caps. Can the polar ice caps be melted to add CO₂ to the atmosphere? Can we add super-greenhouse gases such as NF₃ and CFCCl₃ to initiate warming? Also, allowing for the current solar flux, the rate of melting will require a considerable timescale. The problem divides into two: increasing the temperature and initiating the required chemical changes. There are many suggestions, some even looking plausible, but we have insufficient knowledge of both the structure and composition of Mars to control the warming process and the chemical conversion. Conservative estimates suggest that terraforming may take 10^4 – 10^5 years.

7.10 Atmospheric photochemistry

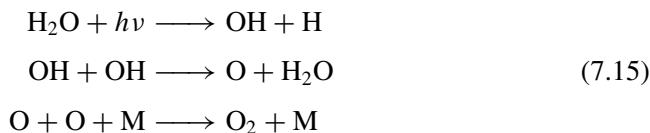
Life on Earth requires the energy from the Sun as a primary energy source but it must be protected from all of the radiation at shorter wavelengths. Radiation shorter than 323 nm can break the C–C bond and this would lead to mutations or complete photolytic destruction of carbon-based life forms. The protection from the short-wavelength radiation is achieved on Earth in two ways: the ozone layer and the photic zone.

The radiation wavelength required to break chemical bonds may be calculated by setting the bond energy equal to the photon energy, $h\nu$, as we saw in Section 5.6. The strongest chemical bond is C≡O and this would undergo photolysis with a threshold wavelength of 105 nm. A significant radiation flux at this wavelength or shorter would break all chemical bonds and prevent any chemistry. The Earth and all planets supporting life have to develop mechanisms for filtering out the short-wavelength radiation.

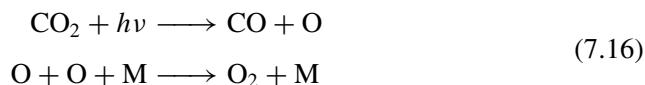
Water and CO₂ photochemistry

The photon flux onto the atmosphere can be absorbed at all wavelengths and the absorption spectrum of water does just that, absorbing radiation in the UV, visible,

infrared and microwave and is responsible for significant global warming. The influx of water from comets suggests that the water concentration of the atmosphere and hence water photolysis may have been a source of non-photosynthetic oxygen and ozone. Water photolysis starts with the following reaction mixture (M in the reaction scheme below is any other atom or molecule, termed a third body):



This mechanism has an analogy with CO_2 photolysis:



Both processes are switched on by the absorption of short-wavelength radiation $\lambda < 240$ nm for H_2O and $\lambda < 230$ nm for CO_2 . On the assumption that H atoms escape from the atmosphere, there is a net gain in oxygen to the atmosphere. Reactions of O atoms and O_2 chemistry would then lead to the formation of a small ozone layer with a low ozone concentration.

Short-wavelength shield

Ozone (O_3) is responsible for absorbing all radiation in the region 240–290 nm arriving on the Earth from the Sun. The classification of the UV radiation is familiar from the back of sunscreen products and is divided into three main groups: UVA 400–320 nm; UVB 320–290 nm; and UVC 290–100 nm. Ozone absorbs radiation in the UVC shortest wavelength region. Radiation in the region 290–320 UVB is also dangerous to organisms and prolonged exposure will result in DNA mutations.

In the present-day atmosphere ozone forms into layers and this was first explained by Chapman who proposed a photolysis mechanism for ozone formation. Chapman's mechanism is a simple steady-state production of ozone and led to the concept of *odd oxygen*. The odd-oxygen reaction scheme is shown in Table 7.4.

Table 7.4 Photolysis mechanism for ozone formation

Reaction		Reaction number	Rate constant	Variation in odd oxygen
$\text{O}_2 + h\nu$	\rightarrow	1	J_1	+2
$\text{O} + \text{O}_2 + \text{M}$	\rightarrow	2	k_2	0
$\text{O}_3 + h\nu$	\rightarrow	3	J_3	0
$\text{O} + \text{O}_3$	\rightarrow	4	k_4	-2

Conventionally in atmospheric modelling, the flux of photons is given the symbol J with the chemical rate constants, k : the constants are defined in Table 7.4. Reactions 1 and 2 account for the layer structure of ozone in the Chapman mechanism because they depend on both the photon flux (Reaction 1) and the number density of all of the other species in the atmosphere, given the symbol M. It is the balance between the photolysis rate, which is faster nearer the top of the atmosphere where the radiation has passed through only a thin molecular density, and the third-body collision rate. Ozone production also requires the energy from the $O + O_2$ association reaction (Reaction 2) to be removed by a third body, which must become faster as the atmosphere becomes thicker. The short wavelengths for the oxygen photolysis are removed at high altitudes by oxygen and therefore are not present at low altitudes. The photolysis also occurs at a slower rate at high altitudes as the O_2 concentration drops off with height. The balance of these two processes forms the ozone layer. Reactions 3 and 4 result in the destruction of ozone, with Reaction 3 the main screening photolysis reaction for shorter wavelengths. Reactions 2 and 3 are very fast and rapidly interconvert O and O_3 .

The odd-oxygen concept can be understood from a look at the differential rate equations for the mechanism:

$$\begin{aligned}\frac{d[O_3]}{dt} &= k_2[O][O_2][M] - J_3[O_3] - k_4[O_3][O] \\ \frac{d[O]}{dt} &= 2J_1[O_2] - k_2[O][O_2] + J_3[O_3] - k_4[O_3][O]\end{aligned}\tag{7.17}$$

The rapid rate of Reactions 2 and 3 means that O and O_3 convert rapidly, so it is useful to think of the sum of O and O_3 as a single species called odd-oxygen. The change in odd-oxygen in all of the reactions is listed against each reaction in Table 7.4. Odd-oxygen is created in Reaction 1 and destroyed in Reaction 4. To a good approximation, the pseudo steady-state concentration of [O] is controlled by the scheme in Equation 7.17 and the reactions involving k_4 can be ignored to give:

$$\begin{aligned}\frac{d[O_3]}{dt} &= k_2[O][O_2][M] - J_3[O_3] = 0 \\ \frac{[O]}{[O_3]} &= \frac{J_3[O_3]}{k_2[O][O_2][M]}\end{aligned}\tag{7.18}$$

The ratio is clearly pressure dependent: in the lower stratosphere $[O_2]$ and [M] are fairly large and J_3 is small (due to absorption above the required wavelengths), so the dominant odd-oxygen species is ozone. At higher altitudes both $[O_2]$ and [M] fall and the photolysis rate increases so that O is the dominant species in the atmosphere. The net flux of radiation in the band 240–290 nm is nearly zero at the surface of the Earth, which is then shielded from this radiation.

Layers of any molecule will form when the rate of production is controlled by photochemistry. These so-called Chapman layers can be described using all of

the approximations for pressure and photon flux with temperature to describe the layer structure. Replacing the atmospheric pressure, p , with number density, n , in Equation 7.12 and substituting the scale height $H = kT/mg$ gives:

$$n = n_0 \exp\{-z/H\} \quad (7.19)$$

Now for the photon flux. As the altitude increases, the atmosphere looks thinner so the light has to pass through a smaller amount of atmosphere dependent on the zenith angle, which is related to the latitude on the planet and the angle that its axis of rotation makes with the plane of the solar system, the season of the year and hence the position of the planet in its orbit. The depth of atmosphere through which the Sun's rays pass is given by $dz \sec \vartheta$. Hence the optical absorption is:

$$dI = I n \sigma (dz \sec \vartheta) \quad (7.20)$$

The combination of Equation 7.19 and Equation 7.20 gives the integrated radiation absorption as:

$$I = I_0 \exp\{-n_0 \sigma H \sec \vartheta \exp(-z/H)\} \quad (7.21)$$

The rate of photolysis, J , depends on the absorption cross-section, σ , the number density, the scale height and the angle, all of which are unique properties of a planetary atmosphere. For the Earth and the Chapman mechanism for ozone the O_3 concentration maximum is 5×10^{12} molecules cm^{-3} and this occurs at 25 km, shown in Figure 7.12, and forms the Chapman layer structure.

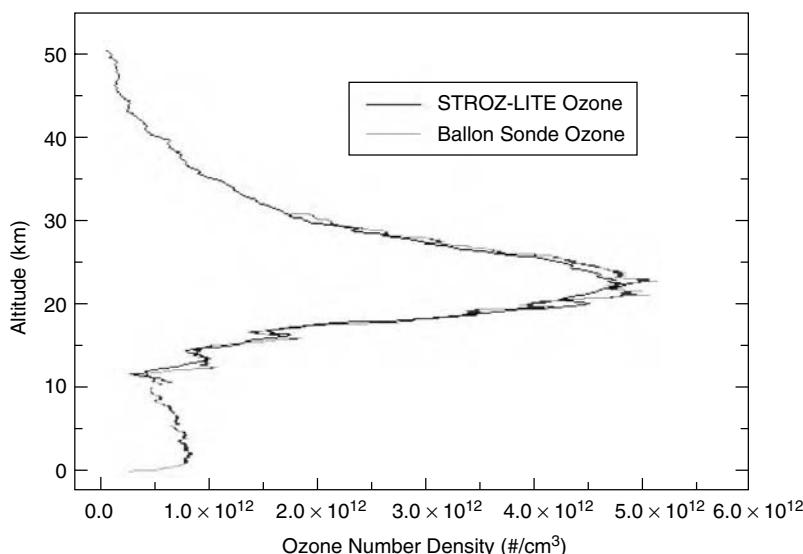


Figure 7.12 Ozone layer profile for the Earth Reproduced by permission of Ozone Profiles at Lauder, NZ; June 1, 1994. (Reproduced by permission of Atmospheric Chemistry and Dynamics, GSFC, NASA)

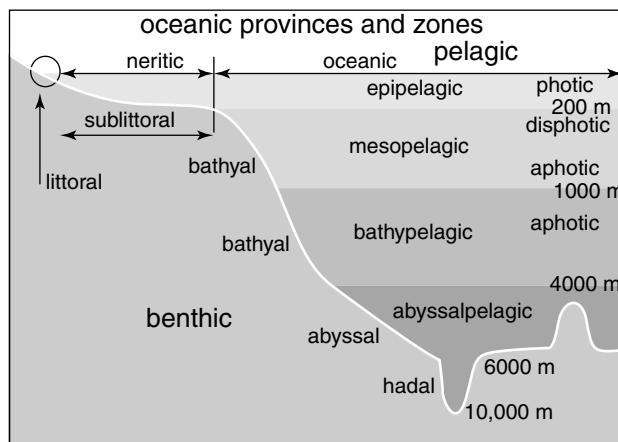


Figure 7.13 Oceanic zones

A short-wavelength shield achieves protection of chemistry on the surface of the Earth or the initiation of chemistry in a prebiotic Earth or Titan. Essential it may be but there is an alternative – the photic zone.

The Photic zone

Sunlight can enter a body of water to a depth defined as the photic zone (Figure 7.13). The lower boundary of the photic zone is the region where the light levels have fallen to 1 per cent of their surface value. The photic zone may be as little as 1 m where the water is unclear (due to particulates for example) or may extend up to 200 m. Within this region phytoplankton are capable of photosynthesis whereas below this region no light penetrates and the oceans are dark. Protection from all radiation is possible below 200 m and will allow chemistry and bond formation to occur.

7.11 Biomarkers in the atmosphere

As soon as life began on Earth, it started to disturb the atmosphere's chemical equilibrium. Oxygen built up continuously in the atmosphere over the next 3000 million years as a waste product expelled by some single-celled organisms. Then, around 600 million years ago, multicellular creatures developed and the oxygen content of the atmosphere rose suddenly. This trend was assisted by plants, which established themselves on the continents around 450 million years ago. Concurrently, evolution produced animals that breathed oxygen, instead of excreting it as a waste product. Without life, the oxygen content of the atmosphere would have sunk to undetectable levels.

Oxygen is a very reactive molecule and, without a continuous source to replenish it, would be used up and disappear from the atmosphere very quickly by reaction with surface materials. Hence detection of oxygen in the atmosphere either by the $0.76\text{ }\mu\text{m}$ absorption band of oxygen or the $9.7\text{ }\mu\text{m}$ absorption band of ozone would be a good biomarker for life. Similarly with methane, which again would be quickly destroyed were it not for a replenishment mechanism (on Earth, this is anaerobic organisms). Early organisms, such as the archaeabacteria, would have respired CO_2 and converted this to methane so an early Earth would have a different atmospheric signature. As with all molecular assignments, however, there may be non-biological mechanisms for the production of ozone in particular. Photodissociation of a significant amount of water vapour would produce some O_3 .

Concepts and calculations

Concepts

<i>Planetary system</i>	Structure of the planets that form around a local star
<i>Planet formation</i>	Formation of a planet by accretion and the fractionation of the molten rock to form a metal core
<i>Earth–Moon capture event</i>	The capture of the moon by collision between the Earth and another large co-orbiting planetesimal with all of the consequences of a close moon: larger tidal rise and fall
<i>Habitable zone</i>	The radiative heating of a planet from the local star to produce a surface temperature on the planet that will allow liquid water. From these calculations, estimates of global warming contributions can be made. A continually habitable zone allows for the variations in stellar flux as it evolves
<i>Extrasolar planets</i>	Planets orbiting their own local star outside our solar system
<i>Atmospheric structure</i>	Variations of temperature and pressure with altitude and the consequences for chemistry
<i>Chapman layers</i>	The formation of photochemically driven layers in the atmosphere, such as the ozone hole: a natural consequence of atmospheric structure
<i>Short-wavelength shield</i>	Protection of prebiotic chemistry from the highly energetic UV radiation

Calculations

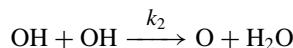
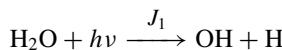
<i>Surface gravity</i>	Calculation of surface gravity as a marker for internal planetary structure
<i>Radiative heating</i>	Habitable and continually habitable zone
<i>Scale height</i>	Variation of temperature and pressure with height
<i>Chapman layers</i>	Steady-state atmospheric networks leading to Chapman layers

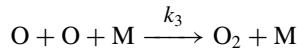
Problems

- 7.1** The radius of an extra solar planet is measured as 7000 km and has a surface gravity of 6.45 m s^{-2} .
- Calculate the mass of the planet.
 - Calculate the mean density of the planet.
 - Comment on the composition of the planet's interior.
- 7.2** Calculate the effective surface temperature of an Earth-like planet with the same diameter, albedo and orbital radius around Aldeberan in Taurus, which has a surface temperature of 7200 K. Is the planet within the habitable zone of this star?
- 7.3** Calculate the extent of the habitable zone around the star Aldeberan.
- 7.4** Pollux, one of the twins, has a surface temperature of 9000 K and a stellar radius of $1.16 R_{\text{Sun}}$. Calculate the reduction in luminosity from the surface of the star when a Jupiter-sized planet passes in front of the star. Is this likely to be detectable?
- 7.5** Calculate the λ_{\max} in the emission spectra of the following planets: (a) Earth; (b) Venus; (c) Jupiter; (d) Neptune.
- 7.6** A probe making a descent onto an extra solar planet measures the local acceleration due to gravity as 10.0 m s^{-2} and the scale height as 11.43 km.
- Rearrange the equation for pressure variation with altitude, given below,

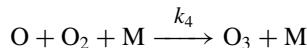
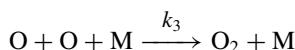
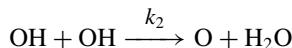
$$p = p_0 \exp\{-mgz/kT\}$$

- to derive an expression for the scale height within the atmosphere.
- Assuming that the local temperature of the planet is 220 K, estimate the composition of the atmosphere and identify a possible molecular composition.
 - What role would this molecular composition play in both radiation trapping and as a short-wavelength shield?
 - Would you consider the molecular composition to be an atmospheric biomarker?
- 7.7** A planet has a mass and radius the same as that of Earth but has a much lower surface gravity.
- Provide a possible explanation for the lower surface gravity.
 - Calculate the mean atmospheric pressure on the surface of the planet if $g_0 = 3.2 \text{ m s}^{-2}$.
 - On climbing a nearby mountain on the planet, the surface pressure is seen to drop to 70 per cent of its surface value at the summit of the mountain 2000 m above the surface. Calculate the scale height of the atmosphere. Assume that the temperature remains constant at 298 K.
- 7.8** Consider the water photolysis mechanism for the formation of oxygen atoms and ozone, shown below:





- (a) Write down an expression for the steady-state concentration of OH molecules.
- (b) If $J_1 = 5.96 \times 10^{-10} \text{ s}^{-1}$ and k_2 has parameters $\alpha = 3.87 \times 10^{13}$, $\beta = 1.69$ and $\gamma = -469$, calculate the OH steady-state concentration for 1 atmosphere of water at 298 K.
- (c) If the scale height of the atmosphere is 11 km, re-calculate the OH steady-state concentration at 80 km.
- 7.9** To investigate the role of water photolysis in ozone production in a prebiotic atmosphere, the following mechanism can be explored:



- (a) Considering the first two reactions only, write down an expression for the steady-state concentration of O atoms.
- (b) The water photolysis reaction scheme in Problem 7.8 has the following reaction parameters: $J_1 = 5.96 \times 10^{-10} \text{ s}^{-1}$, $k_2 \alpha = 3.87 \times 10^{-13}$, $\beta = 1.69$ and $\gamma = -469$. Calculate a steady-state concentration of OH and, using $k_3 \alpha = 1.98 \times 10^{-34}$, $\beta = -0.48$ and $\gamma = -568$, calculate the O atom steady-state concentration for 1 atmosphere of water at 298 K.
- (c) Calculate a steady-state concentration of O_2 using the OH and O steady-state values derived above.
- (d) Calculate the concentration of O_3 produced after 10^9 years if the k_4 constant parameters are $\alpha = 6.6 \times 10^{-35}$, $\beta = 0$ and $\gamma = -510$.

8

Prebiotic chemistry

Introduction

The chemical reactions that prevail on a prebiotic planet must lead to the generation of the molecules required for life or provide a fertile environment in which seeds of life from elsewhere might grow. Prebiotic chemistry does not appear to last too long in the history of the Earth, perhaps only 500 million years, but in this time all of the molecules must be made available for metabolism, propagation of information in an early gene and self-assembly into a membrane for a protocell. The two theories of endogenous or exogenous organic synthesis each require a chemical environment, which on a warm planet within a habitable zone would allow chemical reactions to occur on surfaces in the liquid phase as well as in the gas phase. Chemistry at liquid water temperatures is much more diverse than the ISM, with the potential to harness energy from many sources, either external, such as lightning, impact heating and photochemical, or internal, such as volcanic activity or electrochemistry. The allowed set of chemical reactions is controlled by thermodynamics and kinetics, treating a prebiotic planet as an open system into which molecules and energy are continually being added. Locally, however, thermal equilibrium will prevail. Consider the ‘little warm pool’, Darwin’s suggestion for the vessel containing the primordial soup of molecules. Material will enter the pool from the atmosphere, with gases dissolving in water, molecules being leached out of the mineral surfaces, volcanic activity maybe heating the pool and the activity of the hydrosphere continually replenishing or causing the pool to dry out. The chemistry within this vessel is controlled by thermodynamics producing local chemical equilibria based on temperature, pressure and electrochemical gradient. The resulting combinations of molecules, or something like, produce life.

8.1 Carbon- and water-based life forms

Although life forms could be conceived using any of the chemicals in the Periodic Table, the ability of carbon to string together with itself and other molecules to form polymers is unique and makes it a compelling choice for the primary chemical

structure of organisms. It allows sufficient diversity of chemistry and structure to fit the many roles required of molecules within the cell, not least of which is the propagation of information. The solvent of life certainly does not have to be water – although it probably has to be a liquid. Transport properties in a liquid, including concentrations of species in solution, are much more amenable to the dynamics of an organism. Water is certainly liquid on the Earth within the habitable zone of the Sun, and Europa, one of Jupiter's moons, may be the only other source of liquid water in the solar system, hidden under a thick ice crust several hundreds of kilometres thick. However, *again*, the solvent does not have to be water: carbon-based life forms could exist in liquid hydrocarbons such as might exist on the surface of Saturn's moon Titan. Titan may contain lakes or oceans of mixtures of methane, ethane and longer chain hydrocarbons. Sources of chemical energy to drive prebiotic chemistry and ultimately metabolism have to be utilised and harnessed but this does not rule out non-aqueous life.

Solvent properties

Water is the solvent for life on Earth and it has a number of important properties that which might set the standard for other life-supporting solvents:

1. Water is liquid over a wide temperature range of 0–100°C, although the boiling point may be increased under pressure and the melting point may be decreased by the presence of dissolved salts.
2. The density of ice is less than that of water so that ice floats. Having a frozen ice cap protects life below the ice, whereas a solid phase that sinks means that solid forming near the cold surface will sink, leaving new liquid to freeze and eventually freezing throughout any body of liquid.
3. Water is a polar solvent so has different solvation properties that discriminate between polar and non-polar molecules. Chemical discrimination results in the formation of mixed phases, such as membranes, microenvironments and compartmentalisation.
4. Heat of vaporisation. Water has a very large heat capacity (a large amount of energy has to be removed to lower the temperature by 1°C) and a large heat of vaporisation. This means that the temperature in solution is stabilised by the thermochemical properties of the water as a solvent. All life forms on Earth stabilise their internal environments with respect to temperature and composition so that the internal chemistry or metabolism is kept constant – a process called homeostasis. It would, however, be possible to learn to live in an environment that was fluctuating more wildly and develop a unique evolutionary niche.

The properties of the potential solvents in Table 8.1 might suggest some very interesting possibilities for life based on the range over which the solvent is liquid.

Table 8.1 Melting and boiling points of possible solvents for life

Liquid	Melting point (°C)	Boiling point (°C)	Range for liquid at atmospheric pressure (°C)
H ₂ O	0	100	100
NH ₃	-78	-33	45
CH ₄	-182	-164	18
C ₂ H ₂	-183	-89	94
CH ₃ OH	-94	65	155

Methanol looks particularly promising in this regard, as would all alcohols, many of which are found in the ISM or as processed molecules of comets. Hydrocarbon chemistry in a solvent such as acetylene would be supported over a large liquid temperature range, although the temperature is rather low and the chemistry would have to adapt.

8.2 Spontaneous chemical reactions

Whether a reaction is spontaneous or not depends on thermodynamics. The cocktail of chemicals and the variety of chemical reactions possible depend on the local environmental conditions: temperature, pressure, phase, composition and electrochemical potential. A unified description of all of these conditions of state is provided by thermodynamics and a property called the Gibbs free energy, G . Allowing for the influx of chemicals into the reaction system defines an ‘open system’ with a change in the internal energy dU given by:

$$\begin{aligned} dU &= TdS - PdV + \mu_1 dn_1 + \mu_2 dn_2 + \mu_3 dn_3 + \cdots + \mu_N dn_N \\ dU &= TdS - PdV + \sum_i^N \mu_i dn_i \end{aligned} \quad (8.1)$$

where T is the temperature, S is the entropy, P is the pressure, V is the volume and μ_i is the chemical potential of species i . The internal energy of the system is a measure of the energy available from the surroundings to drive chemical reactions, with entropy being a measure of the order or disorder using the standard definitions. The interesting terms, however, come from the N different chemical components with their own chemical potentials. The chemical potential is the property that drives the formation of equilibria, whether the component exists in the solid, liquid or gas phase. Most reactions in solution occur at locally constant pressure, especially in the laboratory, and the state function enthalpy, H , is introduced, where $dH = dU + PdV$. One further state function is the Gibbs free energy, G , required to provide the neatest link with equilibrium, giving the general result for the change in G as:

$$dG = dH - TdS + \sum_i^N \mu_i dn_i \quad (8.2)$$

The condition for a spontaneous reaction is that the change in the Gibbs free energy must be negative ($dG \leq 0$) – a so-called exoergic reaction.

Chemical potential

The chemical potential and balance of chemical potentials between species control the chemistry at fixed temperature, T , and pressure, P , even when distributed between phases. If a concentration of a species in solution is not at equilibrium with the species in the solid, such as during precipitation, the species will seek to reduce its chemical potential and precipitate. At equilibrium the chemical potential of species in both phases is the same. The chemical potential is thus analogous to the electrical gradient for charge, allowing the charge to move along the potential gradient to lower potential. The chemical potential for a chemical species i is given by:

$$\mu_i = \mu_i^0 + RT \ln a_i \quad (8.3)$$

where μ_i^0 is defined as the standard chemical potential, R is the gas constant, T is the temperature and a_i is the activity of the species. Activity relates to concentration in solution and $a_i = \gamma_i c_i$, where γ_i is the activity coefficient and c_i is the concentration. The activity coefficient allows for the non-ideal behaviour of molecules in real solutions, especially as the concentration increases, near a charged interface for example. The general expression in Equation 8.3 can be extended to gas-phase species relating the fugacity to the partial pressure of the gases present.

Equilibrium

Allowing an open system to close and come to equilibrium implies an extent of reaction – a measure of how close the system has come towards the equilibrium condition. For a single species the extent of the reaction towards an equilibrium concentration is given by:

$$n_i = n_{i0} + v_i \xi \quad (8.4)$$

where v_i is the stoichiometric coefficient of species n_i with initial concentration n_{i0} and ξ is the extent of the reaction running from zero to one mole at equilibrium. The stoichiometric coefficient is defined from a balanced equation of the reaction:



The stoichiometric coefficient for A is a , and for B is b , choosing the convention that for reactants the coefficients are negative ($-a$, $-b$) whereas for products the coefficients are positive.

The change in the composition of n_i is given by $dn_i = v_i d\xi$, which can be substituted into Equation 8.2 to give

$$dG = dH - TdS + \sum_i^N (v_i \mu_i) d\xi \quad (8.6)$$

At a fixed temperature and pressure, the Gibbs free energy of the reaction goes through a minimum and the rate of change of G with extent of reaction is zero – the reaction has reached equilibrium:

$$\left(\frac{dG}{d\xi} \right)_{T,P} = \sum_i^N (v_i \mu_i) = 0 \quad (8.7)$$

This equilibrium condition applies to all reactions involving gases, solids or liquids, including phase changes and interactions with surfaces. The expression in Equation 8.7 is termed the change in the reaction free Gibbs energy, with the symbol $\Delta_r G$. Substituting the expression for the chemical potential in Equation 8.3 gives:

$$\sum_i^N v_i \mu_i^0 = -RT \sum_i^N \ln a_{i,eq} \quad (8.8)$$

where $a_{i,eq}$ are the activities at equilibrium. A little manipulation of Equation 8.8 yields something that can be measured – the equilibrium constant:

$$\sum_i^N v_i \mu_i^0 = -RT \ln \prod_i^N a_{i,eq}^{v_i} \quad (8.9)$$

The equilibrium constant (although a little disguised) is defined as:

$$K = - \prod_i^N a_{i,eq}^{v_i} \quad (8.10)$$

Remembering that the stoichiometric coefficients are negative for reactants and positive for products, the equilibrium constant for the alphabetical reaction (Equation 8.5) is now the familiar:

$$K = \prod_i^N a_{i,eq}^{v_i}$$

$$K = \frac{[C]^c [D]^d}{[A]^a [B]^b} \quad (8.11)$$

The quantities of each reactant and product can be measured as a function of composition to determine the equilibrium constant. Finally, the most important

relationship between the change in standard Gibbs free energy of the reaction and the equilibrium constant gives:

$$\Delta_r G^\circ = -RT \ln K \quad (8.12)$$

A listing of thermodynamic properties determined by a full range of methods enables the $\Delta_f G^\circ$ values to be determined and hence the allowed reactions and equilibrium constants for all reactions. A tabulation of some thermodynamic quantities is found in Appendix C.

Example 8.1

The reality of using thermodynamics is somewhat simpler than the preceding derivations imply. Consider the reaction we have been using in the formation of ozone in the Chapman mechanism:



The enthalpy change for the reaction is given by the difference in the standard enthalpies of formation of the species, which can be found from the table in Appendix C:

$$\begin{aligned} \Delta_r G^\circ &= \sum_i^N v_i \mu_i^0 = (+1)\Delta_f G^\circ(\text{O}_3) + (-1)\Delta_f G^\circ(\text{O}) + (-1)\Delta_f G^\circ(\text{O}_2) \\ \Delta_r G^\circ &= 163.2 - 249.17 - 0 = -114.03 \text{ kJ mol}^{-1} \end{aligned} \quad (8.14)$$

Note that $\Delta_f G^\circ$ for O_2 is zero because this is the energy of formation of ozone in its standard state and is defined as zero. The $\Delta_r G^\circ$ for the reaction is negative and so the reaction is spontaneous. Using Equation 8.12, the equilibrium constant at 298 K can be calculated:

$$\begin{aligned} \Delta_r G^\circ &= -114.03 = -(8.314)(298) \ln K \\ \therefore K &= 1.05 \end{aligned} \quad (8.15)$$

The reaction favours the formation of ozone with a significant equilibrium constant. Appendix C also lists the enthalpies of formation and the standard enthalpy of the reaction $\Delta_r H^\circ$ can be calculated. The answer for the enthalpy calculation is $\Delta_r H^\circ = -106.47 \text{ kJ mol}^{-1}$, showing this to be an exothermic reaction, liberating heat. The entropy change at 298 K can also be calculated because $\Delta_r G^\circ = \Delta_r H^\circ - T\Delta_r S^\circ$, so $\Delta_r S^\circ = 25.4 \text{ J mol}^{-1} \text{ K}^{-1}$, indicating an increase in the entropy of the reaction as it proceeds by creating one molecule from two.

Once the standard Gibbs free energy is known for a reaction, then the effect of composition can also be considered by adding in the concentration-dependent

terms of the chemical potential, using Equation 8.7:

$$\Delta_r G = \sum_i^N v_i \mu_i^0 + RT \sum_i^N v_i \ln a_{i,eq} = \Delta_r G^0 + RT \ln Q \quad (8.16)$$

where Q is the reaction quotient for a chemical reaction under different compositions of reactants and products, defining the reaction quotient to be:

$$Q = \prod_i^N a_{i,eq}^{v_i}$$

$$Q = \frac{[C]^c [D]^d}{[A]^a [B]^b} \quad (8.17)$$

Equation 8.16 is the underlying equation of thermodynamics for all of chemistry involving mixing, such as in the test tube on the bench or in a prebiotic environment.

Example 8.2

Returning to the ozone formation reaction example, when the photolysis reaction is producing O atoms the $\Delta_r G$ for the reaction now changes due to the contribution from the reaction quotient. At a particular point the concentration of O atoms will reach that required by equilibrium and stop. If the O atom concentrations increase further the reaction is no longer spontaneous in the forward direction but reverses. What value of Q stops the reaction from being spontaneous?

$$\Delta_r G = 0 = 114.03 + (8.314)(298) \ln Q \quad (8.18)$$

$$Q = 0.955$$

The reaction quotient corresponds to an excess production of 5 per cent more O atoms than ozone molecules for the reaction no longer to be spontaneous.

It has been a long journey through the equations above and we could have stopped to examine all of the terms but some of the concepts are a little abstract whereas the final answer is useful and we shall now explore the utility of the results and then the wonders of the concepts that can develop with understanding.

Example 8.3

Consider what could be a promising primordial carbon fixation reaction:



The $\Delta_r G^0$ for the reaction, under standard conditions, is obtained by looking up the tabulated standard heats of formation and then using Equation 8.16 to give:

$$\begin{aligned}\Delta_r G &= \sum_i^N v_i \mu_i^0 = \Delta_r G^0 \\ \Delta_r G &= (+1)\Delta_f G(\text{CH}_3\text{OH}) + (-1)\Delta_f G(\text{CO}) + (-2)\Delta_f G(\text{H}_2) \\ \Delta_r G &= (+1)(-134.27) + (-1)(-155.41) + (-2)(0) \\ \Delta_r G &= +21.140 \text{ kJ mol}^{-1}\end{aligned}\quad (8.20)$$

The stoichiometric coefficient for methanol is +1 because it is a product and -1 for CO and -2 for H₂ because they are reactants. The reaction is not spontaneous under standard conditions of temperature and pressure but at 500 K the equilibrium constant is given by Equation 8.12:

$$K = \exp\left(-\frac{21140}{(8.314)(500)}\right) = 6.19 \times 10^{-3} \quad (8.21)$$

The equilibrium constant is small but not insignificant. The problem comes with the kinetics of the reaction and it does not proceed without a catalytic surface, even at 500 K. It remains, however, a good example to consider the extent of the reaction as a function of initial reactant concentrations.

Extent of chemical reactions

The reaction in Example 8.3 demonstrates the idea that some reagents may be limiting the reaction and so the reaction does not reach equilibrium. The extent of the reaction towards equilibrium, ξ , is a measure of the progress of the reaction. Consider that an equimolar concentration of CO and H₂ is passed over a catalyst at 500 K. The extent of the reaction can be determined by using Equation 8.4 and constructing Table 8.2.

The initial partial pressure of each gas in the equimolar mixture is 1 for each of the reactants and zero for the product. At the equilibrium extent ξ , $1 - \xi$ is left of CO and $1 - 2\xi$ is left of H₂: ξ amount of product has been formed. The

Table 8.2 The extent of the CO + H₂ reaction

	CO	H ₂	CH ₃ OH
Initial amount	1	1	0
Equilibrium amount	$1 - \xi$	$1 - 2\xi$	ξ
Mole fraction	$\frac{1}{2}$	$1 - 2\xi/2(1 - \xi)$	$\xi/2(1 - \xi)$

equilibrium constant can be calculated by:

$$K = \frac{4\xi(1-\xi)}{(1-2\xi)^2 1^2} = 6.19 \times 10^{-3} \quad (8.22)$$

Solving the quadratic equation in ξ gives the extent of the reaction $\xi = 0.00155$, from which the mole fractions of the components can be calculated. This highlights a real problem with prebiotic reactions: even if the equilibrium constants are known and the kinetics are favourable, the reaction may only produce small quantities of the required material simply because the concentrations of the reactants are too low.

Dissolved gases – Henry's Law

The pH of the oceans forming the primordial soup is important in controlling the charged nature, or otherwise, of the amino and carboxylic acid species and hence their chemistry. Generating reaction schemes for the prebiotic synthesis of molecules requiring basic conditions will not be relevant if the oceans are acidic. Consider dissolving CO₂ into water, simply written as:



The change in Gibbs free energy, ΔG , for this process is the difference between the solution and gas-phase chemical potentials. Using the chemical potential of each of the species in the gas phase and in solution and Equation 8.16 gives:

$$\Delta_r G = \Delta_r G^\circ + RT \ln Q \quad (8.24)$$

when the system is at equilibrium and the species are in their standard states: 1 atm CO₂ in the gas phase, 1 M solution, $Q = 1$ and $\Delta_r G = \Delta_r G^\circ$. The amount of gas in solution is given by Henry's Law and has (at least to a good approximation) an equilibrium constant given by:

$$P_i = K_i x_i$$

$$K_{\text{CO}_2} = \frac{P_{\text{CO}_2}}{x_{\text{CO}_2}} = 0.167 \times 10^8 \text{ Pa} \quad (8.25)$$

where K_{CO_2} is the partition coefficient, x_{CO_2} is the mole fraction and P_{CO_2} is the partial pressure of CO₂. The amount of dissolved CO₂ is given by:

$$K_{\text{CO}_2} = \frac{P_{\text{CO}_2}}{x_{\text{CO}_2}} = \frac{P_{\text{CO}_2}}{[\text{CO}_2]} \left([\text{CO}_2] + \frac{1000}{18.02} \right)$$

$$[\text{CO}_2]_{(\text{g})} = \frac{P_{\text{CO}_2}[55.5]}{0.167 \times 10^9} \text{ mol l}^{-1} \quad (8.26)$$

The concentrations of dissolved atmospheric gases in the oceans, whether of water or other liquids, can be calculated using the Henry's Law coefficients listed in

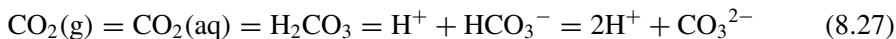
Table 8.3 Henry's Law coefficients for gases dissolved in water and benzene

Gas	K_i (10^9 Pa) for water at 25°C	K_i (10^9 Pa) for benzene at 25°C
H ₂	7.12	0.367
N ₂	8.68	0.239
O ₂	4.40	—
CO	5.79	0.163
CO ₂	0.167	0.0114
CH ₄	4.19	0.0569
C ₂ H ₂	0.135	—
C ₂ H ₄	1.16	—
C ₂ H ₆	3.07	—

Table 8.3. Notice how the polar solvent water supports significantly more dissolved gas than the non-polar solvent. However, gases dissolved in solutions go on to produce other compounds, most importantly protons, that determine the ocean pH or the pH of Darwin's 'little warm pool'.

pH of the oceans

Equation 8.26 predicts a concentration of CO₂ in one litre of water at 25°C of $3.32 \times 10^{-2} \text{ mol l}^{-1}$ at a pressure of 1 bar. The pH of the oceans is related to the amount of dissolved CO₂ but also to the equilibria controlling the formation of carbonic acid and the bicarbonate and carbonate ions:



The equilibrium constants for the formation of bicarbonate and carbonate are given by:

$$K_{a1} = \frac{[\text{H}^+][\text{HCO}_3^-]}{[\text{CO}_2]} = 4.45 \times 10^{-7} \quad \text{p}K_a = 6.35$$

$$K_{a2} = \frac{[\text{H}^+][\text{CO}_3^{2-}]}{[\text{HCO}_3^-]} = 4.69 \times 10^{-11} \quad \text{p}K_a = 10.33 \quad (8.28)$$

With the constraints of mass conservation so that all of the dissolved CO₂ is accounted for, no matter what form it takes, $c = [\text{H}_2\text{CO}_3] + [\text{HCO}_3^-] + [\text{CO}_3^{2-}]$ (typically CO₂(aq) + H₂CO₃ 1%, HCO₃⁻ 93% and CO₃²⁻ 6%) and enforcing an electroneutrality condition gives $2[\text{CO}_3^{2-}] + [\text{HCO}_3^-] + [\text{OH}^-] = [\text{H}^+]$, which sets up five equations with five unknown quantities and to a good approximation [H⁺] is given by:

$$[\text{H}^+] = \sqrt{K_{a1}c + K_{a2}c + K_w} \quad (8.29)$$

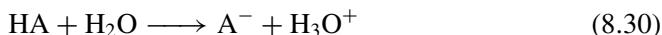
From the solubility of CO₂ calculated from Henry's Law, the pH of the oceans with one atmosphere of CO₂ is 3.95, falling to 3.81 with an increased temperature of 35°C.

The total volume of the world's oceans is 1.35 × 10³¹ l with a total mass of 1.4 × 10³¹ kg compared with a fresh water mass of 1.26 × 10¹⁷ kg. The pH of the oceans, however, is moderated into layers because the oceans are not well mixed. There are essentially three layers – the surface layer, the mixed layer and the deep layer – and the volume of the mixed layer is 2.7 × 10¹⁹ l to a mean depth of 75 m. The solubility of CO₂ enables the precipitation of CO₃²⁻ to be estimated, from which an estimate (although it is fraught with approximations) of the total inorganic carbon in the world's mixed layer is 3.91 × 10¹⁶ kg.

The chemical composition of the atmosphere and of course the CO₂ fixing mechanisms within the oceans mean that the pH of the oceans is nothing like as acidic as pH 4 but is fairly efficiently buffered within the range 7.8–8.4, making it slightly alkaline. The present indicative CO₂ level in the atmosphere is 280 ppm or 0.028 per cent but this was perhaps 1000 times greater in the Archean atmosphere, importantly for the global warming environments. Repeating the calculation above with 5 per cent CO₂ gives a solution pH of 4.8. It would be very tempting to suggest that prebiotic oceans were considerably more acidic than present but the dissolved CO₂ concentration is only part of the story. The effect of mineral composition also must be considered but is hard to detail. What is clear, however, is that the primordial oceans would have acted as considerable buffers to the pH, and life may have had to develop in harsher pH conditions.

Acid–base equilibria

The application of the equations of thermodynamics is more straightforward than it might appear, as demonstrated by acid–base equilibria. Consider the simple dissociation of an acid dissolved in water:



with the equilibrium constant defined as:

$$K = \frac{a_{\text{A}^-} a_{\text{H}_3\text{O}^+}}{a_{\text{HA}}} = \frac{c_{\text{H}^+}^2}{c_{\text{Total}} - c_{\text{H}^+}} \quad (8.31)$$

The activities a_i , of dilute solutions are simply the concentrations of the solutes and the equilibrium constant can be used to determine the pH of a solution when a known amount of acid is dissolved in water. The proton concentration and hence pH is given by the solution of the general quadratic:

$$c_{\text{H}^+} = -\frac{1}{2}K + \sqrt{c_{\text{Total}}K + \frac{1}{4}K^2} \quad (8.32)$$

The degree of dissociation, α , of the acid is given by:

$$\frac{c_{H^+}}{c_{Total}} = \frac{-\frac{1}{2}K + \sqrt{c_{Total}K + \frac{1}{4}K^2}}{c_{Total}} \quad (8.33)$$

From this equation it is easy to calculate the charged nature or otherwise of a species in solution. However, in the context of a large buffered ocean at fixed pH the degree of dissociation is controlled by the buffer pH. For example:

$$\log\left(\frac{c_{A^-}}{c_{HA}}\right) = pH - pK$$

$$\alpha = \left(\frac{c_{A^-}}{c_{Total}}\right) = \frac{c_{A^-}}{c_{HA} + c_{A^-}} = \frac{1}{1 + 10^{pK-pH}} \quad (8.34)$$

The objective of the preceding equilibrium calculation has been to determine the state of a molecule such as an amino acid in the conditions that prevailed on the early Earth. The pH, degree of dissociation and the extent of the reaction all have a direct effect on the population of the species present. Temperature and cooperative effects have not been considered but serve to complicate the problem. Any prebiotic reaction scheme must take account of that troublesome restriction to chemistry – the second law of thermodynamics.

8.3 Rates of chemical reactions

An example of an allowed chemical reaction is the hydrolysis of the dipeptide glycylalanine (Figure 8.1), an example of peptide bond formation present in the 1000s of amino acids making up proteins.

From the standard thermochemical data $\Delta_r G^\circ = (-371.3 - 379.9 + 733.9)$ kJ mol⁻¹ = -17.3 kJ mol⁻¹, corresponding to an equilibrium constant $K = 1.1 \times 10^3$ M⁻¹. This is a worrying result because all peptides in solution at 298 K should spontaneously fall apart to the monomers and hence all proteins are subject to degradation due to spontaneous hydrolysis. Fortunately, the reaction is kinetically hindered, which means that it occurs very slowly. Kinetics always control the rate at which equilibrium is achieved, relating the ratio of the forward and backward rate constants to the equilibrium constant:

$$K = \frac{k_{forward}}{k_{backward}} \quad (8.35)$$

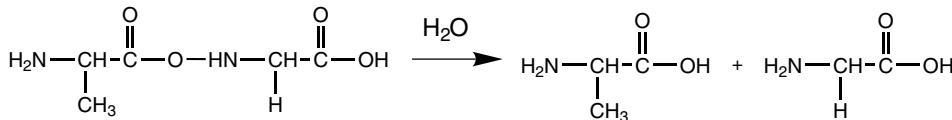


Figure 8.1 Hydrolysis of the dipeptide glycylalanine

The variation of reaction rate with temperature follows the Arrhenius equation, which we have used to study the rate of chemical reactions in the interstellar medium ISM (Section 5.4, Equation 5.9), and can be applied to the liquid phase or reactions occurring on surfaces. Even the smallest increases in temperature can have a marked effect on the rate constants, as can be seen in the increased rate of chemical reactions at body temperature over room temperature. Considering a reaction activation energy that is of the order of a bond energy, namely 100 kJ mol^{-1} , the ratio of the rate constants at 310 K and 298 K is given by:

$$\frac{k(310) \propto 310^{0.5} \exp\left[-\frac{100 \times 10^3}{(R \ 310)}\right]}{k(298) \propto 298^{0.5} \exp\left[-\frac{100 \times 10^3}{(R \ 298)}\right]} = 4.9 \quad (8.36)$$

Chemistry within the body is approximately five times faster than in a test tube at room temperature. The reverse is true, of course, with chemical reactions in liquid methane at 100 K some 1.2×10^{35} times slower than at 298 K. Neutral chemical reactions remain slow in solution at 100 K if they have a significant activation barrier. As with the ISM, chemistry involving breaking of chemical bonds is frozen out at 100 K and has direct implications for chemistry on the surface of Titan, for example.

8.4 Endogenous production of organic molecules

Armed with the combined weapons of thermodynamics and kinetics, we can now look at possible reaction mechanisms in prebiotic chemistry and begin to assess how likely or unlikely the origins of life really were. The chemistry of the prebiotic Earth may have been responsible for generating organic molecules spontaneously; this process is called endogenous production of organic molecules. The starting materials must be derived from the Archean atmosphere or more generally the Archean biosphere. Theories about the origin and evolution of the atmosphere focus on two sources for the gases and volatile materials seen at present: an *internal reservoir* requiring a source of these compounds from within the materials making up the planetesimal; and an *external reservoir* suggesting comets and meteorites as the origin of the biosphere. The origin of the Moon adds a further dimension to this argument. The hypothesis of a collision between the Earth and a Mars-size object resulted in the vaporisation of the collision partner, condensing to form the Moon together with the evaporation of the oceans, atmosphere and much of the surface of the Earth. The biosphere would then have to be replaced after such an event.

The period of emergence of life on Earth is constrained to be between the period 4.0–3.7 Gyr ago, for which there is no fossil record. Urey postulated that all of the planets formed from the same solar nebula and so the early Earth should have an atmosphere with a composition the same as that of Jupiter (known at the time),

i.e. H₂, CH₄ and NH₃ – a reducing atmosphere. Urey and his graduate student Miller set about some interesting experiments to explore the consequences of this. Another geologist suggested that the gases in the early atmosphere resulted from volcanic activity and would be a more oxidising atmosphere of CO₂, N₂, H₂O and lesser amounts of CO and H₂.

The current model for the prebiotic atmosphere is $p(\text{CO}_2)$ 0.2 bar compared with 3.6×10^{-4} bar at present. The majority of the atmosphere was thought to be inert nitrogen and 1 per cent water. Other trace gases such as H₂ and CO were also present. The altitude profile of these species as a result of photochemistry will change, including the formation of low concentrations of O₂ by water photolysis; all of the atmospheric oxygen at present is a product of photosynthesis. The large partial pressure of CO₂ is sufficient to overcome the ‘faint young Sun’ luminosity deficit. Current models of the Sun suggest that its luminosity may be some 20–30 per cent less than at present, which, without significant global warming, would have resulted in the freezing of the Earth’s oceans. The predicted CO₂ levels would have prevented this.

Urey–Miller experiments

Endogenous production of organic molecules from the primordial atmosphere requires a knowledge of the energy sources in the atmosphere. The largest energy source in the atmosphere is the Sun’s radiation, which would have been dominated by wavelengths shorter than 300 nm (no short-wavelength shield). Bond photolysis would then be possible at nearly all altitudes, producing radical species and molecules in excited electronic states. These highly reactive species then form a complex chemical reaction network anywhere and everywhere within the pre-biosphere. The second source of energy in the atmosphere comes from electrical discharges such as lightning. The final significant source of energy is volcanic activity. Lightning strikes produce a breakdown in the air around them, generating a plasma and producing a cocktail of molecules that could initiate prebiotic synthesis.

Miller performed the first prebiotic synthesis experiments with simulated lightning strikes, which produced many prebiotic building blocks, including amino acids and purines. The experimental layout is shown in Figure 8.2. An electrical discharge was struck in a primordial atmosphere that was strongly reducing and contained CH₄, NH₃, H₂O and H₂. The result was a large yield in amino acids (all 20 naturally occurring ones), hydroxyacids, short aliphatic acids and urea. A relatively small set of molecules was produced in high yield rather than a completely random mixture – all of them biologically interesting. A typical organic yield is shown in Table 8.4.

Estimates of the total lightning energy contribution to the atmosphere are placed at 1.8×10^{22} – 1×10^{24} eV m⁻² yr⁻¹. From the experiments it is possible to work out a yield of all amino acids per electron-volt passed into the experiment. The calculation for glycine is 2×10^{-5} molecules eV⁻¹ for a methane environment and

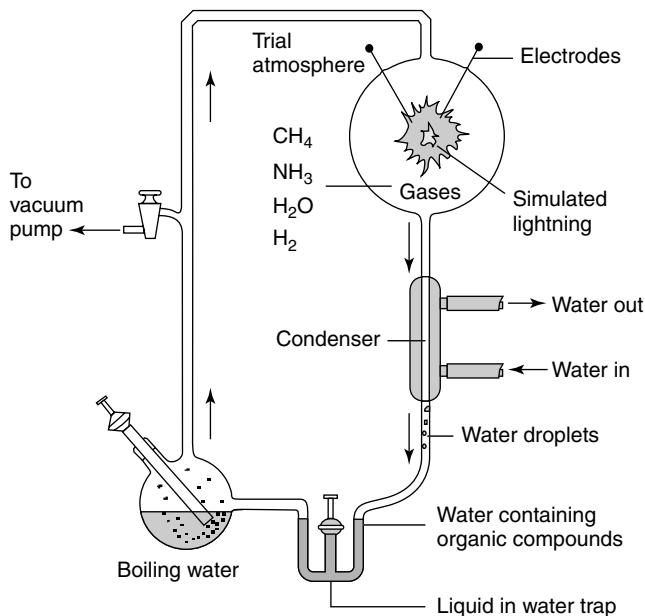


Figure 8.2 Urey–Miller experiments

Table 8.4 Typical yield of a Urey–Miller experiment (Miller (2000))

Compound	Yield (%)	Compound	Yield (%)
Formic acid	4.0	α -Hydroxybutyric acid	0.34
Glycine	2.1	Succinic acid	0.27
Glycolic acid	1.9	Sarcosine	0.25
Alanine	1.7	Iminoacetic-propionic acid	0.13
Lactic acid	1.6	<i>N</i> -Methylalanine	0.07
α -Alanine	0.76	Glutamic acid	0.051
Propionic acid	0.66	<i>N</i> -Methyl urea	0.051
Acetic acid	0.51	Urea	0.034
Iminodiacetic acid	0.37	Aspartic acid	0.024
α -Amino- <i>n</i> -butyric acid	0.34	α -Aminoisobutyric acid	0.007

10^{-8} molecules eV $^{-1}$ for a CO atmosphere. At this rate the production of glycine would have been 0.3–20 nmol m $^{-2}$ per year. This figure may be compared with the energy from cosmic rays based on estimates of the flux of these particles during the early period of the Earth's formation. The cosmic ray energy flux is estimated as 2.9×10^{21} eV m $^{-2}$ yr $^{-1}$, giving rise to 10 μ mol m $^{-2}$ yr $^{-1}$ of glycine for the same equimolar mixture of CO and N $_2$. Both of these may be compared with the energy flux from the Sun at 6.8×10^{28} eV m $^{-2}$ yr $^{-1}$, although no glycine was detected without a nitrogen source such as NH $_3$. Extreme photon energies of $\lambda < 110$ nm would allow N $_2$ photolysis but the estimated glycine yield from this is 0.5 μ mol

$\text{m}^{-2} \text{ yr}^{-1}$ as an upper limit, so cosmic rays appear to be the most important factor in generating glycine. This has implications for precursor formation in Earth and Mars and glycine synthesis in the ISM.

Strecker synthesis

Amino acid formation in the Urey–Miller experiment and almost certainly in the prebiotic environment is via the Strecker synthesis shown in Figure 8.3. This reaction mechanism shows that the amino acids were not formed in the discharge itself but by reactions in the condensed water reservoir. Both HCN and HCO are formed from the bond-breaking reactions of N_2 and H_2O in a plasma, which then react with NH_3 in solution. The C=O group in formaldehyde or other aldehydes is replaced by NH and this undergoes a reaction with HCN to form the cyano amino compound that hydrates to the acid. The Strecker synthesis does not provide stereo-control over the carbon centre and must result in racemic mixtures of amino acids. There is no room for homochirality in this pathway.

If this is an important production reaction for amino acids, what concentrations are required globally to generate the amino acid budget and should this represent a global concentration within the oceans? The concentrations of HCN and NH_4^+ are limiting in this process and HCN in particular is familiar from comet chemistry. Amino acids seem to be synthesised rather easily in the chemistry of a reducing atmosphere or from exogenous delivery.

Changing the initial conditions in the Urey–Miller experiment to favour a primordial atmosphere with CO_2 as the primary carbon source does not produce such a rich mixture of prebiotic molecules and there is a significantly lower yield of amino acids: an argument against an endogenous source of organic molecules.

Purine and pyrimidine synthesis

The organic bases forming the Watson–Crick base pairs in RNA and DNA are the purines adenine (A) and guanine (G) and the pyrimidines thymine (T), cytosine (C) and uracil (U) (Figure 8.4). The base pairs in DNA are A–T and G–C and in RNA they are G–U and A–C.

It has been discovered that adenine can be synthesised by the polymerisation of five HCN molecules in a truly remarkable synthesis. The HCN is produced either photochemically or within a lightning discharge and when dissolved in water may

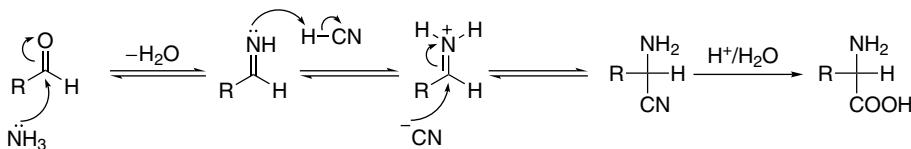
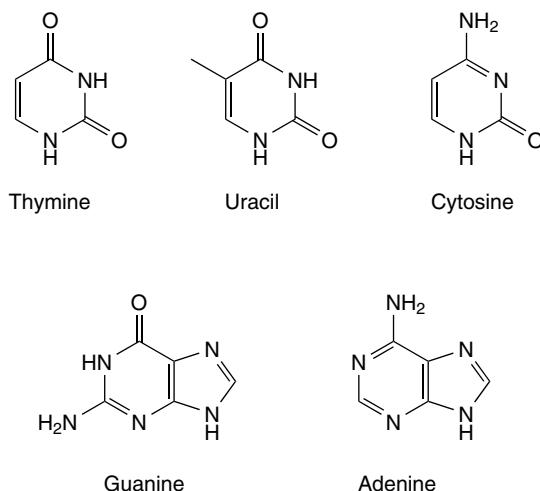
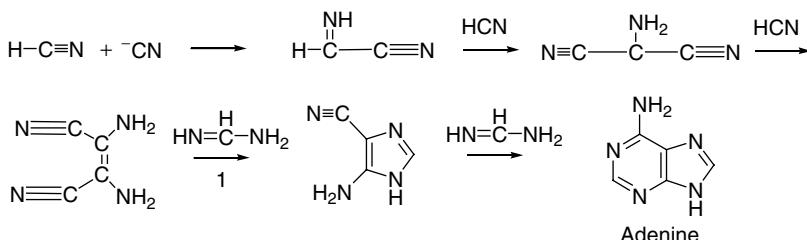


Figure 8.3 Strecker synthesis of amino acids

**Figure 8.4** Bases in DNA and RNA**Figure 8.5** Polymerisation of HCN

then polymerise according to the probable reaction scheme in Figure 8.5. The hardest step is step 1, which may be photochemically assisted. If solutions of ammonium cyanide are refluxed for a few days then the yield of adenine is of order 0.5 per cent. Concentrations of the intermediates are intrinsically rather small but a concentrating mechanism such as the ‘little warm pool’ scenario would provide regular concentration and replenishment of reactants. The other purine, guanine, along with a whole host of other purines not used in biology, can be synthesised from variations of the above aminoimidazole and carbonitrile intermediates.

The synthesis of the pyrimidines involves cyanoacetylene, which is synthesised in good yield from discharge reactions of CH_4 and N_2 . The reaction with cyanoacetylene or cyanoacetylaldehyde, in a concentrating environment, produces cytosine and uracil according to Figure 8.6.

Conspicuously missing from this mixture is thymine, which led to the suggestion that RNA preceded DNA in what has become known as the ‘RNA World’ hypothesis.

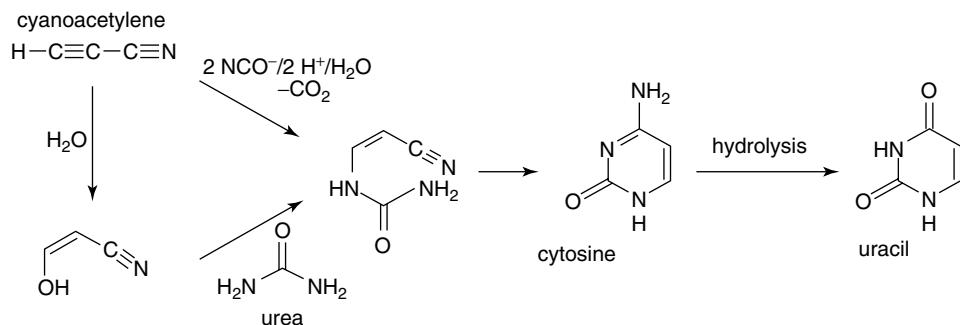


Figure 8.6 Prebiotic synthesis of cytosine and uracil

Formose reaction

The formation of sugars from the reaction of formaldehyde under alkaline conditions was discovered in 1861 and is known as the ‘formose reaction’, although it is not understood fully (Figure 8.7). It requires the presence of suitable inorganic catalysts such as $\text{Ca}(\text{OH})_2$ or CaCO_3 , either of which may be found on a prebiotic Earth. The reaction is autocatalytic and produces over 40 different types of sugars, some rings, some long chains.

The formose reaction leads to the production of many sugars with the exception of ribose, which occurs in very low concentrations and it is hard to see how this could be increased. A variation on the formose reaction starting at the phosphorylated glycoaldehyde intermediate, with the addition of base or in the presence of mineral substrate, produced a surprisingly large amount of ribose but this still

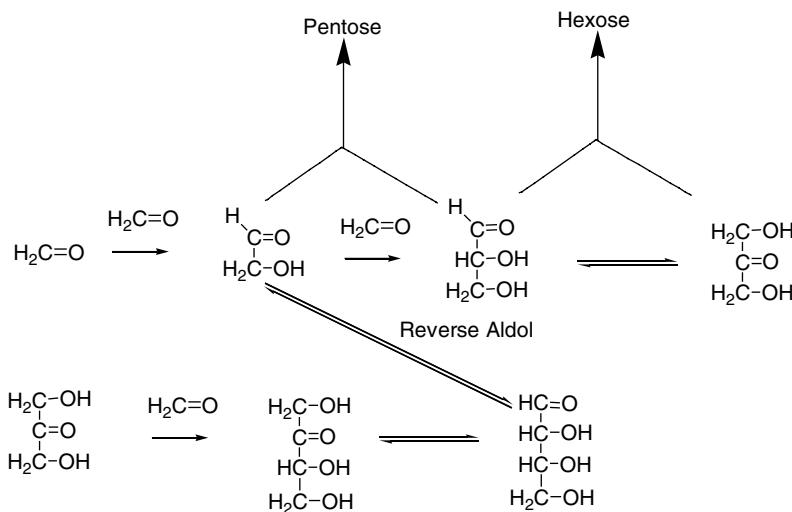


Figure 8.7 The formose reaction

remains a serious problem. Once formed, ribose would have a half-life of less than 73 min at 100°C or 44 years at 0°C. Yet somewhere in the past it became intrinsically and fundamentally linked to the structure of RNA and DNA. A biosynthetic pathway for ribose seems to be high on the list of reactions required for life.

Inorganic phosphate

The last component on the way to forming a large oligonucleotide such as RNA is the presence of inorganic phosphate. Inorganic phosphate, P_i , is the sum of the concentrations $H_2PO_4^-$ and HPO_4^{2-} and should precipitate in nearly all prebiotic mixtures in the form of calcium phosphate. Consequently, there is little or no P_i to be incorporated in the oligonucleotides (Figure 8.8), despite playing an important role in energy storage and conversion in metabolism.

The synthesis of the nucleotides requires both a source of phosphate and a means of activating the phosphate group so that phosphate esters can be formed. There are three possible sources of phosphorus on the early Earth as well as from meteorites (Figure 8.9).

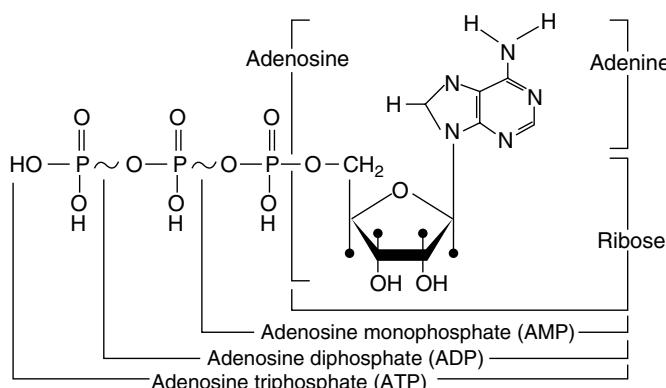


Figure 8.8 Structure of adenosine, AMP, ADP and ATP

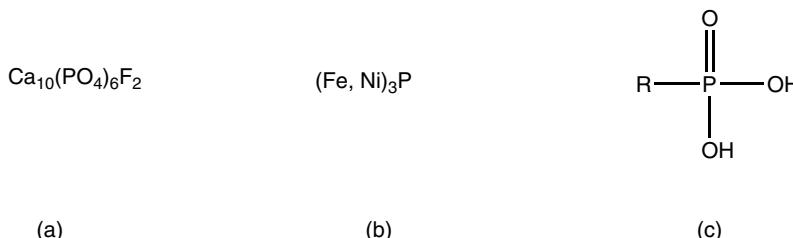


Figure 8.9 Sources of phosphorus: (a) fluorapatite; (b) the mineral schreibersite; (c) alkyl phosphonic acids

Fluorapatite is the only significant phosphorus-containing mineral in the Earth's crust and schreibersite has been found in iron meteorites. The only organic species to be found containing phosphorus in meteorites are the alkyl phosphonic acids. These are at least promising even if they do not contain the P–O–P phosphoester bond unit.

Moderately simple syntheses have been performed for the purines cytosine and uracil but nothing seems to work as a prebiotic synthesis of the pyrimidines. Then adding the sugar ribose to the base makes them nucleosides and one phosphoric acid residue makes it a nucleotide, or specifically a mononucleotide: a rare but curiously important sequence of events in present-day life but perhaps not for prebiotic chemistry and early life forms.

The Ribose riddle

The major missing component for the formation of RNA as an information-bearing molecule is the distinct lack of ribose in any particularly large quantities. The backbone of RNA and DNA has the phosphoester linkage linking the nucleosides – the bases linked with ribose, as seen in Figure 8.10a.

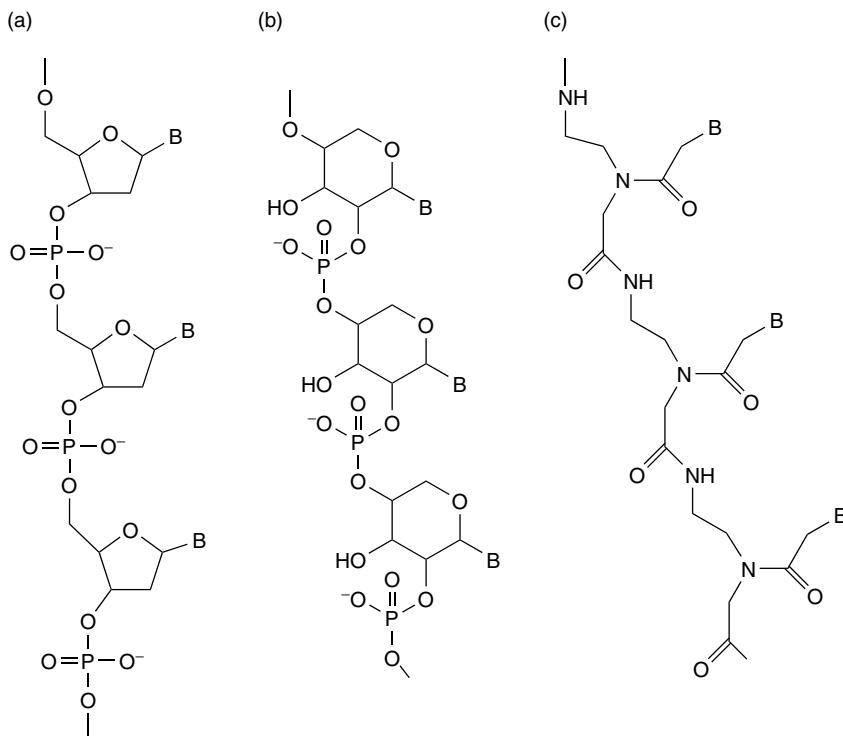


Figure 8.10 DNA and potentially information-bearing oligonucleotide analogues: (a) DNA; (b) pyranosyl analogue of RNA; (c) peptide nucleic acid, PNA

The low yield of ribose in the formose reaction and the low concentrations of phosphorus-containing compounds have led to the search for other molecules to replace the phosphorus–sugar backbone chain in DNA or RNA. The first variation simply replaces the ribose with pyranose, a much larger component of the formose product reaction (Figure 8.10b). The pyranosyl-RNA molecule, however, does not eliminate the phosphorus dependence. Addressing both the ribose and phosphate problem is a peptide-linked template, PNA. This molecule (Figure 8.10c) was initially designed as a potential anticancer agent and in common with the pyranosyl group is capable of template-driven autocatalysis and has been observed with an achiral peptide backbone. An achiral molecule such as PNA could form either left- or right-handed helices and take part in chiral amplification, although some studies have now shown that the different chiral structures inhibit one another. A switch at some stage in development would lead to the chiral selectivity seen in DNA and RNA.

8.5 Exogenous delivery of organic molecules

We have seen from the discussion on meteorites and comets that in a large selection of processes molecules can be prepared on the surface of dust grains, collected within meteorites or within the coma of a comet. The delivery of organic material during the period of heavy bombardment in the Hadean and less dramatically in the Archean could account for a large amount of the carbon inventory on the planet. The survival of the molecular diversity during vaporisation on entry into the atmosphere or on collision on the ground is a serious problem. A high-energy impact would result in the cometary material becoming soot in the impact fireball and so not contributing to prebiotic synthesis.

Some consideration of the quantity of organic material is required. Assuming the impacting object to be the Murchison meteorite, then 1.8 per cent of the incoming mass would be organic material. Most of this, however, is the kerogen-like polymer, leaving less than 100 parts per million as amino acids. An example of this type of impact is stored in the geological record when 65 million years ago the Cretaceous–Tertiary boundary is marked by the collision of a 10-km-diameter object leaving a clear carbonaceous layer at the Earth. The amino acid concentration in the layer however is rather small. Comets may be more promising with a typical inventory of 80 per cent water, 1 per cent HCN, 1 per cent H_2CO as well as CO and CO_2 , CH_4 , etc. A 1-km comet would contain about 2×10^{11} moles of HCN, which if distributed over the entire Earth's surface of, $5 \times 10^{18} \text{ cm}^2$ gives a concentration of order 40 nmol cm^{-2} – rather low. The argument for a concentrating mechanism is compelling although it is unlikely that all of the material would be distributed uniformly over the entire surface of the Earth. The simple sum requires all of the HCN to survive the impact, whereas temperatures as low as 100°C would result in hydrolysis to formate.

The in-fall of dust is also a possible source of carbon. The present rate of in-fall is $5 \times 10^6 \text{ kg per year}$ or $8 \times 10^{-9} \text{ g cm}^{-2} \text{ yr}^{-1}$. All of the above calculations

regarding delivery rates may have been 10^2 – 10^6 times greater during the Hadean and Archean periods of the Earth's formation. These may be important factors if the Archean atmosphere was much less reducing than the Urey–Miller experiments require.

8.6 Homochirality

It is not clear from any of the observations made above on exogenous or endogenous organic molecule delivery where there might be a mechanism for producing one type of molecule over another. Unfortunately, life made such a choice at some stage to produce an important biomarker – homochirality. Louis Pasteur discovered chirality in 1848, noticing that crystals of sodium paratartrate had different structures and each crystal type rotated the plane of polarised light. The rotation of the plane of the electric field of light is termed optical activity and requires a unique three-dimensional structure of a molecule in the same way as the left and right hands are mirror images of one another. The handedness of an amino acid produces two enantiomers that are mirror images of one another (Figure 8.11). The tetravalent nature of carbon leads naturally to the asymmetry and chirality, with the carbon atom at the chiral centre. Each of the groups around the carbon atom must be different and for amino acids this must contain the COOH acid group, the amino NH₂ group and the group R representing all other functional groups. There are many different naming conventions for the different enantiomers but the only measurable quantity is the rotation of the plane of polarised light, giving the L configuration in amino acids. Any sp³ carbon atom can potentially be a centre for chirality, a stereocentre, and in the case of sugars this is the D configuration. The original L and D configurations refer to the rotation of the plane of polarised light to the left (*leavo*) or to the right (*dextro*); other conventions have confused the matter but need not detain us here.

Homochirality is the choice in Nature for one handedness or another. Macroscopic homochirality can be observed in the winding of some species of snail shell

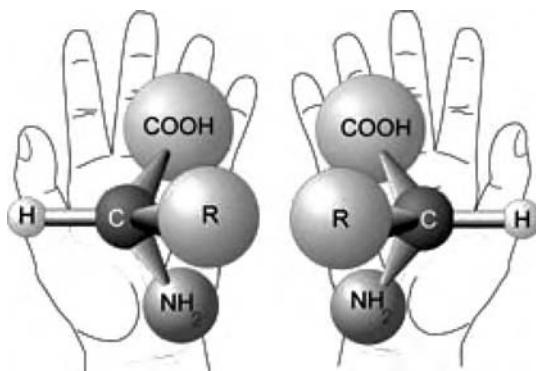


Figure 8.11 Amino acid chirality. (Reproduced with permission from NASA, Ames Research Center)



Figure 8.12 Homochiral snail shell. (Reproduced by permission of Steve Crompton)

(Figure 8.12), the coiling of plants and, unbelievably, the chewing motion of cows, which shows a 10 per cent enantiomeric excess! More importantly, homochirality is fundamental to the lock-and-key three-dimensional nature of interactions in biology. Specific binding of one enzyme to only one substrate requires the correct bonding structure within the binding site of an enzyme in the same way that a right hand must go into a right-hand glove. It was initially postulated that homochirality is the product of living systems and would thus be an important biomarker. Certainly, the lock-and-key amplification of homochirality would favour one enantiomer over another but the question of which-one-first remains. Homochirality would appear to be a necessary prerequisite for life on Earth and the biology we see but this does not have to be a restriction for astrobiology.

There have been two possible general mechanisms proposed for the production of an enantiomeric excess (EE) for molecules: parity violation in the electroweak force and dichroic photodissociation. The former comes from a prediction on the unified model for the electroweak force of Nature that suggests there is an intrinsic lower energy for the L-form of a molecule due to the interactions of the electrons within molecular orbitals. The small energy difference allows for a 1 part in 10^{17} EE in favour of the L-form. More plausible is the dichroic absorption of radiation by molecules. Chiral molecules have a preference for absorbing right or left circularly polarised light, called dichroism, so the rates of photolytic destruction would be different. Experiments in the laboratory have shown a small EE for the photolysis of amino acids in solution at 213 nm (Figure 8.13). The EE is very small, of the order of a few per cent, and the EE is only seen when most of the total amino acid present has been photolysed.

For the dichroic photolysis mechanism to be successful, amino acids must be synthesised and destroyed in an intense circularly polarised radiation field. Day-light shows little or no excess but recent observations at $2.2\ \mu\text{m}$ of the Orion reflection nebula OMC-1 shows polarisations in excess of 17 per cent, although

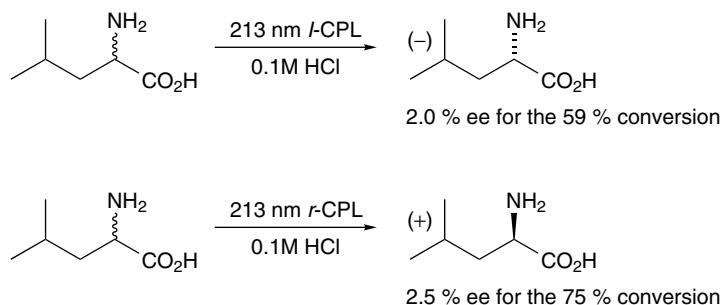


Figure 8.13 Asymmetric photolysis of amino acids

this has not yet been observed in the UV wavelength 200–230 nm required for photolysis. The excess derives from scattering of starlight by irregularly shaped dust particles that show a dichroic scattering. The radiation would then have to destroy one enantiomer over another. The theory gained some exciting support recently with the discovery of an enantiomeric excess of 9 per cent in amino acids extracted from the Murchison meteorite, suggesting an EE within the solar nebula. The exogenous delivery of amino acids from the solar nebula and the dichroic scattering of dust particles may then deliver an EE to the prebiotic Earth. It remains, however, to explain how biology has amplified one enantiomer into a living organism.

The discovery of homochirality on a planet such as Mars *could* be an excellent biomarker and strengthen the argument for life on Mars. With an EE in the solar nebula there should be an EE on the surface of Mars of order 9 per cent but remains of ancient life on Mars would show a greater excess. The interchange of enantiomers occurs naturally in a process called racemisation and for the most labile amino acid, aspartic acid, the half-life for the racemisation is 800 years at 300 K: in 800 years, half of the non-biotic aspartic acid would racemise and the EE would go to zero. In dry conditions, however, the half-life is much longer, perhaps as large as 5×10^4 years at 300 K. Extrapolation of the racemisation rate to 215 K, the equatorial temperature of Mars, extends the half-life further to 3×10^{12} years and to 10^{27} years at 150 K, Martian polar temperatures. Hence, discovery of a considerable EE in the Martian soil would be a strong indicator of ancient Martian life.

Achieving homochirality requires the polymerisation of a structurally unique already homochiral collection of chemicals or some pre-selection based on chirality. This would suggest a catalyst, perhaps a surface of some form. The chemical potential for different enantiomers on surfaces is different for some chiral surfaces such as quartz crystals where homochirality binding preferences are observed. Unfortunately, microcrystals of each of the chiral surface of quartz would have been present in the same local area, amplifying the enantiomers and destroying the EE. The role of chemical reactions on surfaces in a prebiotic Earth has generated many ideas, culminating in the concept of surface genes and organisms.

8.7 Surface Metabolism – ‘clay organisms’

All reactions in cells are catalysed by enzymes, which increase the rate and diversity of chemical reactions and the complexity of molecules inside the cell compared with the building blocks outside. It is this set of self-replicating reactions that points to a chemical definition of life. The source of the first set of catalysts is, however, not known but some have postulated primitive catalysis on the surface of minerals. This would have been a source of the first inorganic templates for replication and even perhaps an inorganic gene. The surface of minerals such as silica (Figure 8.14) becomes negatively charged in water and, depending on the pH of the solution, develops a surface potential of order -120 mV. The negatively charged surface then attracts positive ions as counter ions to balance the negative charge. The positive ions near the surface have an enhanced concentration over the bulk due to the electric field and this is given by the Boltzmann equation:

$$n_s = n_0 \exp\left(\frac{-qe\varphi_s}{kT}\right) \quad (8.37)$$

where n_s is the concentration of ions at the surface, n_0 is the bulk concentration, q is the charge on the ionic species (+1, -1, etc.), φ_s is the potential at the surface, k is the Boltzmann constant and T is the temperature. The layer next to the surface is perhaps a few nanometres in dimension, cell-like dimensions, and the surface potential decays exponentially from the surface. The ionic concentration at the surface for silica is enhanced some 100 times for a fully charged silica surface. The nature of the chemistry, the extent of reaction and the chemical equilibria

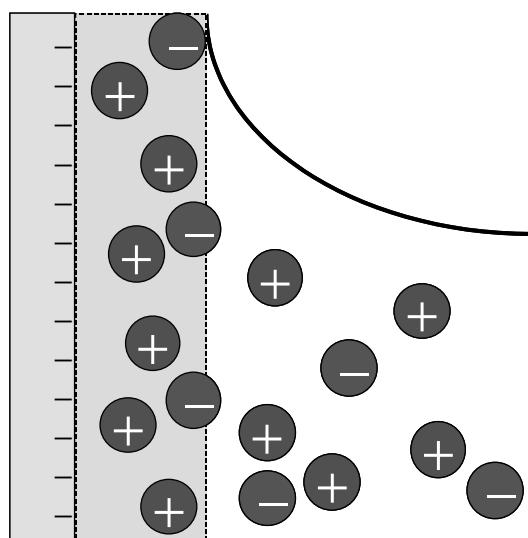


Figure 8.14 Bilayer structure for clay surfaces

are now very different close to the surface compared with the bulk – a unique collection of chemical reactions localised to the mineral surface.

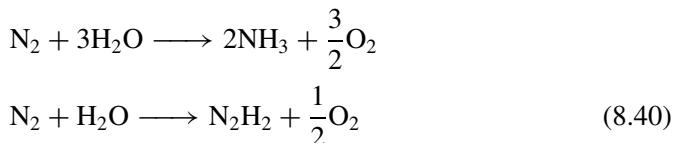
The surfaces also have other catalytic properties, especially associated with the absorption of photons with the resulting production of electrons or radical species. In many mineral structures the electrons can make transitions between the valence band and the conduction band by absorbing visible light. The gap between the levels determines the wavelength of the radiation absorbed, the smallest gap being for TiO₂ with an excitation wavelength of 413 nm. Absorption of light creates an oxygen-surface-active site, which is highly reactive and capable of initiating radical chemistry with a very low activation barrier. Decomposition of H₂O in the presence of CO₂/CH₄ results in the production of OH radicals and formaldehyde H₂CO – the starting material for the Strecker synthesis. Photoconversion of methane produces the methyl radical:



and then radical–radical recombination makes gaseous ethane CH₃CH₃. The carbon chain length increases by repeated radical insertion processes on the surface to produce increasingly longer polymers – perhaps information-bearing molecules. A rich cocktail of compounds results, performing useful reactions such as carbon fixation from CO₂ (Equation 8.39) on Si/TiO₂ surfaces and AgCl/zeolites when irradiated with UV radiation.



Similar N₂ fixation has been seen with (H₂O + N₂)/TiO₂ and other metal-doped mineral surfaces, producing:



The extension to amino acid synthesis on surfaces has also been observed with platinised TiO₂ surfaces, with CH₄, NH₃ and H₂O solutions producing glycine, alanine, glutamine, aspartic acid and serine, all photosynthesised and trapped in the resulting layers near to the surface by their charge.

Taking this one step further, perhaps even an ‘inorganic gene’ may have been provided by clay mineral sources. Earliest clay samples are of a mineral called montmorillonite that consists of sheets of aluminosilicates in which Fe²⁺, Fe³⁺ and Mg²⁺ are substituted for some of the Al³⁺, and Al³⁺ is substituted for Si⁴⁺. The oxygen content of the layers does not change and the alternative valencies allow the production of positive and negatively charged layers. Dramatically, Paecht-Horowitz and co-workers showed that the amino acid adenylate could be polymerised with up to 50 units on the montmorillonite surface in aqueous solution. Similar condensation reactions for carbohydrates on hydrotalcite surfaces have

also been observed, producing polymer chains of carbohydrates that are more specific than the rather random formose reaction. Condensation reactions have been observed for nucleotide polymerisation with the correct 3'-5' coupling but only with activated nucleotide bases.

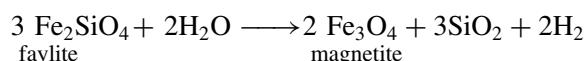
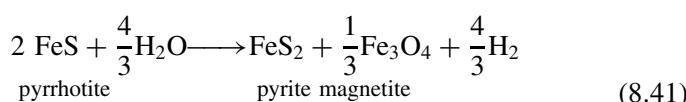
8.8 Geothermal Vents – ‘black smokers’

The current models of the Sun suggest that its luminosity would have been some 20–30 per cent lower than its present value during the early part of the formation of the Earth. After the enormous temperatures of the Hadean period, the early precambrian may have been cooler, requiring prebiotic chemistry to occur below a layer of ice, perhaps heated by volcanic activity such as that found in geothermal vents. A layer of ice several hundreds of kilometres thick may have formed over the entire surface of the early Earth, providing protection from UV radiation and some global warming – conditions such as these may exist on the Jovian moon Europa.

Volcanic geothermal vents provide a warm environment in which chemistry would occur quickly; there would be a rich diversity of molecules, mineral surfaces and molecules in reduced states (Figure 8.15). The reducing nature of this environment provides electrochemical energy for organic synthesis and carbon fixation, replacing the photochemistry and lightning strikes of the Urey–Miller experiments. The first vents were discovered in 1997, containing hot water conditions with temperatures up to 500 K and a new collection of previously unrecognised bacteria.

Seawater is circulated below the sea bed where it is heated by volcanic activity before being re-injected into the sea at high pressure and temperature. The pressure prevents water from boiling until the temperature reaches 725 K. The superheated water dissolves minerals from around the vent that then precipitate as the water temperature cools. This gives the vents their ‘black smoker’ appearance. Many small molecules such as H_2 , H_2S and Mn^{2+} do not precipitate but remain in their reduced formed and are available as electron donors.

The minerals dominating the sea bed that determine the redox state of the environment are pyrite (FeS_2), pyrrhotite (FeS) and magnetite (Fe_2O_3), all with active charged surfaces capable of being templates and supporting chemical reactions. This group of minerals are called collectively the PPM mineral assemblage. The temperature dependence of hydrogen evolution from the surfaces of the PPM minerals in water controls the chemistry of the simple species. The hydrogen-evolving reactions active on the mineral surface are responsible for the redox state of the vent:



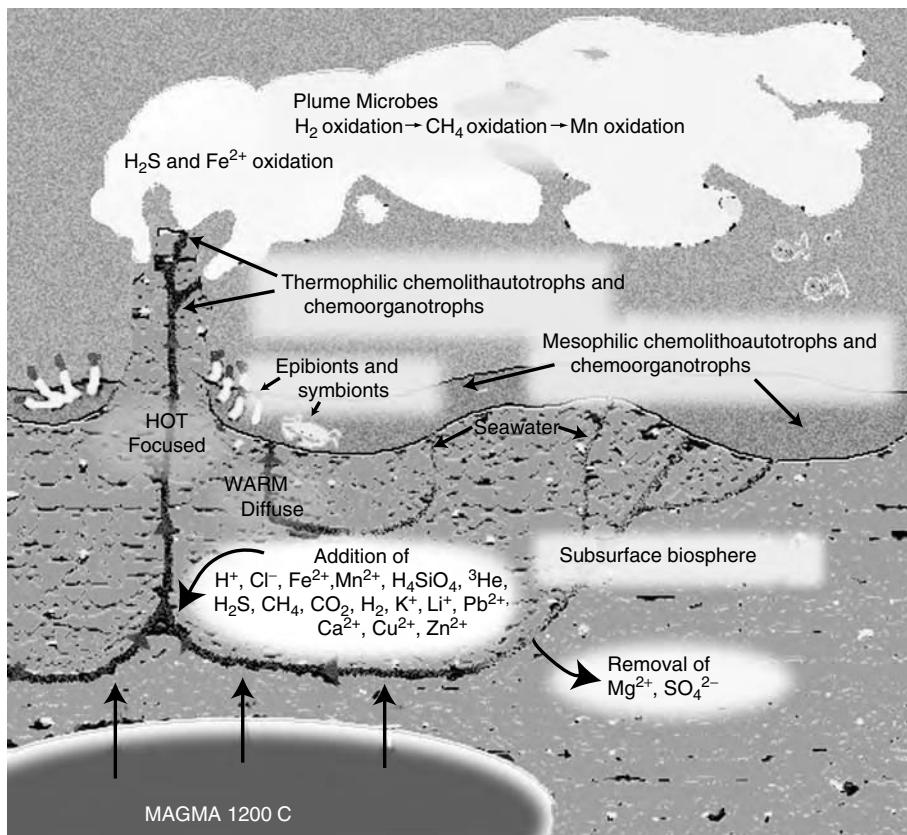
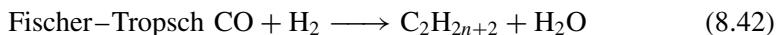


Figure 8.15 Geothermal vent chemical environment. (Reproduced by permission of Anna-Louise Reysenback and Elsevier Science)

At temperatures below 500 K the dominant carbon form is CH_4 , becoming CO_2 above this temperature. This is similar for nitrogen: low temperatures, NH_3 ; high temperatures, N_2 . These conditions are strongly suggestive of the conditions used in the endogenous synthesis of organic compounds in the Urey–Miller experiments. The production of amino acids and sugars via the Strecker and formose reactions continues in the same way only now, however, the temperature is higher and hence rates of reactions are much faster and the yields are greater. Organic species would have adsorbed and desorbed from the mineral assemblage surface and undergone Fischer–Tropsch-type processes, having the effect of fixing carbon into long aliphatic chains via a general reaction scheme:



Other reaction schemes have been suggested based on active FeS producing a more complex array of molecules, including the scheme in Figure 8.16, and others

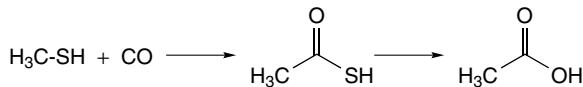


Figure 8.16 Carboxylic acid synthesis

leading to amino acids, which have been observed plentifully in fluid samples from hydrothermal vents.

The local conditions of temperature and pressure, as well as the new energy source in the form of the electrochemical gradient, can all be incorporated into the Gibbs free energy by adding new terms to the chemical potential. Variation of ΔG and ΔH with temperature are all standard thermodynamics, although we will resist the temptation to explore them here.

8.9 RNA World hypothesis

There are many potential molecules and possible routes to the synthesis of biomolecules that might form the basis of a primitive metabolism but thus far we have not addressed the question of information propagation or Darwinian evolution. Information storage must be contained within a sequence, such as words in a sentence or the base sequences within the genetic code, and that requires a polymerisation reaction, which is preferably autocatalytic to reproduce the information accurately. Peptides and nucleotides have this property, although the condensation reaction joining them together needs to be activated.

The salt- and Cu^{2+} -catalysed condensation of peptides provides a very simple polymerisation reaction with remarkable efficiency at 80°C . The proposed mechanism is shown in Figure 8.17 for the dimerisation of glycine. The presence of Cu^{2+} is important in this process and is unlikely to be present in the geothermal vent environment but it does require only small quantities of O_2 to oxidise copper. A better condensation reaction would be autocatalytic and provide a template for future generations – in short, a genetic code.

The theory that life began in an ‘RNA World’ suggests that the first self-replicating system was a set of RNA molecules. The catalytic and information-transferring properties of RNA indicate a possible scenario:

1. Relatively short RNA oligonucleotides were first formed by random events.
2. Some oligonucleotides began to catalyse complementary copies by acting as templates.
3. Chemical replication of oligonucleotides led to the evolution of self-replicating RNA molecules.

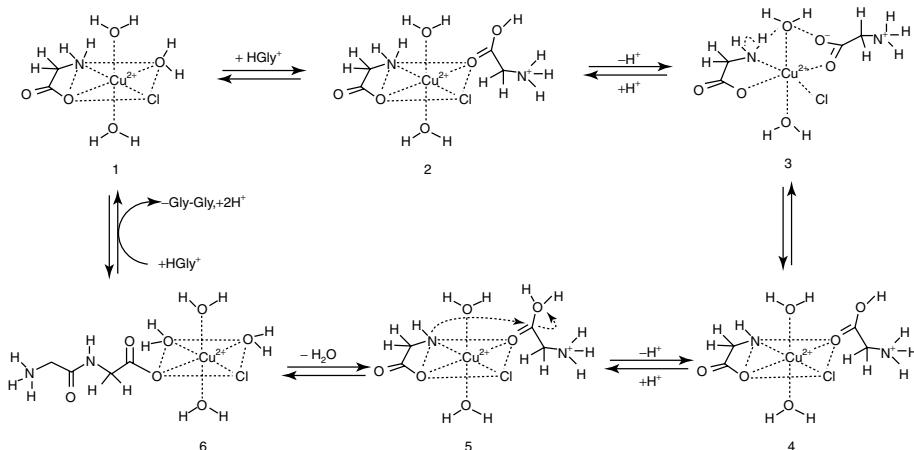


Figure 8.17 Proposed mechanism for the $\text{Cu}^{2+}/\text{NaCl}$ -catalysed peptide condensation reaction

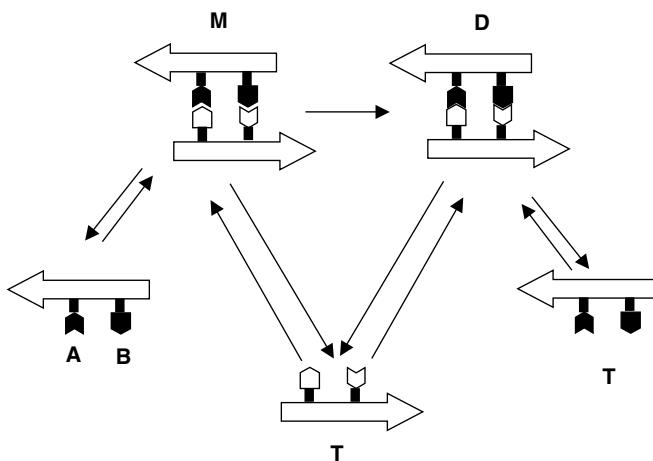


Figure 8.18 Three-step self-replication model

A recent discovery that RNA will act as a self-catalyst, called a ribozyme, leads to a simple three-step model for self-replication – this might include a surface. In the model (Figure 8.18), the template molecule T is self-complementary and is able to act as an autocatalyst. In the first step, it reversibly binds with its constituents A and B, forming the termolecular complex M. The termolecular complex undergoes irreversible polymerisation and becomes the duplex molecule D. Reversible dissociation of D gives two template molecules T, which can initiate new replication. The model preserves the order of the moieties on the template (the direction of the arrow) and the backbone, which may be on the surface

itself or another polymer chain. This has been achieved for nucleotides and phosphate sugars in a test tube without an enzyme. The process is termed *chemical ligation*.

Three pieces of evidence are cited in support of an RNA World. Firstly, some 17 RNA ribozyme catalysts have been discovered that produce a diverse array of organic molecules, including peptide bond formation. Second, the ability to form the peptide bond and build proteins may lead to a complex evolution favoured by the proximity of proto-proteins, producing enhanced reaction efficiency. Finally, RNA is the intermediate in the biosynthesis of DNA, suggesting that it must have preceded DNA in the evolutionary process.

A short look through prebiotic chemistry presents many possibilities that have to be constrained by the second law of thermodynamics and favourable kinetics, but could conceivably have produced all of the molecules required for both metabolism and information propagation. All that remains now is to build a cell and finally an organism.

Concepts and calculations

Concepts

<i>Spontaneous chemical reactions</i>	The use of thermodynamics to determine what reactions are possible in a particular prebiotic environment, including the ideas of chemical potential, equilibrium and extent of reaction
<i>Endogenous organic synthesis</i>	Urey–Miller experiments as a source of prebiotic molecules via the Strecker synthesis for amino acids, HCN polymerisation for purines and pyrimidines and the formose reaction for sugars
<i>Exogenous delivery of organics</i>	Cometary impact and the origin of the oceans loading the full cometary molecular inventory onto the prebiotic Earth
<i>Homochirality</i>	A preferential amplification of one enantiomer over the other due to biogenic activity – a biomarker
<i>Other habitats</i>	Surface metabolism and geothermal vents as habitats with selected reaction networks
<i>RNA World hypothesis</i>	The early synthesis of purines and the observation of ribozyme autocatalysis lead to the idea of RNA preceding DNA

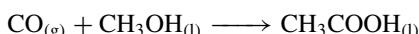
Calculations

<i>Equilibrium</i>	Chemical potential and ΔG with the extent of reaction
<i>Dissolved gases</i>	Henry's Law and the pH of the oceans
<i>Rates of reaction</i>	Temperature dependence of chemistry and the analysis of chemical networks in prebiotic environments

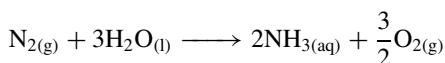
Problems

Thermodynamic data for these problems are found in Appendix C.

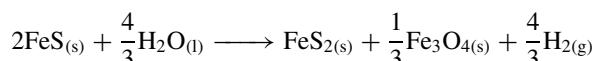
- 8.1** Global warming has appreciable effects on the oceans, leading to ice ages and rising water level. If the specific heat capacity of water $S = 4.186 \text{ kJ kg}^{-1} \text{ K}^{-1}$, the latent heat of fusion is 334 kJ kg^{-1} and the mass of the oceans is $1.35 \times 10^{31} \text{ g}$.
- Assuming that the average temperature of the oceans is 288 K, calculate:
 - The energy removed to cool the oceans to 273 K.
 - The energy removed to freeze the oceans.
 - The energy removed to cause the entire planet to freeze and enter an ice age.
 - Calculate the energy required to melt the Greenland ice cap with a volume of $5.1 \times 10^6 \text{ km}^3$ (density of ice is 946 kg m^{-3}).
 - If the Martian polar ice cap is the same size and made of solid CO₂, calculate the energy required to melt the ice cap if the latent heat of fusion is 196.1 kJ kg^{-1} and the density of CO₂ ice is 1562 kg m^{-3} .
- 8.2** Calculate the dissolved concentrations of N₂ and CO₂ for a prebiotic atmosphere with a species ratio of 90:10 and a total pressure of 1 bar.
- 8.3** The formation of acetic acid by bubbling CO through a solution of methanol follows the reaction:



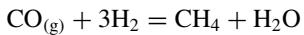
- Calculate $\Delta_r H^\circ$ for the reaction and classify as exothermic or endothermic.
 - Calculate $\Delta_r G^\circ$ for the reaction at 298 K. Is the reaction spontaneous?
 - Calculate the equilibrium constant for the reaction at this temperature.
- 8.4** The nitrogen-fixation reaction:



- has been proposed as a source of amines in a surface synthesis of amino acids.
- Calculate the enthalpy change for the reaction.
 - Calculate $\Delta_r G^\circ$ for the reaction at 298 K and 400 K. Assume that $\Delta_r H$ and $\Delta_r S^\circ$ are constant for both temperatures.
 - Calculate the equilibrium constant at each temperature.
 - Comment on the possible role of a surface in this reaction.
- 8.5** The hydrogen evolution reaction responsible for driving the redox chemistry in geothermal vents is:



- Calculate $\Delta_r H^\circ$ and $\Delta_r G^\circ$ for the reaction at 298 K and 380 K. Assume that $\Delta_r H$ and $\Delta_r S^\circ$ are constant for both temperatures.
- Calculate the equilibrium constant at 380 K.
- Comment on the role of this reaction, as in geothermal vents.

8.6 The equilibrium constant for the reaction:

is monitored as a function of composition. Express the equilibrium constant as a function of the extent of the reaction, ζ , when 1 mole of CO is mixed with 1 mole of H₂.

- 8.7** The pH of the oceans is controlled by a number of factors, including the dissolved CO₂. Use Henry's Law to calculate the total dissolved CO₂ for an atmosphere of 10 per cent CO₂ at a total pressure of 1 bar. Assume that dissolved CO₂ accounts for all of the dissolved carbon (c in Equation 8.29) and estimate the pH of the oceans.

9

Primitive life forms

Introduction

Finally, we must look at the last step in the origins of life – building the first organism. A possible collection of metabolic processes made from prebiotic molecules must localise and compartmentalise into a cell. We must not, however, be restricted by biology but adopt more of the spirit of astrobiology and look for life by a different set of rules. The precise physical chemistry that defines life and the boundaries of the biosphere are no longer clear and precise, even on Earth. Life occurs at high temperatures, with the record holder at 113°C, high pH ($\text{pH} \sim 0$) and high salinity ($\sim 5 \text{ M NaCl}$): these are all organisms that belong to a family called Archaea.

The general problem of astrobiology, however, is always prejudiced by the particular solution to the life problem that has occurred on Earth. It is often said that a second example of life on another world would broaden the astrobiological mind but for the moment we are somewhat conservative in approach. To address this problem, we shall explore some basic characteristics of life as they might apply

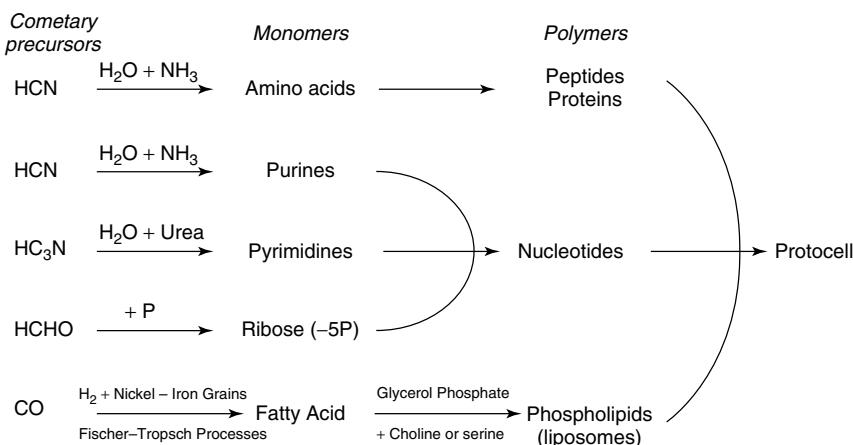


Figure 9.1 Towards life. (Adapted from Deaner 1997, with permission from the American Society for Microbiology)

generally but have been solved in particular ways specifically for Earth. We shall look at all of the extreme solutions to see how bacteria have developed in extreme environments on Earth, even enduring space flight. Perhaps this is the best we can do for the present to prompt the field astrobiologist to look in some strange places for life forms.

The process of building life can be summarised in Figure 9.1 and thus far we have looked at all of the processes building molecules of biological interest from prebiotic chemistry. The roles of these molecules in an astrobiological sense are more general and probably everywhere we might look for life. The prebiotic soup, whether from cometary infall or endogenous synthesis, will produce a collection

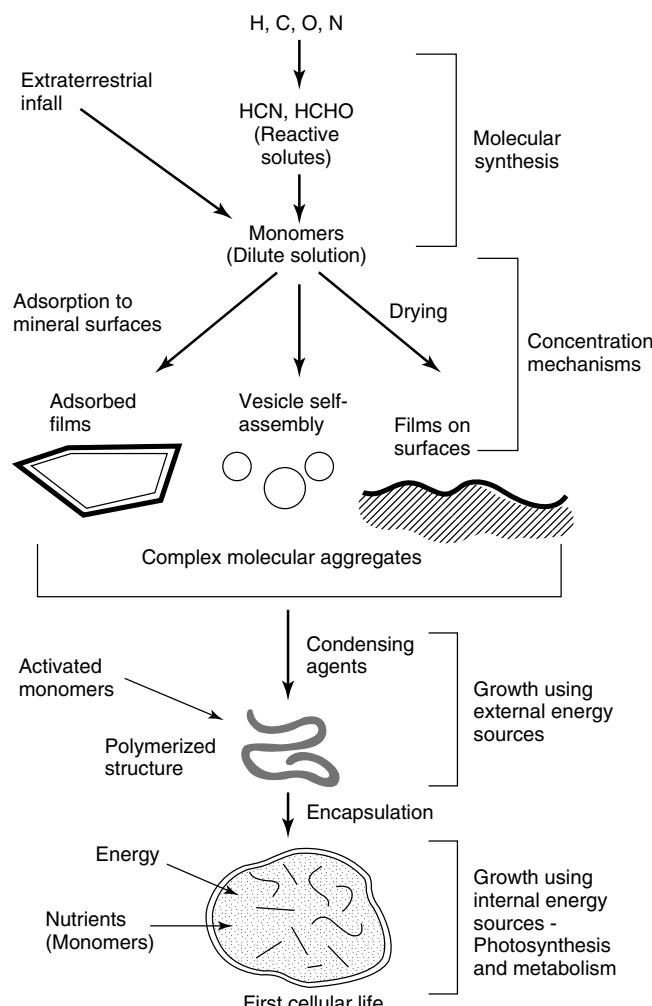


Figure 9.2 Building life. (Adapted from Deamer 1997, with permission from the American Society for Microbiology)

of molecules from which the biological monomers seem to form spontaneously, assuming that you are in the right solvent, on the right planet and near enough to the local star. The monomers themselves have useful properties, especially the purines and pyrimidines with the Watson–Crick base pairing, but there is no need for this to be unique. However, it is important that polymerisation occurs to string the information together. Polymers and a localised optimised environment segregated from the rest of the world form the basis for protocell formation.

A representation of all of the processes can be seen in Figure 9.2 and we must now turn attention to the formation of molecular aggregates on a larger scale: encapsulation and all of the problems associated with maintaining the environment within the encapsulated protocell, pointing towards metabolism and information propagation.

The complex processes of growth, reproduction, collecting nutrition and movement, perhaps even carbon-fixing processes such as photosynthesis, have to be performed to get to the simplest life forms found in the fossil records. There can be no time to dawdle; complex life was formed rapidly perhaps over a period of 100–500 million years, reaching a living form close to something that we would recognise today as a bacterium.

9.1 Self-assembly and encapsulation

The process of forming a cell and a first membrane relies on the general principles of self-assembly. An amphiphilic molecule has a distinct structure that, in the case of water, has a hydrophobic fatty end and a hydrophilic charged end (Figure 9.3). When the concentration of phospholipids reaches a critical concentration, a new self-assembled, aggregated structure called a micelle is formed. Micelles made from lipids are called liposomes and have the general bilayer structure of all biological cell membranes (Figure 9.4). The process is spontaneous and so the ΔG for the assembly is negative. However, there is a considerable increase in the entropy of the system and this must be overcome by the more favourable solvation energies at

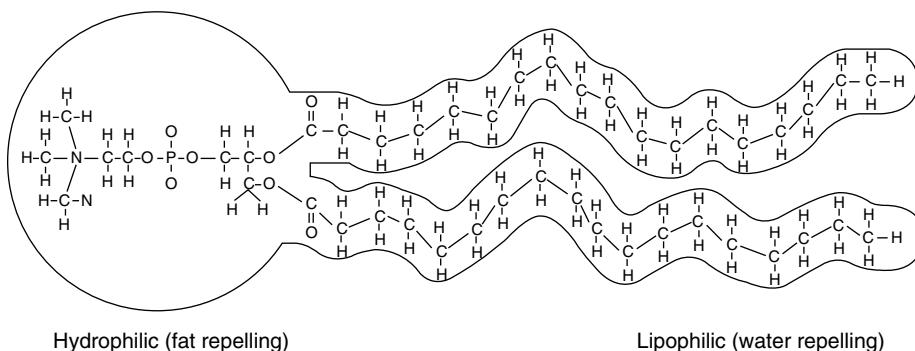


Figure 9.3 Structure of a phospholipid

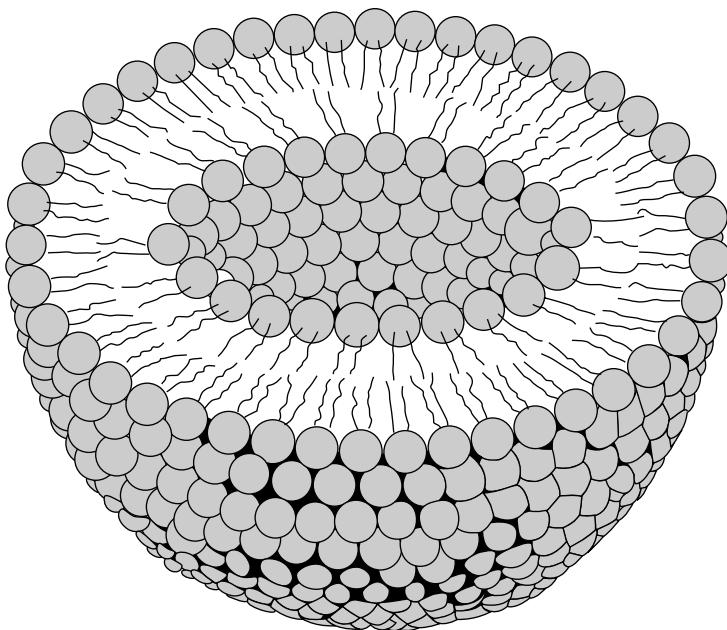


Figure 9.4 Liposome bilayer structure

each end of the molecule: charged phosphate groups interact favourably with the water and the fatty lipid tails interact favourably with each other. The aggregation occurs when the concentration reaches the *critical micelle concentration* (CMC) in solution and it must also be below the Krafts temperature, typically 25°C. Critical micelle concentrations are of the order μM – mM but they also depend on the ionic strength (total ion concentration of the solution) so that if the conditions are too salty the micelles will not form. Micelles of mixed compounds also form with varying CMCs for these species. Water extraction of the kerogen-like material from the Murchison meteorite produces *coercevate* structures in water spontaneously, as can be seen in Figure 9.5.

The conditions required for the formation of coercevates are rather delicate:

- Micelle formation below the Krafts temperature, perhaps as low as 25°C.
- Critical micelle concentrations in the prebiotic environment as high as millimolar.
- Potential for salty solutions with a high ionic strength.
- Harsh conditions in the drying and encapsulation scenarios.

All of these make the formation of the first protocell rather critical. There is also the general problem of membrane transport and how molecules might form within the protocell. The simplest postulated mechanism for the initial inclusion of molecules in protocells is that of encapsulation.

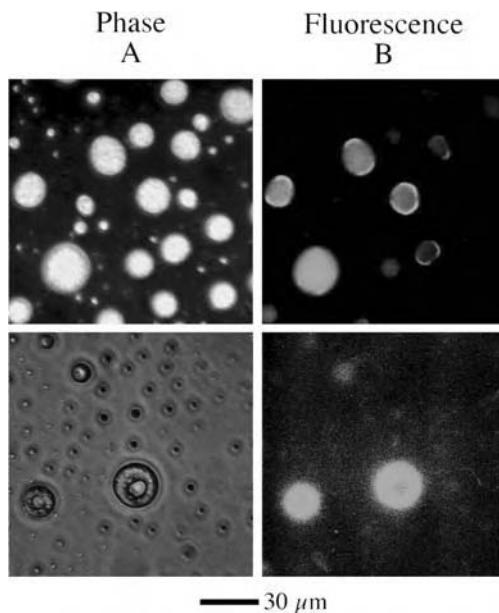


Figure 9.5 Pyranene dye encapsulated in various sizes of vesicles made from an acid extract of the Murchison meteorite. (Adapted with permission from Dworkin, 2001)

Encapsulation in organisms today occurs with active membrane processes but a spontaneous process is required for the first protocell. Encapsulation is depicted in Figure 9.6. Initially, coacervates form in solution spontaneously at the right temperature and concentration and then are subjected to a drying process. They then collapse onto a substrate such as a mineral surface, maintaining some of the membranous structure of the coacervate in the form of lamellae; the coacervates thus have fused with the surroundings. A re-wetting event occurs when the tide comes in, causing the coacervates to reform with the encapsulated molecules. Thus a protocell is born.

The encapsulation results in a chance collection of molecules that then form an autocatalytic cycle and a primitive metabolism but intrinsically only an isolated system of chemical reactions. There is no requirement for the reactions to reach equilibrium because they are no longer under standard conditions and the extent of reaction, ξ , will be composition limited (Section 8.2). Suddenly, a protocell looks promising but the encapsulation process poses lots of questions. How many molecules are required to form an organism? How big does the micelle or liposome have to be? How are molecules transported from outside to inside? Can the system replicate? Consider a simple spherical protocell of diameter 100 nm with an enclosed volume of a mere 125 fL. There is room within the cell for something like 5 billion molecules, assuming that they all have a density similar to that of water. This is a surprisingly small number and is a reasonable first guess for the number of molecules within a bacterium.

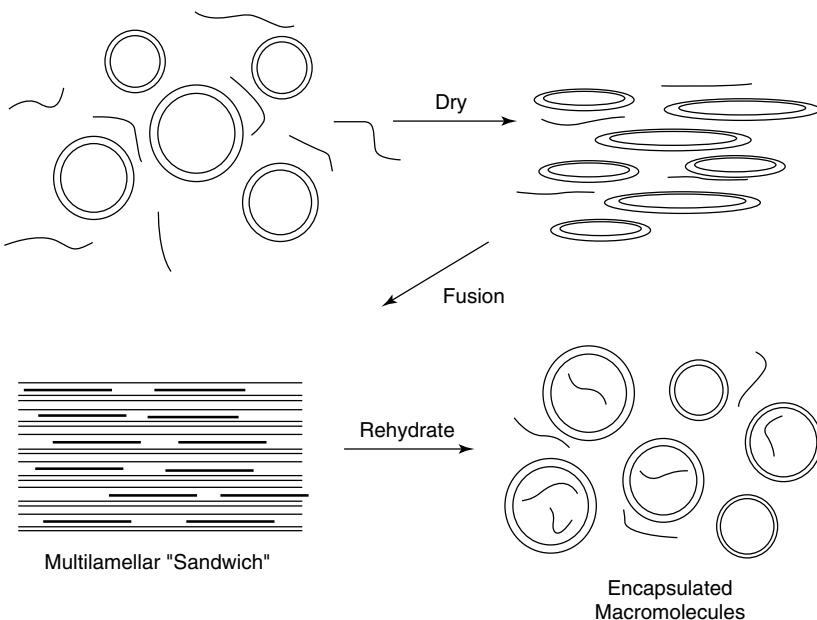


Figure 9.6 Encapsulation. (Adapted from Deaner 1997, with permission from the American Society for Microbiology)

9.2 Protocells

The simplicity of the encapsulation mechanism in Figure 9.6 suggests that molecules will be trapped within coacervates fairly easily and, similarly, smaller coacervates will be trapped within larger ones. The simplest organisms alive on Earth are the prokaryotes (although we shall see how this will change later), derived from the Greek meaning ‘before the nucleus’. A diagram of a typical bacterium is shown in Figure 9.7 and, importantly, there are no regions within the cell that are compartmentalised. This seems unlikely from the encapsulation mechanisms. What is important, however, is the ability to synthesise a strengthened cell wall made from sugars (in the case of Gram-positive bacteria). The cell wall reduces the fragility of the organism but also stops it taking part in any further simple encapsulation process. There must also be a minimal set of genetic material and some mechanism for passing the information stored within the genetic organism from generation to generation.

The isolation of the chemicals within the protocell causes a specific problem. Once the reactions have reached the extent of equilibrium allowed, limited by composition, all chemistry stops until the protocell dries out or brings in new molecules. The drying mechanism could certainly replenish the metabolism, requiring the protocell to dry out in a localised region on a substrate and re-form. The protocell would gain some new molecules and lose some old molecules in the process. The cell metabolism chemistry will only continue away from the surface and in a fully independent way when the membrane transport problem is solved.

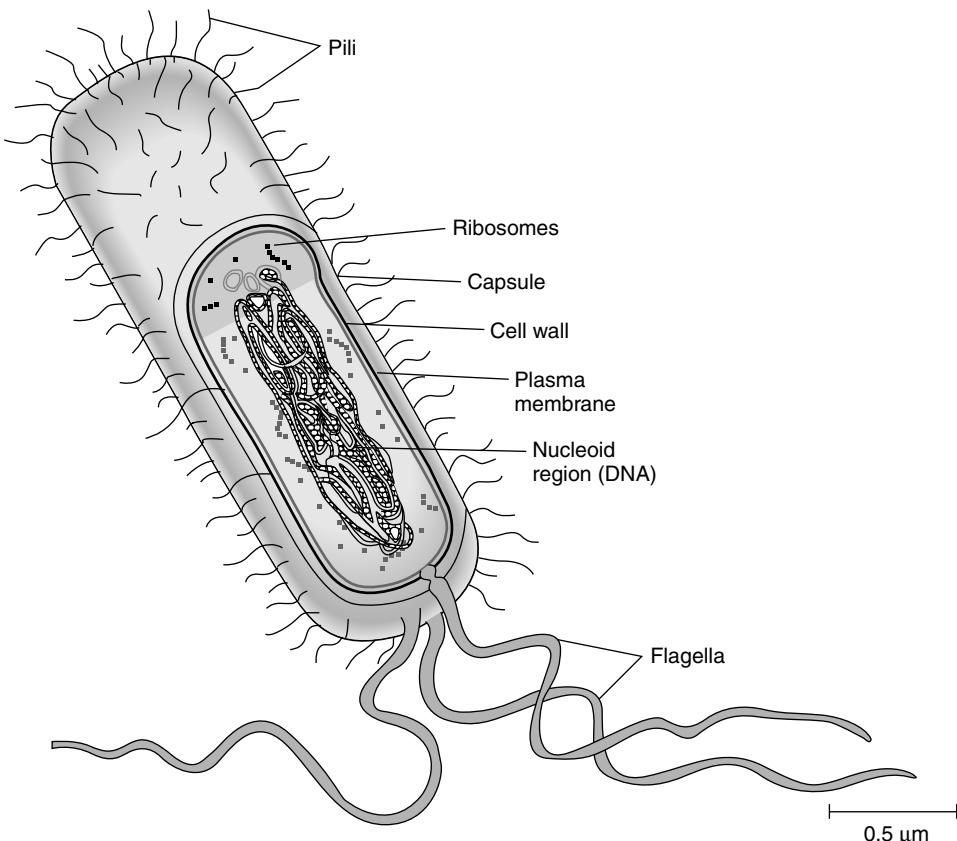


Figure 9.7 Structure of a bacterium

Membrane transport

The transport of chemicals across the membrane in biology takes two forms: active and passive (Figure 9.8). Passive transport is driven by diffusion and osmosis, whether through a membrane structure such as a protein or simply through the membrane itself. Active transport requires chemical energy to be consumed, usually in the form of ATP, to drive conformational changes of proteins thermodynamically uphill in a very specialised and selective way. The complexity of active transport suggests that it is most likely to have evolved after passive transport in the scheme of things. Incorporating prebiotic debris into the membrane or inside the protocell as it formed seems to be a fairly probable event, suggesting perhaps that transmembrane structures may have evolved rather quickly. The energy source for active transport may have taken somewhat longer to evolve.

Once the molecules have been captured inside the protocell, the concept of a concentration gradient and semipermeable membrane becomes important. Near a negatively charged mineral surface, such as silica, the surface concentration

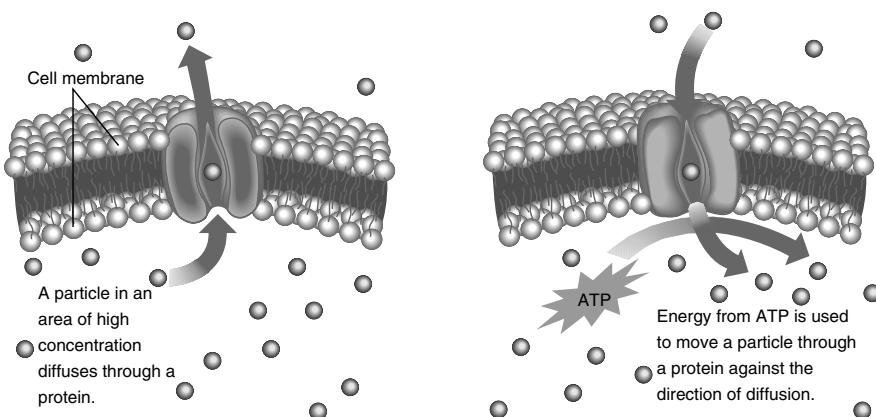


Figure 9.8 Passive and active membrane transport

of positive ions will be enhanced by the surface potential (Section 8.7), so the encapsulation event will have occurred at high positive ion concentration. However, once the protocell formed, it was free to leave the surface, immediately moving to a lower salt environment where osmosis would occur. Does this have implications for a freshwater rather than a seawater origin for life?

Thermodynamically the quantitative treatment of both active and passive processes requires them to be ‘downhill’ or exoergic. The description of chemical potential as a function of mole fraction follows the same form as before for a neutral species (Section 8.2):

$$\mu_A = \mu_A^0 + RT \ln x_A \quad (9.1)$$

where μ_A is the chemical potential of the dissolved substance, μ_A^0 is the standard state chemical potential, R is the gas constant, T is the temperature (K) and x_A is the mole fraction of A . The difference in chemical potential of A in two different places in space (i initial position and f final position), whether along a concentration gradient or across a membrane, is given by:

$$\begin{aligned} \Delta\mu_{f-i} &= (\mu_{A,f}^0 + RT \ln x_{A,f}) - (\mu_{A,i}^0 + RT \ln x_{A,i}) \\ \Delta\mu_{y-x} &= RT \ln \left(\frac{x_{Ai}}{x_{Af}} \right) \end{aligned} \quad (9.2)$$

The final equation is a simplification because the standard chemical potentials are the same at each point.

Example 9.1

Consider a sucrose concentration gradient across a membrane with 50 mM inside the cell and 500 mM outside the cell at 298 K; the transfer from outside to inside is:

$$\Delta\mu_{y-x} = RT \ln \left(\frac{500}{50} \right) = 8.314 \times 298 \ln(10) \quad (9.3)$$

and for the reverse direction

$$\Delta\mu_{y-x} = RT \ln \left(\frac{50}{500} \right) = 8.314 \times 298 \ln(0.1) \quad (9.4)$$

For the first equation corresponding to the transfer of sugar into the protocell from outside to inside the free energy is $+5.7 \text{ kJ mol}^{-1}$, whereas for the reverse process the free energy is -5.7 kJ mol^{-1} . If the membrane will allow the passage of sucrose then the sugar will move from high concentration to low concentration by diffusion. If, however, the membrane pores are too small to allow the sugar to move, then the water will move in a process called osmosis.

Encapsulation of any species within a protocell creates a chemical potential gradient, which will allow molecules to move from inside the cell to outside. If the protocell is to reverse this process and acquire nutrients from the external medium then active transport across the membrane must be achieved with the expenditure of energy. A protocell formed within any surface environment is likely to be at equilibrium with its surroundings during the formation process, but moving away from the surface during the rehydration process will then create the chemical gradient. Although some micellar structures are likely to be impermeable to large molecules such as sugars, the same will not be true for smaller species such as water or salts. A semipermeable membrane leads to the process of osmosis (derived from the Greek for ‘push’).

Osmosis

A semipermeable membrane has a pore structure that will keep large molecules such as formose sugars, amino acids and fats behind the membrane but will allow smaller molecules such as water to pass. A hydrostatic pressure difference results from the presence of more water within the protocell and an osmotic pressure develops (Figure 9.9). On the pure solvent/water side the chemical potential of the water at pressure p is $\mu_w^*(p)$. On the other side of the membrane, the chemical potential of the water is lowered due to the presence of the dissolved molecules. The mole fraction of the water is lower such that $x_w < 1$. Using Equation 9.2 the free energy change associated with water passing into the cell is then negative and spontaneous – a process called osmosis. However, there is an increase in pressure

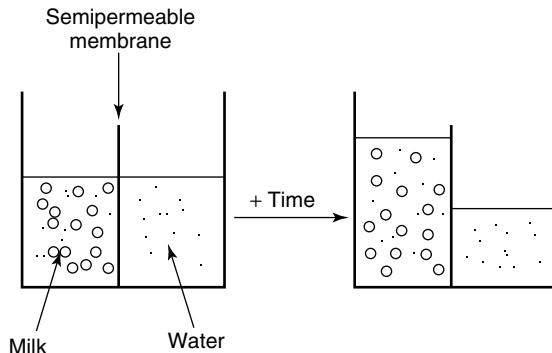


Figure 9.9 Osmotic pressure

inside the cell due to the osmotic flow, called the osmotic pressure, and this is given the symbol π . The total pressure inside the cell is now the sum of the external pressure and the osmotic pressure, $p + \pi$. At equilibrium, the chemical potential of the pure water outside the cell must be the same as the chemical potential of the water inside the cell, which must now include the effect of the internal cell pressure pushing the water out. The equilibrium balance can be written as:

$$\begin{aligned}\mu_W^*(p) &= \mu_W(x_W, p + \pi) \\ \mu_W(x_W, p + \pi) &= \mu_W^*(p + \pi) + RT \ln(x_W)\end{aligned}\quad (9.5)$$

The first equation simply states the balance in chemical potentials inside and outside of the cell. The expression for the chemical potential inside the protocell separates into a term involving the mole fraction and the chemical potential associated with the pressure difference. The work done by the cell in opposing the pressure change, assuming that the cell remains at constant volume, is given below, where the change in pressure is from p to $p + \pi$:

$$\mu_W^*(p + \pi) = \mu_W^*(p) + \int_p^{p+\pi} V_m \, dp \quad (9.6)$$

The integral calculates the work done by the membrane resisting the pressure change in the cell. Assuming that the molar volume V_m is constant, the osmotic pressure is given by:

$$-RT \ln(x_W) = V_m \pi \quad (9.7)$$

This is a complete expression but we can simplify it further, especially in the case of dilute solutions: $\ln(x_w) = \ln(1 - x_B) \approx -x_B$, where x_B is the mole fraction of the dissolved molecules. One further simplification is to take the total volume of the cell to be the number of moles of water, n_w , multiplied by the molar volume of water, to give $n_w V_m = V$. Now we can derive a simple expression for the osmotic pressure:

$$\begin{aligned}
 x_B RT &= V_m \pi \\
 \frac{n_B}{n_w + n_B} RT &= V_m \pi \\
 n_B RT &= (n_B + n_w) V_m \pi \\
 \pi &= \frac{n_B}{V} RT = [B]RT
 \end{aligned} \tag{9.8}$$

This simple equation relating osmotic pressure to the concentration of the dilute solute was discovered by van't Hoff and enables the pressure to be estimated.

Example 9.2

If the typical concentration of sucrose within a cell is 100 mM, calculate the mole fractions of water and sucrose. The molar volume for water is the mass of water in 1 cm^{-3} (from the density of water, 1 g cm^{-3}) converted to moles $1/18 \times 1000$ to give 55.55 M – the concentration of pure water at room temperature.

$$\begin{aligned}
 x_w &= \frac{55.55}{55.55 + 100 \times 10^{-3}} = 0.998 \\
 x_{Sucrose} &= \frac{100 \times 10^{-3}}{55.55 + 100 \times 10^{-3}} = 1.796 \times 10^{-3}
 \end{aligned} \tag{9.9}$$

The mole fraction of water is much larger than that of sucrose and practically for all cell components, so van't Hoff's equation is easily justified for estimates of cellular osmotic pressure. Given $R = 0.08314\text{ bar L}^{-1}\text{ mol}^{-1}$ and a protocell temperature of 298 K, calculate the osmotic pressure.

$$\begin{aligned}
 \pi &= [B]RT \\
 \pi &= 0.1 \times 0.08314 \times 298 \\
 \pi &= 2.48\text{ bar}
 \end{aligned}$$

The osmotic pressure is more than twice the atmospheric pressure and represents a significant challenge for the forming protocell.

Ion transport

The transport of charged species across a membrane requires an additional term to be added to Equation 9.1 to allow for the electric field:

$$\mu_A = \mu_A^0 + RT \ln x_A + zFE \tag{9.10}$$

The new term includes the charge on the solute ($z = +1$ for Na^+ and -1 for Cl^-), the Faraday constant F (96.5 $\text{kJ mol}^{-1} \text{V}^{-1}$) and the electrochemical potential E in volts. The difference between inside and outside now contains a term for the difference in the electrochemical potential:

$$\begin{aligned}\Delta\mu_{f-i} &= (\mu_{A,f}^0 + RT \ln x_{A,f} + zFE^f) - (\mu_{A,i}^0 + RT \ln x_{A,i} + zFE^i) \\ \Delta\mu_{f-i} &= RT \ln \left(\frac{x_{Ai}}{x_{Af}} \right) - zF(E^f - E^i)\end{aligned}\quad (9.11)$$

The two terms in the final expression of Equation 9.11 can compete against one another and with only a small membrane potential the concentration gradient chemical potential can be overcome. Electrochemical energy is thus harnessed to allow nutrients to flow into a cell and increase the local concentration.

Example 9.3

Consider a concentration gradient for NaCl with an internal protocell concentration of 150 mM and an external concentration of order 1 mM, as might be found in a freshwater environment. There is a membrane potential of -200 mV .

Substituting these values into Equation 9.11 gives:

$$\begin{aligned}\Delta\mu_{f-i} &= RT \ln \left(\frac{x_{Ai}}{x_{Af}} \right) - zF(E^f - E^i) \\ \Delta\mu_{f-i} &= 8.314 \times 298 \ln \left(\frac{1}{150} \right) - 1 \times 9.6485 \times 10^4 \times (-200 \times 10^{-3}) \\ \Delta\mu_{f-i} &= -1.24 \times 10^4 + 1.93 \times 10^4 = +6.9 \text{ kJ mol}^{-1}\end{aligned}\quad (9.12)$$

The first term in the final expression indicates that the concentration gradient is downhill for the NaCl to move out of the cell, whereas the electrochemical potential is in the opposite direction: the electrochemical gradient is against the concentration gradient. Substituting these values into Equation 9.11 gives a free energy change of $+6.9 \text{ kJ mol}^{-1}$ at 298 K. The electrochemical potential of Na^+ is clearly higher outside the cell, due to potential gradient, and Na^+ would passively enter the cell.

Calculation of the internal cell potential is a very complicated matter because the electrochemistry of all of the species within the protocell would have to be balanced subject to their composition quotient Q , after which the standard free energy would have to be established from tabulations. The transport of Na^+ would also change this balance, along with the ionic strength of the solution and the stability of the proteins or prebiotic molecules within the protocell. Such non-equilibrium thermodynamics forms the basis of the protocell metabolism. The construction

of a protocell metabolism is one requirement and another is the propagation of information for Darwinian evolution. What is a minimal genome?

Minimal genome

Nanobacteria of dimensions 50–200 nm as proposed for the microfossil structures in the ALH84001 meteorite, are about 1/10 of the normal cellular lengths and consequently have a volume 1/1000 of that found in conventional cells. It has been questioned whether such small volumes could contain the necessary amount of information-bearing material to form a minimal gene set. Although some groups contend the existence of nanobacteria, the minimal gene set is an important question. The current cell size best estimated for the minimal genome is around 100 nm and would contain a genome ten times smaller than the current bacterial genome: a length of 400 kb (kilobases). The smallest genome known today belongs to the bacterium *Mycoplasma genitalium*, which is 580 kb and contains about 480 protein-coding genes. Experiments knocking out the genes from the genome have suggested that 265–350 of the protein-coding genes are required for growth under laboratory conditions. Similar experiments suggest that only 9 per cent of the genome of *Bacillus subtilis* is required, amounting to some 552 kb. An analysis of DNA content as a function of cell volume for current life forms suggests that a minimal cell would be around 200 nm in diameter and contain about 400 kb as a minimal genome. This already complex genome may have evolved from smaller genomes from nanobacteria where cell size and number of genes have been estimated to give a 50 nm cell with 8 genes, a 160 nm cell with 200 genes and a 190 nm cell with 750 genes.

The roles played by the genes in the minimal gene set have been divided into a number of gene classes:

1. Biosynthesis of cofactors
2. Cell envelope
3. Cellular processes
4. Central metabolism
5. Energy production and storage
6. Fatty acid and lipid metabolism
7. Purine and pyrimidine metabolism
8. Regulatory
9. Replication, recombination and repair
10. Transcription
11. Translation

12. Transport (nutrients and elements)

13. Others unknown

These functions already outline a phenomenally complicated organism with information storage and retrieval systems in two stages of translation and transcription and a complete biosynthesis pathway set for all of the structural components of the cell. The last category is an interesting one for molecular biology. The smallest genome organism, *Mycoplasma genitalium*, has 170 genes with unknown function of which 103 could not be disrupted in gene-knockout experiments (otherwise the bacterium is not viable) and are therefore thought to be essential. There remains a lot to be discovered about even the smallest genetic organisms.

The prospect of creating an artificial cell, however, is very promising, with the potential to design an artificial cell for a particular biosynthetic task – growing insulin for example. The current suggestion is that this sort of ‘organism’ will be created within the next 30 years and would have the characteristics detailed in Table 9.1.

The use of artificial cells for biotechnology applications is simplified if the replication and reproduction characteristics are removed so that an artificial cell does one particular job. A liposome structure would continue to uptake molecules from the environment and synthesise proteins or cofactors, but the molecular inventory within the liposome is comparatively small.

Table 9.1 Properties of an ideal minimal cell (according to Andrew Pohorille and David Deamer)

-
- An information-carrying polymer, such as a nucleic acid, must be synthesised by a template-directed polymerisation that occurs in a membrane-bound volume
 - The monomers of the polymer must be provided externally and transported across the membrane boundary to support the replication process. Other small molecules or ions needed for biosynthetic reactions must be delivered from the environment
 - An external source of chemical energy must be available to drive the biosynthetic reactions
 - A catalytic activity must be present that is linked to the replication process, so that variations in replication affect the catalysed reactions. Under such circumstances, variations will change the fitness of the system and lead to evolution. Compartmentalisation of the replicating catalytic system within a membrane-bound volume allows selection of variations, leading to ‘speciation’
 - The boundary membrane itself must be able to grow, either by accumulation of membrane-forming material from the environment or by conversion of precursor molecules into such material
 - There must be a mechanism that allows the assembly to separate into two or more smaller structures during the growth process, and the smaller structures should contain the capabilities of the larger system
 - Catalysis, replication and growth must be well regulated so that none of the processes lags behind or gets far ahead of other processes in the cell
-

9.3 Universal tree of life

Evolution of a species requires the safe passage of information from one generation to the next. This information is encoded in the DNA (or RNA) as genes and represents defined sequences of GATC in DNA. However, during the transcription process external forces such as photolysis may introduce reading errors and these are called mutations. Successful mutations survive and are required for evolution; unsuccessful mutations die – the survival of the fittest. However, there are ‘random’ sequences of DNA between the genes, often called intron or ‘junk’ DNA, and these can adopt more and more mutations without affecting the biological well-being of the organism. Thus, older species have more junk DNA: the crocodile has a huge genome in length, most of which is simply junk but telling us that it is an ancient species.

RNA carries the same junk information as DNA due to the intrinsic link between DNA and RNA. Cells contain many different types of RNA: messenger RNA (mRNA) to transfer information from the DNA molecule; transfer RNA (tRNA) to transport the information from the nucleus to the ribosome where proteins are made; and ribosomal RNA (rRNA), which is RNA that has been modified just prior to production of proteins in the ribosomes. A systematic analysis of the 16S rRNA and 18S rRNA found in nearly all species has provided information on the evolution of junk RNA and a common genetic ancestor. 16S rRNA is some 1500 base pairs long and has been sequenced for many species. The differences in the junk RNA sequences have been used to mark the points of divergence from a common genetic ancestor. Recent divergence means that the sequences are very similar (homologous) whereas species that diverged early in evolution have very different species origins.

A universal tree of life has been derived from this analysis (Figure 9.10) for all species alive on Earth today that enables a ‘genetic distance’ to be determined between species but not necessarily a historical distance.

The tree branches quickly into three domains:

1. Eukarya – cells containing membranes and nuclei (e.g. plants, fungi and animals).
2. Bacteria: no cell membranes and smaller organisms.
3. Archaea: a distinct type of organism with different basic properties and cell wall structures.

Finding the root of the tree is hard, especially as we do not know the sequence of the RNA from the common ancestor. However, there are sequences that are propagated similarly in early organisms and show us which species are closest to the common ancestor. The first major divergence was between the bacteria and the archaea but no living species is close to the root, implying that all life has evolved significantly from the common ancestor. However, all species closest to

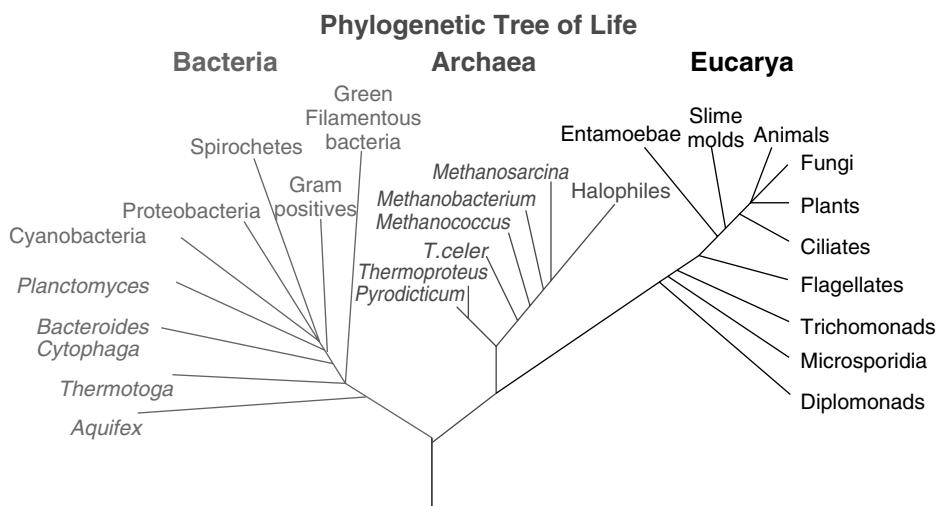


Figure 9.10 Common genetic ancestor. (Reproduced by permission of NASA Astrobiology Institute)

the root of the tree are hyperthermophilic in nature (they live in hot springs) and it is clear that all species alive today had a common hot-spring origin. Nevertheless, the Earth experienced many asteroid impacts in its early formation that would have sterilised all life, evaporating all of the oceans, so there may have been many different origins to life with only the current organisms surviving.

9.4 Astrobiology

The formation of life and how to recognise it is the central problem in astrobiology. All searches for other life forms will be prejudiced horribly by the study of the only example we know to date – life on Earth. Despite attempts to classify the general roles of molecules in definitions of life as biomarkers, there have been many discussions of possible biomarkers, including:

1. Cellular remains.
2. Textural fabrics in sediments that record the structure and/or function of biological communities (e.g. stromatolites).
3. Biogenic organic matter, including hydrocarbons.
4. Minerals whose deposition has been affected by biological processes.
5. Stable isotopic patterns that reflect biological activity.
6. Atmospheric constituents whose concentrations require a biological source.

Biomarkers are one thing but where to look is another. Life on Earth has shown remarkable adaptation to different and extreme environments and we shall take

a short journey around the edge of the biosphere to look at what is out there and to open the mind and even surprise at the diversity of the address book of life on Earth.

Extremophile bacteria

The fairly recent discovery of bacteria living in obscure environments (Table 9.2) has led to a new classification of bacteria: the ‘extremophile’ bacteria, with a diverse array of likes and dislikes. A useful working definition for extremophile is:

Extremophile: an organism that is isolated from an extreme environment and often requires the extreme condition for growth.

Of course what is clear is that extreme is an anthropocentrically derived term because with the bacterium there is nothing unusual about their environment. Similarly, for the astrobiologist in the field, no particular environment should be off-limits for investigation.

Many types of extremophile bacteria have been identified, described by some of the following definitions:

- Alkaliphile: an organism with optimal growth at pH values above 10.
- Barophile: an organism that lives optimally at high hydrostatic pressure.
- Endolith: an organism that lives in rocks.
- Extreme acidophile: an organism with a pH optimum for growth at, or below, pH 3.

Table 9.2 Possible life environments

Characteristics	Type of thermal area	
	Terrestrial	Marine
Location	Mud holes; geothermal power plants; deep terrestrial oil fields	Submarine hot springs; vents or ‘black smokers’; deep submarine oil fields
Temperature	Surface: up to 100°C Depth: above 100°C	Up to 400°C
Salinity	Usually low 0.1–0.5% salt	Usually about 3% salt
pH	0.5–9	5–8.5
Major gases and sulphur compounds	CO ₂ , CO, CH ₄ , H ₂ , H ₂ S, S, S ₂ O ₃ ²⁻ , SO ₃ ²⁻ , SO ₄ ²⁻ NH ₄ ⁺ , N ₂ , NO ₃ ⁻	CO ₂ , CO, CH ₄ , H ₂ , H ₂ S, S, S ₂ O ₃ ²⁻ , SO ₃ ²⁻ , NH ₄ ⁺ , N ₂ , NO ₃ ⁻

- Halophile: an organism requiring at least 0.2 M salt for growth.
- Hyperthermophile: an organism having a growth temperature optimum of 80°C or higher.
- Oligotroph: an organism with optimal growth in nutrient-limited conditions.
- Psychrophile: an organism having a growth temperature optimum of 15°C or lower, and a maximum temperature of 20°C.
- Toxitolerant: an organism able to withstand high levels of damaging agents. For example, living in water saturated with benzene or in the water-core of a nuclear reactor.
- Xerotolerant: an organism capable of growth at low water activity. For example, an extreme halophile or endolith.

With the suggestion that the last common genetic ancestor is a hyperthermophile, the role of temperature on the origins of life is important. The lower temperature limit in water is limited by the phase transition from liquid to ice. This is a problem because the density of ice is lower than that of water and the increase in volume on freezing will cause the cell structure to become disrupted in the same way that pipes burst in the winter. The lower limit for bacterial growth reported so far is –20°C, which is the temperature at which intracellular ice is formed. Adaptation to the cold requires a considerable salt content to depress the melting point of water: the Don Juan Pond in Antarctica, which has a saturated CaCl₂ solution, preserves the liquid phase at temperatures as low as –53°C.

At the other end of the scale is hot water that has an elevated boiling point due to increased local pressure. Boiling occurs when the vapour pressure above a liquid equals the atmospheric pressure, so at sea level water boils at 373 K when the vapour pressure of water is equal to one atmosphere. In the conditions that prevail around geothermal vents, the hydrostatic pressure is much higher than one atmosphere, allowing temperatures as high as 350°C to be maintained. The Archaea *Pyrolobus fumarii* grows at temperatures as high as 113°C and can tolerate 121°C for an hour. There are some real advantages to living at these temperatures, namely the speed at which chemical reactions will occur. There is more chemical energy in the environment to break stronger chemical bonds and increase the rate of chemical reactions. The flipside, however, is that the bonds in the enzymes must also be stabilised against the temperature to maintain their shape and activity. Extensive S–S cross-linking in proteins increases the thermal stability of its three-dimensional structure. The reward is high turnover number and rate of catalysis: an advantage that has not gone unnoticed by the pharmaceutical industry.

Extremes of pH would also cause problems for mammals that have a body pH controlled very close to around 6.7 in the blood, but some microorganisms have developed a tolerance to extreme pH conditions of 0–12.5. The organisms work

hard to maintain their internal cellular pH at 4–9 despite the extreme external conditions. Perhaps the most interesting in terms of the availability of potential environments, however, is the tolerance to extreme pressure. There are organisms in the deepest parts of the oceans that can tolerate pressures of 1100 bar (approximately 1000 atmospheres) and we have only just started looking deeper in the marine bed and rock sediments. Prokaryote bacteria have been found 750 m below the sea floor, suggesting that there is a vast biolithosphere that has yet to be explored deep within the rocks of the Earth. Bacteria may reside deep within planets such as Mars and remain undiscovered. It is always possible that bacteria discovered during the excavation process are in fact efficient, enterprising, terrestrial bacteria that have been transported there during the mission. Similar concerns have been expressed about taking bacteria to the Moon and Mars and beyond. It may well be possible to argue that Mars was not inhabited and never has been, and the arrival of unmanned missions has seeded Earth life forms on these planets.

Bacteria in space

The excitement surrounding the discovery of ‘nanobacteria’ fossil features within the mineralogy of meteorite ALH84001 has raised the prospect of panspermia within the solar system at least. The express delivery of life from Mars to Earth via a meteorite requires the interplanetary bacterial traveller to tolerate the condition of space: extremes of temperature varying from 70 K to 400 K depending on whether you are pointing directly at the Sun or not, and of course the constant solar wind. The solar wind will bombard any bacterial passengers on a meteorite with cosmic particles, which should cause mutation in a living system or, more catastrophically, breakdown of chemical bonds. Particularly interesting among the durable bacteria on Earth is *Deinococcus radiodurans* (Figure 9.11), which is capable of withstanding ionising radiation levels some ten times that of humans or other bacteria. It has developed a DNA structure that is tolerant to the radiation and does not mutate rapidly. Bacteria of this sort could have lived unprotected on the surface of a planet before the short-wavelength shield formed in the atmosphere.

Bacterial spores have been deliberately taken into space, in particular spores of the bacterium *Bacillus subtilis*. Up to 70 per cent of the bacterial spores survive in the short term for approximately 10 days of exposure to space vacuum. The chances of survival in space are increased if the spores are embedded in chemical protection such as sugars, or salt crystals, or if they are exposed in thick layers. For example, 30 per cent of *B. subtilis* spores survived for nearly 6 years in space when embedded in salt crystals and 80 per cent survived in the presence of glucose. *Bacillus subtilis* spores in Dominican amber have remained viable for several million years. Transport to a friendly environment with a source of energy is all that is required for life to be seeded on this planet or the next.

There is a consequence for space travel and planetary exploration associated with the viability of bacteria in space. Survival times of a few years would enable spacecraft to take bacteria from Earth to a target planet such as Mars. The two

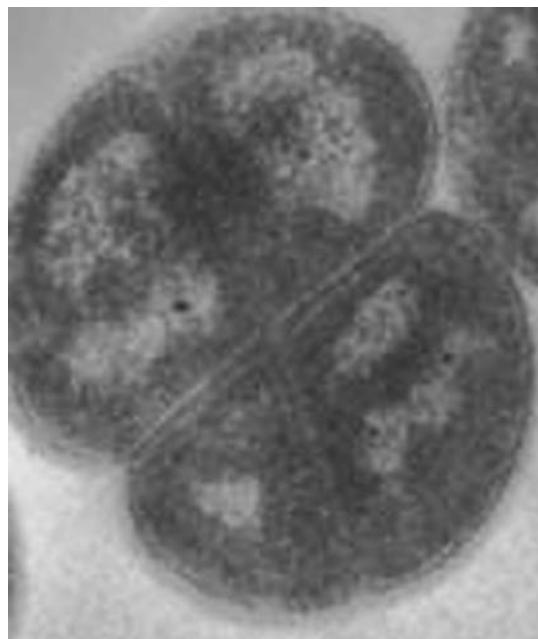


Figure 9.11 *Deinococcus radiodurans*. (Reproduced by permission of Oak Ridge National Laboratory)

Mars rovers, Spirit and Opportunity, could deposit Earth bacteria on the Martian surface, which fortuitously could find an environment in which to colonise. One possible false alarm for Martian life exploration is that evidence is found on the Martian surface of life on Earth. Extreme measures have been taken with the NASA spacecraft to use exposure to the UV radiation from the Sun to sterilise the spacecraft, rotating the various surfaces to face a prolonged exposure, but none of this could guarantee a sterile spacecraft.

Energy sources – autotrophs

The big jump in the evolution of the atmosphere, causing a large increase in the number of organisms on Earth, was the production of oxygen by photosynthesis (Figure 9.12). In the absence of life the oxygen concentration may have been about 5×10^{-9} at its present atmospheric level. The question of what inorganic energy source preceded photosynthesis is important not only to understand the origins of simple organisms but also to identify possible chemistry in the putative environments in which bacteria may be found.

Bacteria that live in the dark must derive the free energy to move electrons around in chemistry and a series of reactions based on the iron pyrite surfaces have been shown to be exoergic (Figure 9.13). The cationic pyrite surface serves as the catalyst for the binding of negatively charged species, perhaps tagged with

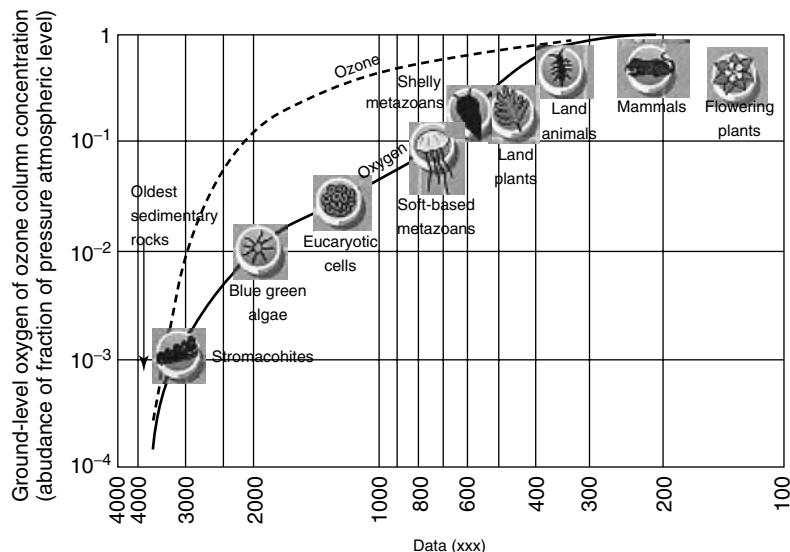
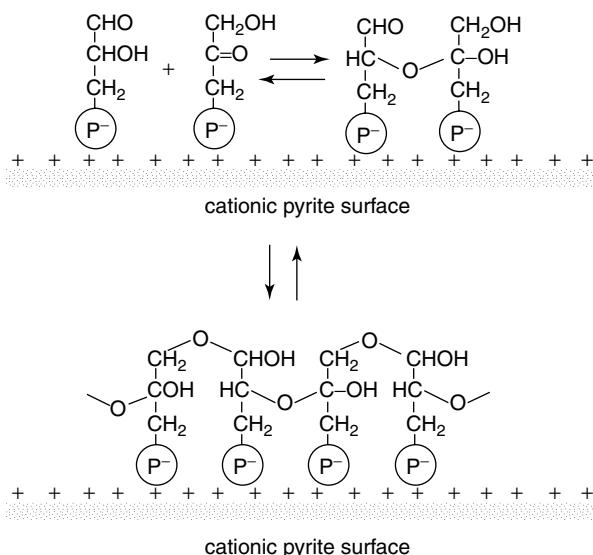


Figure 9.12 Oxygen and ozone concentrations in the Earth's atmosphere and the diversity of life forms. (Reproduced from Wayne, 2000, with permission from Oxford University Press)



1. $\text{CO}_2 + \text{H}_2 \longrightarrow \text{HCOOH}$ $\Delta G^\circ = +30.2 \text{ kJ mol}^{-1}$
2. $\text{Fe}^{2+} + 2\text{H}_2\text{S} \longrightarrow \text{FeS}_2 + 4\text{H}^+ + 2\text{e}^-$ pyrite formation
3. $\text{FeS} + \text{H}_2\text{S} \longrightarrow \text{FeS}_2 + \text{H}_2$ $\Delta G^\circ = -41.9 \text{ kJ mol}^{-1}$
4. $\text{CO}_2 + \text{FeS} + \text{H}_2\text{S} \longrightarrow \text{FeS}_2 + \text{H}_2\text{O} + \text{HCOOH}$ $\Delta G^\circ = -11.7 \text{ kJ mol}^{-1}$

Figure 9.13 Energy-yielding reactions on pyrite surfaces. (Adapted from Deaner 1997, with permission from the American Society for Microbiology)

inorganic phosphate, but also generates free energy from the reduction of hydrogen sulphide. Reaction 1 is the endoergic reduction of CO_2 to formic acid, which requires $+30.2 \text{ kJ mol}^{-1}$ of energy that may have been provided by the reduction of H_2S at the charged Fe^{2+} surface, first suggested by Wachterhauser as a primary autotrophic reaction:



This produces H_2 and has a reaction exoergicity of $-41.9 \text{ kJ mol}^{-1}$, providing both a mineral surface template and a free energy source for the fixation of CO_2 . The reaction scheme for this process, including the role of the mineral surface, is shown in Figure 9.13. Once the energy source has been established, a finite set of chemical reactions, perhaps phosphate linked to the mineral surface, provides a recipe book for a primitive metabolism: the case for a mineral-based metabolism looks increasingly compelling.

Bacterial metabolisms that have derived an independent source of energy are generically called autotrophs. Bacteria in several places around the biosphere have discovered their own autotrophic energy sources (Table 9.3). There are significant problems, however, in estimating the thermodynamics of these reactions due to the states of some of the species involved. Reactions on the surface of minerals require the thermodynamics of the species on the surface to be known, including the role the surface has in breaking or making chemical bonds. We have also seen that once some species are formed they go on to react rapidly with others, taking for example the formation of carbonic acid in the analysis of the pH of the oceans. The restricted concentration of species within the cell adds a further constraint to the available reactions. The autotrophic reactions listed Table 9.3 are good estimates of the primary energy sources but are a long way from understanding the reactions forming the metabolism of an organism. Putting this into context, there are thought to be about 4500 reactions associated with the metabolism of a single cell.

Table 9.3 Autotrophic reactions and reaction enthalpy

Energy-yielding reactions	Type of organism	Estimated $\Delta_r H$ (kJ mol $^{-1}$)
$4\text{H}_{2(g)} + \text{CO}_{2(g)} \rightarrow \text{CH}_{4(g)} + 2\text{H}_2\text{O}_{(g)}$	<i>Methanothermos, Methanopyrus</i>	-169
$\text{H}_2 + \text{S} \rightarrow \text{H}_2\text{S}_{(aq)}$	<i>Acidianus</i>	-40
$4\text{H}_2 + \text{H}_2\text{SO}_4 \rightarrow \text{H}_2\text{S} + 4\text{H}_2\text{O}$	<i>Archeoglobulus</i>	-169
$\text{H}_2 + \text{HNO}_3 \rightarrow \text{HNO}_2 + \text{H}_2\text{O}$	<i>Aquifex</i>	-155
$\text{H}_2 + \frac{1}{2}\text{O}_2 \rightarrow \text{H}_2\text{O}$	<i>Aquifex, Pyrobaculum</i>	-286
$2\text{S} + 3\text{O}_2 + 2\text{H}_2\text{O} \rightarrow 2\text{H}_2\text{SO}_4$	<i>Sulfolobus, Acidianus, Metallosphaera</i>	-1250
$\text{FeS}_2 + 7\text{O}_2 + 2\text{H}_2\text{O} \rightarrow 2\text{FeSO}_4 + 2\text{H}_2\text{SO}_4$		-3280
$\text{H}_2 + 6\text{ Fe(OH)}_3 \rightarrow 2\text{Fe}_3\text{O}_4 + 10\text{H}_2\text{O}$	<i>Pyrobaculum</i>	-560
$4\text{H}_2 + \text{HNO}_3 \rightarrow \text{NH}_4\text{OH} + 2\text{H}_2\text{O}$	<i>Pyrolobus</i>	-81
$2\text{FeCO}_3 + \text{HNO}_3 + 5\text{H}_2\text{O} \rightarrow 2\text{Fe(OH)}_3 + \text{HNO}_2 + 2\text{H}_2\text{CO}_3$	<i>Feroglobus</i>	+807

The adaptation of an organism to life below the surface or out of the light requires metabolic innovation that allows other sources of energy, whether from H₂, Fe²⁺ or Mn²⁺, to reduce or oxidise sulphur or nitrogen and electron donors other than water that have enabled organisms to colonise the deep subsurface. Unfortunately, tolerance to wild variations in conditions also has to be endured, from temperature changes to pH and salinity: requiring a complex organism that may have evolved on an evolutionarily short timescale.

9.5 Microbial Mars

Psychrophiles in Antarctica give some insight into one of the more interesting life questions: life on Mars. Low temperature is not a restriction to the growth or hibernation of bacteria. The Antarctic psychrophiles grow happily at -20°C although the doubling time is 180 days. Similarly, storage of cells within the laboratory at liquid nitrogen temperatures (70 K) still results in viable cells when warmed back to room temperature. The record storage temperature is 10 K, at which dry bacterial spores still survive and remain viable on warming. Temperatures on the surface of Mars are low and also suffer from a very dry atmosphere and intense bombardment from radiation and ultraviolet light. The Pathfinder missions to Mars have provided some truly amazing images of the Martian surface, such as in Figure 9.14, and some insight into the harsh environment. The two roving probes Spirit and Opportunity have been moving around the surface of Mars analysing rock mineralogy and in particular looking for signs of life. The mineral structure is dominated by red iron minerals with water-active surfaces and the possibility of features in rocks that look like regions where water has once flowed. Mars could have been in the habitable zone of the Sun 1 Gyr ago, suggesting that the solvent for life was available.



Figure 9.14 Pathfinder image of the Martian surface. (Reproduced by permission of NASA Pathfinder Mission, Jet Propulsion Laboratory, University of Arizona. (A colour reproduction of this figure can be seen in the colour section)

The current surface temperature has a diurnal (daily) variation from 150 to 298 K with a surface pressure of 5.6 mbar (1/1000 of the Earth's atmospheric pressure). The atmosphere is rather tenuous, constituting 95 per cent CO₂, 2.7 per cent N₂ and 1.6 per cent Ar with little or no water. However, the evidence for the existence of liquid water on Mars is growing with the observation of what looks like runoff geological features on the surface, perhaps from ancient rivers. Water delivered to the surface of Mars by comets in the same way as cometary water was delivered to the Earth would suggest that water was present some 3.5 Gyr ago but little remains on the surface at present: the subsurface is another matter. The interior of Mars may be somewhat warmer than the surface, not suffering from the same diurnal temperature variations, and would provide protection from the Sun's hard UV radiation. The diurnal extremes of the surface temperature variation on Mars are expected to average out at depths as shallow as 15–20 cm, converging on the average temperature of –50°C. The general conditions are then not too different from those in the Antarctic, suggesting that subsurface colonies of bacteria may survive in permafrost layers of water and CO₂ ice in the polar regions of Mars.

Concepts and calculations

Concepts

<i>Protocells</i>	Amphiphilic molecules spontaneously form cell-like structures that, by a process of encapsulation, acquire a chemical inventory
<i>Chemical potential gradients</i>	The chemical potential of a species can vary from place to place, e.g. inside and outside a cell, giving rise to a concentration difference and hence a concentration gradient. The inclusion of ions adds an electrochemical gradient and the two may oppose one another
<i>Membrane transport</i>	Once a membrane has formed, nutrients have to be transported across the membrane either by osmosis or actively
<i>Minimal genome</i>	The minimal gene set for an organism (170 genes) in a minimum volume cell
<i>Universal tree of life</i>	Studies of the genetic history and length of intron DNA lead to a universal tree of life and a common genetic ancestor – a hypothermophile bacterium
<i>Astrobiology</i>	The search for exotic environments in which life has evolved. Drawing inspiration from the extremophile bacteria from the Earth's biota, knowing where to look on another planet is not an easy question

Calculations

<i>Membrane transport</i>	Osmotic pressure and ion transport thermodynamics
<i>Chemical reaction networks</i>	Extending the idea of chemical networks to extreme environments of temperature and pressure to discover autotrophs

Problems

- 9.1** The formula for the surface area of a prolate spheroid such as a cell is given by:

$$\text{Surface area} = 2\pi b^2 + 2\pi \frac{ab}{\varepsilon} \sin^{-1} \varepsilon$$

$$\varepsilon = \sqrt{1 - \frac{b^2}{a^2}}$$

(Evaluate the $\sin^{-1} \varepsilon$ function in radians)

where a and b are the major and minor semi-axes. If the dimensions of the cell are $120 \times 40 \mu\text{m}$, then a is $60 \mu\text{m}$ and b is $20 \mu\text{m}$.

- (a) Calculate the surface area of the cell.
- (b) If the surface area of the amphiphilic molecule is 0.45 nm^{-2} , estimate the number of molecules in the membrane by assuming that it forms a bilayer structure.
- (c) The volume of the spheroid is given by $4/3 ab^2$. Calculate the volume of the cell.

- 9.2** An amphiphilic molecule has a single positively charged head group and is in solution with a concentration of 10 mM . If the critical micelle concentration of the molecule is 25 mM and the Krafts temperature is 25°C :

- (a) Calculate the concentration at a silica surface with a charge of -75 mV .
- (b) Will micelles form at the surface?
- (c) What role might the surface have in the stabilisation of the micelles above the Krafts temperature.

- 9.3** The concentration of formaldehyde outside a protocell is 300 mM whereas an encapsulation event has created a concentration inside of 65 mM .

- (a) Calculate the free energy for the migration of formaldehyde in each direction.
- (b) Will an active transport process be required to move formaldehyde into the protocell?

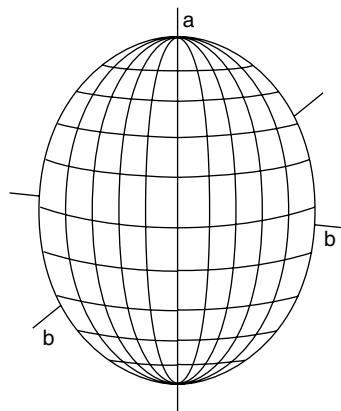
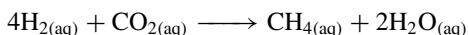


Figure 9.15 Parameters of a Prolate Spheroid

- (c) Comment on the permeability of the membrane structure to formaldehyde.
- 9.4** The protocell concentration of sucrose is 150 mM.
- Calculate the mole fractions of water and sucrose within the cell.
 - Calculate the osmotic pressure within the protocell.
 - How does this change as the temperature is elevated to 70°C in a geothermal vent?
- 9.5** Consider a concentration gradient for NaCl with an internal protocell concentration of 200 mM and an external concentration of 5 mM, as might be found in a freshwater environment.
- Calculate the membrane potential required to prevent the flow of NaCl down the concentration gradient.
 - How does this vary with an elevated temperature of 50°C, as might be found in a geothermal vent?
- 9.6** The methane-metabolising autotrophic bacteria derive energy from the reaction:



- Calculate the $\Delta_r H$ for this reaction, assuming all standard states in the gas phase (g).
- Assume that $\Delta_f H (\text{CH}_4)_{(\text{aq})}$ is not known and estimate the reaction enthalpy in the aqueous phase.

10

Titan

Introduction

The solar system is home to some extreme environments: water conditions here on Earth, sulphuric acid clouds on Venus, the tenuous atmosphere on Mars and the gas giants Jupiter, Saturn, Uranus and Neptune. None offer particular promise as habitats for life no matter how extreme the biology. Mars may have once been more hospitable but probably not now, at least not on the surface. Two other bodies in the solar system are interesting, however: Titan with a nitrogen–methane atmosphere thicker than that on Earth (Figure 10.1) and Europa with a 100-km thick ice layer possibly covering liquid water oceans (Figure 10.2). Titan is the source of much excitement as I write, with the arrival of the Cassini–Huygens mission at the Saturnian system and the recent descent of the Huygens probe to the surface of Titan.

The Huygens probe has made a safe 2 h descent through Titan's atmosphere some 500-km thick, performing measurements of atmospheric composition and weather before landing on the surface where more data were recorded. The results from this mission are immediately exciting because we have the first look at the chemistry and maybe astrobiology of another large object with an atmosphere. The atmosphere is principally nitrogen, as on Earth, with about 3 per cent carbon in the form of methane. The atmosphere is sufficiently dense to allow sound waves to propagate and sufficiently cold for methane and hydrocarbons to condense into oceans. Hence it may be possible to listen to the waves breaking on a snowy methane shore of a hydrocarbon ocean. Sound recordings are now available from the Huygens probe, reporting the very strange sounds of nitrogen and methane winds in another very cold world. Curiously, it would be possible to play music on the surface of Titan with conventional instruments although the pressure and composition of the atmosphere will have some strange effects on the tuning, especially for the ‘wind’ instruments. At the time of writing in the summer of 2005, there is a large data set now being analysed that will provide amazing detail to the chemistry of Titan and generate yet more questions. What is clear already, however, is that we must go back to Titan for a second look.

By contrast, Europa is another set of chemical conditions below a hugely thick ice layer. The surface bears the scars of impact erosion, with debris from the

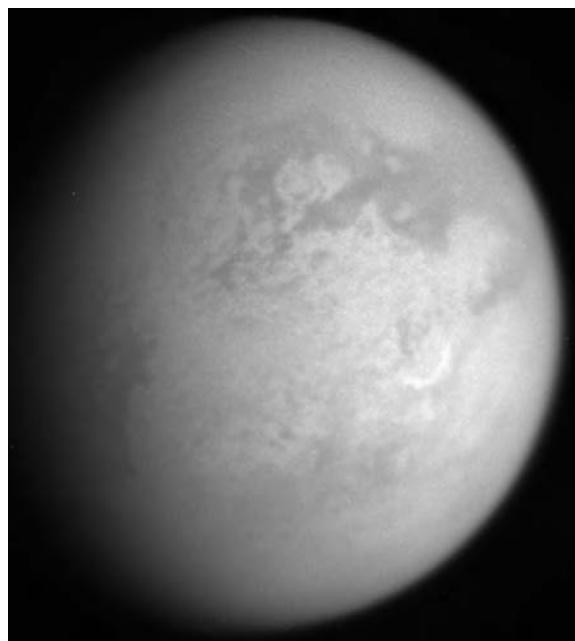


Figure 10.1 This image taken by the Cassini probe is in the infrared centred at 980 nm (Image scale is 7 kilometers per pixel) and shows features on the leading hemisphere of Titan, including the bright, crescent-shaped *Hotei arcus* (right of center), which is also informally called “the Smile” by researchers. (Reproduced by permission of NASA/JPL/Space Science Institute)

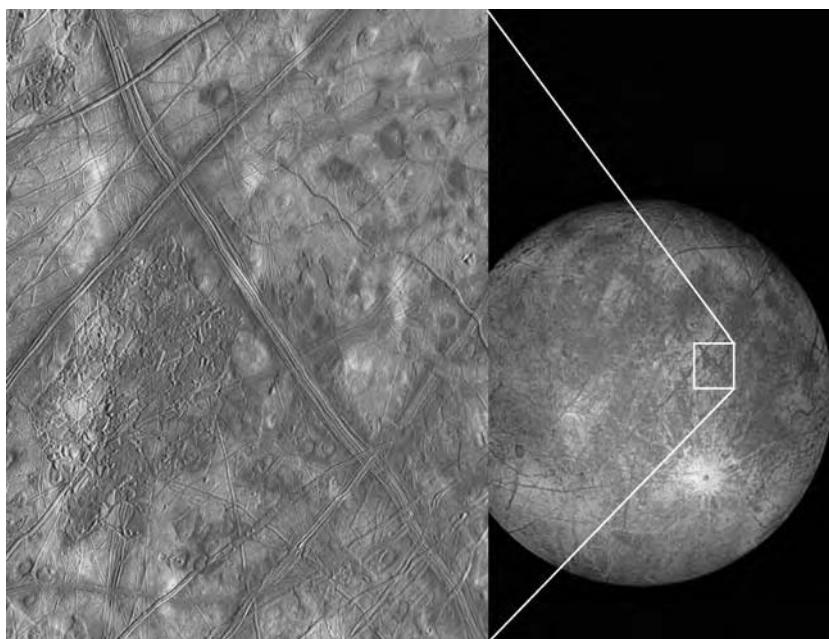


Figure 10.2 Impact-grazed icy surface of Europa. (Reproduced by permission of NASA Gallileo Mission, Jet propulsioin Laboratory, University of Arizona)

collisions providing some colour. The reflection spectra from the surface confirm that it is water ice and this Jovian moon would have formed in the subsolar nebula around the gas giant, causing the deflection of cometary objects towards Earth. A large water hydrosphere may have resulted, which has subsequently frozen. However, the density of ice is lower than that of liquid water so ice floats. With the presence of nearby massive Jupiter, potentially huge tidal forces are present that would heat the water below the 100-km thick ice sheet. Add to this warm wet environment the cometary inventory of organic molecules, and Europa must have some interesting possibilities for water–carbon chemistry. The composition of the oceans is, however, unknown and it seems to be too much of a challenge to drill through 100 km of ice to discover the interior. This makes a trip to Europa to skate on the ice less attractive.

10.1 Physical properties

The first encounter with Titan was a fly-by of the Pioneer spacecraft in 1979, although better data were recovered from Voyager I and II, with Voyager I coming within 4394 km of the surface. Most of what we know about Titan, however, comes from infrared measurements from Earth and in space using the Infrared Space Observatory and the Hubble Space Telescope. All of this has changed, however, with the arrival of the Cassini–Huygens mission to the Saturnian system. For the moment, the pre-Cassini–Huygens physical data of Titan are summarised in Table 10.1. All infrared measurements of the composition of the atmosphere suggest that nitrogen makes up 90–99 per cent of the atmosphere. The remaining 1.7–4.5 per cent (depending on latitude) is made up of hydrocarbons, principally methane. The photochemistry within the atmosphere has led to a complex spectrum in the infrared, identifying long-chain hydrocarbon species containing C, H and N (Figure 10.3) as well as the recent discovery of water by the Infrared Space Observatory.

Table 10.1 Physical data for Titan

Physical property (pre-Cassini–Huygens)	
Mean diameter	5153 km (0.404 Earth diameters)
Mass	(0.0226 Earth masses)
Mean density	1
Distance from Saturn	1.2×10^6 km
Orbital period around Saturn	15.95 days
Orbital period of the Saturnian system around the Sun	29.5 years
Surface temperature	96 ± 4 K
Tropopause (42 km)	71.4 K
Stratosphere (200 km)	~ 170 K
Surface pressure	1496 mbar
Tropopause pressure	135 mbar
Surface gravity	1.35 m s^{-2}

Source: ESA = European Space Agency

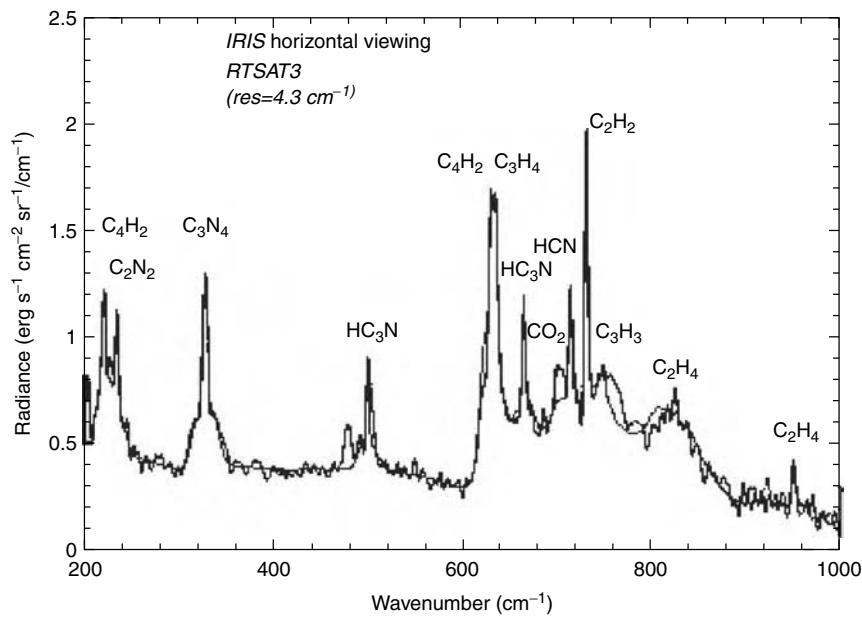


Figure 10.3 Voyager I IRIS spectrum of the northern limb of Titan. (Reproduced by permission of Voyager Missions, JPL NASA)

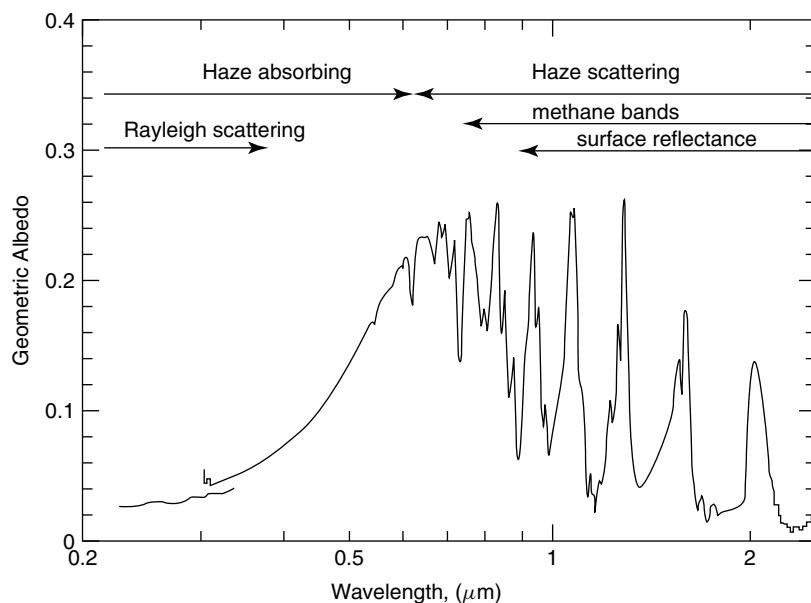


Figure 10.4 Reflection spectrum. (Reproduced by permission of Dr. Chris McKay, NASA Ames)

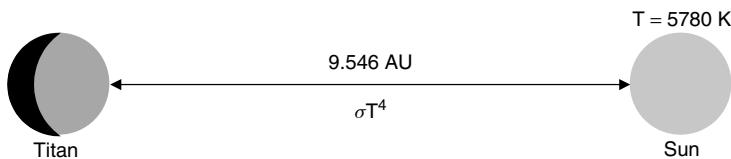


Figure 10.5 Effective surface temperature of Titan

The reflection spectrum of the atmosphere is a measure of the albedo of the planet (Figure 10.4) and, despite the strong methane absorption in the red, Titan’s disc looks orange principally due to scatter from the surface of dense methane–hydrocarbon clouds. Scatter from aerosol particles within the thick clouds obscures the surface of the moon although the radar analysis reveals considerable Chapman layer structure within the atmosphere and some interesting surface features.

A measure of the effective surface temperature of Titan follows from the calculations performed in Section 7.6, allowing both the Sun and Titan to be treated as black bodies (Figure 10.5). The albedo of Titan’s cloud tops suggests that the moon absorbs most of the Sun’s radiation heating the atmosphere and the surface. The effective surface temperature from the energy balance calculation is 82 K, which compares with the measured (average) surface temperature of 92 K, pointing to significant global warming in the atmosphere.

10.2 The atmosphere

Why should such a small moon with a low mass, and hence gravity, be capable of capturing and maintaining such a large atmosphere when other objects such as the comparatively huge Mars cannot? Even comparable moons around Jupiter that have bracketing masses, Ganemede and Callisto, only manage a tenuous atmosphere. Gravitational capture from the Saturn subnebula is unlikely, especially as the Ne/N ratio is 1.1:1 on the Sun whereas on Titan it is more like 0.001:1 and there are also some problems with the D/H ratio. The atmosphere from Titan does not look like it came directly from the solar nebula or the subnebula around Saturn.

The next most likely possibility is cometary delivery of the atmosphere but again there are some problems with the isotope ratios, this time with D/H. The cometary D/H ratios measured in methane from Halley are $31 \pm 3 \times 10^{-5}$ and $29 \pm 10 \times 10^{-5}$ in Hayuatake and $33 \pm 8 \times 10^{-5}$ in Hale-Bopp, whereas methane measurements from Earth of the Titan atmosphere suggest a methane D/H ratio of $10 \pm 5 \times 10^{-5}$, which is considerably smaller than the ratio in the comets. The methane at least in Titan’s atmosphere is not exclusively from cometary sources. Degassing of the rocks from which Titan was formed could be a useful source of methane, especially as the subnebula temperature around Saturn (100 K) is somewhat cooler than that around Jupiter. This would allow volatiles to be more easily trapped on Titan and contribute to the formation of a denser atmosphere. This mechanism would, however, apply to all of Saturn’s moons equally and this is not the case.

The volatile-trapping mechanism has a further problem associated with the temperature. Very volatile molecules such as N₂, CO and CH₄ are not easily trapped in laboratory ice simulation experiments unless the ice temperature is 75 K, which is somewhat lower than the estimated Saturnian subnebula temperature. This has led to the suggestion that the primary source of nitrogen within the Titan surface ices was NH₃, which became rapidly photolysed to produce H₂ and N₂ upon release from the ice. The surface gravity is insufficient to trap the H₂ formed and this would be lost to space. However, the origin of methane on Titan is an interesting question. Methane is a minor component of comets, with a CH₄/CO ratio of <10⁻¹ compared with the present atmospheric ratio of > 10². The D/H ratio is also intermediate between that of comets and the solar nebula, so there must be an alternative source of methane that maintains the carbon isotope ratio and the D/H isotope ratio and explains the abundance on Titan.

Some of the very first observations from the Cassini radar analysis of Titan's surface showed regions where volcanic activity has occurred but with volcanoes erupting water–ammonia ice. This new cryovulcanism seems to suggest that a slurry of water–ammonia ice erupted from below Titan's surface in a similar way to basaltic flow on the Earth. The ammonia-rich slurry would degas ammonia to form nitrogen and hydrogen by photolysis once in the atmosphere. The surface gravity then traps the nitrogen and allows the hydrogen to escape. The presence of ammonia would lower the freezing point of water and decrease the density of the resulting fluid so that it has the buoyancy of water but is much more viscous. These statements require the rather esoteric field of liquid viscosities to be understood in the laboratory properly before being applied accurately to the new problem of cryovulcanism. Having suggested the very plausible mechanism of ammonia photolysis for the origin of the molecular nitrogen, this begs the question: where did all of the ammonia come from?

Current models of Titan's past suggest that the current surface does not tell the full story. It is suggested that below the water-ice crust is a liquid water–ammonia solution that may be as much as 200 km deep. Titan's subterranean ocean is one of four in the solar system: Earth, Titan, Europa and Jupiter's moon Calisto. The warm, electrochemically rich environment of geothermal vents in the Earth's oceans suggests that there may be other thermal vents within the oceans of solar system objects such as Titan where life may have found a small foothold. Titan must have condensed from a local nebula around Saturn that is likely to be rich in ammonia and methane, as has been seen in the atmosphere of Saturn. Interestingly the composition of the gas giants was Urey and Miller's motivation for choosing their nitrogen-rich atmosphere in their initial experiments. As we saw in Chapter 8, these atmospheres produced reasonable yields of prebiotic molecules (possibly with the exception of ribose) and so in a thermal vent environment similar formose and Strecker syntheses may occur. The fraction of NH₃ in the surface ocean may have been as large as 15 per cent by weight and exposed to the surface of perhaps a warmer Sun. The atmospheric composition would reflect the ocean composition and the potential for prebiotic synthesis. The initial temperature of

the surface ocean during formation may have been as large as 300 K (at least according to some speculation) and the oceans would have been ideal for prebiotic chemistry, although perhaps a little restricted by the solar flux at the surface. The NH₃-rich oceans, however, when photochemically processed (somewhat slowly) could explain the N₂-rich atmosphere seen today. Speculation will, in part, be constrained by the atmospheric composition measurements made by the Huygens probe during descent.

The structure of the current Titan atmosphere known prior to the Huygens descent is shown in Figure 10.6, with a temperature profile and chemical/Chapman layer structure very similar to the Earth's atmosphere. The surface temperature is around 92 K and the lapse rate through the troposphere was determined by the Voyager measurements. A temperature inversion occurs at an altitude of 50 km, allowing the stratospheric temperature to rise to 170 K. The small hydrocarbon molecule composition of the atmosphere, principally methane and extended hydrocarbons, is an effective short-wavelength shield with photochemistry inducing the formation of high-molecular-weight hydrocarbons known generally as 'tholins'. A cycle of heating and cooling with convection allows methane to reach its saturated vapour pressure at 1.5 bar to form clouds. The hydrocarbon hydrosphere, called the liposphere (from the Greek *lipos* meaning 'oil'), here completes the cycle, condensing into lakes and oceans of hydrocarbon, hydrocarbon snow and rain.

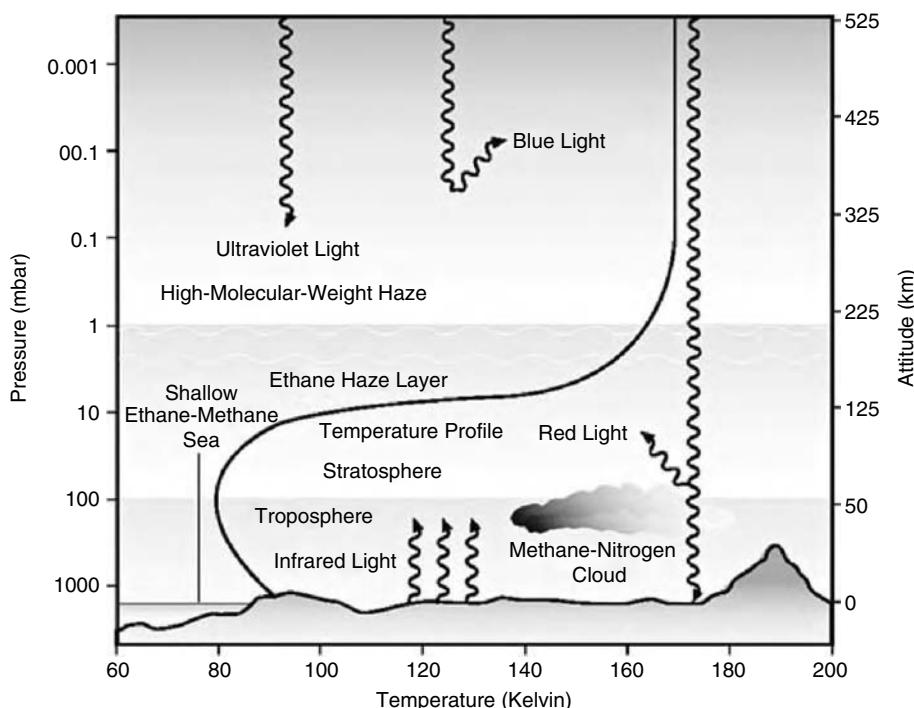


Figure 10.6 Structure of atmosphere. (Reproduced by permission of JPL NASA)

The conditions on Titan, both in the atmosphere and in the oceans, can be investigated using the kinetics and thermodynamics introduced in the modelling of the ISM and the prebiotic Earth, now tuned to the surface temperature and atmospheric temperature conditions on Titan. We have seen previously what happens to reaction rates in the ISM and the atmosphere using the Arrhenius equation but we have not yet extended the concepts of ΔG and thermodynamics to low temperatures.

10.3 Temperature-dependent chemistry

The temperature variation throughout Titan's atmosphere from a cold 92 K at the surface to the somewhat warmer 170 K of the stratosphere suggests a number of different chemistry regimes. The calculation of rates at different temperatures follows the detailed analysis in Section 5.4 and can be applied to the indicative temperatures as follows:

$$\frac{k(92) \propto 92^{0.5} \exp\left[-\frac{100 \times 10^3}{(R 92)}\right]}{k(298) \propto 298^{0.5} \exp\left[-\frac{100 \times 10^3}{(R 298)}\right]} = 3.1 \times 10^{-40} \quad (10.1)$$

$$\frac{k(170) \propto 170^{0.5} \exp\left[-\frac{100 \times 10^3}{(R 170)}\right]}{k(298) \propto 298^{0.5} \exp\left[-\frac{100 \times 10^3}{(R 298)}\right]} = 4.8 \times 10^{-14}$$

Essentially, all reactions that require the formation of a chemical bond with an activation energy of around 100 kJ mol⁻¹ are frozen out at the surface of Titan but are considerably faster in the stratosphere, although still rather slow compared with the rates of reaction at 298 K. Chemistry in the atmosphere of Titan will proceed slowly for neutral reactions but faster for ion–molecule reactions and radical–neutral reactions, both of which have low activation barriers. The Arrhenius equation provides the temperature dependence of rates of reactions but we also need to consider the effect of cold temperatures on thermodynamics and in particular equilibrium.

The extension of thermodynamic calculations to low temperatures requires knowledge of how the equilibrium composition of a mixture, which varies at different temperatures, can be derived from the standard relation between ΔG and the equilibrium constant (Equation 8.12) to give the van't Hoff equation for the variation of the equilibrium constant with temperature:

$$\left(\frac{d \ln K}{dT}\right) = \frac{\Delta_r H}{RT^2} \quad (10.2)$$

The equilibrium constant is fairly easy to measure as a function of temperature in the laboratory and this provides a way of determining $\Delta_r H$. The integrated form

of Equation 10.2 has the following form:

$$\ln K_2 - \ln K_1 = -\frac{\Delta_r H}{R} \left(\frac{1}{T_2} - \frac{1}{T_1} \right) \quad (10.3)$$

A plot of $1/T_2$ versus $\ln K_2$ gives a straight line with slope $-\Delta_r H/R$, however the line will not be straight for all temperatures, especially when they become rather low such as the effective surface temperature of Titan (98 K). The temperature also changes the chemical potential, notably with respect to a phase change, and the material will find a chemical potential gradient towards a phase change and become a liquid, then ultimately a solid. The formation of the solid or liquid phases will remove species from the gas phase and change the equilibrium of a reaction dramatically. The formation of liquid methane, for example, at the surface temperature and pressure on Titan is the favoured phase and this is true for most hydrocarbons. Understanding the phase diagram for a pure species, or worse still mixtures of hydrocarbon species, within the liposphere is a non-trivial thermodynamics problem.

Example 10.1

The gas-phase formation of ozone is via the following reaction:



The equilibrium constant K_2 for the reaction at 298 K is 1.05. Calculate the value of K_2 at 98 K and 170 K; $\Delta_r H$ for the reaction can be calculated from the data in Appendix C and is $-106.47 \text{ kJ mol}^{-1}$. Substituting this into Equation 10.3 gives:

$$\begin{aligned} \ln K_2 - \ln(1.05) &= -\frac{(-106470)}{8.314} \left(\frac{1}{98} - \frac{1}{298} \right) \\ K_2 &= 8.85 \times 10^{-39} \end{aligned} \quad (10.5)$$

Repeating the calculation at 170 K gives $K_2 = 32.41$, which is considerably larger. At the colder temperature the reaction is strongly in favour of the reactants. This suggests that any source of oxygen atoms from water photolysis of cometary material would not form at the surface but the higher altitude layers with different chemistry may indeed form. Unfortunately, as we have seen, the formation of O_2 from water photolysis is too slow by the simplest mechanisms.

What we have deliberately avoided throughout the discussion of the variation of equilibrium with temperature is the idea of phase change. A moment's thought suggests that all molecular species have the potential to change phase from gas to liquid and then from liquid to solid. A precise understanding of this requires

a detailed knowledge of the phase diagram not just for single species but also for mixtures such as the ammonia–water ice at the temperatures and pressures of interest. The application of thermodynamics and kinetics to take account of the possibility of phase changes remains a tough problem in the laboratory.

10.4 Energy balance and the greenhouse effect

Titan is approximately 9.5 AU from the Sun and has a surface albedo of around 0.2. The energy balance calculation (Figure 10.5) suggests that the surface temperature of Titan as a black body should be 82 K, indicating that the atmosphere is responsible for approximately 10 K of global warming, principally from CH₄. The CH₄ greenhouse processes do not appear to be saturated, as can be seen in Figure 10.4 where the methane absorption bands do not absorb all of the radiation at the infrared absorption band wavelengths of the black body emission. The thermal radiation re-radiated from Titan as a black body with a temperature of 82 K has a wavelength maximum given by Wein's Law in the infrared at 34.4 μm. On Earth this radiation would be trapped in the spectrum of polyatomic molecules such as water and CO₂, where their heteroatom constituents have a strong rotational and vibrational spectrum. On Titan the dominant constituent is N₂, which to a very good approximation has no rotational spectrum so that Titan's atmosphere has some thermal windows where radiation from the surface and the atmosphere is lost to space. To summarise, the energy balance for Titan is: approximately 30 per cent of the incoming solar radiation is reflected, 60 per cent is absorbed in the atmosphere and 10 per cent is absorbed by the surface. The comparable figures on Earth are 20 per cent, 35 per cent and 45 per cent, respectively.

Ten per cent of the incoming solar radiation is believed to reach the surface of Titan, heating the ground and the troposphere adjacent to it. The troposphere then becomes unstable with respect to convection, causing the warmer gases to rise in Titan's weak gravitational field. The temperature lapse rate on Titan is controlled by the heat capacity of the atmosphere and the pressure and hence surface gravity, which for Titan is 1.4 K km⁻¹ compared with Earth at 10 K km⁻¹. Voyager measurements from radar occultation confirm these numbers and the general structure of the atmosphere, as in Figure 10.6. The temperature inversion at the tropopause is consistent with heating in this region of the atmosphere due to the absorption of CH₄ and radiation trapping by the aerosol particles. The near-vertical temperature profile for Titan in the stratosphere is very different from Earth and consistent with a high proportion of aerosol particles and organic vapour at these levels. Thus the majority of incoming radiation from the Sun is absorbed high up in the atmosphere. The radiation-trapping property of the atmosphere at all wavelengths is generally referred to as the opacity function and is not well understood for Titan at present. The Huygens data will doubtless change this.

The temperatures in the troposphere allow methane to reach its saturated vapour pressure and hence the potential to condense and produce precipitation. This suggests that regions of the surface may be covered with liquid methane or, more

generally, hydrocarbon molecules produced in the higher altitudes and transported to the surface. The understanding of the infrared opacity function can be improved by allowing for methane supersaturation in the troposphere, leading to clouds. There are some observations consistent with low-level condensates – clouds in the IRIS data consistent with particle sizes of 5–10 μm . Nitrogen is fairly soluble in liquid methane, perhaps with droplets containing about 25 per cent nitrogen, and may aggregate to rather large sizes. The size of raindrops is determined by the aerodynamic properties of the drop within the atmosphere, leading to drops perhaps as large as 6 mm on Earth. The drop-size calculation for Titan is somewhat larger at around 9 mm, although it is not clear whether these will form. The liquid composition of the atmosphere may cause problems for the Huygens probe during descent. Layers of methane ice could build up on the outside of the probe and interfere with the operation of the instruments but this has not yet been reported from the data set so far. The temperature profile is determined by the composition and saturation conditions of the species within the atmosphere and thus is a sensitive test of composition. There are, however, unsolved problems with the composition of the stratosphere associated with the rich photochemistry in this region, which is worth a closer look.

10.5 Atmospheric chemistry

The IRIS spectra of Titan's atmosphere returned from Voyager I show extensive organic composition across the entire disc of Titan (Figure 10.7) and only promises (or threatens) more organic complexity when the measurements from the gas chromatography/mass spectrometry instruments on the Huygens probe are returned.

The spectra contain features from some oxygen-bearing molecules, including CO, CO_2 and H_2O , and an indication of perhaps molecular H_2 (360 cm^{-1}) in the troposphere. Direct absorption of UV photons and cosmic rays produces a network of chemical reactions initiated by radical formation and subsequent reactions with low activation energies and favourable collision parameters to survive the cold 170 K temperatures in the stratosphere. The absorption of UV and far-UV photons also produces a short-wavelength shield for the surface of Titan. A complex network of chemical reactions can be constructed based on photolysis of an N_2/CH_4 mixture that is characterised by the ever-increasing molecular complexity and carbon chain length (Figure 10.8). The larger organic species are not limited by the reaction network and lead to the production of tholins that can condense to form aerosol particles. A laboratory campaign is required to measure the rate constants for the reactions and their UV/IR signatures to piece together this chemical puzzle.

The initial photochemical processes for CH_4 are complex, with five competing reaction channels. The amount of each possible reaction, i.e. the number of products appearing in each reaction channel, is called the branching ratio and is a matter of

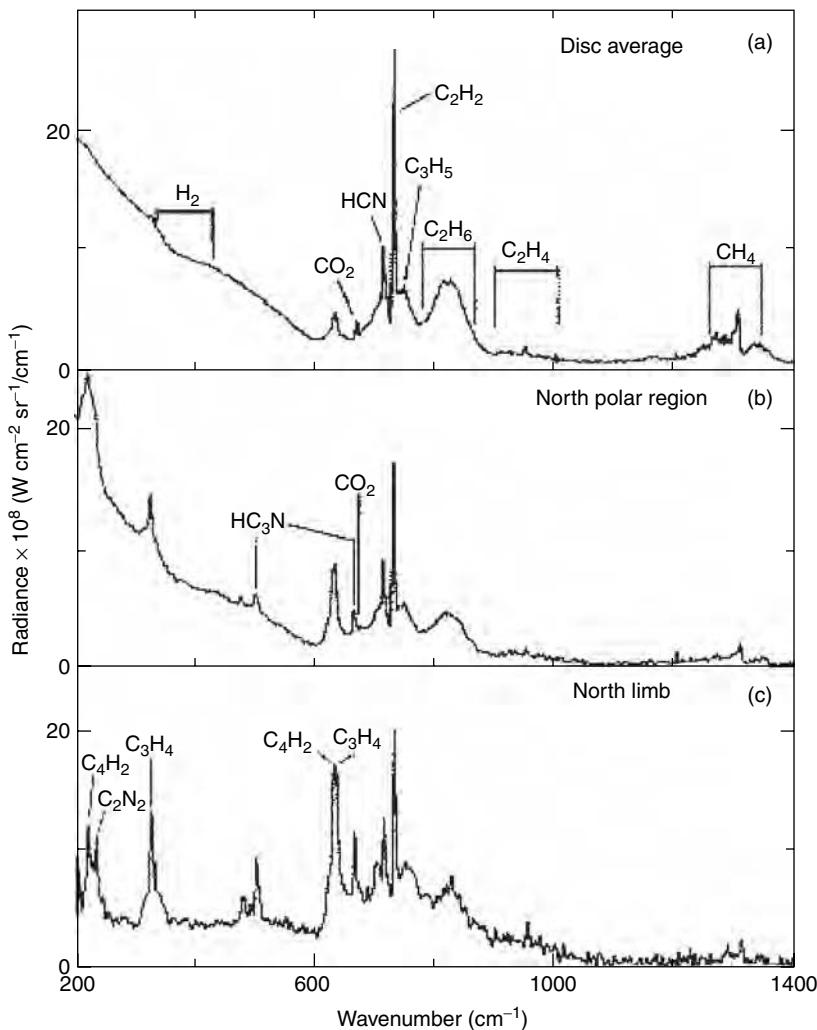
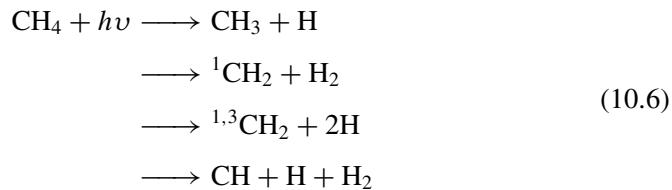


Figure 10.7 Voyager IRIS. (Reproduced by permission of JPL NASA)

fundamental reaction dynamics:



The formation of the ${}^{1,3}\text{CH}_2$ methylene radical in either the triplet or singlet electronic states depends on collisions with other molecules such as N₂ in the Titan

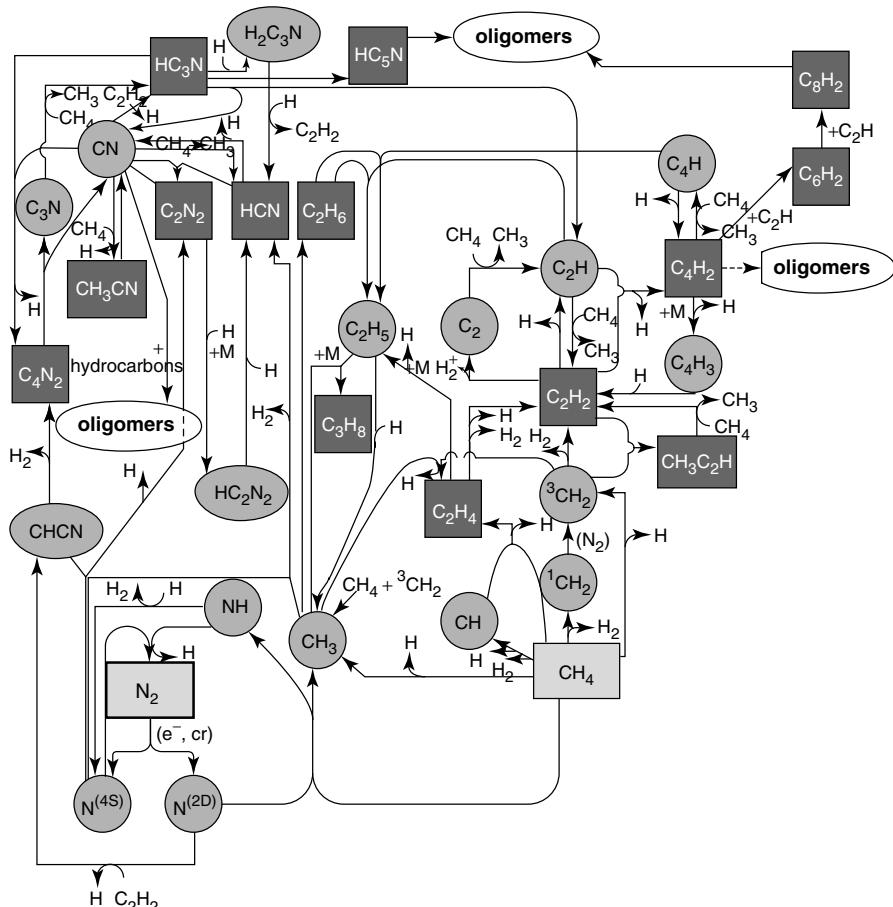
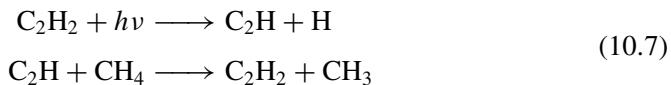
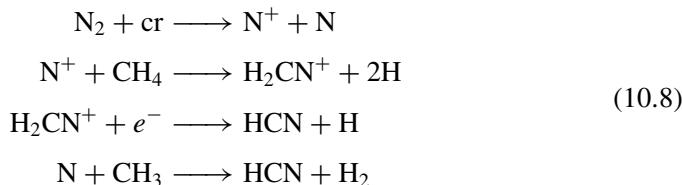


Figure 10.8 Chemical network based on CH_4 and N_2 . (Adapted with permission from Yung *et al.* 1984)

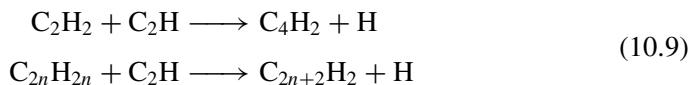
atmosphere. The singlet and triplet designation refers to the number of unpaired electrons. The spin multiplicity is defined as $(2S + 1)$, where S is the total spin: so if one unpaired electron has $S = \frac{1}{2}$, $2S + 1 = 2$, hence a double. Two unpaired electrons has $S = 1$ and $2S + 1 = 3$, hence a triplet. The ${}^1\text{CH}_2$ molecules have the convention ${}^{(2S+1)}\text{CH}_2$ and the singlet has no unpaired electrons with the triplet having two unpaired electrons. The high-energy singlet state is quenched and rapidly becomes ${}^3\text{CH}_2$ in the triplet electronic state. The combination of two methylene molecules to form acetylene (C_2H_2) is the first of the reactions extending the carbon chain length. Photolysis of acetylene leads to the formation of the ethynyl radical, which can go on to react with CH_4 to produce the CH_3 methyl radical:



It is then possible to construct reaction schemes to build all of the hydrocarbon molecules observed to date. The reactions with N₂ require much shorter wavelength photons to break the N≡N triple bond and the chemistry is initiated by cosmic ray (cr) ionisation, with the reactions leading to HCN:

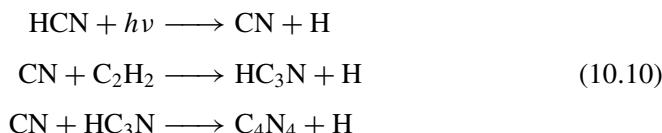


Having initiated the chemistry within the atmospheric model, the elongation of the carbon-chain polyyne species occurs by successive addition of the C₂H radical:



where $n = 1, 2, 3, \dots$. The rate constant for the first reaction in Equation 10.9 has recently been measured at a number of different temperatures from 39 K to 295 K and is found to have the general form $9.53 \times 10^{-11} \exp(30.8/T) \text{ cm}^3 \text{ s}^{-1}$. The rate constant at 180 K in the stratosphere is 15 per cent slower than near the surface at 92 K. The same factors that produced a Chapman layer for ozone and a balance of photon flux, concentration and rate of reaction as a function of altitude will also produce the layers seen on Titan (Figure 10.9). The long-chain hydrocarbons are responsible for the layer structure, with the upper detectable blue layer the result of aerosol particle scatter.

Polymerisation of HCN species is also possible once the initial monomers have been formed by the reactions with nitrogen; HCN polymers have been postulated in many places in the solar system, from the clouds of Jupiter and Saturn to the dark colour of the surface of comet Halley, not to mention its possible role in the formation of the prebiotic molecule adenine. Photolysis of HCN produces CN and then the formation of nitrile polymers:



followed by nitrile polymerisation reactions of the form shown in Figure 10.10.

It is clear that a complex hydrocarbon polymer chemistry must exist in the atmosphere of Titan involving polyyne species, polynitrile species and mixtures of the two, and additional routes to polycyclic aromatic hydrocarbon formation. This presents a significant problem for the gas chromatography/mass spectrometry instruments on the Huygens probe. There should be hydrocarbon fragments, producing perhaps



Figure 10.9 Hazy layers in Titan's atmosphere. (Reproduced by permission of Voyager Project, JPL NASA) (A colour reproduction of this figure can be seen in the colour section)

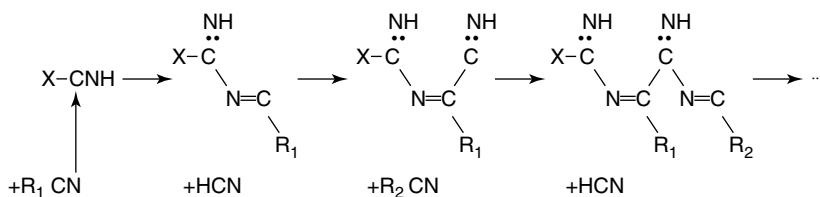


Figure 10.10 Nitrile polymerisation

characteristic cracking profiles as seen in the μ L²MS data from meteorites, but identification of specific species will be very hard indeed.

The diversity of chemical species, particularly polymers, suggests that it would not be difficult to find an information-bearing molecule for a life form on the surface of Titan, although the chemistry for metabolism at the surface temperatures would be slow. The bigger problem may be finding an amphiphilic molecule that will dissolve in the surface bodies of lipofluid to form micelles. A protocell on Titan would still require compartmentalisation and the spontaneous way for this to occur, as we have seen, is through micelle formation. The data from the Huygens probe on both the atmospheric structure and surface characteristics should improve our understanding.

10.6 Astrobiology on Titan

One of the most striking first images returned from the Huygens probe is the mosaic of three images of Titan's surface shown in Figure 10.11, with what appears to be a river delta perhaps flowing into a lake. The contrast in structure in the upper portion of the picture with the very smooth appearance of the lower half suggests a shoreline of a lake. Full analysis of the data, particularly the reflectance spectrum, should strengthen the interpretation of the pictures. A further intriguing possibility is to use sound waves and hence sonar to listen to the waves breaking on the shoreline and to probe the structures below the liquid surface.

The average density of Titan (1.9 g cm^{-3}) suggests an equal mix of silicates and ice, and Cassini measurements of the gravitational field will enable the possibilities of a deep liquid layer to be assessed. The current models assume that Titan accreted rapidly from the cold temperatures of the subnebula around a core of ice and rock, followed by a chondritic layer and finally more ice. Heating from radioactive decay should ensure the sinking or turnover of the core and produce a layered structure of ices and clathrates. The current model is shown in Figure 10.12, suggesting a liquid layer perhaps made up of the ammonia–water–ice components of cometary impact and certainly containing dissolved methane to form clathrates. The high-pressure ices are formed in the layers due to their different densities and conductivities, all

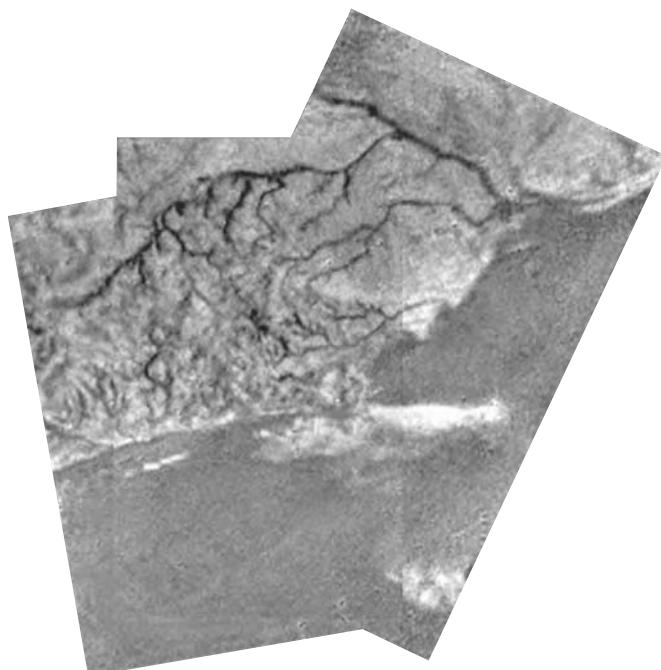


Figure 10.11 Mosaic of river channel and ridge area on Titan during descent. (Reproduced by permission of the European Space Agency, NASA and the Jet propulsion Laboratory, University of Arizona)

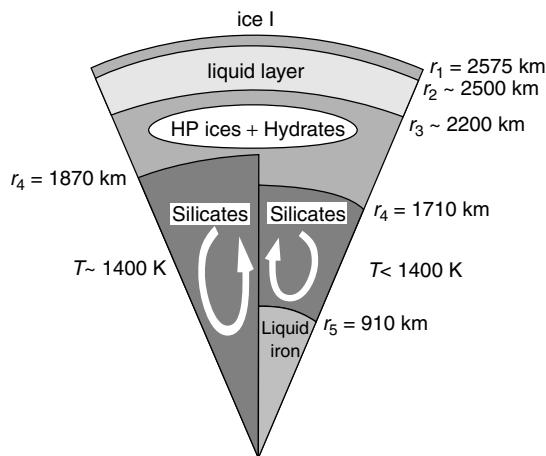


Figure 10.12 Possible internal structure of Titan based on a chondritic core (HP = high pressure). (Reproduced from O. Grasset by permission of Elsevier)

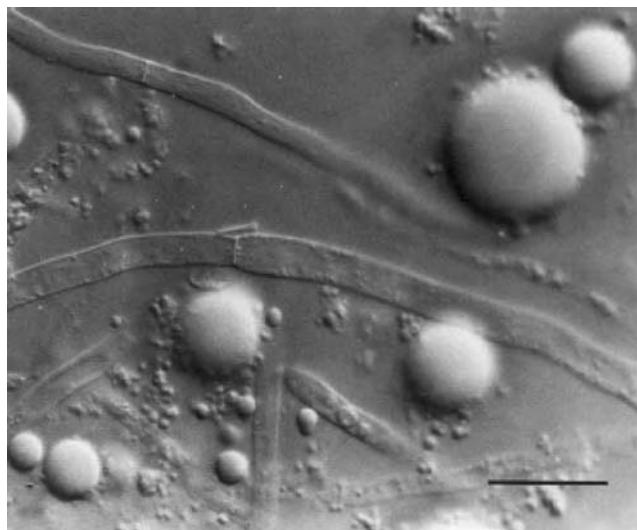


Figure 10.13 Hyphae of *Fusarium alkanophilum* from light hydrocarbons producing important amounts of metabolic water from the substrate. In cultures for more than 90 days, drops of water associated to the soluble degraded products were observed. Scale = $10.5 \mu\text{m}$ (Light Microscope image). (Reproduced from Marcano, 2002, by permission of Elsevier)

of which suggest that Titan either has or once had a well-mixed internal structure. Tidal heating from the presence of Saturn will cause the pressure in the liquid layer to rise, allowing it to force through the outer ice layer that will be encountered by the Huygens probe on landing. Volcanic activity, possibly replenishing the surface with ammonia and water ice, suggests a warmer environment with the liquid state being maintained under the pressure – a lipothermal vent.

Biology on Earth has adapted to living in hydrocarbon environments, with a notable example being the fungus *Fusarium alkanophyllum* (Figure 10.13), which is capable of growth in long-chain non-aromatic hydrocarbon media. The organism is resistant to UV radiation and is viable in the presence of only small amounts of water and CO₂. The small water droplets seen in Figure 10.12 are products of the organism's metabolism, which has evolved to produce water from the hydrocarbon substrate.

The fungus has been shown to grow under extreme conditions, including exposure to UV light, and it responds by producing UV-protective indole pigments that are orange/red in colour. The single most striking fact for cell biology, as well as for astrobiology, is that transmission electron microscopy observations suggest that it does not have a plasma membrane. Instead it modifies the cell wall, made of sugars and proteins, to allow for rapid transport of material from the hydrocarbon medium and for water droplets to be excreted. This incredible adaptation suggests that life on Titan, especially subsurface, could be possible. The astrobiology does not look impossible, with a plentiful supply of organic polymers polymerizing from acetylene and HCN to form long-chain species, and the mechanisms of encapsulation being possible, if a little slow.

Concepts and calculations

Concepts

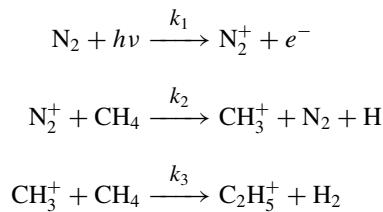
<i>Titan – the test case</i>	Physical and chemical properties of the only other body in our solar system with a significant atmosphere
<i>Atmospheric structure</i>	Composition of N ₂ and CH ₄ with a complex hydrocarbon/Chapman layer structure
<i>Cryovulcanism</i>	Origin of the nitrogen in Titan's atmosphere may be due to the photolysis of ammonia erupting in water-ice slurries from volcanoes at low temperatures
<i>Atmospheric chemistry</i>	Temperature dependence dictates that all neutral chemistry is frozen out at the surface temperature of 92 K. What of radical-driven chemical networks? Stratospheric chemistry at 170 K
<i>Liposphere</i>	The circulation of hydrocarbon-based materials in Titan's atmosphere, which is analogous to the hydrosphere (water cycles) on Earth
<i>Tholins</i>	Long-chain organic polymers responsible for the blue haze in the atmosphere
<i>Astrobiology of Titan</i>	A diversity of environments for life broadens the mind to microbial life forms on Titan

Calculations

<i>Kinetics</i>	Refining kinetics for the conditions on Titan
<i>Equilibrium</i>	Variation of equilibrium constants with temperature – van't Hoff equation
<i>Scale height</i>	The presence of an atmosphere leads to the analogous concepts of scale height
<i>Steady state</i>	Simple calculations for the reactions at the beginning of the processes for tholin formation

Problems

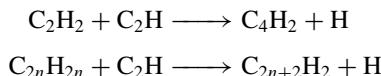
- 10.1** The surface gravity for Titan has been measured as 1.35 m s^{-2} . Given that the radius is 2575 km, estimate the mass of Titan.
- 10.2** Assuming that Titan is a sphere, calculate its mean density and comment on its internal composition.
- 10.3** The effective surface temperature of Titan is 98 K and the atmosphere is nearly completely nitrogen. Calculate the scale height for Titan's atmosphere. Comment on the diurnal variation of the scale height and the variation with altitude.
- 10.4** Use Wein's Law to estimate the black body radiation maximum for Titan, expressing your answer in microns and cm^{-1} . Comment on the spectroscopy of molecules in this region of the electromagnetic spectrum. Identify species in Titan's atmosphere that may be photoactive at these wavelengths.
- 10.5** The photoproduction of CH_3^+ in Titan's atmosphere initiates a sequence of ion–molecule reactions. The mechanism is as follows:



Derive a steady-state concentration for CH_3^+ in this reaction scheme and calculate the steady-state concentration given the following data:

$k_1 = 1.7 \times 10^{-7} \text{ cm}^{-3} \text{ s}^{-1}$, $k_2 = 9.12 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$, $k_3 = 1.10 \times 10^{-9} \text{ cm}^3 \text{ s}^{-1}$, total pressure of the atmosphere is 1 bar, 90 per cent N_2 and 10 per cent methane.

- 10.6** The C_2H polymerisation reaction shown below results in the formation of long-chain hydrocarbon species that form part of the liposphere.



The rate constant for the general insertion reaction is $9.53 \times 10^{-11} \exp(30.8/T) \text{ cm}^3 \text{ s}^{-1}$. Calculate the rates of reaction at 170 K and 98 K and comment on likely layer formation mechanisms.

Glossary of terms and abbreviations

21-cm Line The hyperfine transition in the hydrogen atom associated with the proton and electron nuclear spins in parallel and anti-parallel configurations. This is the primary detection line for H atoms.

67P/Churyumov-Gerasimenko The comet that the Rosetta mission is scheduled to encounter and then send a probe to land on its surface.

Accretion disc The collection of matter around a star that will form into a solar system.

Achondrite A meteorite subclass that does not contain chondrules – small globules of once-molten mineral.

Activation energy The energy required to start a chemical reaction and overcome an energy barrier.

Albedo The amount of radiation reflected from a surface across all wavelengths.

ALH84001 A meteorite found in the Allan Hills in Antarctica in 1984.

Amphiphilic A molecule with a hydrophobic head and a hydrophilic tail – or the other way around.

Archaea Period of time in the evolution of the Earth 3.8–2.5 billion years ago.

Astrobiology The study of life on other planets.

Astrochemistry The study of chemistry away from the Earth.

Atmospheric windows Regions of the electromagnetic spectrum that allow radiation to pass through the atmosphere to reach the ground.

Balmer series The series in the spectrum of atomic hydrogen with lower level $n = 2$ and all other possible levels $n = 3, 4, 5, \dots$

Balmer temperature The temperature of a star determined by the presence of the Balmer series and hence the population of the $n = 2$ level in H atoms on the photosphere of the star.

Beer–Lambert Law The empirical law relating the amount of radiation absorbed by a sample. This law assumes that the sample is a weakly absorbing species.

Big Bang The description of the start of the Universe that is part of the Standard Model.

Big Bang nucleosynthesis The energy-matter conversion at the beginning of the Universe forming the elements H, He and Li directly.

Binary stars A system of two stars that co-orbit one another about a central point.

Biogenic structures Structures in fossils or rock formations that may have been formed by biological activity.

Biomarker A molecule such as ozone in the atmosphere of a planet that might suggest the presence of life.

Birth lines The track of stars as they evolve from red giants onto the main sequence of the Herzprung–Russell diagram.

Black body An object that emits and absorbs all wavelengths with the same, equal efficiency.

Black hole A sufficiently heavy mass the will not allow a photon and hence light to escape as a result of its gravitational attraction.

Blue shift The apparent shift to shorter wavelengths of a light source moving towards an observer.

Boltzmann distribution The law describing the partitioning of the available energy from the local environment among the available levels of a molecule, atom or, in general, a system.

Bond energy The energy required to break a chemical bond and pull a molecule apart.

Brackett series Progression in the spectrum of the hydrogen atom starting with $n = 4$ and present in the infrared.

B/V ratio The ratio of the Blue and Visible regions of the spectrum from which the temperature can be derived assuming a black body.

Carbon insertion reaction The addition of a carbon atom to an existing carbon chain, usually by a radical reaction mechanism.

Carbon–nitrogen–oxygen cycle CNO fusion cycle in stars greater than 1 M_\odot that produces odd-mass nuclei ^{13}C , ^{15}O and ^{15}N .

Carbonaceous chondrite A meteorite containing once-molten globules of rock called chondrules that are surrounded by carbon-containing species.

Cepheid variable Stars that have a periodic variation in their luminosity with a direct relation between the luminosity and the period. These stars are important distance markers.

Chapman layers Layers of molecules forming in an atmosphere as a result of the balance between photolysis and the concentration of the reactants.

Charge transfer reaction A chemical reaction between a charged species and a neutral species where the charge changes location.

Chemical networks The collection of connected chemical reactions that form the description of a complex chemical system. Applications include giant molecular clouds, atmospheres and cellular metabolism.

Chemical potential The change in Gibbs free energy associated with the compositional change of each species.

CHON Carbon, hydrogen, oxygen and nitrogen species found in the ice mantles of dust grains.

Chondrites Meteorites containing once-molten droplets of rock called chondrules.

Chondrules Once-molten droplets of rock that are thought to derive from the protoplanetary disc of the forming solar system.

Circumstellar medium The environment around a star that has chemistry driven by the photon flux or radiation field.

Coercevate The enclosed lipid structure that forms from an acid extraction of kerogen-like material in the Murchison meteorite.

Collision number The number of times per second, on average, that molecules collide with one another.

Collisional dissociation The collision between a molecule and another species with sufficient energy to break one or more bonds in the target molecule.

Coma The central region around a comet immediately surrounding the nucleus.

Comet A ‘dirty snowball’ object in a solar system in orbit around a local star.

Constellation A collection of stars that when viewed from Earth appear to form a shape in the night sky.

Continually habitable zone A region around a star where a planet can maintain a surface temperature above the melting point of water.

Cosmic microwave background radiation Fossil radiation surviving from the Big Bang with a black body temperature of 2.725 K.

Cosmic ray A highly energetic (fast-moving) particle emitted from a star.

Cosmology The study of the Universe, past, present and future.

Cycle of star formation The collapse of a giant molecular cloud forms a star; nuclear synthesis within the star produces more elements; the star ages and ultimately dies in a supernova event; elements are thrown into the interstellar medium to form a giant molecular cloud.

Differential rate equation The equation relating the rate of change of concentration with time of a reactant or product.

Diffuse interstellar bands (DIBs) Unassigned spectral features that are seen along many lines of sight, especially towards reddened stars.

Diffuse interstellar medium The most tenuous regions of space with perhaps one molecule per cm³.

DNA Deoxyribonucleic acid – the information-bearing molecule in most (but not all) cells.

Doppler shift The apparent change in wavelength of a transition associated with the relative motion of the observer and the source: red-shifted, moving away; blue-shifted, moving towards.

Drake equation The factors that might effect the evolution of an intelligent civilisation collected in one equation.

Dust Micron-sized particles of silicate in the interstellar medium responsible for short-wavelength scatter – dust particles become covered with water-ice mantels.

Einstein coefficients The measure of the rate of a transition, whether spontaneous (Einstein A coefficient) or stimulated (Einstein B coefficient).

Enantiomer One stereo-structure around a chiral centre – the left- or right-hand configuration of a molecule.

Encapsulation The process of drying out a liposome and capturing molecules in the rehydration process.

Endogenous organics Organic molecules made on the surface of a planet as part of a prebiotic chemistry, perhaps generated by Urey–Miller syntheses.

Equilibrium The balance of forward and backward rates of reactions to produce a steady-state concentration of products and reactants.

Exogeneous organics Organic molecules delivered to a planet that may seed life in prebiotic chemistry: organic molecules delivered by meteorites or comets.

Expectation value Quantum mechanics only gives a probability of a quantity having a certain value – the expectation value.

Extent of reaction A measure of the progress towards equilibrium: equilibrium may not be achieved if the concentration of one of the reactants is too low.

Extinction The removal of radiation from a path by scatter or absorption processes.

Extrasolar planets Planets around stars other than our Sun.

Extremophile A bacterium that has adapted to life in an extreme environment perhaps of temperature, pressure, salinity, etc.

Flux The number or amount of substance passing a point or through an area: the number of photons leaving a star.

Formose reaction A prebiotic reaction producing sugars but little ribose.

Galaxy A collection of stars located close in space, such as our spiral galaxy, the Milky Way.

Geological time The passage of time recorded in the rocks, usually by radioisotope dating.

Geothermal events Underwater volcanic activity produces a continuous stream of hot water, usually associated with organisms that have adapted to the environment.

Giant molecular cloud (GMC) A region of space with a larger molecular density of $\sim 10^6 \text{ cm}^{-3}$ and a rich chemical composition. The GMC may also contain young stellar objects.

Greenhouse effect Atmospheric gases trap long-wavelength radiation emitted by a planet, resulting in an increased surface temperature.

Habitable zone A region around a star that maintains the surface temperature of a radiatively heated planet above the melting point of water.

Hadean A geological period of time from the formation of the Earth lasting for 800 million years, characterised by huge bombardment of the young planet.

Hale-Bopp Comet from the Oort Cloud with a short period, discovered in 1995 when it was 7AU from the Sun.

Half-life The period over which the concentration of a radionucleus falls to half of its initial concentration.

Halley The comet whose passage around the Sun has been observed 30 times from 239 BC to 1986. The orbital period is 75, with aphelion outside Neptune's orbit and perihelion at 0.59 AU.

Harmonic oscillator The representation of the vibrational motion of a molecule as a perfect spring.

Helium flash A rapid burst of nuclear reactions in the hydrogen-shell burning phase of stellar evolution.

Henry's Law The relation between mole fraction of dissolved gas in solution and partial pressure of the gas above the solution.

Herbig–Haro object A bright object associated with young stellar object, probably due to a region of ionisation associated with high-speed polar jets.

Herzprung–Russell diagram A graphical plot of stellar intensity versus photosphere temperature showing that observed stars fall into classes: main sequence, red giants, supergiants and white dwarfs.

Homochirality The favoured expression of one optical isomer of a molecule over the other – a biomarker.

Hubble constant The numerical constant, H , that relates the red shift of a galaxy to its distance, d .

Hydrophilic Water-loving molecules or region of a molecule that attracts water.

Hydrophobic Water-hating molecules or region of a molecule that repels water.

Hyperthermophile A bacterium that has adapted to an environment of extreme temperature, up to 100°C.

Impact frustration The collision of a large meteorite with a planet that results in the extinction of some or all species on the planet, causing the sterilisation of life. An example would be the extinction of the dinosaurs.

Interstellar ice The formation of ice layers or mantles on the surface of interstellar dust grains formed by the adsorption of O and H separately, forming water ice on the surface.

Interstellar medium (ISM) The tenuous medium between the stars with molecular densities as low as 1 molecule cm⁻³ rising to 10⁶ molecules cm⁻³ in giant molecular clouds. Temperatures may be as low as 10–40 K.

Inverse Square Law The drop of intensity of radiation (for example) with the square of the distance $1/(4\pi d^2)$.

Ionisation energy The energy required to remove an electron from a molecule.

Ion–molecule reaction Reactions between ions and molecules with characteristically low activation energies that are important in the interstellar medium.

Iron catastrophe Melting and fractionation of the interior of a planet, resulting in a molten core in which the density of iron causes it to sink to the core of the planet and generate a magnetic field.

Iron Class of meteorite dominated by iron and nickel in composition.

Jeans mass The mass of giant molecular cloud that is required for it to collapse spontaneously under the force of gravity.

Kerogen The carbonaceous material found around minerals and chondrules in meteorites such as the Murchison meteorite.

Kuiper Belt A source of comets in the solar system at a distance of 100 AU from the Sun.

Linewidth The spread in wavelengths or frequencies associated with a transition in an atom or molecule. There are three contributions: natural linewidth associated with the lifetime of the transition; pressure broadening associated with the presence with the other molecules nearby; Doppler broadening associated with relative motion of the molecule and light source.

Liposome A collection of molecules with hydrophobic tails and hydrophilic heads so that the heads and tails aggregate together to form a spherical structure in solution.

Local group The collection of galaxies, including the Milky Way and Andromeda, that form part of the Virgo cluster as part of the local supercluster of galaxies.

Local thermal equilibrium (LTE) The assumption that all molecules within a particular environment have acquired their fair share of the local energy.

Luminosity The photon flux from the photosphere of a star.

Luminous arc An apparent arc of light around a massive object that is causing a gravitational lensed image of galaxies behind. Can be hundreds of light-years in length.

Lyman series The series of the hydrogen atom spectrum with $n = 1$ as the starting level.

Magellanic clouds Small and large Magellanic Clouds are two irregular galaxies that are the nearest neighbours to the Milky Way and are visible in the southern hemisphere, reported by Magellan in his voyages.

Main sequence star The majority of stars, 92 per cent, that fall on the leading diagonal of the Herzprung–Russell diagram.

Maser The microwave equivalent of a laser that represents some of the strongest transition intensities in the radio region of the electromagnetic spectrum, e.g. the water maser at 22.235 GHz.

Meteor A meteoroid entering the Earth's atmosphere at speeds greater than 70 km s^{-1} that burns up completely before it hits the ground.

Meteorite A meteoroid entering the atmosphere that survives the journey through the atmosphere to land on the ground and become a 'find'.

Meteoroid A particle of interplanetary debris that can enter the atmosphere of a planet to become either a meteor or a meteorite.

Microwave spectroscopy The study of the interaction of microwave radiation with the rotational motion of a molecule. Microwave transitions are not restricted to molecules and can occur in atoms whenever the separation between the energy levels is in the microwave region of the spectrum.

Milky Way The spiral galaxy containing the Sun.

Minimal genome The smallest number of genes required to produce a viable organism. The smallest known genome is *Mycoplasma genitalium*, containing about 480 protein-coding genes.

Moment of inertia The rotational equivalent of mass required to overcome the rotation of a body, depending on the mass and distance of the mass from the axes of rotation.

Moon-capture event The collision of a mars-sized object with the Earth that resulted in the formation of the moon.

Murchison meteorite A carbonaceous chondrite meteorite landing 100 miles north of Melbourne in a town called Murchison.

Neutrino deficit Subatomic particles predicted to be released by the nuclear reactions on the Sun and should be detected on Earth. The number of neutrinos observed on Earth is much less than predicted by the models of solar nuclear fusion.

Neutron star A very compact dense stellar remnant that is stabilised with respect to collapse by the degenerate neutron gas composition of the star.

Nova A temporary flash in the brightness of a star associated with extreme nuclear activity. Compare this with a supernova, which results in the destruction of a massive star.

Nuclear fusion The release of energy associated with the joining or fusing of two nuclei. Nuclear fusion for nuclei lighter than Fe liberates energy.

Nuclear fission Splitting nuclei such as ^{238}U to produce daughter nuclei and an excess energy. This process can occur for nuclei heavier than Fe.

Oort Cloud A source of comets in the solar system at a distance of 50 000 AU.

Operator The mathematical tool required to extract the information from a wavefunction.
The total energy of the system is extracted using the Hamiltonian operator.

Osmosis The movement of water against a sugar concentration gradient through a semi-permeable membrane.

Panspermia The presence or transport of seed molecules for life in all parts of the Universe.

Parsec A unit of distance resulting from a stellar parallax of 1 arcsecond, equal to 3.26 light-years.

Paschen series The series of the hydrogen atom spectrum with $n = 3$ as the starting level.

Pfund series The series of the hydrogen atom spectrum with $n = 5$ as the starting level.

Photic zone The first 200 m of the oceans through which life penetrates.

Photodissociation Breaking of a chemical bond due to interaction with a high-energy photon.

Photoionisation Removing an electron from an atom or molecule due to interaction with a high-energy photon.

Photon-dominated region (PDR) The region around a bright star where photochemical processes dominate the ion–electron, plasma-like state of matter.

Planck length A distance of 1.6×10^{-31} m over which gravity becomes quantised.

Planck time Time required for light to travel the Planck length, 10^{-43} s.

Planck's Law The relation between energy of a photon and its frequency, $E = h\nu$.

Planetary nebula A cloud of glowing, ionised gas ejected by a star as it begins to die.

Polycyclic aromatic hydrocarbons (PAHs) Organic molecules containing many fused six-carbon rings building up from benzene.

Protocell A structure probably forming spontaneously from amphiphilic molecules into a cell-like structure.

Proton–proton cycle The nuclear reaction sequence combining four hydrogen atoms to form one helium nucleus.

Protostar A star early in the evolutionary process, shortly after the initial gravitational collapse.

Purines One of the two types of bases in the nucleic acids adenine and guanine.

Pyrimidines One of two types of bases in the nucleic acids cytosine and thiamine.

Radiation field The region around a star containing photons – the radiation field of photon intensity decreases further from the star.

Radiative association Collision of two atoms or molecules to form a new chemical bond stabilised by radiating the energy of bond formation to produce the molecule in its ground state.

Radiative heating Absorption of radiation by a planet from its local star to create an effective surface temperature.

Radiodating The use of the radionuclide decay half-life to determine the age of a specimen.

Rate constant The constant of a reaction relating the rate of the reaction to the concentration of the reactants.

Rayleigh criterion The rule-of-thumb for the spectral resolution of two transitions.

Red giant A star that has finished core H atom burning that now continues in the shell of the star, resulting in an increase of the star radius and a drop in it luminosity.

Red shift The shift of a transition wavelength to longer, redder wavelengths associated with the relative separating distance between the object and an observer.

Reddening The preferential removal of short-wavelength, blue radiation from a star, leaving it looking redder.

Relativity Einstein's theory invoking a constant speed of light, which divides into special relativity and general relativity, the latter describing gravity.

Ribose The sugar that attaches to the DNA bases to form part of the DNA backbone helical structure. The sugar is not made in the formose reaction process.

Rigid rotor The simple model of molecular rotation assuming that two atoms are connected by rigid bonds when rotating.

RNA Ribonucleic acid analogue of DNA.

RNA World hypothesis The recognition that RNA can self-replicate leads to the idea that information propagation and the genetic code first started with RNA.

Rosetta mission The mission to orbit and send a probe to land on the comet 67P/Churyumov-Gerasimenko.

Rybozyme RNA self-catalysed replication structure.

Scale height The decay constant assuming an exponential decay in pressure and temperature with height in a planet's atmosphere.

Schwarzschild radius The radius of a sphere into which a mass must be confined if it is to be a black hole.

Selection rules The quantum mechanical rules that govern whether a transition between quantum states is allowed or forbidden.

SETI Search for Extra Terrestrial Intelligence – a survey of the radio spectrum looking for signs of communication from other intelligent life.

Short-wavelength shield The protection of the surface of a planet from dangerous short-wavelength radiation from the local star. On the Earth, the ozone layer shields the surface, as does the photic zone.

Snow line The distance from the Sun at which ice and snow can form on the surface of a meteorite or comet.

Solar nebula The giant molecular cloud from which the Sun formed.

Spatial resolution The ability of a telescope or other optical system to resolve sources separated in space – the eye can reveal features some 200 km apart on the surface of the moon.

Spectral mapping Using a known transition in a molecule, such as the 155 GHz transition in CO, to map the column density or concentration of the molecule within a giant molecular cloud.

Standard Model The model for the formation of the Universe and the unification of forces – the current, flawed theory of everything.

Steady-state approximation The balance of rates of formation and destruction of chemical reactions, which, when it equals zero, gives a steady-state concentration of species in the mixture.

Stefan–Boltzmann Law The relation between temperature and luminosity of a star.

Stellar magnitude The astronomical definition of stellar luminosity divided into six classes or magnitudes initially associated with the visibility of stars with the naked eye.

Stony Class of silicate-based meteorites.

Stony-iron Class of silicate–iron-based meteorites.

Stratosphere The layered region of the Earth's atmosphere above 50 km.

Strecker synthesis The prebiotic synthesis of amino acids.

Super string theory The theory of cosmology that attempts to unite gravity and the other forces of nature.

Supernova The explosion of a star resulting in the destruction of a star. Important in the formation of heavier nuclei such as ^{238}U .

Surface gravity The measure of the acceleration due to gravity at the surface of a planet: on Earth this is 9.8 m s^{-2} .

Terraforming Adding chemicals to the atmosphere of a planet such as Mars to initiate a global warming process, raising the surface temperature to a habitable level.

Triple-alpha process The nuclear process resulting in the fusion of three He nuclei or alpha particles to form ^{12}C .

Tropopause The point of temperature inversion in the atmosphere at 10–15 km when the temperature stops falling and begins to rise.

Troposphere The layer of the atmosphere closest to the Earth' surface rising to the tropopause.

T-Tauri stars Stars early in their evolution life cycle that throw off a dust jacket in the form of polar jets and begin to shine.

Uncertainty principle Quantum mechanics restricts the knowledge of certain pairs of variables, notably time and energy and position and momentum, so that complete

knowledge of one variable produces no knowledge about the other. There is an uncertainty in the knowledge of the two variables.

Universal tree of life The categorisation of species on Earth due to its genetic material producing a genetic distance from a common genetic ancestor.

Watson–Crick base pairs The combination of bases that spontaneously pair (A–T and G–C) to form the basis of the genetic code.

Wavefunction The quantum mechanical description of a system such as an atom or molecule. Information about the system is derived by operating on the wavefunction with the appropriate operator.

Wave-particle duality The description of matter as a wave or a particle is possible and neither is preferred. Particle-type questions get particle-type answers and wave-type questions get wave-type answers.

White dwarf The compact remnant of a low-mass star that resists further collapse by the internal degenerate electron gas.

Young stellar object (YSO) Young stellar object – a protostar that is beginning to shine but at low temperature so it has a spectrum with a maximum in the infrared.

Zeeman effect The interaction of energy levels of an atom or molecule with an external magnetic field. The Zeeman interaction changes the energy of the p-orbitals in the atom.

Appendix A

Constants and units

Table 1. Constants

Constant	Value	Units	Symbol
Speed of light (<i>in vacuo</i>) – exact value	2.99792458×10^8	m s^{-1}	c
Planck's constant	6.62608×10^{-34}	J s	h
Boltzmann's constant	1.38065×10^{-23}	J K^{-1}	k
Stefan–Boltzmann constant	5.6696×10^{-8}	$\text{W m}^{-2} \text{ K}^{-4}$	σ
Wein's constant	2.8979×10^{-3}	m K	
Gravitational constant	6.6682×10^{-11}	$\text{N m}^2 \text{ kg}^{-2}$	G
Permittivity of free space	8.8542×10^{-12}	$\text{J}^{-1} \text{ C}^2 \text{ m}^{-1} (\text{F m}^{-1})$	ε_0
$4\pi\varepsilon_0$	1.11265×10^{-10}	$\text{J}^{-1} \text{ C}^2 \text{ m}^{-1}$	
Proton mass	1.6726×10^{-27}	kg	m_p
Neutron mass	1.6749×10^{-27}	kg	m_n
Electron mass	9.1095×10^{-31}	kg	m_e
Electron charge	1.6022×10^{-19}	C	e
Atomic mass unit	1.66054×10^{-27}	kg	u
Avogadro constant	6.02214×10^{23}	mol^{-1}	N_A
Bohr radius	5.29177×10^{-11}	M	a_0
Rydberg constant	1.09737×10^5	cm^{-1}	R
Faraday's constant	9.64853×10^4	C mol^{-1}	F

Table 2. Units used in addition to standard SI units

Unit	Value	Symbol
<i>Length</i>		
Astronomical unit	1.496×10^{11} m	AU
Light-year	9.5×10^{15} m	ly
Parsec	3.26 ly, 3.08×10^{16} m	pc
<i>Time</i>		
Second	1	s
Minute	60 s	min
Day	86 400 s	day
Month	Lunar: 29 days, 12 hours, 44 minutes	month
Year (common)	365.2422 days	yr
Year (sidereal)	365.256 days	

Table 3. The SI prefixes

Prefix	Power factor	Symbol for metres
Exa	10^{18}	Em
Peta	10^{15}	Pm
Tera	10^{12}	Tm
Giga	10^9	Gm
Mega	10^6	Mm
Kilo	10^3	km
Milli	10^{-3}	mm
Micro	10^{-6}	μm
Nano	10^{-9}	nm
Pico	10^{-12}	pm
Femto	10^{-15}	fm
Atto	10^{-18}	am
Zepto	10^{-21}	zm

Conversion factors

1 eV	=	1.60218×10^{-19}	J
		86.485	kJ mol^{-1}
		8065.56	cm^{-1}
1 atm	=	101.325	kPa (defined)
		760	Torr (defined)
1 cm^{-1}	=	1.9864×10^{-23}	J

Appendix B

Astronomical data

Star	Constellation	Type of star	Distance (light-years)	Mass (solar masses)	Radius (solar radius)	Surface temperature (Fahrenheit)	Intrinsic brightness (relative to sun)	Magnitude
Sirius A	Canis Major (Great Dog)	Main sequence	9	2.3	2.5	16 000	35	-1.47
Arcturus	Boötes (Herdsmen)	Red giant	36	4.5	20	7200	100	-0.06
Vega	Lyra (Lyre)	Main sequence	26	3	2.5	18 000	50	0.04
Capella	Auriga (Charioteer)	Red giant	46	3.5	13	9000	100	0.05
Rigel	Orion (Hunter)	Blue supergiant	800	20	36	27 000	60 000	0.14
Procyon A	Canis Minor (Little Dog)	Main sequence	11	1.5	2	12 000	6	0.37
Betelgeuse	Orion	Red supergiant	600	20	1000	3600	14 000	0.41
Altair	Aquila (Eagle)	Main sequence	16	2	1.5	14 000	10	0.77
Aldebaran	Taurus (Bull)	Red giant	70	5	20	7200	100	0.86
Spica	Virgo (Virgin)	Main sequence	260	15	3	45 000	2000	0.91
Pollux	Gemini (Twins)	Red giant	35	4	8	9000	30	1.16
Deneb	Cygnus (Swan)	Yellow supergiant	1600	14	60	18 000	30 000	1.26

Data taken from Appendix 5 of Holliday (2004).

Note: Mass, radius and brightness are given in solar units. For example, Sirius A has 2.3 solar masses, is 2.5 times the size of the Sun and is intrinsically 35 times brighter than the Sun. 1 Solar mass = 2×10^{30} kg = 330 000 Earth masses; 1 solar radius = 700 000 km = 110 Earth radii.

Appendix C

Thermodynamic properties of selected compounds

Quantity	$\Delta_f H^\circ$ (kJ mol ⁻¹)	$\Delta_f G^\circ$ (kJ mol ⁻¹)
C ₁₂ H ₂₄ O ₁₂ (sucrose)	-2222	-1543
C ₂ H ₆ (g)	-84.68	-32.82
C ₆ H ₁₂ O ₆ (β -D-glucose)	-1268	-910
CaCO ₃ (s) (argonite)	-1207.1	-1127.8
CH ₂ (NH ₂)COOH (glycine)	-532.9	-373.4
CH ₃ NH ₂ (g) (methylamine)	-22.97	+32.16
CH ₃ COOH _(g)	-484.5	-389.9
CH ₃ OH _(g)	-200.66	-161.96
CH ₃ OH _(l)	-238.66	-166.27
CH ₄ (g)	-74.81	-50.52
CH ₃ (g)	+138.9	
CH ₂ (g)	+392.0	
CH _(g)	+595.8	
C ₃ H ₆ (g) (propene)	20.42	62.78
CN ⁻ (aq)	+150.6	+172.4
CO _(g)	-110.53	-137.17
CO(NH ₂) _{2(s)} (urea)	-333.51	-197.33
CO ₂ (aq)	-413.80	-385.98
CO ₂ (g)	-393.51	-394.36

Quantity	$\Delta_f H^\circ$ (kJ mol ⁻¹)	$\Delta_f G^\circ$ (kJ mol ⁻¹)
CO_3^{2-} (aq)	-667.14	-527.81
$\text{D}_2\text{O}_{(\text{g})}$	-249.20	-234.54
$\text{D}_2\text{O}_{(\text{l})}$	-294.60	-243.44
Fe^{2+} (aq)	-89.1	-78.90
Fe_2O_3 (haematite)	-824.2	-742.2
Fe^{3+} (aq)	-48.5	-4.7
Fe_3O_4 (magnetite)	-1118.4	-1015.4
FeS	-100.0	-100.4
FeS_2	-178.2	-166.9
H_2CO_3 (aq)	-699.65	-623.08
$\text{HCHO}_{(\text{g})}$	-117.3	-113
$\text{H}_2\text{O}_{(\text{g})}$	-241.82	-237.13
$\text{H}_2\text{O}_{(\text{l})}$	-285.83	-237.13
H_2S (aq)	-39.7	-27.83
H_2S (g)	-20.63	-55.56
H_3PO_3 (aq)	-1277.4	-1018.7
H_3PO_3 (s)	-1279.0	-1119.1
$\text{HCN}_{(\text{g})}$	+135.1	+124.7
$\text{HCN}_{(\text{l})}$	+108.87	+124.97
HCO_3^- (aq)	-691.99	-586.77
NH_3 (aq)	-80.29	-16.45
NH_3 (g)	-46.11	-16.45
NH_4^+ (aq)	-132.51	-79.31
$\text{O}_{(\text{g})}$	+249.17	+231.73
O_3 (g)	+142.7	+163.2
$\text{OH}_{(\text{g})}$	+38.95	+34.23
PO_4^{3-} (aq)	-1277.4	-1018.7
SiO_2 (s)	-910.94	-856.64

Answers to problems

Chapter 2

2.1 (a) $\lambda_{\max}(\text{Betelgeuse}) = 806 \text{ nm}$; $\lambda_{\max}(\text{Aldebaran}) = 403 \text{ nm}$; $\lambda_{\max}(\text{Rigel}) = 107 \text{ nm}$; (b) $F(\text{Betelgeuse}) = 9.52 \times 10^6 \text{ W m}^{-2}$; $F(\text{Aldebaran}) = 1.52 \times 10^8 \text{ W m}^{-2}$; $F(\text{Rigel}) = 3.01 \times 10^{10} \text{ W m}^{-2}$; (c) $L(\text{Betelgeuse}) = 5.9 \times 10^{31} \text{ W}$; $L(\text{Aldebaran}) = 3.71 \times 10^{29} \text{ W}$; $L(\text{Rigel}) = 2.37 \times 10^{32} \text{ W}$.

2.2 (a) $T = 13\,007 \text{ K}$; (b) $\lambda_{\max} = 223 \text{ nm}$; (c) $E = 8.9 \times 10^{-19} \text{ J}$, 537 kJ mol^{-1} ; (c) $34\,900 \text{ W m}^{-2}$

2.6 Betelgeuse = 184.0 pc No, Aldebaran = 21.5 pc Yes and Spica = 79.8 pc Yes.

2.7 $L(\text{Spi})/L(\text{Rig}) = 0.29, -1.25$; absolute magnitude of Rigel is 1.12 and Spica –0.13.

2.8 $f(\text{moon})/f(\text{sun}) = 1.318 \times 10^{-11}$ and $f(\text{Moon}) = 1.8 \times 10^{-8} \text{ W m}^{-2}$.

2.9 Abs Mag (Capella) = –6.36; Abs Mag (Deneb) = –5.15.

2.10 54.2 ly.

Chapter 3

3.1 $l = 38.7 \text{ ly}$.

3.2 656.473 nm: blue-shifted (ascending limb) 656.447 nm and red-shifted (descending limb) 656.499 nm.

3.3 (a) 50.6 km s^{-1} ; (b) 275 km s^{-1} .

3.4 $B = 1.58 \times 10^{20} \text{ J}^{-1} \text{ m}^3 \text{ s}^{-2}$.

3.5 (a) $I = 2.8988 \times 10^{-46} \text{ kg m}^{-2}$; (b) $r = 112.8 \text{ ppm}$.

3.6 $T = 43.9$ K.

3.7 $v = 1.303 \times 10^7$ m s $^{-1}$.

3.8 656.532 nm, 486.321 nm, 434.215 nm and 410.333 nm.

3.9 Ionisation energy = 1311.46 kJ mol $^{-1}$.

3.10 $\mu(^1\text{H}^{35}\text{Cl}) = 1.613 \times 10^{-27}$ kg, $\mu(^1\text{H}^{37}\text{Cl}) = 1.616 \times 10^{-27}$ kg.

3.12 $r = 125.96$ pm, 655.8 GHz.

3.13 $^1\text{H}^{35}\text{Cl}$, $\sigma_e = 3012$ cm $^{-1}$, $^2\text{H}^{35}\text{Cl}$ $\sigma_e = 2156$ cm $^{-1}$.

3.14 $^1\text{H}^{35}\text{Cl}$ $\nu_e = 2990$ cm $^{-1}$, $x_e = 1.766 \times 10^{-2}$.

3.15 (a) $k(\text{CO}) = 1367.1$ N m $^{-1}$; (b) $^{13}\text{C}^{16}\text{O}$ $\nu_e = 1940.9$ cm $^{-1}$; (c) $^{12}\text{C}^{18}\text{O}$ $\nu_e = 1948.4$ cm $^{-1}$.

Chapter 4

4.1 (a) 86 200 K; (b) 94 314 K.

4.2 37 126 K.

4.3 3.49×10^{29} W.

4.4 (a) 3.6×10^{24} W; (b) 5.3 kg m $^{-3}$.

4.5 8.5×10^{-5} .

4.6 $\Delta\lambda = 3.11 \times 10^{-4}$ nm.

4.7 35 Mly.

4.8 (a) 1.48×10^{-27} m; (b) 1.35×10^{-25} m; (c) 1.5×10^{-24} m.

4.9 1.5×10^3 kg.

Chapter 5

5.1 7.25×10^{-4} ly.

5.2 7.3×10^{-3} ly.

5.3 (a) 1.5 years; (b) 3471 s, nearly one per hour.

5.4 19 400 K

5.5 (a) 1.7×10^{-9} cm $^{-6}$ s $^{-1}$ at both temperatures; (b) 3.61×10^{-19} cm $^{-6}$ s $^{-1}$ at 20 K and 8.63×10^{-20} cm $^{-6}$ s $^{-1}$ at 100 K.

5.6 (b) $k_1(20\text{ K}) = 1.9 \times 10^{-111}$ cm $^{-6}$ s $^{-1}$, $k_1(100\text{ K}) = 7.76 \times 10^{-31}$ cm $^{-6}$ s $^{-1}$, $k_2 = 1.9 \times 10^{-10}$ cm $^{-6}$ s $^{-1}$ at both temperatures, $k_3 = 1.4 \times 10^{-9}$ cm $^{-6}$ s $^{-1}$

at both temperatures and $k_4 = 1.2 \times 10^{-9} \text{ cm}^{-6} \text{ s}^{-1}$ at both temperatures; (c) $1.04 \times 10^6 \text{ cm}^{-3}$ at both temperatures.

Chapter 6

6.1 (a) $\Delta T = 6.12 \times 10^6 \text{ K}$; (b) ΔE (fusion) = $1.44 \times 10^7 \text{ J}$; (c) ΔE (vap) = $3.181 \times 10^9 \text{ J}$.

6.2 (a) $t = 3.08 \text{ Myr}$; (b) 2.91 Myr .

6.3 4.9 zL .

6.4 1–2860 comets.

6.5 Aphelion night-time $2.59 \times 10^{19} \text{ cm}^{-3}$, aphelion daytime $2.04 \times 10^{19} \text{ cm}^{-3}$; perihelion night-time $2.59 \times 10^{12} \text{ cm}^{-3}$, perihelion daytime $1.9 \times 10^{12} \text{ cm}^{-3}$.

Chapter 7

7.1 (a) $4.78 \times 10^{24} \text{ kg}$; (b) 3.326 g cm^{-3} .

7.2 1413 K.

7.3 14.4–26.8 AU.

7.4 $1.25 \times 10^{25} \text{ W}$.

7.5 (a) $10.0 \mu\text{m}$; (b) $3.93 \mu\text{m}$; (c) $17.5 \mu\text{m}$; (d) $40.1 \mu\text{m}$.

7.6 (b) 16.1 amu CH_4 .

7.7 (b) 0.387 bar; (c) $H_s = 2001$; (d) $m = 5.41 \times 10^{-25} = 325 \text{ amu}$.

7.8 (b) $[\text{OH}]_{\text{ss}} = 8.87 \times 10^{10} \text{ mol cm}^{-3}$; (c) $2.33 \times 10^9 \text{ mol cm}^{-3}$.

7.9 (b) $[\text{O}]_{\text{ss}} = 2.10 \times 10^{12} \text{ mol cm}^{-3}$; (c) $7.68 \times 10^{12} \text{ mol cm}^{-3}$; (d) $1.85 \times 10^8 \text{ mol cm}^{-3}$.

Chapter 8

8.1 (a) (i) $8.47 \times 10^{35} \text{ kJ}$, (ii) $4.51 \times 10^{36} \text{ kJ}$, (iii) $5.36 \times 10^{36} \text{ kJ}$; (b) $1.61 \times 10^{21} \text{ kJ}$; (c) $1.56 \times 10^{21} \text{ kJ}$.

8.2 $x\text{CO}_2 = 3.4 \text{ mM}$, $x\text{N}_2 = 0.58 \text{ mM}$.

8.3 (a) $-135.3 \text{ kJ mol}^{-1}$; (b) $-86.5 \text{ kJ mol}^{-1}$; (c) $K = 1.04 \text{ M}^{-1}$.

8.4 (a) $205.5 \text{ kJ mol}^{-1}$; (b) $185.2 \text{ kJ mol}^{-1}$; (c) $K = 0.915$ at 298 K and 0.936 at 400 K.

8.5 (a) $\Delta_r H^\circ = +30.1 \text{ kJ mol}^{-1}$; $\Delta_r G^\circ = 11.8 \text{ kJ mol}^{-1}$; (b) $\Delta_r G^\circ = 53.4 \text{ kJ mol}^{-1}$, $K = 0.985$.

8.7 pH 2.20.

Chapter 9

9.1 (a) $1.23 \times 10^{-8} \text{ m}^2$; (b) 2.7×10^{10} or 45 fmol; (c) 32 pL.

9.2 185 mM.

9.3 -3.8 kJ mol^{-1} and $+3.8 \text{ kJ mol}^{-1}$.

9.4 (a) 0.997, 3×10^{-3} ; (b) 3.72 bar; (c) 4.28 bar.

9.5 (a) -94.7 mV ; (b) -110.6 mV .

9.6 (a) $-164.9 \text{ kJ mol}^{-1}$; (b) -233 kJ mol^{-1} .

Chapter 10

10.1 $1.342 \times 10^{23} \text{ kg}$.

10.2 $1.88 \times 10^3 \text{ kg m}^{-3}$.

10.3 21.6 km.

10.4 $29.59 \mu\text{m}$ and 337.9 cm^{-1} .

10.5 $3.11 \times 10^{12} \text{ mol cm}^{-3}$.

10.6 $1.14 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$ at 170 K and $1.30 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$ at 98 K.

Bibliography

This is a highly abbreviated list of all the work I consulted during the writing of this book, principally because the astrochemistry is still scattered among the research literature. This provides a problem for the student seeking further reading on a particular topic because the next place to look is at the frontier of the subject. I would encourage such exploration and I hope I have provided enough background information in astronomy, chemistry, astrophysics, atmospheric chemistry and astrobiology to make the journey worthwhile.

Books

- Atkins P. W. 2003. *Physical Chemistry* (7th edn). Oxford University Press: Oxford.
- Brack A. 2000. *The Molecular Origins of Life*. Cambridge University Press: Cambridge.
- Coustenis A. and Taylor F. 1999. *Titan: The Earth-Like Moon*. World Scientific: London.
- Faraday Discussion 109. 1998. *Chemistry and Physics of Molecules and Grains in Space*. Royal Society of Chemistry: London.
- Fry I. 1999. *The Emergence of Life on Earth*. Free Association Books: London.
- Holliday K. 2004. *Introduction to Astronomy*. Wiley: Chichester.
- Jakosky B. 1998. *The Search for Life on Other Planets*. Cambridge University Press: Cambridge.
- Kitchen C. R. 1987. *Stars, Nebulae and the Interstellar Medium*. Adam Hilger: Bristol.
- Kuhn H. and Försterlong H.-D. 2000. *Principles of Physical Chemistry*. Wiley: Chichester.
- Morrison D., Wolff S. and Fraknoi A. 1995. *Abel's Exploration of the Universe* (7th edn). Saunders College Publishing: Philadelphia.
- Raff L. M. 2001. *Principles of Physical Chemistry*. Prentice Hall: London.
- Seinfeld J. H. and Pandis S. N. 1998. *Atmospheric Chemistry and Physics*. Wiley: Chichester.
- Snow P. D. 1991. *The Dynamic Universe* (4th edn). West Publishing: Eagan, MN.
- Wayne R. P. 2000. *Chemistry of Atmospheres* (3rd edn). Oxford University Press: Oxford.
- Whittet D. C. 1997. *Planetary and Interstellar Processes Relevant to the Origins of Life*. Kluwer Academic: Dordrecht.
- Zeilik M. 2002. *Astronomy: The Evolving Universe* (9th edn). Cambridge University Press: Cambridge.

Papers

This abbreviated list of papers highlights one of the problems for astrochemists: the results important to the subject are scattered around the literature of many disciplines. The way forward is to consult some of the review papers and follow the literature trails, so I have collected some of the papers with a broader perspective of the subject but this is certainly not complete. The references are given in the format of author, date title, journal title, volume and page number.

- Becker L. *et al.* (1999). The origin of organic matter in the Martian meteorite ALH84001. *Earth and Planetary Science Letters* **167**: 71–79
- Bockelee-Morvan D. and Crovisier J. (2002). Lessons of Comet Hale–Bopp for Coma Chemistry: Observations and Theory. *Earth, Moon and Planets* **89**: 53–71.
- Bogard D. D. and Johnson P. (1983). Martian gases in an Antarctic meteorite, *Science*, **221**: 651–654.
- Bogard D. D. and Garrison D. H. (1998). Relative abundances of argon, krypton, and xenon in the Martian atmosphere as measured in Martian meteorites. *Geochimica et Cosmochimica Acta*, **62(10)**: 1829–1835.
- Brack A., Horneck G. and Wynn-Williams D. (2001). Exo/astrobiology in Europe. *Origins of Life and Evolution of the Biosphere* **31**: 459.
- Charnley S. B., Rodgers S. D., Kuan Y. J. and Huang H.-C. (2002). Biomolecules in the Interstellar Medium. *Advances in Space Research* **30**: 1419.
- Chyba C. F. and McDonald G. D. (1995). The origin of life in the solar system: current issues. *Annual Reviews of Earth and Planetary Science* **23**: 215.
- Chyba C. F. and Philips C. B. (2001). Possible ecosystems and the search for life on Europa. *Proceedings of the National Academy of Sciences* **98**: 801.
- Courtina R. and Kimb S. J. (2002). Mapping of Titan's tropopause and surface temperatures from Voyager IRIS spectra, *Planetary and Space Science* **50**: 309–321.
- Davis W. L. and McKay C. P. (1996). Origins of Life: a comparison of theories and applications to Mars. *Origins of Life and Evolution of the Biosphere* **26**: 61–73.
- Deamer D. W. (1997). The first living systems: a bioenergetic perspective. *Microbiology and Molecular Biology Reviews* **61**: 239.
- Des Marais D. J and Walter M. R. (1999). Astrobiology: exploring the origins, evolution and distribution of life in the Universe. *Annual Reviews of Ecological Systems* **30**: 397.
- Dworkin J. P. (2001). Self-assembling amphiphilic molecules: synthesis in simulated interstellar/precometary ices. *Proceedings of the National Academy of Sciences* **98**: 815.
- Dworkin J. P. *et al.* (2001). Self-assembling amphiphilic molecules: Synthesis in simulated interstellar/precometary ices. *Proceedings National Academy of Sciences* **98(3)**: 815–819
- Ehrenfreund P. and Charnley S. B. (2000). Organic Molecules in the Interstellar Medium, Comets and Meteorites: A Voyage from Dark Clouds to the Early Earth. *Annual Review of Astronomy and Astrophysics* **38**: 427–483.
- Ehrenfreund P. (2002). Astrophysical and astrochemical insights into the origin of life. *Reports on Progress in Physics* **65**: 1427.
- Ehrenfreund P., *et al.* (2003). Physics and chemistry of icy particles in the universe: answers from microgravity. *Planetary and Space Science* **51**: 473–494
- Emeline A. V. (2003). Abiogenesis and photo-stimulated heterogeneous reactions in the interstellar medium and on primitive Earth. Relevance to the genesis of life.. *Journal of Photochemistry and Photobiology C* **3**: 203.

- Engrand C. and Maurette M. (1998). Carbonaceous Micrometeorites from Antarctica (Invited Review). *Meteorites and Planetary Science* **33**: 565–580
- Ferriere K. M. (2001). The interstellar environment of our galaxy. *Reviews of Modern Physics* **73**: 1031.
- Field P. A. (2001). Review: Protein function at thermal extremes: balancing stability and flexibility, *Comparative Biochemistry and Physiology Part A* **129**: 417–431.
- Flynn G. J., Keller L. P., Feser M., Wirick S. and Jacobsen C. (2003). The origin of organic matter in the solar system: evidence from the interplanetary dust particles. *Geochimica et Cosmochimica Acta* **67**: 4791.
- Fraser H. J., McCoustra, M. R. S. and Williams, D. A. (2002). Astrochemistry : The molecular universe. *Astronomy & Geophysics*, **43(2)**: 2.10–2.18.
- Gibb E. L. *et al.* (2000). An Inventory of Interstellar Ices toward the Protostar W33A1, *Astrophysical Journal*, **536**: 347–356.
- Gibson Jr E. K. *et al.* (2001). Life on Mars: evaluation of the evidence within Martian meteorites ALH84001, Nakhla and Shergotty, *Precambrian Research* **106**: 15–34.
- Goldsmith P. *et al.* (2006). Spectral Mapping of the Taurean Molecular Cloud (in preparation)
- Grasset O. (2000). On the internal Structure and Dynamics of Titan. *Planetary and Space Science* **48(1)**: 617–636.
- Guillemot J.-C., Bouyahi M. and Riague E. H. (2004). Prebiotic, planetary and interstellar chemistry starting from compounds detected in the interstellar medium. *Advances in Space Science* **33**: 81.
- Horneck G. (2000). The microbial world and the case for Mars. *Planetary and Space Science* **48**: 1053–1063
- IrvineW. M. (1999). The Composition of Interstellar Molecular Clouds. *Space Science Reviews* **90**: 203–218.
- Joergensen J. K., Schöier F. L. and van Dishoeck E. F. (2002). Physical structure and CO abundance in low-mass protostellar envelopes. *Astronomy & Astrophysics* **389**: 908.
- Khare B. N. *et al.* (2001). Solid Organic Matter in the Atmosphere and on the Surface of Outer Solar System Bodies. *Adv. Space Res.* **27(2)**: 299–307.
- Kobayashi K. *et al.* (2001). Formation of Bioorganic Compounds in Simulated Planetary Atmospheres by High-Energy Particles or Photons. *Adv. Space Res.* **27(2)**: 207–215.
- Kuan Y.-J., Charnley S. B., Huang H.-C., *et al.* (2004). Searches for interstellar molecules of potential prebiotic importance. *Advances in Space Research* **33**: 31.
- Laufer D., Notesco G., and Bar-Nun A. (1999). From the Interstellar Medium to Earth's Oceans via Comets – An Isotopic Study of HDO/H₂O. *Icarus* **140**: 446–450
- Le Teuff Y. H., Miller T. J. and Markwick A. J. (1999). The UMIST database for astrochemistry. *Astronomy & Astrophysics Supplement Series* **146**: 157.
- Lissauer J. J. (1999). How common are habitable planets?. *Nature* **402**: C11.
- Luck R. E. and Lambert D. L. (1992). The chemical composition of Magellanic Cloud cepheids and nonvariable supergiants. *Astrophysical Journal Supplement Series* **79**: 303.
- MacKay D. D. S. and Charnley S. B. (2001). Phosphorus in circumstellar envelopes. *Monthly Notices of the Royal Astronomical Society* **325**: 545–549
- Marcanoa V. *et al.* (2002). Growth of a lower eukaryote in non-aromatic hydrocarbon media C12 and its exobiological significance. *Planetary and Space Science* **50(7–8)**: 693–709.
- McClendon J. H. (1999). The origin of Life. *Earth-Science Reviews* **47**: 71.
- McKay C. P., Coustenis A., Samuelson R. E. *et al.* (2001). Physical properties of the organic aerosols and clouds on Titan. *Planetary and Space Science* **49**: 79.

- Millar T. J., Farquhar P. R. A. and Willacy K. (1997). The UMIST database for astrochemistry 1995, *Astron. Astrophys. Suppl. Ser.* **121**: 139–185.
- Miller T. J. (1995). The chemistry of complex molecules in interstellar clouds. *International Journal of Mass Spectrometry and Ion Processes* **149/150**: 389.
- Miller S. L. (2000). The Endogenous Synthesis of Organic Compounds in *The Molecular Origins of Life*, Editor BrackA. CUP (2000).
- Miller T. J and Freeman A. (1984). Chemical modeling of molecular sources I – TMC-1. *Monthly Notices of the Royal Astronomical Society* **207**: 405.
- Navarro-González R. et al. (2001). Production of Hydrocarbons and Nitriles by Electrical Processes in Titan's Atmosphere, *Adv. Space Research* **27**(2): 271–282.
- Nealson K. H and Conrad P. G. (1999). Life: past, present and future. *Philosophical Transactions of the Royal Society of London B* **354**: 1923.
- Orgel L. E. (1998). The origins of life – a review of the facts and speculations. *Trends in Biochemical Sciences* **23**: 491.
- Owen T. C. (2000). On the origin of Titan's atmosphere. *Planetary and Space Science* **48**: 747–752.
- Pratap P. et al. (1997). A Study of the Physics and Chemistry of TMC-1. *Astrophysical Journal* **486**: 862–885.
- Rae J. G. L., Bell N., Haquist T. W., Pilling M. J. and Ruffle D. P. (2002). Reduced networks governing the fractional ionization in interstellar molecular clouds. *Astronomy & Astrophysics* **383**: 738.
- Reysenbach A.-L. and Cady S. L. (2001). Microbiology of ancient and modern hydrothermal systems. *Trends in Microbiology* **9**: 79.
- Ruffle D. P. and E. Herbst (2001). New models of interstellar grain chemistry – II. Surface chemistry in quiescent cores. *Monthly Notices of the Royal Astronomical Society* **322**: 770–778.
- Sandford S. A. (1996). The inventory of interstellar materials available for the formation of the solar system. *Meteorites and Planetary Science* **31**: 449.
- Sandford S. A. (2002). Interstellar processes leading to molecular deuterium enrichment and their detection. *Planetary and Space Science* **50**: 1145–1154.
- Setter K. O. (1999). Extremophiles and their adaptation to hot environments. *FEBS Letters* **452**: 22–25
- Smith D. and Spanel P. (1995). Ions in the terrestrial atmosphere and in interstellar clouds. *Mass Spectrometry Reviews* **14**: 255.
- Smith I. W. M. et al. (2004). Rapid neutral–neutral reactions at low temperatures: a new network and first results for TMC-1 *Monthly Notices of the Royal Astronomical Society* **350**: 323.
- Snyder L. E. (1997). The Search for Interstellar Glycine. *Origins of Life and Evolution of the Biosphere* **27**: 115–133.
- Teyssier D., Fossé D., Gerin M., et al. (2004). Carbon budget and carbon chemistry in photon dominated regions. *Astronomy & Astrophysics* **417**: 135.
- Tielens A. G. G. M. and Charnley S. B. (1997). Circumstellar and Interstellar Synthesis of Organic Molecules. *Origins of Life and Evolution of the Biosphere* **27**: 23–51.
- Vacher R., Le Duc E., and Fitaire M. (2000). Clustering reactions of HCNH^+ , $\text{HCNH}^+ (\text{N}_2)$ and $\text{HCNH}^+ (\text{CH}_4)$ with ethane: application to Titan atmosphere. *Planetary and Space Science* **48**: 237–247
- Williams D. A. and Herbst E. (2002). It's a dusty Universe: surface science in space. *Surface Science* **500**: 823.

- Winnewisser G. and Kramer C. (1999). Spectroscopy between the stars. *Space Science Reviews* **90**: 181.
- Yi-Jehng Kuan *et al.* (2004). Searches for interstellar molecules of potential prebiotic importance. *Advances in Space Research* **33**: 31–39
- Yung Y., Allen M. and Pinto J. (1984). Photochemistry of the atmosphere of Titan: comparison between model and observations. *Astronomy & Astrophysics Supplement Series* **55**: 465.
- Zubko V. G. (1999). On the model of dust in the small Magellanic Cloud. *Astrophysical Journal* **513**: L29.

Index

- A_v , 122
 τ -Bootes, 208
 H_2^+ , 124
18S rRNA, 273
21-cm line, 79
22.235 GHz water maser, 78
51-Pegasi, 207
67P/Churyumov-Gerasimenko, 187
- Acasta gneisses, 165
Accretion, 159
Acetylene polymerisation II, 138
Acid–base equilibria, 235
Activation energy, 128
Adiabatic lapse rate, 213
Albedo, 202, 203, 291
ALH84001, 165, 168, 170, 173, 178
Alkaliphile, 275
Amino acids, 238
Formation, 143
asymmetric photolysis, 247
Andromeda Galaxy (M31), 3, 31, 106
Aphelion, 183
Archaea, 273
Arrhenius equation, 125, 237
Astrobiology, 260, 274, 283
Titan, 302, 305
Atmospheres
Archean, 202
absorption, 71
chemistry, 305
inner planets, 210
- primordial, 238
Surface pressure, 210
structure, 221, 305
windows, 52
- Atomic structure, 82
H atom Absorption spectrum, 57
Heavy element atomic spectra in stars, 99
Lyman, Balmer, Paschen, Brackett and Pfund series, 58
Selection rules, 59
Spectroscopy, 82
Autumnal Equinox, 27
- Bacillus subtilis, 179
Bacteria, 273, 264
Balmer series, 58, 98
Balmer temperature, 110
Beer–Lambert Law, 82
Big Bang Theory, 1, 34
Big Bang Nucleosynthesis, 2
Binary stars, 103
Binding energy of atomic nuclei, 91
Biogenic fossils, 190
Biogenic structures, 177
Biomarker, 8
Biosynthesis, 154
Black Body Radiation, 15, 291
Black body radiation laws, 38
Black hole, 110, 106
Blue shift, 48
Bond energies, 134

- CAIs, 165
 CNO cycle, 94
 Carbonaceous chondrites, 173
 Cassini–Huygens, 289, 302
 Celestial coordinates, 26
 Cepheid variable stars, 104
 Chapman layer, 216, 218, 221,
 300
 odd oxygen, 216
 Oxygen and ozone, 278
 ozone, 213
 Ozone layer, 218
 UVA, UVB, UVC, 216
 Chemical network, 145, 154, 146,
 297
 Chemical potential, 228, 283
 Chlorophyll, 19
 CHON, 151
 Chondrules, 162
 Circumstellar medium, 121
 CNO cycle, 110
 CO, 61–64
 CO₂ photolysis, 216
 Coercevate, 172, 262
 Collision rate, 125
 colour excess, 122
 Comet chemistry, 157, 180, 190
 Collisions, 185
 Coma, 181, 183
 dust tail, 180
 Hale-Bopp, 181, 187
 Halley, 181, 291, 300
 Halley, Hyakutake, 187
 ion tail, 180
 Kuiper belt, 180
 Molecular inventory for Hale-Bopp,
 184
 Oort Cloud, 180, 195
 surface synthesis, 185
 Common Ancestor, 273
 Cosmic Background Explorer, 20
 Cosmic rays (cr), 136
 Cosmology, 36, 38
 Cretaceous–Tertiary, 245
 Critical Micelle Concentration (CMC),
 262
 Cryovulcanism, 292, 305
 Detection of hydrogen, 79
 21-cm line, 79
 Atom hydrogen atom, 57
 Deuterium enrichment, 149, 187
 Diffuse interstellar bands, 80
 Diffuse interstellar medium, 120
 Dissolved gases – Henry’s Law, 233
 DNA, 6, 244
 Doppler Effect, 20, 103
 Broadening, 46
 Shift, 82, 206
 Profile, 48
 Drake equation, 9, 35
 Dust, 154
 Grains, 140
 Structure, 140
 Surface reactions, 142
 Earth, 158
 Earth–Moon, 199, 221
 Formation, 197
 Effective temperature, 203
 Einstein coefficients, 45
 Elastic scatter, 44
 Electronic spectroscopy, 82
 Encapsulation, 263
 Endogenous organic synthesis, 237, 256
 Endolith, 275
 Energy and mass, 90
 Equilibrium, 228
 Escape velocity, 158
 Eukarya, 273
 Europa, 287
 Exogenous delivery of organics, 256
 Extrasolar planets, 206, 221
 Extraterrestrial astronomy, 55
 Extremophile bacteria, 275
 First cellular life, 260
 First point of Aries, 26
 Force constant, 73, 149
 Formose, 242, 292
 Fusarium alkanophyllum, 304
 Galaxies, 3, 31, 38
 Genetic code, 6
 Geological time, 165, 190, 200

- Giant molecular clouds (GMCs), 114, 121, 151
 OMC, 247
 TMC, 118, 146
Glycine, 67
Gravitational lensing, 36
Great Bear, 29
Greenhouse Effect, 212
- Habitable zone, 221
Hadear, 197, 199
Hale-Bopp, 181, 187
Halley, 181, 291, 300
Halley, Hyakutake, 187
Hayuatake, 291
Haze layers, 300
Helium flash, 93
Herbig Ae, 90
Herzprung–Russell diagram, 88, 110
HII regions, 116
Homochirality, 8, 256
Hubble constant, H , 33
Hubble Space Telescope, 56, 76, 195
Huygens probe, 287, 293, 300, 303
Hyperthermophile, 8, 276
- Impact frustration, 13, 197
Index Catalogues (IC), 31
Infrared astronomy, 71
Inorganic phosphate, 243
Intensities of rotational transitions, 70
Interaction of radiation with matter, 82
Interplanetary organic material, 178
Interstellar medium ISM, 117, 185
 molecules detected, 117
 Conditions in the ISM, 154
 Glycine, 67, 151
 Organic synthesis in the interstellar ice, 143
 Prebiotic molecules, 151
 Reactions, 130–131
Ionisation energies, 135
- Kerogen, 164, 171
Krafts temperature, 262
Kuiper belt, 180
- Landé factor, 101
lapse rate, 212
Large Magellanic Cloud, 3
Light harvesting, 38
Line shape, 46
Liposphere, 293, 305
Lipothermal vent, 303
Little warm pool, 199
Local Group, 20, 34, 35
 Large and Small Magellanic Clouds (LMC and SMC), 3, 35, 106
Local thermal equilibrium (LTE), 126
Luminosity, 16
Luminous arc, 36
- Magnitude, 22
 Absolute, 26
 apparent, 122
Main sequence, 89
 lifetime, 93
MASER, 77, 78, 82
Mass extinction, 11
Membrane transport, 283
Meteorite, 157, 158, 190
 ALH84001 *see above*
 Allende, 165
 Analysis, 190
 Carbonaceous chondrites, 173
 Kerogen, 171
 Meteor, 157
 Meteoroid, 158
 Leonid meteor, 158
 Murchison, 171
 Stony, 161, 163
 Stony-iron, 161, 163
 Stony, 161
 Methanol Maser, 78
 Micelles, 261
 Microbial Mars, 281
 Microprobe laser desorption laser ionisation mass spectrometry (μ L2MS), 169
 Microwave and millimetre wave astronomy, 60
 Molecular identification, 68
 Microwave spectroscopy, 60, 82
 microwave spectrum, 68
 Mid-ocean tidal rise, 199

Milky Way, 3, 20, 31, 33, 81, 106
Minimal genome, 283
Mineral surface, 265
Molecular processing, 154
Moments of inertia, 65
Mycoplasma genitalium, 272

Natural linewidth, 46, 47
New General Catalogue (NGC), 31
Nitrile polymerisation, 300
non-elastic scatter, 44
Non-radiative decay, 78
Novae, 95
Nuclear synthesis in heavy stars, 94
Nucleosynthesis processes, 96

Oort Cloud, 180, 195
Orbital velocity, 158
Organic material, 176
Origin of Life, Theories, 10
Origin of the elements, 97
Orion nebula, 57, 68, 89
Osmosis, 267, 268

Parallax, 24
Pathfinder, 281
Perihelion, 180, 183
pH of the oceans, 233, 234
Phospholipid, 261
Photochemistry, 133
 Photodissociation, 134
 Photoionisation, 135
Photon-dominated regions, 121
Planck length, 37
Planck tension, 37
Planck time, 37
Planck's Law, 38
Planet formation, 221
Planetary atmospheres, 209
Planetary chemistry, 193
Planetsimals, 159
PNA, 244
Polycyclic aromatic hydrocarbons, 136–7
Population inversion, 77
Pressure broadening, 46, 47
Protocells, 264, 283
Proton–proton cycle, 110
protostar, 85

Purine and pyrimidine synthesis, 240
pyrite surface, 278
Pyrolobus fumarii, 276

Quantum mechanics, 41

r-process, 96
Radiation trapping, 211
Radiative decay
 Isotopes and daughter products dating, 166
 lifetime, 77
 decay pathways, 168
 radioactive clocks, 167
Rates of chemical reactions, 123, 236, 154
Bimolecular reaction, 125
rate constant, 123
temperature dependence, 125
Steady-state approximation, 127, 154
 total overall order of the reaction, 123
Rayleigh criterion, 54
Red shift, 48
Reddening, 121
Reduced mass, 61, 149
Right ascension, 26
RNA World hypothesis, 244, 253–256
Ro-vibronic spectrum, 60
Rosetta mission, 181, 187, 210
Rotational constant, 61

s-process, 96
Saturated adiabatic lapse rate, 213
Scale height, 211
Short-wavelength shield, 221
Snow line, 186, 190
Solar System
 Solar nebula, 190
Spectral mapping, 81, 82
Spectroscopy, 60, 82
 Selection rules, 45, 82
 Spatial Resolution, 54
 Infrared (IR), 82
 IR stretching frequencies, 73
 spin multiplicity, 299
 Transition moment, 46
Spontaneous chemical reactions, 256
Spontaneous emission, 44

- Stars
 Be stars, 90
 Becklin–Neugebauer (B-N) object, 89
 Binary stars, 103
 Birth lines, 89
 Classification, 86
 Constellations, 26
 Cycle of star formation, 108, 110
 Evolution of heavier stars, 95
 Exotic stars, 102
 Helium flash, 93
 Herbig Ae, 90
 Herzprung–Russell diagram, 88, 110
 neutron star, 95
 Signs of the Zodiac, 28
 Supergiants, 89
 T-Tauri, stars, 90, 116, 158, 186
 white dwarf, 89, 95, 102
Steady-state approximation, 127, 154
Stefan–Boltzmann Law, 16
Steric factor, P , 126
Stimulated absorption, 44
Stimulated emission, 44, 78
Stoichiometric coefficient, 228
Stratosphere, 213
Strecker syntheses, 240, 292
Super string theory, 37
Supernova, 96, 97, 110

T-Tauri, stars, 90, 116, 158, 186
Taurus molecular cloud (TMC), 114
Telescopes, 52, 82
 James Clerk Maxwell Radio Telescope, 56

Jodrell Bank Radio, 56
UK Infrared Telescope, 71, 76
Temperature profile, 212
Terrestrial planets, 193
The Standard Model, 1
Thea, 199
Titan, 287–305
 atmosphere, 293, 297
 internal structure, 302
 Physical data, 289
 Tholins, 297, 305
TMC, 118, 146
Transition intensities, 51
Transition moment, 46
Triple-alpha process, 93, 110
Tropic of Cancer, 27
Tropic of Capricorn, 27
Troposphere, 213

UMIST database, 126
Universal tree of life, 273, 283
Urey-Miller, 238
Ursa Major, 28, 29

Vernal Equinox, 26, 27
Visible astronomy, 76
Visible extinction, 121, 154
Volatiles inventory, 209

Water–ammonia ice, 292
Watson–Crick base pairs, 6, 261

Zeeman effect, 101–102, 110, 123
zero point energy, 73, 149
Zodiac, 28

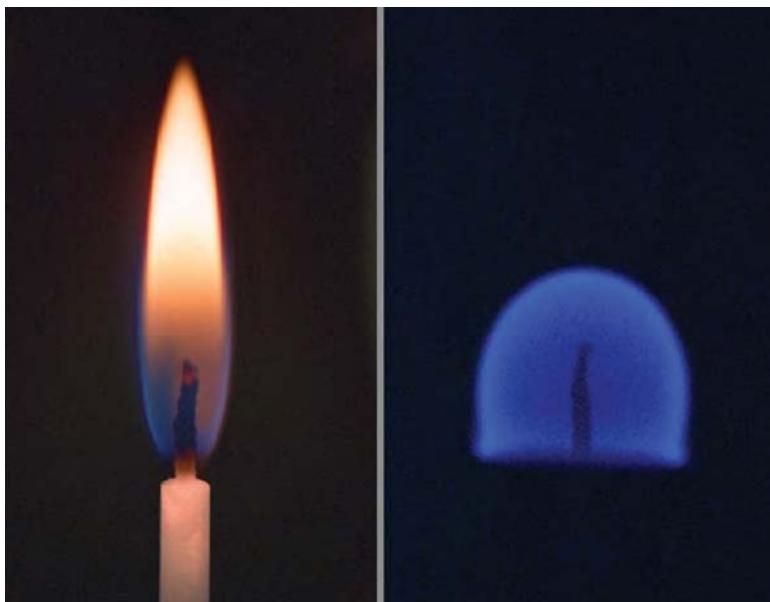


Figure 1.1 Two species of candle flame – dead or alive? The flame on the left is on Earth and the flame on the right is burning under zero gravity. (Reproduced from photos by courtesy of NASA)



Figure 1.4 Hyperthermophile bacteria at Prismatic Lake in Yellowstone National Park. (Reproduced from a photo of Prismatic Lake by courtesy of National Park Service, Yellow Stone National Park)



Orion Constellation This is a mosaic picture of 18 (6x3) frames. Additional four frames are added around the Barnard Loop region to increase S/N. Vignetting of the telescope is corrected by flat fielding, using proprietary software written in IDL. Distortion correction and initial colour match is done in Registar, with final colour correction and mosaic processed using PhotoShop. (Reproduced by permission of Wei-Hao Wang, Institute for Astronomy at University of Hawaii)



Figure 2.12 Andromeda Galaxy. (Reproduced from photos by courtesy of NASA)



Figure 3.13 UK Infrared Telescope situated close to the summit of Mauna Kea, Hawaii, at an altitude of 4092 m. (Reproduced with permission by UKIRT)

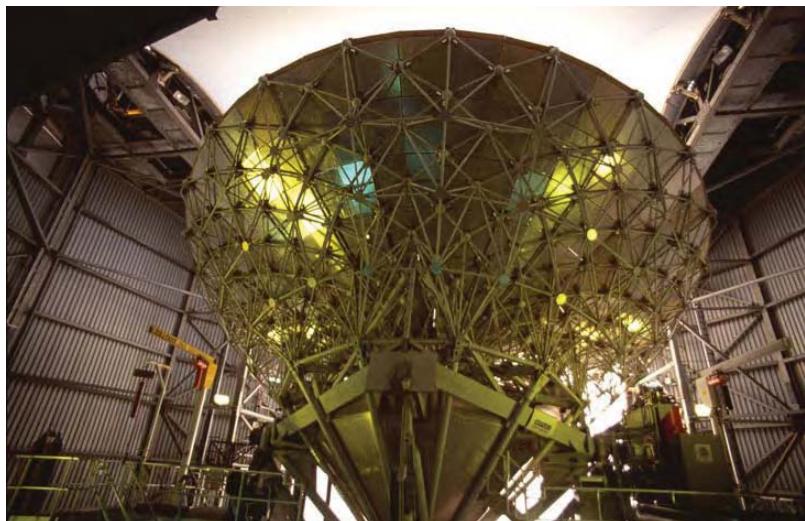


Table 3.2 The collection dish of the JCMT telescope on Hawaii. The 15 m diameter JCMT is the largest astronomical telescope in the world designed specifically for sub-millimeter astronomy. (Reproduced by permission of the James Clerk Maxwell Radio Telescope)

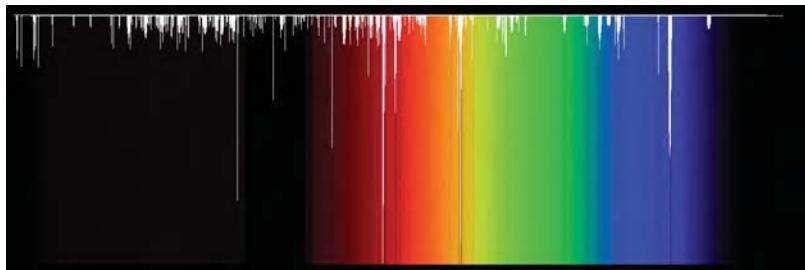


Figure 3.19 Diffuse interstellar bands. (Reproduced by permission of Peter Jeniskins, SETI Institute)

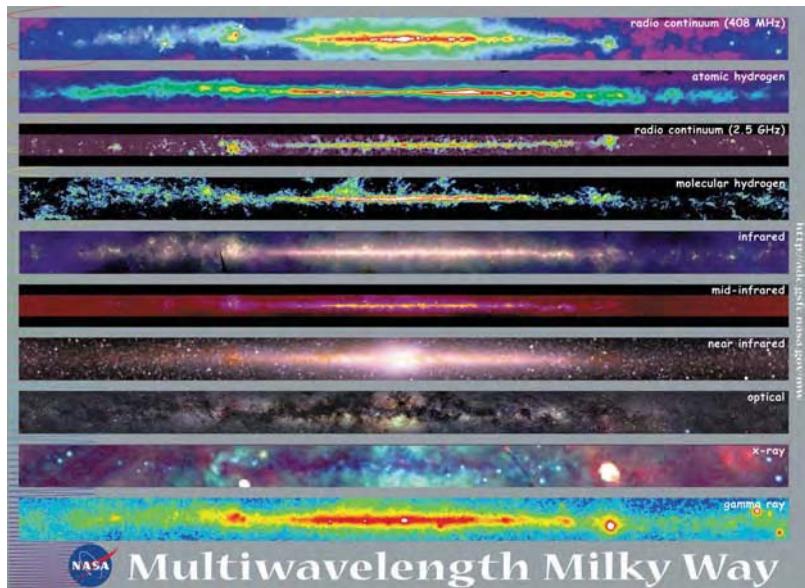


Figure 3.20 Multiwavelength view of the Milky Way. (Reproduced by permission of Dave Leisawitz, Observational Cosmology Laboratory, NASA Goddard Space Flight Center)

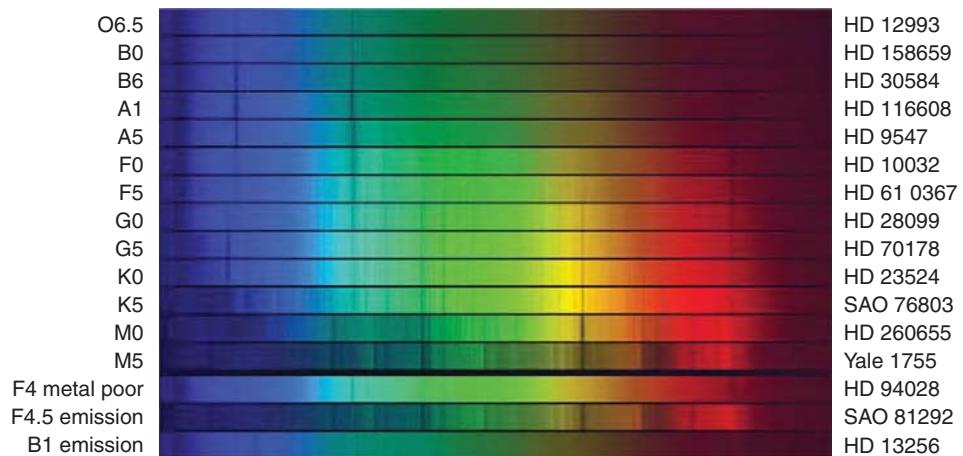


Figure 4.2 Spectra from different stellar classes OBAFGKM. (Reproduced by permission of KPNO 0.9-m Telescope, Aura, NOAO, NSF)

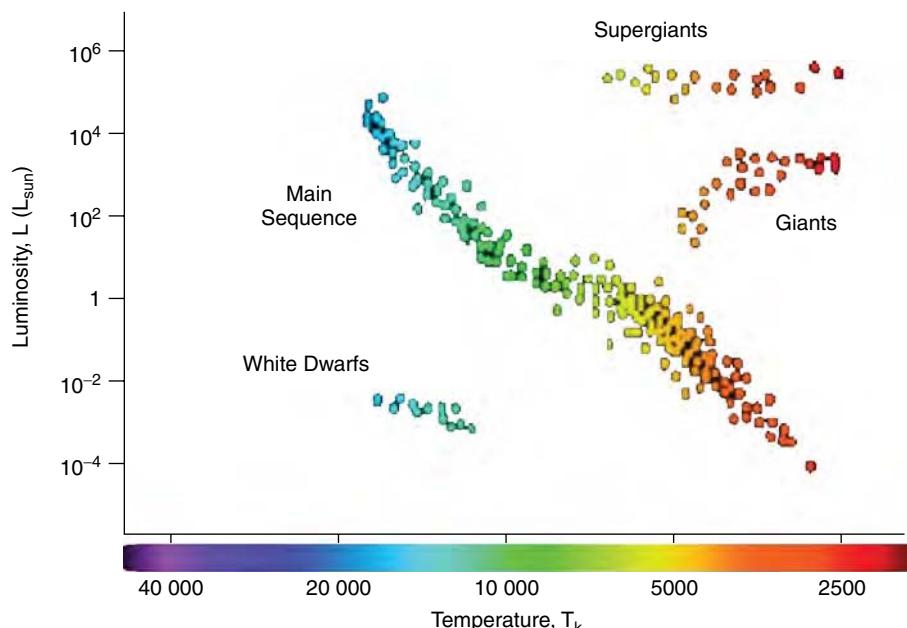


Figure 4.3 Herzprung–Russell diagram. (Reproduced by permission of D.Padgett (IPAC/Caltech), W.Brandner (IPAC), K.Stapelfeldt (JPL) and NASA)

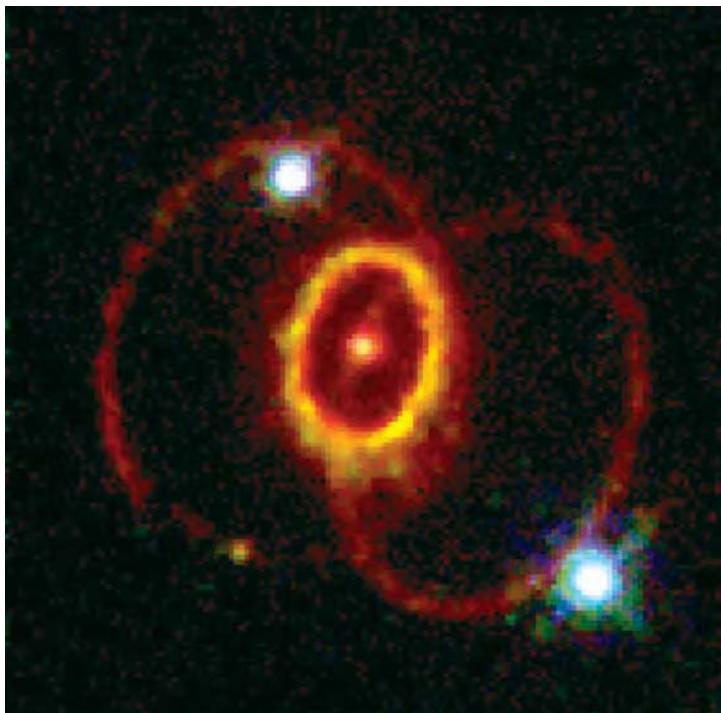


Figure 4.9 Supernova 1987A captured by the HST wide Field Planetary Camera 2. (Reproduced with permission from P.Challis, R.Kirshner (Harvard Smithsonian Centre for Astrophysics) and B. Sugarman (STScI) and NASA.)

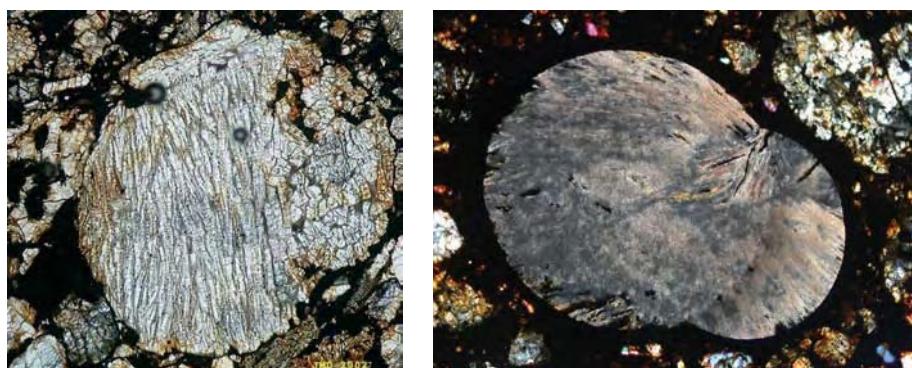


Figure 6.3 Chondrules. (Reproduced by permission of J. M. Derochette)



Mars Iron Meteorite (see page 161) NASA's Mars Exploration Rover Opportunity has found an iron meteorite on Mars, the first meteorite of any type ever identified on another planet. The pitted, basketball-size object is mostly made of iron and nickel. Readings from spectrometers on the rover determined that composition. Opportunity used its panoramic camera to take the images used in this approximately true-color composite on the rover's 339th Martian day, or sol (Jan. 6, 2005). This composite combines images taken through the panoramic camera's 600-nanometer (red), 530-nanometer (green), and 480-nanometer (blue) filters. (Reproduced with permission of NASA/JPL/Cornell)



Figure 6.17 Tail structures of Hale-Bopp (left) and Halley (right) (1997/3/10, 4:38–5:00, Mt. Ho-Huan, Taiwan). (Reproduced by permission of Wei-Hao Wang, Institute for Astronomy, University of Hawaii)



Figure 9.14 Pathfinder image of the Martian surface. (Reproduced by permission of NASA Pathfinder Mission, Jet Propulsion Laboratory, University of Arizona)



Figure 10.9 Hazy layers in Titan's atmosphere. (Reproduced by permission of Voyager Project, JPL NASA)