Oxygen and the Evolution of Life

Heinz Decker · Kensal E. van Holde

Oxygen and the Evolution of Life

Professor Dr. Heinz Decker
Institut für Molekulare Biophysik
Johannes Gutenberg-Universität Mainz
Jakob Welder Weg. 26
55128 Mainz, Germany
hdecker@uni-mainz.de

Kensal E. van Holde
Distinguished Professor Emeritus
Dept of Biochemistry and Biophysics
Oregon State University
Corvallis OR 97331
USA
vanholde@asbmb.org

Cover illustration: Different oxygen transport (respiratory) proteins developed after the oxygen
concentration increased some billion years ago: earthworm hemoglobin (*red*), arthropod hemocyanin
(scorpion), mollusc hemocyanin (cephalopod) (*front cover*, *clockwise*) and the myriapod hemocyanin
(*back cover*); see also Fig. 5.8. The molecules artwork are courtesy of Jürgen Markl, Institute for
Zoology, Johannes Gutenberg University Mainz.

Cover design: WMXDesign GmbH, Heidelberg, Germany

Printed on acid-free paper

Preface

This book has a curious history. It evolved, like its subject, from a much simpler beginning. Both the authors have had long-standing common interests in the proteins and processes of oxygen transport in animals. During a sabbatical year that KvH spent in the laboratory of HD, our discussions broadened to encompass the much deeper question as to how oxygen transport, and indeed oxygen utilization, were related to the evolution of life. As we considered the geological and paleontological evidence, it became clear that changes in the earth's atmosphere and biological evolution have been, and continue to be, interrelated in complex and fascinating ways. Furthermore, these relationships have important implications for human health and humanity's future.

Thus, the book grew outward from its original focus on oxygen transport, sometimes into areas in which we must confess less confidence than we would like. But, we must ask the reader's indulgence, for we feel that the fascination of the whole story such that it is vital to try to tell it.

Some readers may find Chapter 1 daunting, with too much dry chemistry. Skip it if you wish! Although we feel that it provides a useful background for the rest of the book, most of the following Chapters can be read intelligently without this material.

Mainz, Germany   Heinz Decker
Corvallis, OR, USA   Kensal E. van Holde

Contents

Abbreviations

AD	Anno Domini (years after the start of this epoch)
AIF	Apoptosis activating factor
ATP	Adenosine triphosphate
BYA	Billion years ago
cGMP	Cyclic guanisylmonophosphate
DOPA	Dihydroxyphenylalanine
EDRF	Endothelium-derived relaxing factor
FU	Functional unit
GSH	Glutathione
INF	Interferon
IPCC	Intergovernmental Panel on Climate Change
IR	Infra red
ITP	Inositol phosphate
MYA	Million years ago
NADH	Nicotinamide adenine dinucleotide (reduced)
OBP	Oxygen binding proteins
OTP	Oxygen transport proteins
PAL	Present dioxygen level
PDE	Phosphodiesterase
ROS	Reactive oxygen species
SOD	Superoxide dismutase
TNF	Tumor necrosis factor

Chapter 1
Oxygen, Its Nature and Chemistry: What Is so Special About This Element?

Contents

1.1   A Brief Introduction to Oxygen

It would seem that an introduction to oxygen is unnecessary, for we deal with it and depend upon it every moment of our lives. Oxygen is to us the essential stuff of the air we breathe. We are *aerobic* animals who obtain energy by oxidizing foodstuffs. As such, we are wholly dependent on oxygen for life – go without it for a couple of minutes and we panic and may even suffer irreversible brain damage. In a few more minutes, we perish. Animal metabolism depends upon oxygen for almost all of its energy-generating processes. Yet this was not always so. Early in the history of the Earth, there was essentially no *free* oxygen anywhere, although oxygen has always been one of the most abundant elements on Earth. In the early Earth, virtually all oxygen was bound in compounds, mainly water and silicate rocks. Primitive microbes managed life without free oxygen. Examples of this less efficient *anaerobic* metabolism still persist, such as bacteria that live in oxygen-poor environments. Remarkably, just as most life today depends on oxygen, so also the Earth's supply of free oxygen depends, in turn, on life. Virtually all of the free oxygen in our

environment comes from plant photosynthesis, and it was the evolutionary invention of this process, nearly 3 BYA, that turned the initially anaerobic world into our present aerobic one. Our Earth is the only planet in the solar system exhibiting significant amounts of free oxygen, which may signify that Earth is the only one on which life (or at least advanced life) has evolved.

The introduction of oxygen into an anaerobic world brought problems for the then-existing organisms, for many of the by-products of oxygen metabolism are toxic substances. Chemical defenses had to be erected against these; we still find them in our own chemistry today. On the other hand, certain organisms evolved aerobic metabolic pathways, much more efficient than the anaerobic ones. These were the ancestors of all animals and higher plants.

There are still deeper reasons why oxygen is essential to life. Water is the ideal milieu for biological processes and structures, and oxygen is essential to water. Furthermore, the element hydrogen is required in almost all organic compounds and structures, but free hydrogen is easily lost into space from a small planet like ours. It is only by virtue of the binding of hydrogen by oxygen to form water that there remains any significant amount of this vital element on Earth. Without binding to oxygen almost all hydrogen would have been lost ages ago.

This book will explore the history of oxygen, from its genesis in stars to its role in reshaping the Earth and its creatures. We will find its history is entwined with evolutionary and geological history in remarkable and often unexpected ways. But to understand this, it is best to consider first some fundamental properties of this intriguing element, properties have allowed it to play its unique note and that which stem directly from its atomic structure. Thus an introduction to oxygen is necessary.

1.2   Atomic Structure of Oxygen: Chemical Bonding Potential

What an element can do, what compounds it will form, and what properties it has depends on its atomic structure. We begin our analysis of oxygen with the atom's core, the nucleus. The number of protons in a nucleus gives its atomic number and its positive charge. Add the number of neutrons and you have the atomic mass.

The nucleus of the most common isotope of oxygen contains eight protons and eight neutrons, and thus has an atomic number of 8, and 16 atomic mass units. It is designated in conventional shorthand as ^{16}O. There exist other isotopes (mainly ^{17}O and ^{18}O) differing in numbers of neutrons, but they are found in nature in very small amounts. With eight positively charged protons, one needs eight negative electrons to make a neutral atom. Quantum-mechanical theory tells us how these electrons must be distributed in the space around the nucleus. This is not in the circular "orbits" depicted in the earlier atomic theories (and often still in popular illustrations). Rather, according to quantum mechanics, we can only describe the electron distributions in terms of "orbitals," regions in space where the electrons are most likely to be found. There are strict quantum-mechanical rules regulating how

orbitals can be filled up as we add electrons to a nucleus. The orbitals available for the lowest energy state are described as follows. There is a lowest energy orbital, closest to the nucleus, called 1s which is a spherically symmetrical could about the nucleus. Further from the nucleus is a symmetric 2s orbital, and then four so-called 2p orbitals. These latter are asymmetric and directional as pictured in Fig. 1.1a. A fundamental rule is that each orbital can accept no more than two electrons, and these pairs must be of opposite spin. Originally spin was interpreted as it sounds like, a "spinning" of the electron but a quantum mechanical interpretation would simply emphasize different responses to a magnetic field. Each electron has only two possibilities for its spin, designated $+$ or $-$. We use here only a few general concepts from quantum mechanics. A clear, but more sophisticated discussion is found in Tinoco et al. (2002).

Now we have enough information to describe the possible electronic structures of the oxygen atom. With eight electrons to distribute, we first put two in the 1s orbital, two in the 2s, and have four left for the 2p's. In general the lowest overall energy is obtained by pairing electrons of opposite spin, so in forming the ground state (the lowest energy state) we fill only two of the 2p orbitals, leaving two empty. Note that in forming bonds with other atoms through the $2p_x$, $2p_y$ and $2p_z$ orbitals band angles will be close to 90° and thus awkwardly forced. To relieve that strain, oxygen and some other atoms such as carbon, at least when forming compounds, actually rearrange orbitals somewhat. The 2s and 2p orbitals get mixed or "hybridized" to make four sp^3 "hybrid" orbitals that are aimed toward the corners of a tetrahedron as shown in Fig. 1.1b. There are six electrons to put into this set (two 2s
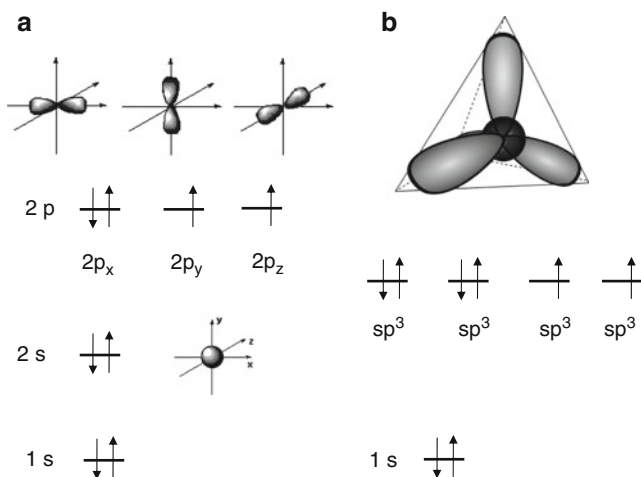
Fig. 1.1 Orbitals for oxygen. (**a**) Lowest energy atomic orbitals for oxygen; here are depicted (not to scale) the 2s and 2p orbitals; those that are available to oxygen. The 1s orbital is spherical and concentrated closer to the nucleus than the 2s. The ground-state occupancy by electrons is indicated by the *arrows* denoting spin. (**b**) Hybrid Orbitals. sp^3 hybridization. Four orbitals are produced by a "mixing" of one 2s and three 2p orbitals pointing to the four edges of a tetrahedron

electrons and four 2p electrons). Two orbitals will have spin-paired electrons, and two will each have one unpaired electron. These sp^3 orbitals point in the direction toward the four corners of a tetrahedron (Fig. 1.1b) with bond angles of about 109°.

With these simple rules, we are in a position to explain the most important aspects of oxygen chemistry. First, note that when an atom has only partially filled orbitals, it is almost always energetically favorable to fill them. With the oxygen atom, this can be done in two different ways. First, oxygen may simply gain two electrons from some other atom (a metal M, for example) to form an ionic compound in which oxygen exists as the oxide ion, O^{2-}. For example:

$$M + O \rightarrow M^{2+} + O^{2-}$$
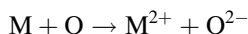
Alternatively, oxygen may share two electrons with another atom or atoms, in covalent bonds. This is what happens when oxygen combines with hydrogen to form water, as shown in Fig. 1.2a. The angle between the two oxygen–hydrogen bonds is 104.5°, being slightly different from the value expected for a tetrahedron (109.5°) as a consequence of electron–electron repulsion between the two pairs of electrons.

The covalent bonds of oxygen are quite stable, and much of Earth's chemistry is explained by this fact. For example, the abundance and stability of the silicates such as quartz, that make up much of the Earth's crust depends on the strength of the covalent Si–O bond and the vast amount of water depends on the O–H bonds.

Oxygen can form covalent bonds with a number of elements, but exceptionally important for life are those with carbon. The enormous variety of these "organic"
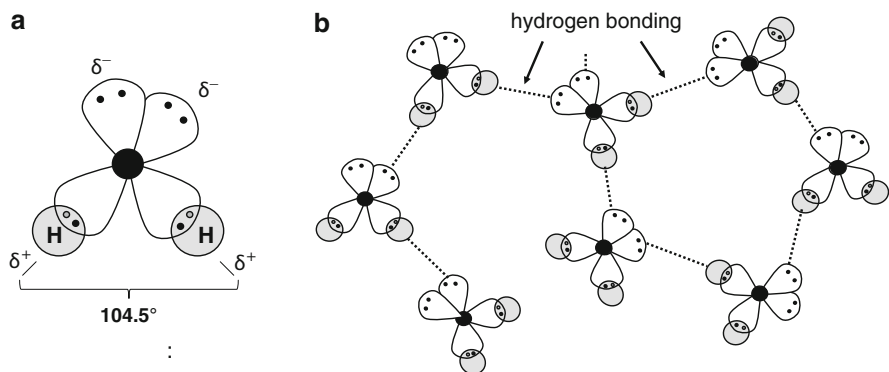
Fig. 1.2 (**a**) Water structure. The electron structure of an individual water molecule: The nonbonded electron pairs of the two orbitals can act as hydrogen acceptors. The oxygen atom (O) in the center is shown in *black*, hydrogen (H) in *gray*. The *symbols* δ^- and δ^+ indicate partial charges on the two sides of the molecule. The angle between the two hydrogen binding orbitals is 104.5° instead of 109° in a tetraedric state sp^3 hybridisation. (**b**) Hydrogen bonding in water between water molecules. Each molecule acts as both a hydrogen donor and a hydrogen acceptor, allowing clusters of water molecules to form (Mathews et al. 2000)

compounds is enriched by the numerous possibilities for O–C bonding, as in the atomic groups:

hydroxyl: $-C - OH$

carbonyl: $-C = O$

ether: $-C - O - C -$,

etc.

Equally important for biological functions is the strong tendency of the oxygen atom to form noncovalent *hydrogen bonds*. This is a consequence of the existence of the two filled sp^3 orbitals on the oxygen atom. Even when it is making covalent bonds with other atoms via the half-filled orbitals, the "lone pair" electrons in the remaining two sp^3 orbitals will still strongly attract protons on other molecules (see Fig. 1.2b). These hydrogen bonds play a major role in forming the structures of proteins, nucleic acids, and water (see below).

All of these properties of oxygen are an inevitable consequence of the physical laws of our universe and the subatomic structure of the oxygen atom. As we shall see in Chap. 2, the existence of oxygen atoms is in turn a necessary result of the evolution of the universe.

1.3   The Dioxygen Molecule

Virtually all of the oxygen in the air we breathe is present as the diatomic molecule O_2 which is correctly called *dioxygen*. This is an extremely stable molecule, in which the atoms are held together by very strong covalent bonding. In elementary chemistry, covalent bonding is described in terms of electron sharing between atoms. This is basically correct, but we need a more detailed and sophisticated picture, to understand the peculiar properties of O_2.

To describe the electron distribution in a covalent bond in quantum-mechanical terms, we need to invoke the concept of *molecular orbitals*. These orbitals are not only constructed from the atomic orbitals of the atoms involved, but they also take into account electron sharing between partners – the essence of a covalent bond. There are two classes of such orbitals – those that arise from overlap and merging of atomic orbitals (bonding orbitals), and those in which the atomic orbitals repel one another (antibonding orbitals) (see Fig. 1.3). Finally, the geometry of molecular orbitals falls into two major classes (for small atoms). Those that lie along the axis between the two nuclei are called sigma (σ)-orbitals, and those that lie parallel to, but off this axis are pi (π)-orbitals. Thus, the water molecule pictured in Fig. 1.2a is held together by two sigma bonding orbitals formed from hydrogen 1s orbitals and $2sp^2$ hybrid orbitals of the oxygen.
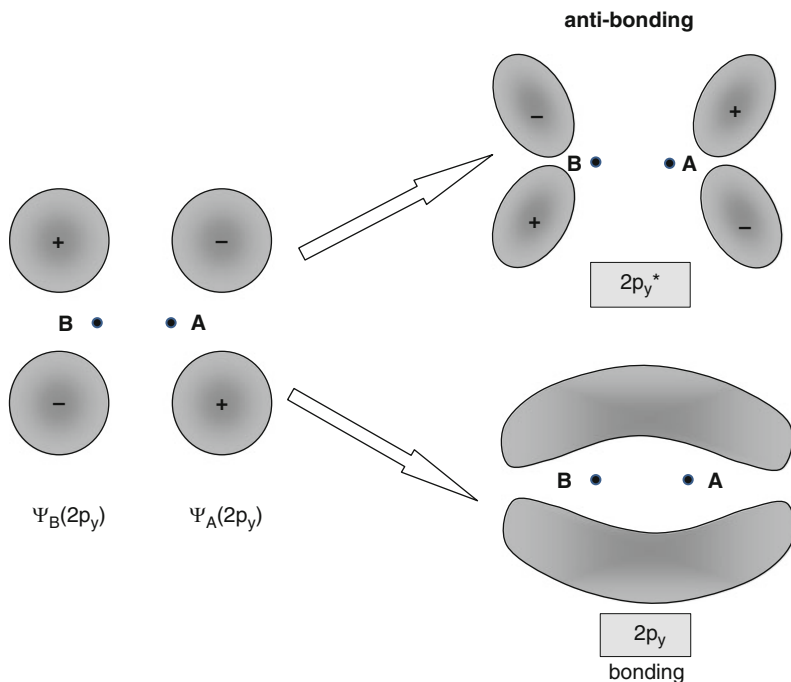
Fig. 1.3 Formation of bonding and antibonding π orbitals. The particular orbitals can be described by a function Ψ which represents the electron distribution in space

With this very brief introduction we can look in more detail into the electronic structure of the O_2 molecule. There is no magic microscope to reveal this, rather all has been deduced from many careful experiments and theoretical calculations. The picture that emerges is shown in terms of an "energy level diagram" in Fig. 1.4. The two oxygen atoms together carry 16 electrons. Four of these are in σ(1s) orbitals, and thus, yield no net bonding. This leaves twelve electrons in the outer shell. The 2s electrons form one bonding and one antibonding orbitals, and thus contribute no net bonding. Two of the 2p electrons form a $\sigma(2p_x)$ bonding orbital, and four more form two π bonding orbitals. This leaves two more electrons. They could be distributed in a number of ways, but in the oxygen "ground state" (the lowest energy state) they exist unpaired in two different antibonding π orbitals (see Fig. 1.4). The spins can add + or −, or cancel. These three possibilities (+, 0, −) yield a "triplet state" for the molecule. To emphasize this we will sometimes designate molecular oxygen in its ground state as $^3O_2$. Now we can calculate the net number of bonding electrons. In sum: of the twelve p electrons discussed above, eight are in bonding orbitals, four in antibonding. This leaves a net excess of four bonding electrons, which corresponds to two "classical" covalent bonds, in the traditional representation of the oxygen molecule as O=O.

The existence of two unpaired electrons in a molecule is very unusual and gives triplet oxygen some unique properties. For one, it means dioxygen must be

Fig. 1.4 Schematic molecular orbital energy level diagram for the molecule O_2 in its ground state. The relative energy levels of the σ and π electrons are schematically shown in the bonding and antibonding levels
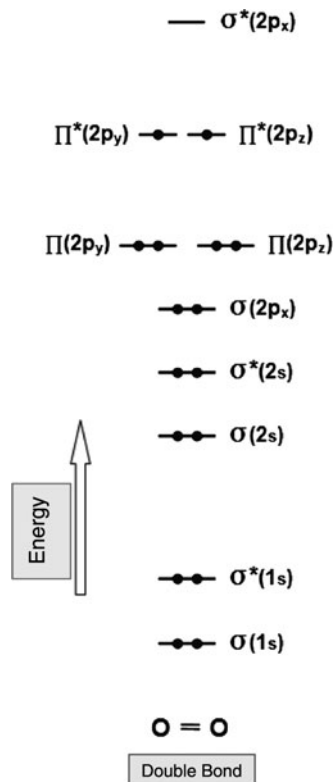
paramagnetic, and therefore attracted to the poles of a magnet. This was in fact discovered by Michael Faraday in 1845! Second, it tends to make ground state (triplet) oxygen less reactive than one might expect. The reason for this is a bit complicated. The rate at which a molecule such as oxygen can react with another molecule depends on how easily the "transition state" (an intermediate state of the two interacting molecules on the path to completion of the reaction) can be formed. The transition state often involves one molecule temporarily accepting a pair of electrons from the other. That can be easy if the ground state of the acceptor contains an empty orbital which can be shared temporarily with a filled orbital on the other reactant. But with triplet oxygen, the accessible orbitals are each half filled, and neither can accept an electron pair. Unless the other reactant also has an unpaired electron (which we said was rare) transitions are difficult and reactions are slow.

This is actually fortunate for us, for if reactions with oxygen were generally rapid, they would be uncontrolled. Our oxygen – based metabolism depends on the fact that the presence of catalysts favors particular desired oxidation reactions, and oxygen is not wasted in fruitless consumption (see Chaps. 3 and 4). Furthermore, the dioxygen molecule can persist in the atmosphere for long periods, which more reactive molecules such as Cl_2 cannot.

It is possible, by the introduction of a small amount of energy, to shift the orbital distribution of electrons to remove the unpairing by shifting both π^*2p_y electrons into one orbital. This produces what is called "singlet oxygen" designated $^1O_2$. The singlet state has zero net electron spin (all spins are paired) and is therefore not paramagnetic. Furthermore, singlet oxygen is highly and rapidly reactive, because it has an unoccupied orbital, and so does not suffer the same inhibition in forming transition state complexes as does triplet oxygen. As we shall see (Chap. 3), this has important consequences when living creatures have to deal with dioxygen. We provide here a brief view of the chemistry of some reactive forms obtained from dioxygen.

1.4 Reactive Oxygen Species

A number of reactive oxygen derivatives can result from the reaction of the singlet and triplet states of dioxygen with themselves or with other compounds. Only a handful of these are of importance in living systems. Their chemical properties and generations are briefly introduced here; their biological significance will be considered in detail in Chap. 3, and some of their medical consequences in Chap. 8.

1.4.1 Superoxide $^1O_2^{-*}$

Triplet oxygen can easily accept an electron resulting in a radical superoxide ($^1O_2^{-*}$) with a negative charge and singlet state, since one of the π^*2p orbitals is now filled with an electron pair (For nomenclature we shall use "1" indicating the singlet state, the asterisk "*" the radical property).

$$^3O_2 + e^- \rightarrow {}^1O_2^{-*}$$
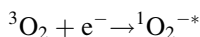
Interestingly, Linus Pauling predicted, as early as 1931, the existence of superoxide, based entirely on quantum mechanical considerations. This radical, however, is not itself very harmful in biological systems and does not cause much oxidative damage.

The main reaction of superoxide is to react with itself and hydrogen to produce hydrogen peroxide and triplet oxygen,

$$2\,{}^1O_2^{-*} + 2H^+ \rightarrow H_2O_2 + {}^3O_2$$
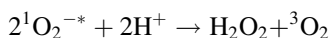
This superoxide dismutation can occur spontaneously or can be catalyzed by the enzyme superoxide dismutase.

1.4.2 *Hydrogen peroxide (H_2O_2)*

Reduction of superoxide ($^1O_2^{-*}$) by addition of an electron delivers first another activated form of oxygen which is termed peroxide ($^3O_2^{2-}$). When the negative charge of -2 is neutralized by two protons the product is hydrogen peroxide (H_2O_2). Although H_2O_2 is not very reactive, it is a precursor of the very reactive and damaging hydroxyl radical (HO^*). Thus, superoxide can also be considered a precursor of (HO^*). This can occur if superoxide acts as a reducing agent by donating one electron to reduce a metal such as ferric iron (Fe^{3+}). In a second step the reduced ferrous iron Fe^{2+} promotes the breaking of the oxygen–oxygen bond of hydrogen peroxide (H_2O_2) to produce a hydroxyl radical (HO^*) and a hydroxide ion (HO^-). The overall process, called the Fenton reaction proceeds as follows:

(superoxide radical) $^1O_2^{-*} + Fe^{3+} \rightarrow$ (peroxide) $^3O_2^{2-} + Fe^{2+}$

$^3O_2^{2-} + 2H^+ \rightarrow$ (hydrogen peroxide) H_2O_2

$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} +$ (hydroxyl radical) $HO^* +$ (hydroxyl ion) HO^-
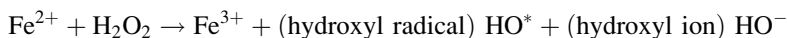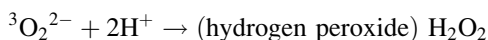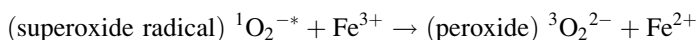
The hydroxyl radical (HO^*) can now react with superoxide $^1O_2^{-*}$ forming reactive singlet oxygen ($^1O_2^*$). Alternatively, the hydroxyl radical can react with many substances in the cell, with accompanying damage.

Another reaction is termed the Haber–Weiss reaction:

$^1O_2^{-*} + H_2O_2 \rightarrow {}^3O_2 + OH^- + OH^*$
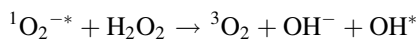
This chain reaction is biologically dangerous because it is readily catalyzed by common metals, and produces highly reactive substances.

1.4.3 *Peroxyl radical (ROO^*)*

The highly reactive hydroxyl radical HO^* can add to a substrate R (e.g., a carbon compound) forming a radical HOR^*, which could also further react with a ground-state triplet oxygen to produce a peroxyl radical (ROO^*).

$HO^* + R \rightarrow HOR^*$

$HOR^* + {}^3O_2 \rightarrow HOROO^*$
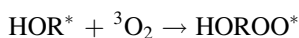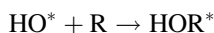
The various oxygen radicals have different lifetimes between 10^{-10} seconds and a few seconds depending on their reactivities (Table 1.1). All of these reactions,

Table 1.1 Lifetime of radicals: the stability of the various oxygen species can be described is by their lifetimes (Sies and Stahl 1995)

Nitritoxiradical (NO[*])	10^{-10} s
Hydroxyl radicals (HO[*])	10^{-9} s
Alkoxyl radical (RO[*])	10^{-6} s
Peroxyl radical (ROO[*])	7 s
Singlet oxygen ($^1O_2$)[*]	10 s
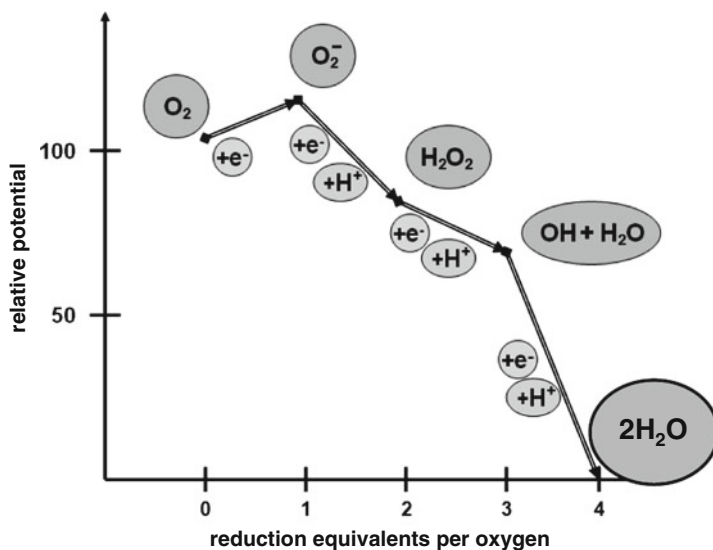
Fig. 1.5 ROS reactions-redox potentials of oxygen species. The stepwise reduction of dioxygen to water is indicated (Elstner 1990)

producing some highly reactive species, are summarized in Fig. 1.5. We shall return to a more detailed consideration of these reactions and their biological consequences in Chap. 3, and some of their consequences for human medicine in Chap. 8.

1.5 Ozone

There exists a second molecular form of oxygen called ozone (O_3). The ozone molecule involves π-orbitals that extend over all three oxygen atoms and σ-bonding orbitals that connect adjacent oxygen atoms to the central oxygen atom. This accounts also for the overall triangular shape of the molecule (Fig. 1.6a).

Ozone is formed when dioxygen is exposed to certain high energy sources, notably ultraviolet light or electrical discharge. The latter explains the acrid odor of ozone we notice during thunderstorms and around high-voltage equipment. Ultraviolet light must have wavelengths shorter than about 250 nm to produce ozone. This reaction involves first the splitting of the dioxygen molecule into two

Fig. 1.6 Ozone. Three
oxygen atoms form the ozone
molecule
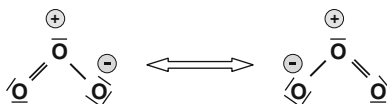
oxygen atoms; either of these can then add to an O_2 molecule to make an O_3 molecule. In nature this reaction occurs only above about 20 km above the Earth's surface. A concentration of 10^5–10^6 molecules ozone/cm^3 is measured. At lower altitudes the short-wavelength UV light from the sun is completely filtered out by O_2 absorption, and thus cannot form ozone.

Thus, ozone is being continuously generated in the stratosphere. It is also consumed there by another photochemical reaction driven by longer wavelength UV light which cleaves O_3 back to $O_2 + {}^1O_1{}^*$, producing an excited singlet state oxygen radical. The ozone absorption band for this cleavage centers at about 255 nm. Because of these opposing reactions, ozone in the stratosphere should come to a steady-state value which is sufficient to prevent much light of wavelengths of below about 300 nm from reaching the Earth's surface. This is fortunate for life, for light between 200 and 250 nm is able to destroy covalent bonds and therefore damage essential biomolecules. Indeed, UV radiation in this wavelength range is strongly absorbed by proteins and nucleic acids, with very deleterious results. In earliest times, life must have been confined to subsurface regions in the sea or land until enough O_2 appeared in the atmosphere to generate an ozone "shield".

Note that the ozone formation reaction depends on the concentration of oxygen. A consequence is that the ozone "shield" lies at around 20–30 km above the Earth's surface. At higher altitudes there is not enough oxygen to form much ozone, and at lower levels there is not enough short wavelength UV light penetrating to generate much. In addition, some long-wavelength light gets through to lower elevations and destroys ozone.

Ozone produces a second kind of protective effect through chemical "cleaning" of the atmosphere: The hydroxyl radical is most important for this, since it converts many compounds to water soluble forms, which will come down to Earth in rainfall. The reaction for HO* formation starts with

$$O_3 + h\nu \rightarrow {}^1O_1{}^* + O_2,$$
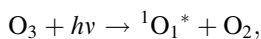
with ${}^1O_1{}^*$ being an excited oxygen radical in a singlet state. Together with water this reacts to form two hydroxyl-radicals:

$$^1O_1{}^* + H_2O \rightarrow 2\ HO^*$$
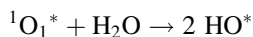
This radical can react with many atmospheric contaminants. For example with nitrogen oxide NO_2 it yields nitric acid HNO_3, which will fall down to as acid rain.

The term "ozone shield" is appropriate, because the protective effects are caused by a definite layer of ozone in the atmosphere.

Ozone and atomic oxygen are extremely reactive, so that the ozone shield is very vulnerable to reactive molecules introduced into the stratosphere. Before human industrial activity, this was uncommon. But more recently we have become the source of damage to the shield. Nitric oxides from jetliners, chlorine from chlorinated hydrocarbons, and many other sources now threaten this protection. Fortunately, through international cooperation, the use of halogenated hydrocarbons has been severely limited in recent years.

1.6   Water

The next sections include a brief description of two oxygen compounds – water and carbon dioxide. Both have played major roles in the evolution of life on Earth. A great deal of the Earth's oxygen is contained in water. About 70% of Earth's surface is covered by water and these oceans have long served as the major habitat of life. Organisms themselves consist of between 60 and 95% of water. Thus, water is fundamental to life. Water has particular and unusual properties due to the special electronic structure of the water molecule, which in turn is the consequence of the electronic structure of oxygen.

This structure has a general consequence that water molecules tend to associate together over a wide temperature range. For example because of the fact that the filled sp^3 orbitals of the oxygen lie on one side of the water molecule and the two protons are bound to the other side, a strong electric dipole is established. Thus, water molecules attract one another by dipole–dipole interaction. Even more important: water molecules also interact with each other by the stronger hydrogen bridges (Fig. 1.2b). These have a major influence on the properties of water, for water molecules form large flickering clusters held by hydrogen bonds (Fig. 1.7). The average lifetime of the water clusters is calculated to be between 10^{-10} and 10^{-11} s. The size of these clusters depends on the temperature (Frank and Wen 1957; Nemethy and Scheraga 1962). Up to about 250 water molecules are associated in the average clusters at temperatures close to the melting point and about 60 at 25°C.

This clustering explains the high viscosity of water at low temperature and its rapid decrease with increasing temperature. The lesser stability of biomolecules at higher temperatures is also largely a consequence of their interaction with water clusters. The interaction with water through hydrogen bonds is important for the stabilization of biomolecules such as proteins in solution, when they are "masked" by water molecules. The water forms hydration shells around the biomolecules, stabilizing their 3D structures. A proof for this is the uptake and release of bound water molecules by a protein when it switches between different conformations as observed, for example, when the cooperative oxygen carriers hemoglobins and hemocyanins switch between a low or high affinity state (Müller et al. 2003; Hellmann et al. 2003). An additional stabilization is due to the fact that the
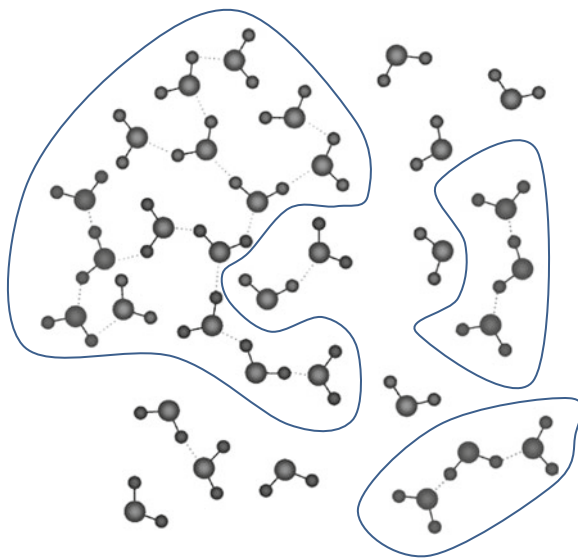
Fig. 1.7 Flickering clusters of water molecules. The water molecules form clusters and break the hydrogen bonds again within 10^{-11} s (Nemethy and Scheraga 1962)

hydration shell is "crosslinked" by clusters. Above about 50°C thermal fluctuation hinders the formation of clusters. As a consequence the stability of most proteins is reduced and they unfold easily under such conditions.

As temperature is lowered, the clustering of water molecules due to hydrogen bonding increases until at 0°C ice is formed. Here, the cluster size is essentially infinite, and every water molecule sits in the center of a tetrahedron in which four other water molecules are bound through hydrogen bonds. This is energetically very favorable for water, but it is far from close packing (Fig. 1.8). When balls are most densely packed, one ball coordinates with 12 other balls and 74% of the space is occupied. However, in ice, only 42% of the volume is occupied by water molecules. Thus, ice has a lot of empty space; in fact it is less dense than water and thus floats on top. For the solid form of a compound to be less dense than the liquid form is very unusual. This unusual behavior is also fortunate for life. If ice were denser than water, the oceans and lakes would long ago have frozen from the bottom, leaving only a thin band of cold water, even in the warmest climates. Thus, water with the highest density at (4°C) will always be found well below the ice shield in a lake, providing space for organisms to survive. Freezing of organisms is usually fatal, for formation of ice crystals will destroy the cells.

Incorporation of ions in water or blood has a major impact on fluid properties. Ions destroy the local water clusters by forming water shells around themselves. These water–ion clusters may either stabilize the structure of biomolecules or unfold them.

ice   water
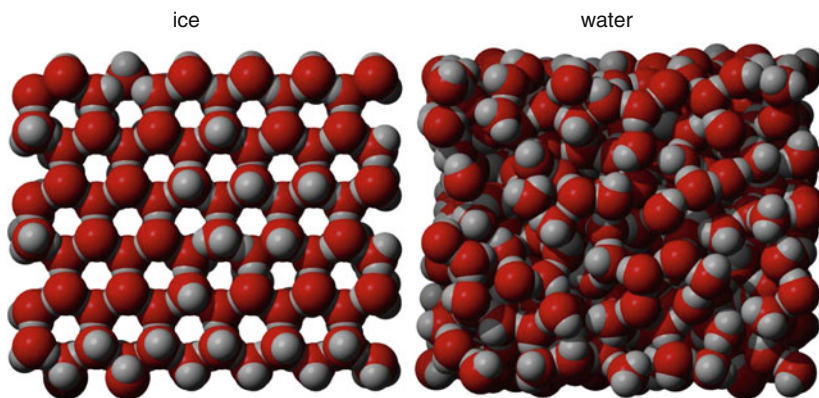
Fig. 1.8 The structure of ice and water. The oxygen atoms (*red*) and hydrogen atom (*gray*) are drawn as "spacefilling" models to illustrate how much free space there is in ice between the atoms. Note that water molecules are more crowded. This explains why frozen water needs more volume (Courtesy of Hermann Hartmann)

Water also possesses another feature which is important for life. Each biomolecule has a net charge, some positive, some negative which should lead to association between opposite charge types. With its strong dielectric property, water is able to lower the electrostatic interaction between the macromolecules by about 100-fold from the value it would have in a vacuum. Thus, biomolecules such as proteins will not cling together even when they are in a crowded neighborhood.

Because of the strong interaction between its molecules, water has a whole host of other properties that have been adventitious for life. For example – the unusually wide temperature range for the stability of liquid water (0–100°C) as well as the high heat capacity of water (4.25 J g^{-1} K^{-1}), have guaranteed that much of the Earth's oceans have remained liquid over the eons despite major variations in temperature. If this were not so, life could not have persisted.

The degree of dissociation of water into positive protons and negative hydroxyl ions is described by the pH-value which is the negative logarithm of the concentration of protons:

$$pH = -\log [H^+].$$

Thus, the higher the pH value, the lower the proton concentration and therefore the degree of dissociation of water. This behavior of water depends strongly on the temperature, the higher the temperature the lower the pH. Since many organisms adapt their body temperature to that of their environments, the pH value of the body will also change. In order to maintain the optimum in the metabolic process, nature must have evolved strategies to optimize the properties of all biomolecules in an organism despite such changes.

1.7   Water Vapor in the Atmosphere

In later chapters we will be much concerned about the composition and properties of the Earth's atmosphere, and how it has changed over geological times. One important constituent is water vapor "dissolved" in the atmosphere.

The content of water in the atmosphere as vapor depends on the temperature. The warmer the atmosphere the more water is dissolved. Evidence for huge amounts of gaseous water can be seen in the morning as dew. In addition, atmospheric water contributes significantly to keep the Earth warm, for water vapor is a strong "greenhouse gas". The surface temperature of the Earth without an atmosphere would be about $-18°C$ but in fact averages $15°C$ with its present atmosphere. Of this temperature increase about $20.6°C$ is due to water vapor, $12°C$ to CO_2 and $0.4–2.4°C$ to anthropogenic gases from human activity, which includes CO_2, CH_4, N_2O, halogenous carbonic compounds and aerosols. The fact that the amount of water in the atmosphere increases with temperature allows the possibility of a "runaway greenhouse effect" (see Chap. 6).

1.8   Carbon Dioxide

In addition to water a second oxygen compound has played a major role, both in the development of life and the evolution of the Earth's climate. This is carbon dioxide, CO_2. It is a very stable compound, involving covalent bonds between carbon (which has four half-filled sp^2 orbitals) and oxygen. Carbon dioxide is the major product of complete oxidation of carbon and carbon compounds by molecular oxygen. It constituted a major constituent of the primeval atmosphere (see Sect. 2.3). It is the primary carbon source for photosynthetic organisms and a primary product of aerobic metabolism. As we shall see, levels of CO_2 in the atmosphere have varied widely over the Earth's history, with often dramatic effect.

The property of CO_2 that is responsible for these climatic effects is its strong absorption in the infra-red region of the spectrum (see Fig. 9.3). Much of the energy that the Earth receives from the sun lies in the visible and near-ultraviolet region of the spectrum, to which the Earth's atmosphere (including CO_2) is quite transparent. As this incoming sunlight heats the Earth's surface, it is reradiated as infra-red radiation. Most of the atmosphere's gases are transparent to infra-red, but CO_2 is not, and acts in the same way that a greenhouse does to retain energy (see Houghton 1997). This "greenhouse effect" as we shall see in later chapters, has had a profound influence on Earth's climate over the ages. When CO_2 in the atmosphere is high, the Earth is warm, and when it is low it can be cold to the point of global glaciations. Ozone and water vapor are also "greenhouse gases".

There is another feature of carbon dioxide important for life. It dissolves very easily in water and reacts spontaneously to form bicarbonic acid (H_2CO_3) which dissociates to hydrogen carbonate ions (HCO_3^-) almost completely. The reactions, as well as the relative amounts at equilibrium, are given in the following:

$CO_2 + H_2O \rightarrow H_2CO_3$ (bicarbonic acid), 1% of CO_2

$H_2CO_3 \rightarrow H^+ + HCO_3^-$ (hydrogen carbonate ion), 95% of CO_2

$HCO_3^- \rightarrow H^+ + CO_3^{2-}$ (carbonate ion), 4% of CO_2

Thus, CO_2 dissolved in water acts as a buffer which is present in the ocean as well as in the blood of animals. In addition carbonate CO_3^{2-} is also crucial for many animals to form a protective shell of calcium carbonate.

1.9 Solubility of Gases in Water

The availability of gases in seas and lakes depends on the solubility in water. Biologically important gases show different solubilities, which are temperature dependent as shown in Table 1.2. The lower the temperature the higher is the oxygen content, a fact of importance for animals in arctic seas (see Chap. 5).

To assume a state of equilibrium between bodies of water and the atmosphere is only valid for the first meter in depth but equilibrium data still give a useful indication of aqueous environment. The equilibrium is on the side of dissolved CO_2 by a factor of 3,000. For example a concentration of 33.4 mmol CO_2 dissolved in water will be in equilibrium with a concentration of 0.01 mmol in air at $25°C$. Thus, an abundance of CO_2 can be quickly dissolved in water. In addition, CO_2 dissolves much better in seawater than in fresh water. With respect to the distribution of gases in deep seas, we note that the diffusion of gases in water is very low (see Chap. 5). Thus an active convection of the upper water with the lower water is necessary to provide gas mixing. However, it takes a lot of time – about 1,000 years to mix the top 1,000 m.

1.10 Hydrolysis and Dehydration: Central Water Reactions
in Biology

In a sense, water and its chemical properties lie at the heart of the most important biological structures and processes. The major macromolecular constituents of all cells are three kinds of polymers – polypeptides (proteins), polynucleotides (nucleic acids) and polysaccharides (carbohydrates). Each of these polymers is made up from a certain class of monomers, as shown in Fig. 1.9.

Table 1.2 Concentration of oxygen and carbon dioxide in water at 1 atm and pH 8.0

T (°C)	Gas concentration (mol m^{-3})			
	Oxygen		Carbon dioxide	
	Fresh	Sea	Fresh	Sea
0	0.45	0.36	0.023	0.019
20	0.28	0.23	0.012	0.010
40	0.20	0.17	0.007	0.006

$$- R_{n-1} - O - R_n - O - R_{n+1}\, O -$$

hydrolysis $(+ H_2O)$ ↓↑ $(-H_2O)$ dehydration

$$- R_{n-1} - OH \qquad HO - R_n - O - R_{n+1}\, O -$$
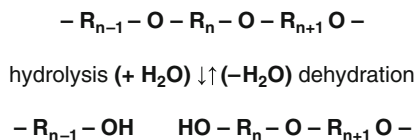
Fig. 1.9 Formation of macromolecular chains by dehydration and cleavage by hydrolysis. The "residues" (R's) can be sugars for polysaccharides, amino acids for proteins, and nucleotides for nucleic acids, but the principle is the same in all cases. In proteins, an amide bond is formed, but water molecule is still removed

In each case the polymer can be considered as formed from the monomers by a condensation reaction resulting in a removal of water (see Fig. 1.9). The actual reactions in vivo are much more indirect, but the overall process can be considered as the removal of one molecule of water from between two monomers. The opposite reaction, in which the polymers are broken down into monomers by addition of water molecules, is called hydrolysis.

Thus, water enters in a peculiar way into the formation and degradation of major cellular constituents. Remarkably it is also involved in the most fundamental process by which the cell stores and utilizes energy: the formation of ATP (adenosine triphosphate) from ADP (adenosine diphosphate).

$$ADP + PO^{3-}_4 \leftrightarrow ATP + H_2O$$
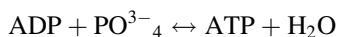
ATP formation is basically a dehydration reaction and requires the input of metabolic energy. The breakdown in hydrolysis is a major way to provide energy for biological processes.

Almost every source of energy a cell can utilize is stored in the dehydration reaction forming ATP. Almost every way the cell uses energy is through ATP hydrolysis. The actual reactions are usually much more indirect, but the essential reaction is that given above. This central role of dehydration and hydrolysis suggests that these processes trace back to the very origin of life itself.

1.11   Redox Reactions

Many important biological processes especially the generation of metabolic energy involve redox reactions: oxidation and reduction processes. The latter case describes the gain of electrons by a molecule resulting in a decrease in the oxidation state. Oxidation is the reverse reaction describing the loss of electrons by a molecule which results in an increase in oxidation state. Thus, the substance which loses electrons is oxidized and increases its oxidation number. This substance is called the reducing agent. The substance which gains electrons is reduced and reduces its oxidation number. This substance is called the oxidizing agent. Thus, redox reactions deal with the transfer of electrons from one reactant to another. This also means that when there is oxidation, there also is reduction.

Table 1.3 Representative reduction potentials

Reduction reaction	$E^{o, a}$ (volts)	$E^{o', b}$ (volts)	Comment
$Na^+ + e^- \rightarrow Na$	−2.714		Very strong reductant
$H^+ + 2e^- \rightarrow H_2$	0	−0.413	Taken as standard for E^o
$NADP^+ + H^+ + 2e^- \rightarrow NADPH$		−0.324	
$NAD^+ + H^+ + 2e^- \rightarrow NADH$		−0.320	
$(GS)_2 + 2H + 2e^- \rightarrow 2\ GSH$		−0.100	GS = glutathione
$FAD + 2H^+ + 2e^- \rightarrow FADH_2$		+0.031	
$O_2 + 2H^+ + 2e^- \rightarrow H_2O_2$	+0.69	+0.295	
$Fe^{3+} + e^- \rightarrow Fe^{2+}$	+0.771		
$O_2 + 4H^+ + 4e^- \rightarrow H_2O$	+1.229	+0.816	
$Cl_2 + 2e^- \rightarrow 2Cl^-$	+1.359		Very strong oxidant

[a] E^o is the chemist's standard, with unit activity in all species
[b] $E^{o'}$ is the biochemist's standard, with $[H^+] = 1 \times 10^{-7}$ M (pH = 7)

But what causes these processes? The capability for oxidation and reduction is described by the redox potential given in $\Delta E^{o'}$ (volt). The more negative the redox potential the stronger is the reduction power. Electrons flow from a redox pair of a more negative potential to the redox pair with less negative or even positive potential. Biologically important examples are given in the Table 1.3. Note that oxygen has a very high reduction potential: it is a powerful oxidizing agent.

As we will see in later chapters, the biological importance of redox reactions is to store and release biological energy. Photosynthesis involves the reduction of carbon dioxide into sugars and the oxidation of water into molecular oxygen (see Chap. 4).

$$6\ CO_2 + 6\ H_2O + light\ energy \rightarrow C_6H_{12}O_6 + 6\ O_2$$
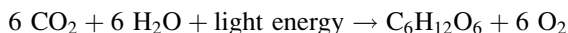
The input of light energy is essential because oxidation of water to O_2 is so unfavorable.

The reverse reaction, respiration, oxidizes carbon compounds

$$C_6H_{12}O_6 + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O$$
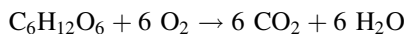
For the proton transfer complex cofactors are involved, $NAD^+/NADH$ (nicotinamide adenine dinucleotide) and the $NADP^+/NADPH$ (nicotinamide adenine dinucleotide phosphate). They recycle another cofactor, $FAD^+/FADH$ (flavin adenine dinucleotide). The balance of $NAD^+/NADH$ and $NADP^+/NADPH$ in a biological system such as a cell or organ describes their redox state (see Chaps. 4 and 8)

References

Elstner EF (1990) Der Sauerstoff: Biochemie, Biologie, Medizine. BI-Wiss-Verl, Mannheim, 3-411-14001-1

Frank HS, Wen WY (1957) Structural aspects of ion-solvent interaction in aqueous solutions: a suggested picture of water structure. Discuss Faraday Soc 24:133–140

Hellmann N, Raithel K, Decker H (2003) A potential role for water in the modulation of oxygen-binding by tarantula hemocyanin. Comp Biochem Physiol A Mol Integr Physiol 136:725–734

Houghton J (1997) Global warming: the complete briefing, 2nd edn. Cambridge University Press, Cambridge

Mathews CK, van Holde KE, Ahern K (2000) Biochemistry, 3rd edn. Addison-Wesley-Longman, San Francisco, CA

Müller G, Fago A, Weber RE (2003) Water regulates oxygen binding in hagfish (*Myxine glutinosa*) hemoglobin. J Exp Biol 206:1389–1395

Nemethy G, Scheraga HA (1962) Structure of water and hydrophobic bonding in proteins. 1. A model for the thermodynamic properties of liquid water. J Chem Phys 36:3382–3401

Sies H, Stahl W (1995) Antioxidant functions of vitamins E and C, beta-carotene and other carotenoids. Am J Clin Nutr 62:1315S–1321S

Tinoco F, Sauer K, Wang IC (2002) Physical chemistry, chapter 9, 4th edn. Prentice Hall, Uppersaddle, NJ

Chapter 2
A Brief History of Oxygen

Contents

2.1 Cosmic History of the Elements

Where did oxygen come from? Remarkably, that atom of oxygen you have just breathed had its origin in the heart of an ancient star. To understand this, one has to make an imaginary journey back to the creation of the universe, the "big bang," more than 12 BYA. We shall avoid details of physics, and simply describe a reasonable scenario that is accepted by most physicists today. Although happenings within the very first minutes of the universe remain very speculative most scientists agree with the following (see for example Müller and Lesch 2005; Allday 1999). After the big bang, the universe expanded exponentially by a factor of 10^{29} within about 10^{-33} s. These are inconceivably large and small numbers (and very rough estimates). At that point, the universe contained only the very elementary particles such as gluons, leptons, and quarks from which all other particles can be made, at a temperature of about 10^{27} K. After 10^{-6} s (1 μs), the infant universe had further expanded and cooled to about 10^{12} K and the basic particles of matter – neutrons, protons, and electrons – had formed from the elementary particles. Thus, very early,

the nucleus of the lightest element, hydrogen, was born but no other elements existed. Generation of other light nuclei could then follow in the hot plasma, by a process referred to as "hydrogen burning" (Müller and Lesch 2005; Truran and Heger 2004). Two protons would fuse to form a deuterium nucleus one proton being converted to a neutron by emission of a positron (see Fig. 2.1). The collision of deuterium with another proton resulted either in tritium (^{3}H, an isotope of hydrogen with one proton and two neutrons) or (more importantly) helium 3 (an isotope of helium with two protons and one neutron). This could be followed by a fusion of two helium 3 nuclei to produce one helium 4 nucleus with the accompanying release of two protons. These protons could then continue the cycle. These reactions are summarized in Fig. 2.1.

The helium 4 nucleus (^{4}He), containing two protons and two neutrons is very stable, and was the first stable product after hydrogen. A few ^{4}He and ^{3}H fused to form ^{7}Li. At this point (about 1 h after the big bang) the temperature had cooled to where hydrogen burning no longer continued. The matter in the universe now consisted mainly of a few kinds of nuclei: about 75% hydrogen, 24% ^{4}He, 0.001% ^{3}He, and traces of deuterium and ^{7}Li. In order to form complete atoms from these nuclei the temperature had to drop to about 3,000 K, which required about 400,000 years. At this temperature, photons no longer have enough energy to strip off electrons (ionize atoms) and the nuclei could catch and hold electrons to become atoms. Up to this point, no stars had formed: the universe consists of almost entirely of a dispersed gas of hydrogen and ^{4}He.

Then for about 200 million years (the "dark age of the universe"), there existed no stars, while gravitation gradually condensed the great clouds of hydrogen and helium (Müller and Lesch 2005). As yet, no further elements had been created: With time the clouds contracted and the pressure within these gas balloons increased to about 200 billion atmospheres which also drove the temperature to 10–40 million Kelvin allowing hydrogen burning to resume. As the clouds contracted and heated, they became stars, which began producing light due to the fusion of the hydrogen nuclei to helium (For insight into how this may have proceeded, see Yoshida et al. 2008). These ancient stars must have been very different from our sun. Many had masses of 100–1,000 solar masses and consisted only of hydrogen and helium. A very few surviving low mass examples of such

$$p + p \rightarrow d + e^+$$
● ● ●○

$$d + p \rightarrow {}^3He$$
●○ ● ●●○

$$^3He + {}^3He \rightarrow {}^4He + 2p$$
●●○ ●●○ ●●○○ ● ●
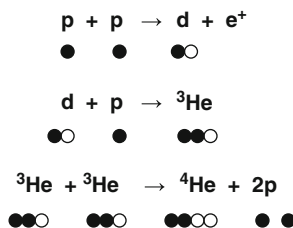
Fig. 2.1 Hydrogen burning. This is the first nucleosynthetic process in new stars. The principal product is ^{4}He. Protons released in step 3 are recycled in steps 1 and 2. Symbols: p = proton (*filled circle*); n = neutron (*open circle*); d = deuterium; e+ = positron (after Müller and Lesch 2005)

ancient hydrogen–helium stars have been found (see for example Schneider et al. 2006). As the hydrogen burning to helium continued, stars gained cores dense in ^{4}He. Compression of this nucleus raised temperatures to about 200 million Kelvin. At this point, a whole new series of nuclear fusion reactions (termed "helium burning") became possible. The most important are:

$$3\,{}^4\text{H} \rightarrow {}^{12}\text{C}$$

and

$$^{12}\text{C} + {}^4\text{He} \rightarrow {}^{16}\text{O}$$
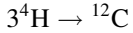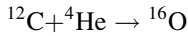
Nitrogen can be formed under similar conditions by the CNO cycle (Fig. 2.2), using ^{12}C and ^{16}O. The last stage in the helium burning process requires higher temperatures that could be produced only in more massive stars. These, then, were the main sources of oxygen. In a remarkable coincidence, three of the elements that are essential for life (^{12}C, ^{16}O, and ^{15}N) were synthesized in high yield in this first burst of elemental synthesis. With the already existing hydrogen, much of the potential basis for life had already been established. Early stars which had progressed through this stage of energy generation met one of two fates, depending on their mass. Those smaller than about 1.8 solar masses upon exhausting ^{1}H and ^{4}He supplies shrunk into white dwarfs. More massive stars became unstable as their dense cores contracted, and in most cases exploded as supernovae.

Such supernovae are of the utmost importance to us, for the exploding stars strewed the elements they had synthesized (including oxygen) about the universe. These products then became included in the composition of the next generation of stars. These were able, in turn, to synthesize higher elements, beginning with the H,
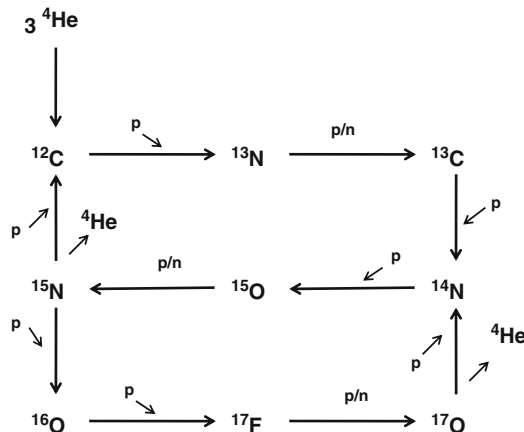
Fig. 2.2 The CNO cycle This process, starting with ^{16}O or ^{12}C, is responsible for most of the stellar synthesis of nitrogen isotopes. For clarification, we have introduced two nomenclatures: p for proton addition, p/n for proton to neutron conversion (Müller and Lesch 2005). Thus, only a few steps are necessary to form the important elements for life

Table 2.1 Successive element burning stages in the evolution of a massive star

Fuel	Main products	Secondary products	T (10^9 K)	Duration (year)	Main reaction
H	He	^{14}N	0.037	8.1×10^6	$4H \rightarrow {}^4He$ (CNO cycle)
He	O, C	^{18}O, ^{22}Ne s-process	0.19	1.2×10^6	$3{}^4He \rightarrow {}^{12}C$ $^{12}C + {}^4He \rightarrow {}^{16}O$
C	Ne, Mg	Na	0.87	9.8×10^2	$^{12}C + {}^{12}C \rightarrow \ldots$
Ne	O, Mg	Al, P	1.6	0.60	$^{20}Ne \rightarrow {}^{16}O + {}^4He$ $^{20}Ne + {}^4He \rightarrow {}^{24}Mg$
O	Si, S	Cl, Ar, K, Ca	2.0	1.3	$^{16}O + {}^{16}O \rightarrow \ldots$
Si	Fe	Ti, V, Cr, Mn, Co, Ni	3.3	0.031	$^{28}Si \rightarrow {}^{24}Mg + {}^4He \ldots$ $^{28}Si + {}^4He \rightarrow \ldots$

The table gives burning stages, main and secondary products (ashes), typical temperatures and burning timescales for a 20 solar mass star, and the main nuclear reactions (Truran and Heger 2004). (…) indicates more than one product of the double carbon and double oxygen reactions, and a chain of reactions leading to the building of iron group elements for silicon burning. Formation of higher elements, in massive stars, is based on further reactions starting with iron

He, C, O, N – rich mix with which they had been endowed. [For details of the processes see for example Truran and Heger (2004) and Rauchfuss (2005) as well as Table 2.1.] Various kinds of burning result in different distributions of elements. A "neon burning" happening in an explosive fashion at around 2×10^9 K led after many and complex steps to the biogenic element phosphorus ^{31}P. Since ^{31}P accumulates as a by-product of neon burning only about 2.5% of the important bioelement phosphorus is present in the universe. Stars with large mass (>4 sun masses) will eventually form iron and nickel centers via reactions involving magnesium and silicon. Table 2.1 summarizes and simplifies some of the very complex sets of reactions that can occur at successively higher temperatures as such stars burn their resources. Although the elements most essential for life (C, H, O, N) are products of the early stages or stellar nucleosynthesis, other critical elements such as iron, calcium, magnesium and phosphor needed these advanced processes for synthesis.

2.1.1 The Sun and Solar System

Our sun and solar system are condensates from a huge gas-dust cloud with an initial low density of about 10^8–10^{10} particles per m^3. This corresponds to a very high vacuum (one that would be difficult to achieve in even the best laboratory). The temperature was about 15 K. The mix of elements must have approximated the present composition of the universe as given in Fig. 2.3 although some hydrogen and other small atoms may have been lost.

As a consequence of the combined effects of gravitation and rotation, the primeval dust/gas cloud contracted into a disc like structure. At its center was the sun, with 99% of the total mass. In the surrounding disc small microparticles aggregated via microagglomerates to "planetesimales," most with dimensions of
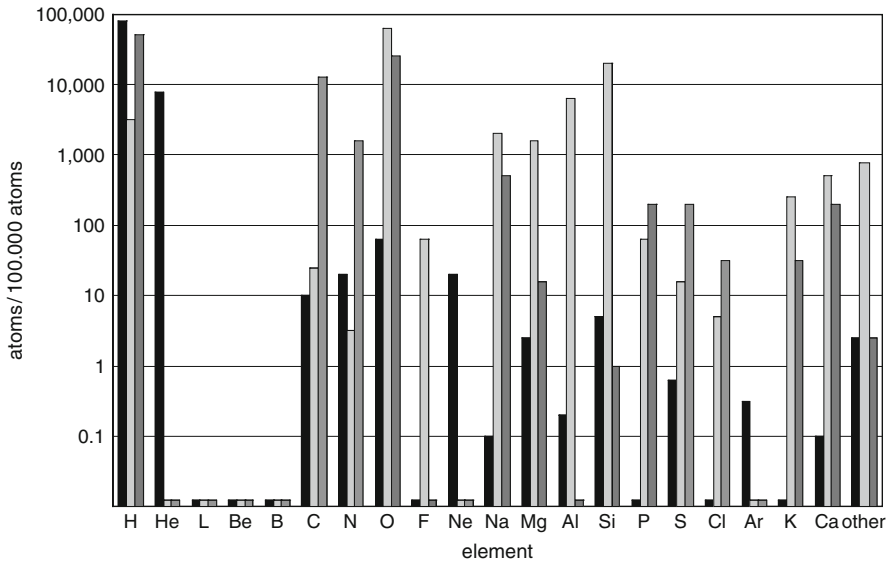
Fig. 2.3 Elemental composition of the universe, the Earth's crust and the human body amounts are expressed as numbers of atoms of each element per 100,000 atoms. *Black* denotes abundance in the universe, *light gray* in the Earth's crust, and *dark gray* in the human body. A logarithmic scale is used. On a linear scale, H and He would be seen to greatly dominate in the universe, O and Si in the Earth's crust, and H, C, N, and O in the body (Mathews et al. 2000)

up to a few meters but including some larger "asteroids" (Rauchfuss 2005). While the sun was formed about 8 BYA, the major formation of planetesimales has been calculated to be about 4.5 BYA (Wetherhill 1981). By collision between planete-simales larger and larger objects were formed, which finally formed the four compact and dense rocky "terrestrial" planets: Mercury, Venus, Earth, and Mars (Heuseler et al. 2000). Beyond Mars, the large gas planets such as Jupiter and Saturn had gravitation sufficient to retain gaseous hydrogen and helium as well. They exhibit a composition similar to the sun. Between Mars and Jupiter there is a broad zone, the asteroid belt. Here about 50,000 planetesimales circle, hindered from forming a planet by strong gravitational influence of the huge Jupiter (Weigert et al. 1996).

2.2 Formation of Earth

It is generally agreed that the age of Earth is of about 4.5 billion years. Between this time and 3.8 BYA, an interval termed the Hadean Eon (see Fig. 2.4), Earth may have first become liquefied by asteroid bombardment, and then slowly became cooler. The planet formed an outer thin crust, while the iron-nickel core separated from the mantle. There is recent evidence that this may have been completed by about 4.3 BYA, a surprisingly short period after the formation (O'Neil et al. 2008).
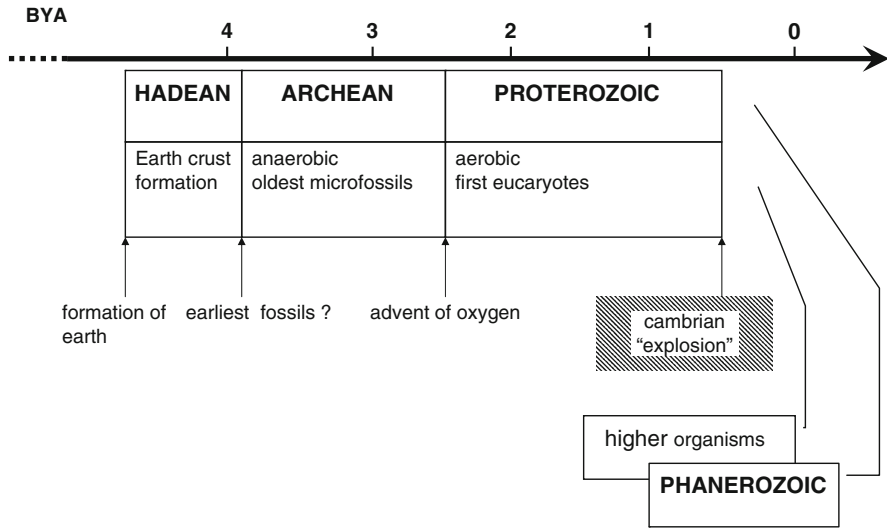
Fig. 2.4 Major periods in Earth's history. Important biological events are marked

However, the Earth was still being bombarded by massive meteorites, so that the surface was continuously reworked and even possibly remolten one or more times. Thus, the oldest solid objects serving us for information about the early solar system are not from Earth but from meteorites and the moon.

Both hydrogen and helium gas were the most abundant elements in the universe and in the solar nebula, and must have been present as a large fraction of that material which became part of primeval Earth. Today both helium and free hydrogen are very minor components of Earth or its atmosphere. What happened to these gases?

It is generally accepted that in the early stages of Earth formation the solar wind and heat of the sun blew away much of the light gases such as hydrogen, helium, methane and ammonia (see Seki et al. 2001). This is also true for much of the water vapor which could not condense on the hot Earth. Thus, mainly silicates and other minerals were retained and the actual atmosphere of the early Earth was created by the bombardments by the planetesimales as described below and outgassing of the interior (Press and Siefer 1995). The present occurrence of elements on Earth is given in Fig. 2.3

Earth as first condensed from planetesimales could not have been solid, but must have been molten due to several sources of heat energy. The radioactivity of the long-lived radioactive isotopes of uranium, thorium, and potassium (^{238}U, ^{235}U, ^{232}Th, and ^{40}K) produce heat in the interior of the Earth. The kinetic energy of captured planetesimals would further contribute to the heating of the surface of Earth as they collided with it. A planetesimal with a speed of about 11 km s^{-1} would deliver the same amount of energy as the same mass of TNT (trinitrotoluene). However, these sources would not alone explain the melting process. According to the homogenous aggregation model, the proto-Earth consisted of matter build

up mostly by iron, nickel, and silicate. As the concentration of heavy and dense elements such as iron and nickel migrated to the center of Earth they delivered an enormous amount of gravitation energy, contributing to the melting. As a consequence of this, the lighter elements are today found in the mantle (Rauchfuss 2005; Wills and Bada 2000) and iron and nickel in the core.

It is believed that about 4.5 BYA the moon was created by the condensation of particles being blasted out of the crust and mantle of Earth by an impactor of the size of Mars, about 20% of the mass of Earth (Münker et al. 2003). This explains why the moon is similar in the elemental composition to the crust and mantle of Earth.

The existence of a moon has been important for life, since it slowed down the rotation period of Earth from about 6 h in Hadean times to 24 h today, which has an influence on the temperature and movement of the water masses and therefore on the evolution of life.

2.3 The Primordial Environment

The history of Earth has been divided into several eons (Fig. 2.4). The Hadean ranges from the formation of the Earth to the first possible evidence for life, the Archean from then until the advent of atmospheric oxygen and the Proterozoic to the explosion of diverse animal forms, about 0.5 BYA. Subsequent time is termed the Phanerozoic.

2.3.1 Atmosphere of the Early Earth

Today, the atmospheres of the terrestrial planets are remarkably different (Table 2.2). The present Earth has 78% N_2 and 21% O_2 and only traces of other gases, including CO_2. In comparison to the other three planets lying close to the sun this is an exception. Mercury has almost no atmosphere, but Venus contains CO_2 up to more than 95% in a dense atmosphere. Mars has a very thin atmosphere containing mostly CO_2. Except for the Earth, a high percentage of N_2 is only found on Titan, a moon of Saturn. As we will discuss below, it is now thought that the atmosphere of the early Earth was close to that of present Venus and has gone through a complete change.

Table 2.2 Comparison of the atmospheres of near-Earth planets (Hunten 1993)

Planet	Total pressures (bars)	CO_2 (%)	N_2 (%)	O_2 (%)
Venus	92	96.5	3.5	<1
Earth	1.01	0.033	>78	21
Mars[a]	0.006	95.3	2.7	<1

[a]Other gases as Argon (1.6%) and CO (<0.1) occur as well

Sometime near the end of the Hadean, life arose. It is likely that this was not a "point event" that can be precisely located, but rather a process, repeated many times and with many failures until a self-replicating structure emerged. At any event, the first potential microfossils, which if real must already represent development of previously existing self-replicating structures, are found at about 3.8 BYA. This date is taken as the end of the Hadean and beginning of the Archeaneon (Fig. 2.4). However, highly creditable microfossils are only found after about 3.5 BYA (see Wacey 2009, for a critical analysis). At that point, Earth had been stabilized and possessed a small crust whose thickness was about 0.5% of the radius of Earth. As today, the crust consisted of large plates floating on the semifluid mantle. Along the connection lines between the plates many volcanoes were found, exhaling gases such as CH_4, H_2, CO_2, H_2O vapor and NH_3 and small amounts of H_2S from the interior of Earth. These materials must have contributed much to the atmosphere of the early Earth.

Rauchfuss (2005) has carefully compared two theories on the creation of the early atmosphere. According to one, it must have been derived from the solar nebula, as found in the atmosphere on Jupiter and Saturn, enriched in the strongly reducing gases (hydrogen, methane, ammonia, and water). However, doubts are based on two facts: the small Earth could not have held hydrogen for any significant period and the volcanoes-exhalation observed today consists mostly of water and CO_2. This should also represent the composition of the exhalation of volcanoes on early Earth (Quenzel 1987). According to Joyce (1989), the composition of atmosphere depended on whether the atmosphere was established before or after the creation of the iron-rich core: Contact with metallic iron before the nucleation would have resulted in a strongly reducing atmosphere with CH_4, H_2, H_2O, and CO. After nucleation, the redox state would depend on the ratio Fe^{2+}/Fe^{3+} and probably result in a weak reducing atmosphere with H_2O, CO_2, and CO and almost no CH_4 or H_2. This is at the moment the best accepted model for the atmosphere of very early Earth.

Even though they may still be released to some extent by outgassing, light gases such as hydrogen or helium do not remain in the atmosphere in significant amounts, because they have escaped from the Earth. The temperature and the atomic mass determine whether a particle can reach escape velocity. The rate at which particles will leave the gravitational attraction of the Earth depends on the mean speed of the particles. H_2 and He have the greatest speed, and they will most likely escape from the Earth. However, much of the hydrogen present in the early Earth was bound by oxygen to form water and to carbon and nitrogen as well. Therefore, almost all He was lost, but much of the hydrogen (in covalent compounds) was saved.

Thus, with time, an initial atmosphere composed mainly of free hydrogen and helium was replaced by denser gases: mostly N_2, CO_2, and water vapor and small amounts of H_2, NH_3, CH_4, and H_2S released from volcanoes. This latter generation continues even today. On the other hand the Hadean atmosphere contained almost no free dioxygen. The generally reducing conditions assured that even the small amounts of O_2 produced by ultraviolet photolysis of water vapor would be immediately reduced, for example by oxidizing methane to CO_2 or ammonia to N_2.

$$2 \, H_2O_{vapor} + \text{energy rich UV-light} \; \rightarrow 2 \, H_2 + O_2$$

$$2 \, O_2 + CH_4 \rightarrow CO_2 + 2 \, H_2O$$

$$4 \, NH_3 + 3 \, O_2 \rightarrow 2 \, N_2 + 6 \, H_2O$$
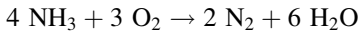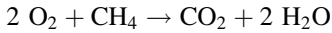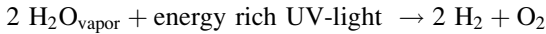
Thus, although the element oxygen was indeed present in abundance (see Fig. 2.3), it was almost entirely locked up in oxides, including hydrogen oxide, (water), carbon dioxide, and silicon dioxide. The surface temperature of the Earth must have been very high during the early part of the Hadean, as the crust formed despite continuous meteorite bombardment. As the bombardment decreased, the crust cooled, eventually to that point when liquid water could exist at the surface. Then, evaporation of water would also contribute to the cooling.

2.3.2 Water on the Earth' Surface: The Origin of Oceans

It is difficult to decide exactly when water, in the form of oceans or lakes, was first present on Earth's surface. There are several reports of stromatolites, which are known today as bacterial mats formed in shallow water with an age of 3.5 billions of years. Observation of 3.8 billion year old sediments gives an upper limit, while earlier rocks that might give evidence have been highly modified.

What was the source of fluid water on Earth? For a long time it was believed that the hydrosphere was exclusively created by volcanic activity. However, Delsemme (1992) has summarized arguments that most water in the oceans has an exogene origin delivered by comets and meteorites, since Earth itself was formed from dust of low water content. Comets consist of more than 40% water. In support of this idea, the ratio of deuterium to hydrogen in Halley's comet was determined to be $0.6–4.8 \times 10^{-4}$, which is in the range of that determined for the water in the oceans (Robert 2001) and meteors (Chyba and Sagan 1997). According to Rauchfuss (2005) the amount of the first ocean formed in this way contributed 20–70% of the ocean today. However, this estimation is very insecure due to the influence of UV induced photodissociation of water, producing hydrogen which can escape to space.

2.3.3 The First Greenhouse Effect

CO_2 was a major constituent of the early atmosphere and played a vital role in determining surface temperature. On the basis of solar radiation alone one might expect a cold early Earth. The sun, 4 BYA, produced only about 70% of the radiant energy today (Sagan and Mullen 1972) most likely due to the lower ratio of He/H in

the sun at that time. But a high CO_2 and CH_4 level in Earth's atmosphere could produce a strong "greenhouse" effect (Kasting and Howard 2006). Atmospheric CO_2 content is estimated to have been about 100–1,000 times higher at that time than today (Owen et al. 1979), as a result of the greater tectonic activity in the early Earth. These levels would compare to those of Venus today. Estimates are difficult, but many believed that average temperatures of about 80°C at 4.5 BYA may have been typical of the latter part of Hadean era (Kasting and Howard 2006). Later, for the Archaean, temperatures have been calculated ranging from more than 50 to 80°C. It is important to note that the extremes lie below the boiling point of water. This is very important for the subsequent history of Earth. If the oceans had boiled, putting all of the water into the atmosphere, a runaway greenhouse would have occurred. It is believed that this is what happened on Venus.

2.4   Life: Its Origins and Earliest Development

Sometime, between 3.8 BYA and 3.5 BYA or possibly even earlier, life began its existence on the anoxic Earth. Exactly when, we can probably never know. The question may not even have a clear meaning, for the earliest divisions between the nonliving and the living would probably be hard to define or for us to recognize. Further, it is possible that this origin occurred many times, only to be wiped out by the catastrophic meteorite bombardments of the late Hadean. It has been generally but not universally accepted that putative fossil microorganisms, dating from about 3.5 to 3.8 BYA represent early life. The latter value allows the remarkably short period of only about 600 million years between the solidification of the crust and the appearance of recognizable organisms. There must, however, have been an extensive evolution from the first self-replicating systems to produce structures that we can recognize as similar to existing microorganisms. Recently the date of 3.8 BYA has been called to question (Fedo et al. 2006), and some now hold that the earliest unambiguous signs of life stem from about 3.5 BYA (see Schopf 2006; Wacey 2009). This leaves a more comfortable 1 billion years for preceding evolution.

Nevertheless, when we consider the evolutionary distance from the nonliving to the living, even a billion years seems scanty. Some have attempted to avoid the problem by invoking the old hypothesis of panspermia – that life was brought to Earth in an already functioning form from elsewhere in space. (see Martin and Line 2002 for a careful discussion). The idea seems less farfetched now that we know that meteoric materials from the moon and Mars actually have reached Earth. Indeed, there have been claims of material of possible biological origin in a Martian meteorite, although these have been disputed (see Thomas-Keprta et al. 2002 for discussion). In any event, the idea is useful in solving the present dilemma only if the "mother" world had a much longer period for life's gestation than Earth. No obvious candidate is known, although Mars which will have cooled more quickly than Earth may have some advantage in that respect.

Wherever and whenever life first arose, we must face the awesome question of how, and in what environment, this could have happened. To even approach the question, we must ask what defines life.

Living organisms are self-replicating, but they are more than that. They are entities that can in one way or another tap energy from a chaotic surrounding to build and replicate highly organized structures. They utilize free energy from some external source to temporally decrease entropy within a confined space. Yet (and this is most important!) they do not replicate *exactly*. The replication involves errors (mutations) which are the progenitors of evolution. This must have been especially true for the earliest "organisms"; to survive and perpetuate (which some did) they must have had to rapidly adapt to a host of different environmental conditions. So, if we try to imagine the protoorganism, we must think of a small, somehow contained self-copying system driven by some external energy source. Somehow a periodicity from the environment must have been imposed on it, so that it could go through cycles of replication. Cycles of day and night, or tides might have played this role. Despite a general prejudice to the contrary, we do not feel that life originated in the ocean, even in the presence of hydrothermal vents. The conditions there too strongly favor hydrolysis of the polymeric molecules essential to life. A periodically moist and dry environment would have been much more favorable (see van Holde 1980).

There is probably no area in science more devoid of testable or even plausible hypotheses than the "origin of life problem." For obvious reasons, we have no physical evidence of the very earliest life, so are obliged to guess at the substances involved, the nature of the reactions, or the physical milieu in which they occurred. Clearly, at some point the polymeric substances that now dominate life – the nucleic acids, the proteins, and the polysaccharides must have come into being, encapsulated within lipid membranes or other containers. But which was first? Current thinking favors the nucleic acids, based on their potential for self replication. We know that DNA carries the information from generation to generation. But while DNA stays in the cell nucleus, RNA, also a nucleic acid, transfers a copy of a particular part of the DNA from the cell nucleus to the cytosol. This serves as a matrix for the synthesis of proteins, composed of amino acids. Thus, proteins express the potential of the genetic information, as, for example, in catalysis by enzymes or oxygen transport in the case of hemoglobin. The surprising discovery that some RNA molecules can also function as enzymes ("ribozymes") has led to the concept of an "RNA world" as life's first aspect. But where did the ribonucleotides come from, and how was their polymerization (and sequence replication) driven? There exist a few suggestions for possible modes of abiotic production of some precursor molecules, but the gaps, even in speculation are enormous. (For an overview of the problem, see Orgel 2004.)

However, whenever the first life appeared on Earth, it seems that dioxygen played no role in the process. Despite some possible production of O_2 via photolysis of water, the atmosphere and all bodies of water, before about 2.5 BYA were almost completely anoxic, as testified by a variety of criteria (see Holland 2006, for discussion). Indeed, it is probably fortunate that this was so, for dioxygen and its

products might have been forbiddingly toxic to primitive metabolisms that were not yet able to evolve defences. Later, as dioxygen levels slowly rose, the bacteria that faced this threat had sufficiently sophisticated metabolisms and evolutionary flexibility to deal with it (see Chap. 3).

We picture then, by at least 3 BYA an essentially anoxic world, populated, at least within parts of its oceans, by anaerobic microorganisms. There must have been a long, slow process of development, from the simplest self-replicating systems to organisms that possessed the ability to selectively utilize available energy sources to synthesize membranes, proteins, and nucleic acids, all in a more-or-less coordinated fashion. It may well be that there were a large number of lines of development, but the basic unity of life today indicates that only one was competitively successful. In the next billion years, this line was able to develop an energy-gathering system that would irrevocably change the entire world.

2.5 A Billion Years of Life Without Dioxygen: Anaerobic Metabolism

As described above, the first strong evidence for the existence of life on Earth is found in rocks about 3.5 billion years old. The Earth at that time shows no evidence for significant amounts of free dioxygen – less than 10^{-5} compared to the values today – either in the atmosphere or in the oceans. Indeed, it is not until about 2.5 BYA that evidence for appreciable dioxygen is found. This means that for a billion years all organisms had to survive and evolve using one or another form of anaerobic metabolism. In all that time no major change in lifestyle appeared. However, we must not assume that evolution was in stasis; the organisms existing at 3.5 BYA must have been very primitive, and increasing sophistication in metabolism must have evolved during the succeeding billion years. Indeed, as we shall see, the evolution of oxygenic photosynthesis, one of the great breakthroughs in metabolism, must have occurred sometime in this period. Nevertheless, for reasons we will discuss below, any dioxygen so generated did not begin to accumulate in the environment until the end of the Archean, 2.5 BYA.

2.5.1 Some Principles of Metabolism

Before discussing possible modes of primitive anaerobic metabolism, some general comments on metabolism are necessary. For any organism to grow, become more complex, and reproduce itself, a source of available energy or "free energy" is required. According to thermodynamics, free energy (G) involves both enthalpy (H), and entropy (S) and temperature (T)

$$G = H - TS.$$

Or, for a process at constant T

$$\Delta G = \Delta H - T\Delta S.$$

H is basically a measure of the energy of a system. Entropy can be thought of as a measure of the randomness or disorder of a system. For life processes to be spontaneous and ongoing, the change in G must be negative. But growth and replication of a cell involve an increase in order and complexity, and hence entropy decreases and therefore ΔS is negative and so the product $(-T\,\Delta S)$ is positive (since T must always be positive). So, to make ΔG negative, ΔH must be negative – energy must be expended to drive living processes. Fundamentally, what a living organism does is to use a source of energy to reverse, within the volume that represents itself, the normal increase in entropy that proceeds through the universe. This energy can be derived from a multitude of metabolic pathways that have evolved over the ages to meet the different circumstances of different organisms. They all have, however, one feature in common, which must be ancient indeed. All pathways of reactions for generating metabolic energy lead, often very indirectly, to the formation of adenosine triphosphate (ATP), usually from adenosine diphosphate (ADP) and inorganic phosphate (HPO^{3-}), abbreviated P_i:

$$ADP + P_i \rightarrow ATP + H_2O$$
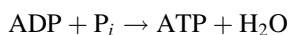
The ATP thus formed in cells acts as a universal "energy currency" because its hydrolysis back to ADP + P_i (the reverse of the above reaction) yields a large amount of free energy that can be coupled to drive a multitude of processes These include such diverse reactions as the replication of DNA, the synthesis of proteins, pumping substances through membranes, cell motility, etc.

The many different reactions that can drive ATP synthesis are in the most part oxidation/reduction reactions in which electrons are transferred from a reduced substance (called the donor) to reduce another molecule (the electron acceptor) (see Sect. 1.11). In much of the metabolism of organisms in an dioxygen-rich world like ours, the ultimate acceptor is O_2, which has very high electron affinity or we can say, high oxidizing power. By accepting protons along with the electrons, O_2 is reduced to water, H_2O. The electron donors are many, and some will be described in the next chapter. This referred to as *aerobic metabolism*.

However, the question we must consider here is the following: In the dioxygen-free environment of the Archaean, what pairs of electron donors and acceptors were available to drive metabolism, and might have been used by Archaean organisms? We can gain some hints by examining those organisms that exist today in wholly anoxic environments. These "strict anaerobes" as they are called, often cannot even tolerate traces of dioxygen and may well represent direct descendents of Archaean bacteria. There are a wide variety of such metabolisms. Many are fermenters, using organic compounds in their environment (Fig. 2.5). Some of the compounds used today are consistent with what we know of the Archaean environment.

a

b
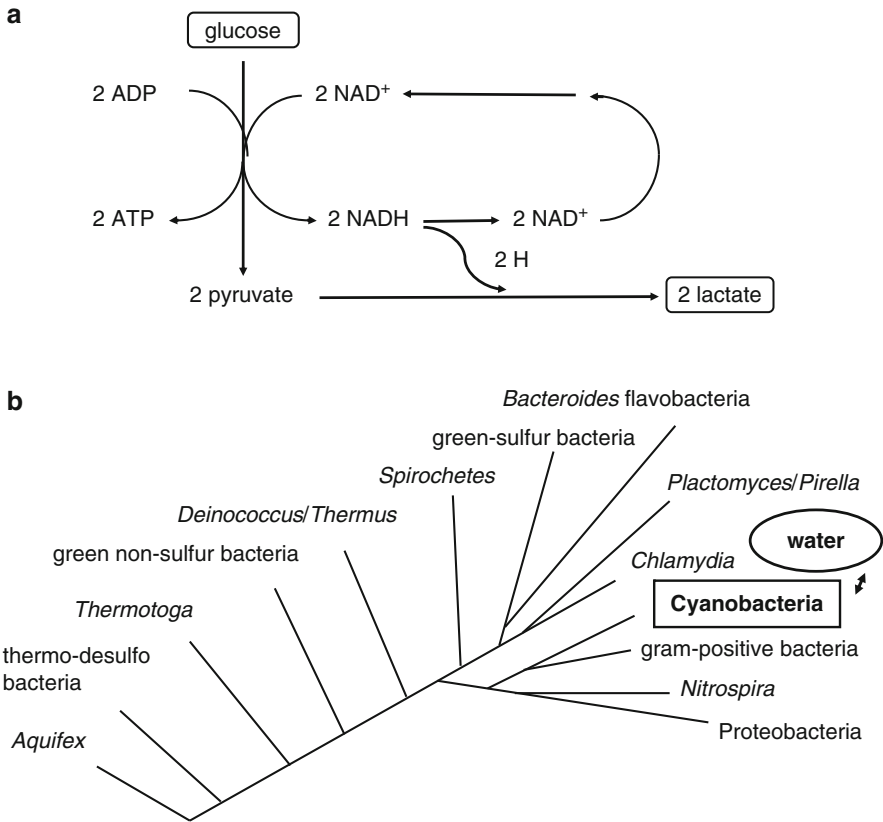
Fig. 2.5 Examples of anaerobic metabolism. (**a**) Fermentation by *Lactobacillus*. All metabolic pathways start with glucose. Fermentation degrades glucose to pyruvate gaining 2 ATPs. The produced H^+ ions are neutralized by pyruvate to lactate. (**b**) Phylogeny of bacteria. Several bacteria families evolved diverse mechanism to fix CO_2 and synthesize basic carbon biomolecules (CH_2O). Most of them used diverse anaerobic pathways for gaining necessary energy. One organism, *Aquifex*, even needs small amounts of dioxygen to survive. The only bacteria that evolved mechanisms to use water as electron donor are cyanobacteria (Courtesy of Helmut König)

Many different bacteria (Fig. 2.5b) including more than 84 different thermophylic organisms evolved on early Earth. The necessary energy for synthesizing complex biomolecules was provided by a wide variety of oxidation/reduction reactions. Some reduce CO_2 by H_2, others used a variety of sulphur-based reaction, such as:

$$FeS + H_2S = FeS_2 + H_2$$
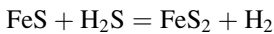
The hydrogen is either used in proton pumps to produce ATP (see Chap. 4) or was used by other primitive anaerobe organisms such as *Pyrodictium occultum* on early Earth to produce H_2S in a hydrogen/sulfur autotrophic process:

$$H_2 + S \rightarrow H_2S$$
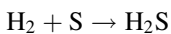
H_2S serves as an electron source at times when dioxygen was not available.

Among the many different bacteria only one bacterium family, the cyano-bacteria, learned to use water as an electron donor producing oxygen (Fig. 2.5b). Beside that another bacterium line (*Aquifex*) needs dioxygen although only in very low concentrations. Thus, only two classes of bacteria have the capability to handle dioxygen. Nevertheless, several strict anaerobic organisms evolved mechanism to defend against dioxygen and derivatives; for example methane bacteria have the enzyme catalase to neutralize H_2O_2. This is a strong indication that trace amounts of dioxygen were already present in the primitive Earth.

One characteristic of all of these anaerobic pathways is a relatively low energy efficiency, in the sense that little ATP is produced per mole of input substrate. Consider, for example, the well-studied fermentation of glucose to produce lactic acid (Fig. 2.5a). This yields only 2 mol of ATP for each mole of glucose consumed. In contrast, we shall see in the next chapter that the aerobic metabolism of 1 mol of glucose will produce 18 mol of ATP! A second major handicap of the metabolic pathways available to primitive anaerobes was the scarcity of raw materials, particularly reductants. While hydrogen and hydrogen sulfide were certainly avail-able on the primitive Earth, they were generally restricted to very limited environments such as places of volcanic activity. There were probably small amounts of various organic compounds such as glucose remaining in the early oceans from prebiotic syntheses, but they were surely highly diluted, and in ever diminishing supply. Archaean life faced starvation, and might ultimately have vanished from Earth, had there not occurred an evolutionary leap of the greatest significance – the development of photosynthesis. This momentous step transformed life, and indeed, the whole Earth. We now trace its origins and effects.

2.6   The Invention of Photosynthesis

When we think of photosynthesis today, we usually think in terms of the oxygenic photosynthesis that is common to all contemporary plants and algae, and some bacteria. This process uses the energy of sunlight (*hv*) to fix carbon dioxide into carbohydrates via the (vastly oversimplified) reaction:

$$CO_2 + H_2O + hv \rightarrow [CH_2O] + O_2$$
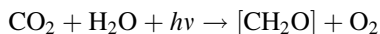
[CH_2O] signifies many different kinds of carbohydrate, which can then be metabo-lized by the organism or by organism that eat the photosynthesizers. This is certainly not descriptive of the first photosynthesis, for using water as a reductant here requires a very sophisticated chemical mechanism. Indeed, even today many microorganisms utilize other, more facile reductants such as H_2S. A list of a few such reactions is shown in Table 2.3. The importance of these processes, some of which surely were operative in the early Archaean, is that in addition to fixing carbon into the cell, they also provide a source, though somewhat inefficient, for the

Table 2.3 Reductants and several reactions for several types of photosynthesis (Matthews et al. 2000)

Organisms	Reductant	Reaction
Plants, higher Algae Cyanobacteria	H_2O	$CO_2 + H_2O \rightarrow [CH_2O] + O_2$
Green sulfur bacteria	H_2S	$CO_2 + 2H_2S \rightarrow [CH_2O] + H_2O + 2S$
Purple sulfur bacteria	HSO_3^-	$CO_2 + H_2O + 2HSO_3^- \rightarrow [CH_2O] + 2SO_4^{2-} + 2H^+$
Nonsulfur bacteria[a]	H_2	$CO_2 + 2H_2 \rightarrow [CH_2O] + H_2O$

[a]Many other reductants, such as lactate, are used by some

generation of ATP from ADP. This was enormously important on the early Earth, for as we have seen, the first anaerobes must have been surviving on ever-diminishing energy sources. The unlimited supply of solar energy would have alleviated, at least in part, this limitation to life.

Exactly how these first photosynthesizers arose is of course unknown, but it must have been through the generation of a chlorophyll pigment associated in some way with an electron transport chain. There is evidence for such anoxygenic photosynthesis as early as 3.5 BYA. Despite their clear advantages over nonphotosynthetic anaerobic bacteria, these early photosynthesizers were still burdened by the limitation of restricted resources. A glance at Table 2.3 shows that the reductants employed were substances that would have been found only in selected habitats, or more widely in great dilution. The incredible importance of oxygenic photosynthesis, as shown by the equation above, arises from the fact that it uses as reactants just the two substances most abundant in the Archean world – water and carbon dioxide. As was pointed out in an earlier section, the levels of CO_2 in the Archaean atmosphere (and presumably in the ocean) were many-fold higher than today.

The advent of oxygenic photosynthesis appears to have been correlated with the evolution of cyanobacteria, a species still thriving on Earth (see Fig. 2.5b). This occurred approximately 2.7 BYA (Fig. 2.6), and marks a turning point in Earth's history (Canfield 2005; Holland 2006). It must not be assumed that this was an easy evolutionary step. To utilize a molecule of water as the source of electrons for reduction, more energy is required than can be provided by a single photon of visible light. With far blue or ultraviolet light single photosynthesis is theoretically possible. However, this would be inefficient because it would make use only of a small part of solar radiation. Anoxygenic photosynthesizers, using more facile reductants, could achieve the necessary electron transfers with a one-step process, but to utilize water requires two excitation steps involving absorption of two photons of visible light in two coupled "photosystems." In chloroplasts of higher plants and algae, these systems are embedded in "thylakoid membranes" which divide into a lumen surrounded by a stroma (Fig. 2.7). Cyanobacteria are rather like free-living chloroplasts, and are most likely their origin (see Chap. 4). The mechanism of producing dioxygen from water (which is actually quite complex) is schematized in a very simplified way in Fig. 2.7. In one step (photosystem II) electrons are stripped from water, free dioxygen is evolved, and protons are released:

$2\ H_2O \rightarrow 4\ H^+ + 4\ e^- + O_2$
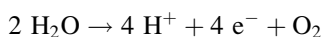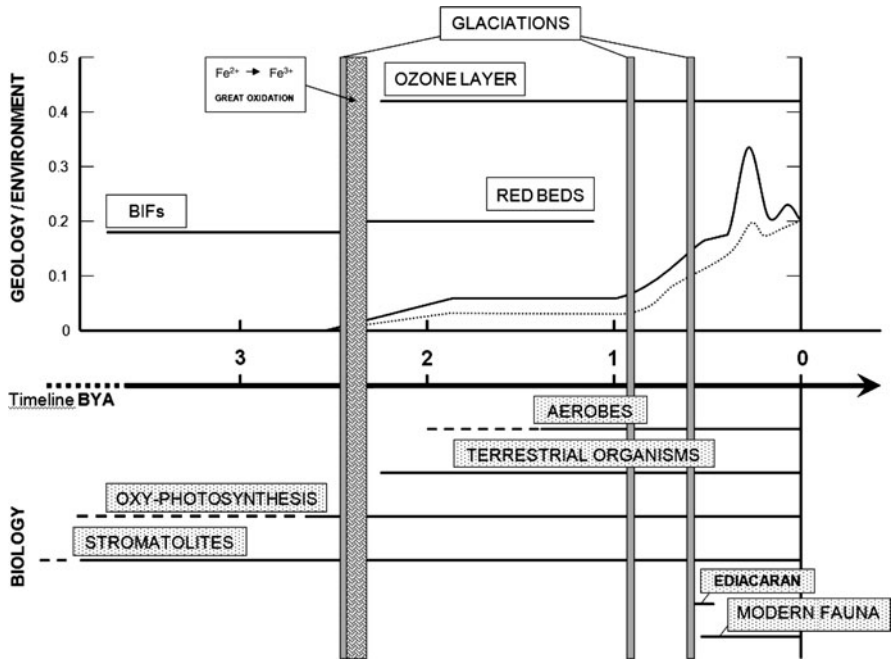
Fig. 2.6 The rise of dioxygen during geological history, and some consequences. The figure was adapted from Holland (2006) BIFs: precipitation of iron from seawater by oxidation Red Beds: Terrestrial iron deposits that have been oxidized, usually by oxidizing atmosphere. Major glaciation periods are indicated. The two *heavy* and *dotted curves* represent upper and lower estimates of dioxygen levels in the atmosphere, according to Holland (2006). Proposed oxygen levels, as fraction of atmospheric gases, are shown in scale to right, top

The protons, in passing through a membrane via special protein complexes, can drive the generation of ATP from ADP (see Chap. 4). The electrons are passed to photosystem I, where a further light absorption excites them to allow reduction of the universal cellular oxidant, nicotinamide adenine dinucleotide phosphate ($NADP^+$)

$$4\ e^- + 2\ H^+ + 2\ NADP^+ \rightarrow 2\ NADPH$$

NADPH can then be utilized in what are termed as the "dark reactions" of photosynthesis to reduce carbon dioxide to various carbohydrates. As we shall see, this remarkable process not only saved life from extinction through starvation, but also literally remade the world.

A continual product of oxygenic photosynthesis is dioxygen, and its gradual accumulation in the ocean and atmosphere had remarkable effects. Next to the origin of life itself, the development of oxygenic photosynthesis is probably the most important step in all of evolution.
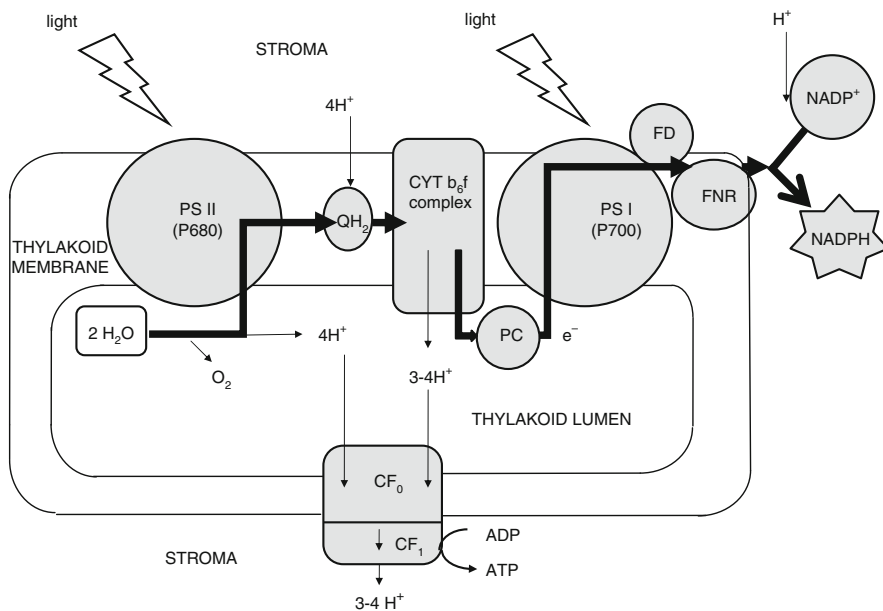
Fig. 2.7 The mechanism of oxygenic photosynthesis photosystems I and II (PS I and PS II) and the cytochrome b_6f complex are physically separate protein complexes embedded in the thylakoid membrane. The *heavy line* indicates electron flow. Electron transfer from PS II to cytochrome b_6f is by diffusion of reduced plastoquinone (QH_2) in the membrane lipid. Transfer from b_6f to PS I is mediated by plastocyanin (PC), soluble in the lumen. The protons added to the thylakoid lumen during the light reactions pass through the thylakoid membrane via ATP synthase complexes (CF_0–CF_1). The ATP synthase particles are in the CF_1 subunit and face into the stroma, so that ATP is generated in that compartment. Reduction of $NADP^+$ also occurs on or near the stromal surface of the membrane via FNR (ferredoxin NADP reductase) and FD (ferredoxin) (Mathews et al. 2000)

There is not complete agreement as to when this step occurred. There is good evidence that cyanobacteria and oxygenic photosynthesis were present by about 2.7–2.9 BYA. However, there are those who hold for a considerably earlier date (see Tice and Lowe 2004). Since the cyanobacteria must have evolved from anoxygenic photosynthesizers, and that evolution must have been complex, we can probably set the date with some confidence as certainly later than 3.5 BYA. In any event, the ultimate consequences were dramatic.

2.7 How Oxygenic Photosynthesis Remodeled the Earth

2.7.1 The First Rise of Dioxygen

Around 2.7 BYA, primitive oxygenic photosynthesis began to deliver molecular dioxygen to the Earth's oceans. For hundreds of millions of years this had

only small effects on the dioxygen concentration in the oceans or atmosphere (see Fig. 2.6). This is not surprising, for all evidence indicates that the primeval seas contained immense amounts of metals in their reduced states, especially ferrous iron. This is supported by the existence from the early archaean of "banded iron formations" corresponding to precipitation of ferric oxides. Formation of these ceases about when free dioxygen first appears in the oceans. The subsequent accumulation of O_2 in the atmosphere led deposition of Fe^{3+} in the form of Fe_2O_3, forming immense terrestrial "red beds" beginning at about 2.3 BYA (Fig. 2.6; Rye and Holland 1998). Only when the oceanic reservoirs of reduced metals were exhausted could the concentration of free O_2 begin to rise. It seems to have done so precipitously at about 2.2 BYA, in what has been referred to as the "great oxidation event." Even then, the level attained was no more than about 0.05 atm, about 1/40 of today's value. Why it occurred in this way is not known, but we should consider the possibility that the process was in a sense "autocatalytic." Recall that one result of a dioxygen-containing atmosphere is the formation of an ozone layer that can protect against UV radiation at the surface. An ozone layer, produced by oxygenation could have allowed photosynthesizers to begin to populate the very uppermost layers of the waters, or even the surface, where sunlight was most abundant. Indeed, there is evidence that some organisms, presumably photosynthetic, began to occupy land surfaces as early as 2.2 BYA (Gutzman and Benkes 1998). Occupation of such productive new niches could have resulted in a burst of dioxygen production. Then, a steady-state appeared to have been reached, with O_2 levels not increasing to present values until about 0.5 BYA (Fig. 2.6). The reasons for this long (~1 billion year) plateau and subsequent rapid increase at about 1.0–0.6 BYA remain unclear. A complex combination of oxygen sources and sinks has been proposed (Canfield 2005; Lambert and Donnelly 1991). But nevertheless it was a time in which further universal changes occurred in metabolism – the development of aerobic metabolism (see Chap. 4).

2.7.2 Effects on Life: An Ecological Catastrophe?

When dioxygen began to accumulate in the previously anoxic ocean, the anaerobic biota faced an unprecedented crisis. As explained in Chap. 1, dioxygen can give rise to a number of reactive oxygen species (ROS) that can be extremely toxic. These are toxins that the anaerobes had never faced in significant amounts, and against which they had little or no defense. Even today, the so-called "strict anaerobes" cannot tolerate even trace amounts of dioxygen. One can only conclude that there must have been mass extinctions, with those anaerobes which could not adapt surviving only in select anoxic niches, as they do today. Other organisms must have developed chemical defenses against ROS; these will be detailed in Chap. 3. As we shall see, the reaction to an oxygenated environment went much further than defense, with the evolution of a whole new, and more efficient metabolism. This

in turn, was probably essential for the evolution of eukaryotes like us. We owe everything to the lowly cyanobacteria!

2.7.3 Effects on the Earth

It was not only the biosphere that experienced major changes upon the rise of dioxygen. The chemistry of the oceans and atmosphere were radically altered. As noted above, oxidation states of a number of elements in sea water were changed. A thoughtful overall assessment has been provided by Frausto da Silva and Williams (2001) and is schematically indicated in Fig. 2.8. Especially important are the changes in the availability of iron (as mentioned above) and copper. Iron, in the ferrous state is very water-soluble and must have been readily available in the anoxic primitive ocean, but copper should not have been, especially in the presence of sulfides. Cuprous sulfide is so extremely insoluble that Cu^+ concentration has been estimated to have been as low as 10^{-17} M (Frausto da Silva and Williams 2001). Upon oxygenation of the oceans, the situation reversed: iron was precipitated as Fe_2O_3, and copper became solubilized as Cu^{2+}. It is of interest that strict anaerobes, having been evolved in the early Earth period, even today use no copper proteins, whereas these are common in aerobic organisms.
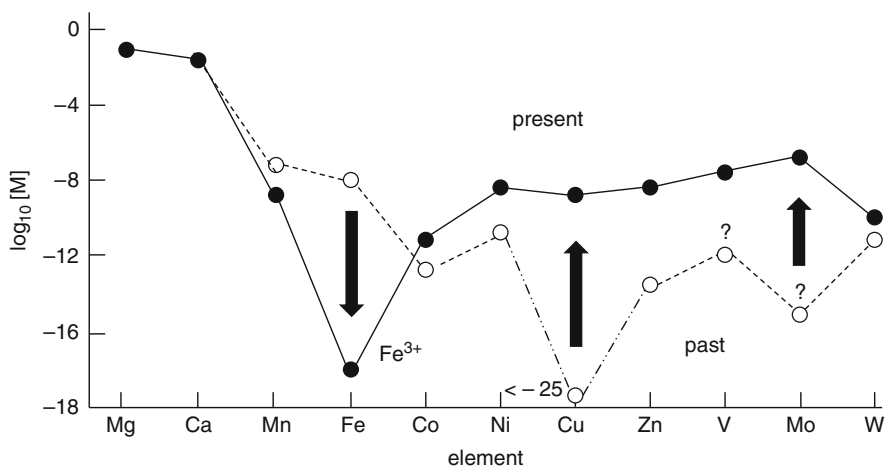
Fig. 2.8 Oxidation states of metals in the sea and water under anoxic and oxic condition logarithms of concentrations of free elements [M] in the primitive sulphide sea (*open circles*) and in the aerated sea today (*closed circles*). Note that very ancient dioxygen carriers used iron, whereas some of the younger ones employ copper. Iron was used in primitive organisms complexed in a heme group. Copper enzymes evolved after dioxygen was present. Molybdenum was not used early, while Tungsten occurs in Archea. Zinc became important later, in eukaryotes (adapted from Frausto da Silva and Williams 2001)

The first evidence for major, world-wide glaciations is found in the early Proterozoic, about 2.3 BYA (see Fig. 2.6). A strong case can be made that this was a consequence of atmospheric oxygenation (Kopp et al. 2005) although it seems to somewhat predate the great oxidation event. Recall that the early sun was considerably weaker than today, and that the warm climate of the Archaean was probably maintained by a greenhouse effect produced by carbon dioxide plus a smaller amount of methane (Owen et al. 1979). Oxygenation would have wiped out the methane rather quickly and accelerated photosynthesis could have significantly decreased the CO_2 level. Together, these may have reduced the greenhouse effect to a level that would permit glaciations. Fortunately for us all, there would have been a corrective feedback effect. Diminished photosynthesis, due to the covering of surfaces with ice, would decrease the CO_2 draw-down, whereas continued volcanism would have replenished the greenhouse gases. From this point in history forward, we shall see that the biosphere and geosphere are intimately interdependent. Indeed, two more periods of world-wide glaciations occurred between 0.8 and 0.6 BYA, again in synchrony with changes of the oxygen levels (see Fig. 2.6 and Chap. 6). These "snowball Earth" episodes are accompanied by evidence of much reduced biological activity, followed by bursts of evolution.

References

Allday J (1999) Quarks, leptons, and the big bang, 2nd edn. Institute of Physics, Bristol, UK

Canfield DE (2005) The early history of atmospheric oxygen. Ann Rev Earth Planet Sci 83:1–36

Chyba CF, Sagan C (1997) Comets as a source of prebiotic organic molecules for the Early Earth. In: Thomas PJ, Chyba CF, Mc Kay CP (eds) Comets and the origin and evolution of life. Springer, Berlin, p 147

Delsemme AH (1992) Cometary origin of carbon, nitrogen and water on the Earth. Orig Life Evol Biosph 21:279–298

Fedo VM, Whitehouse M, Kamber B (2006) Geological constrains on detecting the earliest life. Philos Trans R Soc Lond B Biol Sci 361:851–867

Frausto da Silva FIJR, Williams RSP (2001) The biological chemistry of the elements, 2nd edn. Oxford University Press, Oxford, UK

Gutzman I, Benkes NI (1998) Earliest laterites and possible evidence for terrestrial vegetation in the early Proterozoic. Geology 26:263–266

Heuseler H, Jaumann R, Neukum G (2000) Zwischen Sonne und Pluto. BLV, München

Holland AD (2006) The oxygenation of the atmosphere and oceans. Philos Trans R Soc Lond B Biol Sci 361:903–915

Hunten PH (1993) Atmospheric evolution of the terrestrial planets. Science 254:915–919

Joyce GF (1989) RNA evolution and the origins of life. Nature 338:217–244

Kasting JF, Howard MZ (2006) Atmospheric composition and climate on the Early Earth. Philos Trans R Soc Lond B Biol Sci 361:1733–1742

Kopp RE, Kirshvink IL, Hillburn IA, Cody Z (2005) The Paleoproterozoic snowball Earth: a climate disaster triggered by evolution of oxygenic photosystems. Proc Natl Acad Sci USA 102:11131–11136

Lambert IB, Donnelly TH (1991) Atmospheric oxygen levels in the Precambrian: a review of isotopic and geological evidence. Palaeogeogr Palaeoclimatol Palaeoecol 97:83–91

Martin A, Line MA (2002) The enigma of the origin of life and its timing. Microbiology 148:21–27

Mathews C, van Holde KE, Ahem K (2000) Biochemistry, 3rd edn. Addison-Wesley-Longman, Reading, MA

Müller J, Lesch H (2005) Vom Urknall zum roten Riesen – die Entstehung der. chemischen Elemente. Chem unserer Zeit 39:100–105

Münker C, Pfänder J, Weyer S, Büchl A, Kleine T, Mezger K (2003) Evolution of planetary cores and the Earth – Moon system from Nb/Ta systematics. Science 301:84–87

O'Neil J, Carlson RW, Francis D, Stevenson RK (2008) Neodymium-142 evidence for Hadean Mafic crust. Science 321:1828–1831

Orgel LE (2004) Prebiotic chemistry and the origin of the RNA world. Crit Rev Biochem Mol Biol 39:99–123

Owen T, Cess RD, Ramanathan V (1979) Enhanced CO_2 greenhouse to compensate for reduced solar luminosity on early Earth. Nature 277:640–642

Press F, Siefer R (1995) Allgemeine Geologie. Spektrum-Verlag, Heidelberg

Quenzel H (1987) Die Entwicklung der Erdatmosphäre. In: Wilhelm F (Hrsg) Gang der Evolution. Beck Verlag, München

Rauchfuss H (2005) Chemische Evolution und der Ursprung des Lebens. Springer, Berlin. ISBN 10-3-540-23965-0

Robert F (2001) The origin of water on Earth. Science 293:1056–1058

Rye R, Holland HD (1998) Paleosols and the evolution of atmospheric oxygen: a critical review. Am J Sci 298:621–672

Sagan C, Mullen G (1972) Earth and Mars: evolution of atmospheres and surface temperatures. Science 177:52–56

Schneider R, Ferrara A, Salvaterra R, Omukai K, Bromm V (2003) Low-mass relics of early star formation. Nature 422:869–871

Schopf JW (2006) Fossil evidence of archean life. Philos Trans R Soc Lond B Biol Sci 361:869–885

Seki K, Elphic RC, Hirahara M, Terasawa T, Mukai T (2001) On atmospheric loss of oxygen ions from Earth through magnetospheric processes. Science 291:1939–1941

Thomas-Keprta KL, Clemett SJ, Bazylinski DA, Kirschvink JL, McKay DS, Wentworth SJ, Vali H, Gibson EK, Jr RCS (2002) Magnetofossils from ancient Mars: a robust biosignature in the martian meteorite ALH84001. Appl Environ Microbiol 68:3663–3672

Tice MM, Lowe DR (2004) Photosynthetic microbial mats in the 3, 416-Myr-old ocean. Nature 431:549–552

Truran IW, Heger A (2004) Origin of the elements. In: Davis AM (volume ed) Treatise in geochemistry, vol I. Elsevier, Amsterdam

van Holde KE (1980) The origin of life: a thermodynamic critique. In: Halvorson HO, van Holde KE (eds) The origins of life and evolution. Alan Liss, New York

Wacey D (2009) Early life on Earth: a practical guide. Springer, New York

Weigert A, Wendker JH, Wisotzki L (1996) Astronomie und Astrophysik, 4 Aufl. VCH, Weinheim

Wetherhill G (1981) Spektrum der Wissenschaften August:107

Wills C, Bada J (2000) The spark of life. Perseus, Cambridge, MA

Yoshida N, Omukai K, Hernquist L (2008) Protostar formation in the early universe. Science 321:669–671

Chapter 3
Coping with Oxygen

Contents

3.1   The Impact of Oxygenation on an Anaerobic World

Sometime before 2.7 BYA, a new and biologically toxic substance began to appear in the environment. Biologically produced dioxygen, O_2, probably first began to accumulate in small pools or layers above cyanobacterial mats. These photosynthesizers must have already developed ways to at least partially deal with dioxygen and, with greater difficulty, the reactive oxygen species (ROS) derived from it (see Chap. 1 and below). But for primitive anaerobes in the vicinity, these new substances must have been especially toxic. Nevertheless, it is clear that they evolved ways to cope with the new threats. One way was to simply avoid dioxygen altogether. There is evidence for a long period in which the deep ocean remained anoxic even after dioxygen began to accumulate in surface layers and in the atmosphere (Kasting 1991). Thus, anaerobes were able to hide and "buy time" to evolve defenses. Some still use such refuges today, in deep oceans and lakes. The existence in these early ages of various chemical sinks (such as iron Fe^{2+}) for

oxygen meant that the overall concentration in the oceans did not rise to significant levels for hundreds of millions of years (Fig. 2.6). In that period, organisms developed many and varied ways to cope with ROS. These are the major topics of this chapter.

3.2 Production of Reactive Oxygen Species

In order to understand how organisms learned to fight the toxicity of dioxygen one has to know how ROS are created and what makes them so dangerous. Their basic chemistry has been discussed in Chap. 1, so we shall emphasize here the biological aspects. How does dioxygen give rise to ROS in vivo?

To reiterate a point from Chap. 1: dioxygen is paramagnetic with two unpaired electrons with parallel spin. This constitutes a barrier to many chemical reactions. However, O_2 can be reduced in the presence of paramagnetic transition metals such as iron and copper in a series of steps. The stepwise reduction of O_2 to H_2O produces several reactive species including hydrogen peroxide (H_2O_2), the super-oxide anion radical $^1O_2^{-*}$, and the hydroxyl radical OH^* (Fig. 3.1). Hydroxyl radicals are the most dangerous, for they react with almost every biomolecule. Results from reactions of OH^* include: amino acid modifications such as
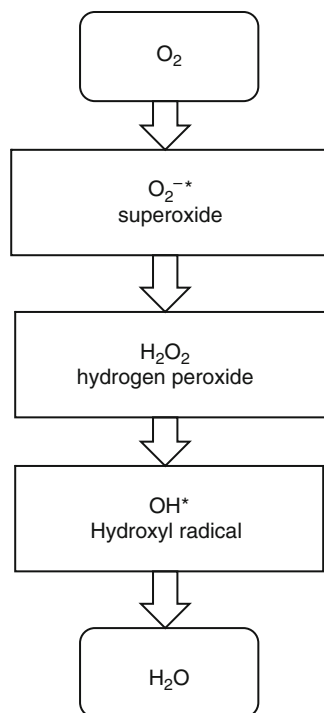
Fig. 3.1 ROS production accompanying the stepwise reduction of dioxygen to water. Note that an electron is added at each step

meta-tyrosine and ortho-tyrosine from phenylalanine, initiation of lipid peroxida-
tion, and oxidation of nucleic acids. In the earliest ages of O_2 production, ROS were
probably produced abiotically in the sea by a variety of adventitious metal-catalysed
reactions. We have noted in Chap. 2 that the anoxic primeval ocean must have been
rich in ferrous iron, an excellent catalyst. There are a number of metals in addition
to iron, including copper, chromium, vanadium and cobalt which are capable of
redox cycling in which a single electron may be accepted or donated by the metal.
This action catalyzes reactions that produce reactive radicals. Two important
reactions should be remembered from Chap. 1: the Fenton's reaction and the
Haber–Weiss reaction. In these reactions hydroxyl radicals are produced from
hydrogen peroxide in the presence of ferrous iron. All enzymes that either produce
ROS or deal with them contain one of these metals. On the other hand proteins have
been evolved which complex these metals, so as to control the charge transfer.
Without these protecting enzymes, organisms will encounter the problem known as
"oxidative stress" (see Chap. 8), which is an imbalance between oxidants and
antioxidants in favor of the oxidants. This may potentially lead to the damage of
biomolecules and therefore be dangerous for life.

In higher organisms, however, most of the free radicals are not obtained from
the environment, but are created in *mitochondria*; the power plants of cells where
dioxygen is reduced to produce energy (see Chap. 4). Within mitochondria, there
seems to be a competition between the useful metabolic process of transporting the
electrons along the respiratory chain or losing them to the immediate environment,
generating reactive compounds (Hamanaka and Chandel 2010; Scherz-Shouval and
Elazar 2010). Up to 2% of the electrons going through the respiratory chain end up
forming superoxides or other radicals. A critical point here is the lifetime of radicals
that are a part of the electron transport chain (see Table 1.1). Those that are slow in
transferring an electron to the next step in the chain stand a significant chance of
transferring it to form an ROS instead.

Another site of radical production is in *peroxisomes* which are organelles
specialized in the oxidation of fatty acids: During β-lipid degradation, electrons
are transferred from a cofactor $FADH_2$ to molecular oxygen. The superoxide
anions created react to form hydrogen peroxide via enzymes like superoxide
dismutase (SOD) (Fig. 3.1). Paradoxically, higher organisms seem to be most
susceptible to ROS. This is certainly true for humans (see Chap. 8). For example,
hydroxyl radicals have a major role in arterioscleroses by destroying lipoproteins
and the endothelial tissue (van Kuijk and Dratz 1987). But hydroxyl radicals can
also cause damage due to lipid peroxidation (Fig. 3.2) of membranes (Stadtman
1992). Interestingly, only unsaturated fatty acids undergo lipid peroxidation,
while saturated fats do not. On the other hand, it should also be noted that
dioxygen is needed to synthezise unsaturated fatty acids – commonly known as
the "good" fatty acids. Why is it so important to have unsaturated fatty acids?
Most likely they establish the particular properties of membranes – the fluidity
and permeability for several substrates in contrast to the stiff membranes made of
saturated fatty acids.
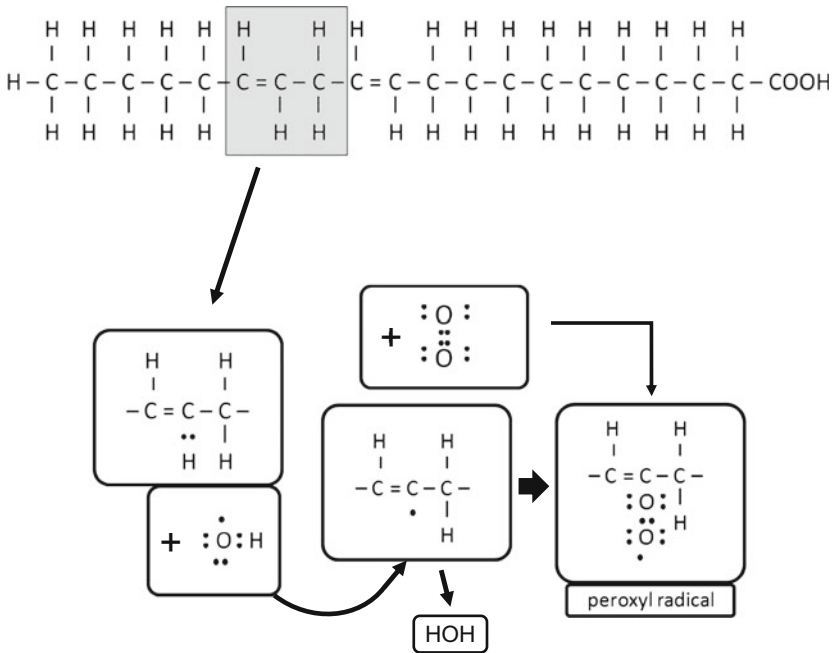
Fig. 3.2 An example for lipid peroxidation: oxygen radicals damage cellular membranes. Unsaturated fatty acids are common targets

Not only macromolecules but also cofactors such as the important metabolic reductant NADPH are vulnerable to ROS. For example, singlet oxygen reacts very rapidly with NADPH resulting in hydrogen peroxide and $NADP^+$ (Bodaness and Chan 1977).

We have introduced a few very important reactions whereby ROS are created and in which way they attack biological targets. There are several other pathways for producing ROS in smaller amounts: via xanthine oxidase, NADPH oxidases, cytochromes P450, and monoxygenase oxidases.

The creation of ROS is not limited to animals but also occurs in plants. During the transfer of high energy electrons within the photosynthetic process, some electrons can escape the chain to produce ROS (Krieger-Liszkay 2005). Here, a class of antioxidants found only in plants, a special class of carotenoids, was evolved to provide protection.

Despite their deleterious nature, ROS in physiological concentrations have become important for life. Through evolution, organisms have learned how to use them for intra- and intercellular signal transmission and regulation of cellular redox states. The remarkable fact is that during evolution higher organisms became dependent on ROS. ROS started to play important roles in cell signaling. Thus, to maintain proper cellular homeostasis, a balance must be held between ROS production and consumption (see Chap. 8). We shall also discuss these aspects later in this chapter.

The problems which anaerobes and aerobes have with ROS are fundamentally different. It seems that early anaerobic life on Earth had to cope with ROS created in the environment; either by inorganic photolysis or emerging photosynthezisers, whereas aerobes have had to deal primarily with the self-generated ROS. Utilizing dioxygen may provide a lot of energy, as we shall see in Chap. 4, but it also carries a considerable danger. Different strategies were evolved in the form of enzymes and chemical compounds to reduce this danger.

3.3 Coping with Reactive Oxygen Species

Most of the life on Earth with which we are familiar evolved in the presence of dioxygen, and had to adapt to this potentially dangerous substance. This was accomplished by evolving a large battery of antioxidant systems. Some of these systems are present in all life forms, from bacteria to mammals, indicating the appearance of at least traces of dioxygen early in the history of life. Indeed, photosynthetically generated oxygen may have appeared in local sites long before the first noticeable rise in atmospheric dioxygen about 2.5 BYA (Chap. 2). Even earlier, traces of dioxygen generated by inorganic photolysis of water must have been present. Here we will concentrate on a few antioxidation systems and their functions. There are two kinds of antioxidants – small molecules that scavenge oxygen by reacting with ROS and enzyme systems that detoxify them. There are also systems that combine small molecules and specific enzymes.

3.3.1 Scavenger Molecules

There exists a class of readily oxidized molecules that can scavenge ROS before they cause damage to the various biological molecules, or prevent oxidative damage from spreading, e.g. by interrupting the radical chain reaction of lipid peroxidation (Fig. 3.2). There are two major groups of such antioxidants. One type is soluble in water, and therefore found in cytoplasm and blood. The other group occurs in lipids and helps protect cellular membranes. An excellent review with over 150 references is provided by Wikipedia (http://en.wikipedia.org/wiki/Antioxidant).

Ascorbic acid, better known as vitamin C (Fig. 3.3), is a monosaccharide occurring in almost all organisms. A deficiency causes a serious disease called scurvy. Unfortunately we humans have to obtain vitamin C from food since we cannot synthesize it. Vitamin C is the major water-soluble antioxidant. It is very effective in scavenging a wide array of ROS and free radicals. It seems to protect low density lipoproteins (LDL) against oxidation even more effectively than vitamin E (see Fig. 3.3 and below). In fact, it recycles vitamin E radicals and vice versa (Fig. 3.3). In the cell there is also an interplay with glutathione (GSH) (see
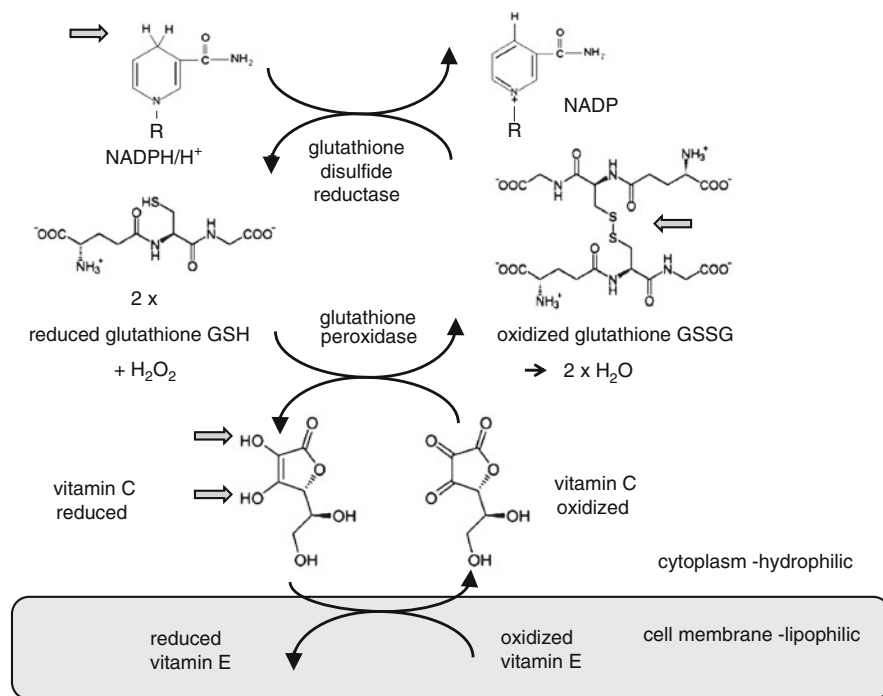
Fig. 3.3 Interplay of vitamin E, vitamin C, glutathione and NADP. The *arrows* point to the important sites of reactions. Note that H_2O_2 (or organic peroxides) can oxidize glutathione resulting in the glutathione redox cycle

below), which reduces ascorbic acid (Ramsey and Sharpless 2006; Wells et al. 1990). Ascorbic acid is also important in plants and occurs in high concentrations. It is a substrate for the antioxidant enzyme ascorbate peroxidise, thereby being important for oxidative stress resistance (Shigeoka et al. 2002). Vitamin C became very famous through the work of Linus Pauling, who suggested consuming several grams per day to delay aging. In which way and to what degree vitamin C is really involved in such protection remains under discussion.

Alpha-Tocopherol or vitamin E is the most prominent among several isoforms of tocopherol. Although it occurs in most organisms, in humans it is by far the most abundant lipid-soluble antioxidant, being present in most membranes. It protects cell membranes against oxidation by reacting with lipid radicals in the lipid peroxidation reaction (Herrera and Barbas 2001). Vitamin E taken each day is thought to substantially lower the risk of heart disease. It seems that vitamin E protects LDL against oxidation, which is a critical step in the development of atherosclerosis. The oxidized form of vitamin E can be recycled to the reduced form by oxidation of ascorbate in the cytoplasma (Fig. 3.3) or retinol or ubiquinol in the membrane (Wang and Quinn 1999). It is still uncertain as to which other protective functions vitamin E may be involved in, such as neuroprotection (Sen et al. 2006).

Beta-carotenoids are typically found in the plant membranes, where they serve for the protection of the photosynthetic apparatus against ROS created by the high energy electrons activated by the captured light. They include lipid-soluble beta-carotene (a vitamin A precursor) and related substances called carotenoids, and alpha-carotene or lycopene (the red color in tomatoes), lutein, and zeaxanthine. It seems that carotenoids are much weaker antioxidants than vitamin E. Thus, whether they are also effective in humans is still a matter of controversy.

Resveratrol. Natural antioxidants are very popular in contemporary society. Among these is a polyphenol named resveratrol. Its presence, especially in red wine, has led to a recommendation to justify drinking wine. Whether, in fact, resveratrol is actually acting as an antioxidant in the human body is a matter of discussion (Sardi 2004). Plants evolved a very effective system using resveratrol to counter oxidative damage especially in parts which are exposed to air and sun light. It appears also to serve as a protection against infections by fungi or parasites. Resveratrol is especially found in high concentrations in the skin of berries. As a consequence, it occurs in white grape drinks up to 200 μg/l and in red grape drinks up to about 1,000 μg/l: However, much higher concentrations are found in wine, being concentrated to up to 50 mg/l during the wine making process. How much resveratrol has to be taken to produce observed effects? It was shown that it causes a longer life in animal experiments, but extrapolating the data to human suggest that one would have to drink more than 30 glasses of red wine per day to have the same concentration as in the study. Obviously this would not be good for the liver. In addition, the action of resveratrol in humans is not really understood despite current intense investigation.

Many other antioxidants could be mentioned, but the principal is clear; each reacts with either dioxygen or ROS. Another way to deal with such threats is via dedicated enzyme systems. Enzymes have the advantage over scavenger molecules that they are not consumed as they deactivate ROS. One enzyme molecule can be used as a catalyst over and over, but a scavenger such as ascorbic acid is either lost or must be regenerated after reaction.

3.3.2 Enzymes for Detoxification of ROS

There are a number of enzymes distributed widely over the whole range of organisms that appear to exist mainly to detoxify ROS. Here, we shall use as an example superoxide dismutase (SOD). A wider view can be found in the book by Lane (2002). SODs are found in all aerobic cells and some extracellular fluids (Fridovich 1995, 1997). They catalyze the conversion of two molecules of superoxide into hydrogen peroxide and dioxygen (see Chap. 1). The benefits here are that H_2O_2 can in turn be broken down into water and dioxygen by another efficient enzyme (catalase) and even if not, hydrogen peroxide will not cause immediately damage since it is substantially less toxic than superoxide. SOD accelerates the detoxifying reaction roughly 10,000-fold over the non-catalyzed reaction. It was recently

Table 3.1 Superoxide dismutases

Type	Organisms	Localization
FeSOD	Prokaryote	Cytosol
	Few plants	Mitochondria
MnSOD	Prokaryote	Cytosol
	Eukaryote	Mitochondria
NiSOD	Prokaryote	Streptomyces
Cu,ZnSOD	Eukaryotes (some bacteria)	Cytosol
Cu,ZnSOD	Eukaryotes	Extracellular

There are three groups of enzymes that exhibit significant sequence/structure similarities, and are presumed to have a common origin

discovered that the gene for SOD is defective in patients with amyotrophic lateral sclerosis (ALS), better known as Lou Gehrig's disease. The patients suffer muscle atrophy and paralysis.

As shown in Table 3.1, there are several kinds of SODs. All are metal-containing enzymes with bound iron, manganese, copper or zinc donating the essential anti-oxidant activity. SODs must have been evolved early since they are present in all aerobic organisms (Fridovich. 1981). SODs are inducible enzymes - exposure of bacteria or vertebrate cells to higher concentrations of dioxygen results in rapid increases in the concentration of SOD (see Chap. 8). The different forms of SOD fall into several seemingly independent classes. The Fe-SOD and Mn-SOD are found mainly in prokaryotes or in eukaryotic mitochondria. They are structurally quite similar. The Cu/Zn-SODs are primarily eukaryotic proteins, one class of which is cytosolic, the other is extracellular. They appear to be unrelated to the Fe- or Mn-SOD's classes. Mn-SOD occurring in mitochondria of eukaryotic cells presumably has been transferred via endosymbiosis from an aerobic bacterium. The Cu/Zn-SOD has been detected in the cytoplasma of eukaryotic cells and some gram-negative bacteria (Fridovich 1997). It has to be distinguished from the Cu/Zn-SOD occurring in the extracellular fluids. It seems that most aerobic bacteria do not have this SOD, perhaps because Fe-SOD and Mn-SOD were adequate for the same purpose. To further complicate the matter, nickel-containing SOD has been found in prokayotes such as the bacterium *Streptomyces*. Both, Fe-SOD and Mn-SOD (as well as Ni-SOD), are thought to be evolved before aerobic photosynthesis was present. Although this collection sounds confusing, its occurance can be understood by evolution. First of all, it implies that some very low levels of oxidants were present even in the primitive environment, perhaps from UV photolysis of water. Recall from Chap. 2 that iron and manganese were prevalent in the ancient anaerobic ocean, while copper and zinc became abundant only after the great oxidation event. This may explain why FeSOD and MnSOD are found in the more ancient prokaryotes and mitochondria that derived from them, while Cu/Zn-SOD occurs in higher bacteria and in the younger cytosols of eukaryotes.

Another prominent enzyme, *catalase*, is needed to finish the work of SOD. It is found in the peroxisomes of eukaryotic cells when it degrades hydrogen peroxide to water and dioxygen. Its importance can be deduced from a special property

evolved: catalase belongs to that class of enzymes with the most efficient catalytic activity. It should be noted that strictly anaerobic methane bacteria – still present on Earth – evolved catalase to survive oxidative conditions.

3.3.3 Antioxidant Enzyme Systems

As we already have seen, some antioxidants do not work alone. Often cascades or cycles are involved consisting of several partners such as an enzyme, cofactor and vitamin. The most prominent of such is the GSH system (Fig. 3.3), which consists of several steps but in summary catalyses the following reaction which essentially detoxifies H_2O_2:

$$NADPH + H^+ + H_2O_2 \rightarrow 2\ H_2O + NADP^+$$
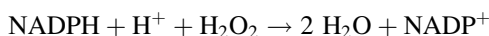
This system consist of GSH, glutathione-disulfide-reductase and glutathione peroxidise (Hashida et al. 2000).

Glutathione (GSH) (Fig. 3.3) is present in almost all cells and may well be the most important intracellular defense molecule against damage by ROS. It occurs in all known aerobic organisms. It is a peptide of three amino acids: glutamyl-cysteinyl-glycine. The cysteine provides an exposed free thiol group (SH) which can transfer two electrons to ROS neutralizing them. Thus, this very reactive thiol group provides an excellent target for radical attack. Reaction with radicals oxidizes two GSH molecules to form a disulfide-linked dimer, denoted as GSSG (see Fig. 3.3). But the reduced forms GSH are then regenerated from GSSG in a redox cycle involving glutathione-disulfide-reductase (named GSSG-reductase; Forni and Willson 1986) consuming NADPH as electron donors. For the regeneration another cofactor is needed: $FAD/FADH_2$, a part of the glutathione-disulfide-reductase.

$$GSSG + FADH_2 \leftrightarrow 2\ GSH + FAD^+$$
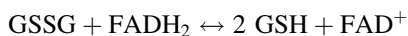
$FADH_2$ will be regenerated from FAD^+ by $NADPH/H^+$

$$FAD^+ + NADPH + H^+ \rightarrow FADH_2 + NADP^+$$
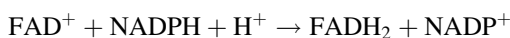
It seems likely that GSH plays a major role in maintaining the cell's redox state, since high concentrations (about 5 μM) is found in cells due to the presence of active glutathione-disulfide-reductase (Fig. 3.3). The ratio of GSH/GSSG in cells is higher than 10:1 which seems to be enough to neutralize the dangers from ROS.

A confirmation for the importance of the gluthation system was recently shown. Programmed cell death (apoptosis) was induced when GSH was withdrawn from cells (Seiler et al. 2008).

Apoptosis can be induced by lipid peroxides which activate the apoptosis activating factor (AIF). It seems that with increasing oxidative stress more and

more GSH and glutathione peroxidases are used to fight various oxidation processes and therefore are now not available for reducing lipid peroxidases. Thus, suicide of a cell seems to be more advantageous than allowing a conversion to a tumor cell. Interestingly, additional vitamin E seems to inhibit the apoptosis.

Glutathione peroxidises as a part of the GSH system posses a selenocysteine at the active site in which selenium replaces the usual sulphur. Like catalase these enzymes catalyse the degradation of hydrogen peroxide: they also reduce organic peroxides to alcohols, providing another route for eliminating toxic oxidants (Fig. 3.3).

Thioredoxin, a small oxidoreductase ($M_r = 12$ kDa) exists in many isoforms, and is found in all known organisms. The key to its function lies in a conserved motif of a short peptide with four amino acids, cysteine–glycine–proline–cysteine. It can cycle between a reduced state with two –SH groups provided by the cysteines which can again neutralize ROS by transferring electrons to yield an oxidized form carrying a disulfide bridge. The reduced form is then recovered after reduction by an enzyme *thioredoxin reductase* using NADPH as an electron donor.

Aging is thought to be connected with the accumulation of oxidized proteins (Stadtman 1992). The major candidates for oxidation within proteins are the sulphur containing amino acids, cysteine and methionine (see Chap. 8). In eukaryotic cells proteasomes are mostly responsible for degradation of the oxidized proteins in the cytosol and nucleus. In addition, specific enzyme systems (the glutaredoxin/glutathione/reductase system and the thioredoxin/reductase system) evolved to reverse the oxidation that has formed disulphide bridges. Also important are methionine sulfoxide reductases which are able to reduce methionine sulfoxide back to methionine within the proteins. This may fail when the concentration of hydrogen peroxide becomes so high that even methionine sulfoxide reductase as well as catalase will also be oxidized. A consequence of this, graying hairs of elderly people, will be discussed in Chap. 8.

When were these defense mechanisms evolved? Because some of these mechanisms are found in all organisms, these must have arisen before the potential for aerobic respiration (see Chap. 4) had appeared. But the early existence of aerobic photosynthesis together with the evolution of the key enzymes like cytochrome oxidase when there was not significant O_2 in the environment suggests an alternative evolution (Castresana and Saraste 1995). It is not generally accepted, however, that aerobic metabolism itself evolved at such an early date. These early defenses against adventitious dioxygen could then have been adapted by the first creatures living in truly aerobic environments.

3.4 How to Avoid Reactive Oxygen Species?

The best strategy to avoid ROS is to avoid their production. This has meant different things at different periods in evolution. At the earliest times, before dioxygen was a significant component in the environment, the major source would have been photolysis of water by high-energy UV radiation. An easy way for organisms to

avoid high energy radiation is not to live at the surface of the ocean but in a depth of at least several meters. On land the best strategy is to find shadowed places or/and to use the day and night cycle. This problem disappeared by the time (see Fig. 2.6) the ozone layers formed. But by then, there was of course, significant photosynthetically produced dioxygen, in the oceans and atmosphere.

Alternatively membrane-bound enzymes like cytochrome oxidase, coupled with an NADH, might have reduced O_2 right at the cell surface. Cytochrome oxidase and cytochrom c appear to be ancient, but how old we do not know. During the earliest periods of rising dioxygen (say 2.9–2.5 BYA) avoidance of dioxygen by anaerobes would probably have been easy. Dioxygen was being generated by photosynthesizers only in restricted environments – mainly shallow ocean fringes and perhaps terrestrial ponds. Vast regions on Earth, especially in deep oceans must have remained anoxic, if for no other reason than that oxygen diffuseing into them from the above-mentioned sources would have been instantly reduced by the abundant reducing agents present (Fe^{2+}, H_2S, etc).

Even today, there exist extensive environments such as some deep ocean region that are essentially anoxic and harbor mainly anaerobic microbes. Such must have persisted throughout geological history. It seems likely that ancient anaerobes developed oxygen-sensing chemotaxis to avoid oxygen-rich regions (see below). Later, facultative aerobes may have (and still do) used such sensors to seek oxygen (see Sect. 3.6).

3.5   Evolving Defense Strategies

As worldwide O_2 levels inexorably rose, stronger defences were required. One, of course, was the evolution of antioxidants and enzyme systems to destroy ROS, as described above. It is very difficult to date antioxidants, but enzymes such as Fe-SOD and Mn-SOD appear to be very ancient. However, it is becoming clear that a wide variety of alternative defence strategies have evolved over the eons.

3.5.1   Aggregation for Defense

Experiments have suggested another important way of avoiding ROS. Ma and Eaton (1992) were puzzled by the fact that the high diffusivity of H_2O_2 should mean that bacteria like *Escherichia coli* should not be able to protect themselves against H_2O_2, even though they contained high concentrations of catalase. They discovered that dilute suspensions of *E. coli* were, in fact vulnerable. However, in concentrated cultures where bacteria clumped together, there was surprising resistance. Apparently, the "surface" members in the aggregate of cells can protect the "interior" members from ROS. This raises the remarkable suggestion that protection against ROS might have been one of the evolutionary driving forces towards

the development of multi-cellular eukaryotes! If so, oxygen has played an even larger part in our evolution than hitherto suspected.

3.5.2 Melanin

Another way to avoid the problem of ROS is to metabolize dioxygen before it can form radicals. On Earth, phenols were already available in great amounts. Using dioxygen, monophenols could easily be hydroxylated to diphenols such as DOPA or dopamine, by enzymes such as phenoloxidase (tyrosinase, catecholoxidases, laccase, and other oxidases) (Fig. 3.4). Dioxygen is split, with one oxygen atom being attached to the benzoic ring at various positions. Tyrosinase for example attaches an oxygen atom in the intimate neighborhood (*o*-position) of the first bound oxygen (Matoba et al. 2006; Decker et al. 2006). The other oxygen atom is incorporated into water. After oxidation the result will be a reactive quinone, a benzol ring with two aggressive keto groups. Because of their severe reactivity with, for example pathogenic invaders, they are considered as a part of the innate immune system (Bogdan 2007; Jiang et al. 2007). On the other hand, quinones can spontaneously polymerize to form polymeric nets called melanin. The color of melanin depends on the incorporation of a cysteine derivate which results in reddish pheomelanin, alternatively dark eumelanin is formed, when it is not incorporated (Fig. 3.4).
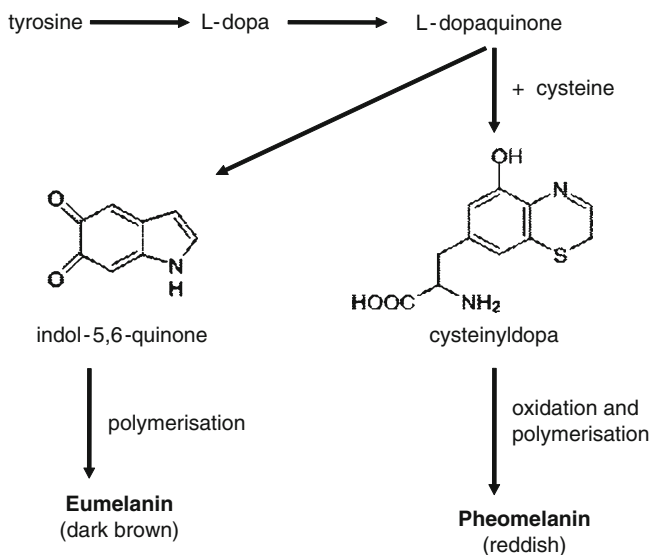
Fig. 3.4 Scheme of melanin synthesis. The conversion of tyrosine to L-dopaquinone is catalysed by the enzyme tyrosinase. Eumelanin colors the hair brown to black, while incorporation of cysteine results in a more reddish pheomelanin

Melanin is also needed for the sclerotization of arthropods, forming a resistible cuticule after molting. In wound healing in arthropods melanin would close open breaks in the exoskeleton. Should bacteria or other pathogens invade such organisms melanin will encapsulate them as a part of the primary immune response (Claus and Decker 2006). These phenomena are also found in fungi, plants and fruits. When an apple is attacked by a worm, melanin nets are formed to guide the worm out of the apple, bypassing the core. Thus, the apple seeds are protected.

Melanin has been conserved during evolution. In humans, melanin occurs in the retina of the eye, colors the hair and protects the skin from energy rich UV light (Oetting et al. 2003; Plonka et al. 2009; Wood et al. 2009). In the retina, melanin protects this area containing very high oxygen concentration from the potential formation of oxygen radicals. Mutation of the conserved environment of the active site would eliminate the function of tyrosinase to form melanin resulting in albinism (Schweikardt et al. 2007). Once formed, melanin can still neutralize ROS as a strong antioxidant due to its large amounts of double bonds. Melanin is an example of a self-stabilizing co-evolution between different strategies. Here metabolizing oxygen may either result in the synthesis of antioxidants or in shields against radiation.

Another enzyme belonging to this protein family is laccase, which in plants facilitates the binding of oxygen to the other free position of monophenols (Claus 2008, 2009). The result will also be a polymer net such as lignin, a major constituent of woody tissues. It appears that due to resulting advantages, phenol oxidation appeared early in evolution and spread to every kingdom. In all cases metal ions had to have been involved, in most cases these were iron or copper.

3.5.3 Oxygen Transport Proteins Prevent Creation of Oxygen Radicals

A more physiological strategy in animals is to keep the dioxygen concentration low in and around cells. Normally very small organisms take up dioxygen by diffusion. In animals with dimensions above about 1 cm, a circulation system with oxygen transport proteins (OTP) is needed (see Chap. 5). The functions of OTP have become optimized to deliver dioxygen, but not provide too much. In other words, only as much dioxygen will be released from OTP to the blood to diffuse into the cell as the cell needs.

It is clear that the OTPs themselves are fashioned so as to allow the transport of considerable amounts of dioxygen without the creation of ROS. The most common OTP's are hemoglobins and hemocyanins (see Chap. 5). How do they prevent dioxygen from forming radicals? Hemoglobins carry heme groups within a protein matrix which is almost open to the surrounding. In its center there is an iron atom which binds onto one side the dioxygen molecule in a well protected cavity. A very short channel protects the iron and its bound dioxygen against the influence of the environment. Thus, even though erythrocytes carry much dioxygen, it is not free to react and form ROS. In the case of hemocyanin, the active site which involves a pair

of copper atoms is protected by the protein matrix as well as via a special folding motif of the protein (see Chap. 5).

Despite the protective function of the OTP, oxidation at the binding site sometimes occurs, forming the so-called "met" form of the protein. In the red blood cells of animals, which maintain a high concentration of dioxygen, a very active reduction system (met-hemoglobin reductase) is present to keep the concentration of met-form low.

3.6 Reactive Oxygen Species as Cellular Signals

So far, we have presented ROS as substances which, over life's history, had to be protected against. But evolution has ways of even turning unfavorable aspects to advantage. A number of ROS have come to play a signaling role in various biological systems in both plants and animals (Hancock et al. 2001; Finkel and Holbrook 2000). For example, ROS are involved in inducing programmed cell death (apoptosis) or necrosis. They also play roles in regulating the expression of genes and activation of cell signaling cascades. We shall center on H_2O_2 as an example of a widely used signal. In cells where it is needed, it is synthesized via enzymatic oxidation of NADPH in two steps

$$2\ ^1O_2^{-*} + NADPH \rightarrow 2\ ^1O_2^{-*} + NADP^+ + H^+$$
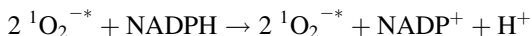
The crucial enzyme here is NADPH oxidase. As we have learned previously superoxide can be reduced to hydrogen peroxide by SOD in the second step:

$$2\ ^1O_2^{-*} + 2\,H^+ \rightarrow H_2O_2 + O_2$$
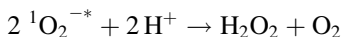
H_2O_2 has two advantages as a signal molecule. First, it is relatively stable (Table 1.1). Second, it is readily diffusible in cells and tissues and through membranes as well. Thus, it can act as a messenger to modulate both intracellular and intercellular processes. In many cases, this is through the regulation of gene transcription. Gene regulations proceed through the action of transcription factors, proteins which interact with the nuclear DNA. In some still unknown way, H_2O_2 appear to activate certain transcription factors, or favor their transport into the nucleus. More detail on the physiological significance of ROS signaling will be presented in Chap. 8.

3.7 Dioxygen as a Signal: Oxygen Sensor

The varying responses and needs of organisms for dioxygen have led to the evolution of molecules that act as dioxygen sensors. For strict anaerobes, including Archean organisms, a sensor for avoidance of dioxygen (a chemiotactical sensor) or for generation of antioxidants is most appropriate. For facultative aerobes or

aerobes, sensors can act positively by helping the organism to seek dioxygen. This may have been an important function in the early years of aerobic metabolism, when dioxygen was scarce and unevenly distributed. Those aerobes that could sense regions of high O_2 concentrations would have a metabolic advantage. There are even some sensors that act to concentrate microbes in regions of optimum dioxygen concentrations - neither too low nor too high (Szurmant and Ordal 2004).

Most dioxygen sensors are membrane proteins comprised of two portions – a dioxygen-binding domain which is often (but not always) of the hemoglobin type, coupled to an effector domain, often an enzyme such as a kinase which can start a phosphorylation cascade. It has even been suggested that all present-day hemoglobins evolved from such molecules (Vinogradov et al. 2007; see Chap. 5). Binding of dioxygen to the heme in the first domain causes a conformational switch; this is transmitted to the enzymatic domain, activating or inactivating it, frequently initiating a cascade of signals that modify cell motility.

Another way was evolved by the bacterium *E. coli*. Under O_2-limiting conditions, a fumarate and nitrate reduction transcriptional regulator, FNR, contains a $[4Fe–4S]^{2+}$ cluster, generating a dimeric protein controlling the activation of the appropriate genes. Exposure to O_2 results in conversion of the cluster into a $[2Fe–2S]^{2+}$ form leading to dissociation of the active dimer into inactive monomers. Thus, bacteria evolved the ability to sense environmental O_2 for protection and choosing the appropriate respiration, anaerobe or aerobe (Unden et al. 1995; Crack et al. 2008; Green et al. 2009).

Oxygen sensing in higher organism often serves very different functions. Frequently, the necessity is to maintain optimal intracellular oxygen levels. For example, low levels of blood oxygen (hypoxia) in humans trigger a complex response in which specialized kidney cells are stimulated to produce erythropoietin. This travels to bone marrow cells and stimulates the release of more cells carrying hemoglobin, countering the hypoxia (see, for example Semenza 2008; see Chap. 8).

3.8 Summary: Reactive Oxygen Species and Life

Summarizing the major ideas in this chapter, we find that important for proper understanding of the role of oxygen in life is the fact that ROS can be both dangerous and useful substances. Nature seems not to have completely rejected these potentially toxic molecules since dioxygen has become abundant on Earth. Rather, ways were evolved to control them and to even make them essential for life. The consequence today is that ROS can play various, even opposing roles during different cellular processes. However, the benefits and the harms potential in ROS molecules dictate a careful balance between their production and controlling them by antioxidants and specific enzymes to maintain almost homeostatic conditions. Otherwise oxidative stress would result in a severe threat to the cells of even multicellular organisms (see Chap. 8).

The problem of dealing with ROS is fundamentally different in anaerobes and aerobes. In the former, avoidance or detoxification is the major strategy. Aerobes, however, have a built-in source of ROS – the oxidative metabolism to be discussed in the following chapter. To cope, they must regulate ROS. For a long time ROS were considered only to be highly toxic by-products of that metabolism that are not useful to the organism (see above). Scandalios (2002) has summarized the discussion during recent years between those who consider ROS to be negative "by-product" molecules and those who favor the idea that ROS became useful molecules for higher organisms. Furthermore, as we shall see in the next chapter the advantage of aerobic metabolism is so great that it overrides the disadvantage of ROS production. This means that during evolution the overall benefits of ROS creation outweighed the negative aspects.

References

Bodaness RS, Chan PC (1977) Singlet oxygen as a mediator in the hematoporphyrin-catalyzed photooxidation of NADPH to NADP+ in deuterium oxide. J Biol Chem 252:8554–8560

Bogdan C (2007) Oxidative burst without phagocytes: the role of respiratory proteins. Nat Immunol 8:1029–31

Castresana J, Saraste M (1995) http://www.ncbi.nlm.nih.gov/pubmed/8578586?ordinalpos=2&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSumEvolution of energetic metabolism: the respiration-early hypothesis. Trends Biochem Sci 20:443–448

Claus H (2008) Laccases and their occurrence in prokaryotes. Arch Microbiol 179:145–150

Claus H (2009) Laccase, Wiley encyclopedia of industrial biotechnology. Wiley, New York

Claus H, Decker H (2006) Bacterial Tyrosinases. Systematic and Applied Microbiology 29:3–14

Crack JC, Jervis AJ, Gaskell AA, White GF, Green J, Thomson AJ, Le Brun NE (2008) Signal perception by FNR: the role of the iron-sulfur cluster. Biochem Soc Trans 36:1144–1148

Decker H, Schweikardt T, Tuczek F (2006) The first crystal structure of tyrosinase: all questions answered? Angew Chem Int Ed Engl 45:4546–4550

Forni LG, Willson RL (1986) Thiyl and phenoxyl free radicals and NADH. Direct observation of one-electron oxidation. Biochem J 240:897–903

Fridovich I (1981) The biology of superoxide and of superoxide dismutases – in brief. Prog Clin Biol Res 51:153–172

Fridovich I (1995) Superoxide radical and superoxide dismutases. Annu Rev Biochem 64:97–112

Fridovich I (1997) Superoxide anion radical (O_2^{*-}) superoxide dismutases and related matter. J Biol Chem 272:18515–18517

Finkel T, Holbrook NJ (2000) Oxidants, oxidative stress and the biology of ageing. Nature 408:239–247

Green J, Crack JC, Thomson AJ, LeBrun NE (2009) Bacterial sensors of oxygen. Curr Opin Microbiol 12:145–151

Hancock JT, Desikan R, Neill SJ (2001) Role of reactive oxygen species in cell signalling pathways. Biochem Soc Trans Part 29(2):345–350

Hashida K, Sasaki K, Makino N (2000) Interactions of nitric oxide and oxygen in cytotoxicity: proliferation and antioxidant enzyme activities of endothelial cells in culture. Free Radic Res 33:147–156

Hamanaka RB, Chandel NS (2010) Mitochondrial reactive oxygen species regulate cellular signaling and dictate biological outcomes. Trends Biochem Sci 35:505–513

Herrera E, Barbas C (2001) Vitamin E: action, metabolism and perspectives. J Physiol Biochem 57:43–56

Jiang N, Tan NS, Ho B, Ding JL (2007) Respiratory protein-generated reactive oxygen species as an antimicrobial strategy. Nat Immunol 8:1114–1122

Kasting JF (1991) Box models for the evolution of atmospheric oxygen: an update. Glob Planet Change 97:125–131

Krieger-Liszkay A (2005) Singlet oxygen production in photosynthesis. J Exp Bot 56:337–346

Lane N (2002) Oxygen: the molecule that made the world. Oxford University Press, New York. ISBN 0-19-860783-0

Ma M, Eaton IW (1992) Multicellular oxidant defense in unicellular organisms. Proc Natl Acad Sci USA 89:7924–7928

Matoba Y, Kumagai T, Yamamoto A, Yoshitsu H, Sugiyama M (2006) Crystallographic evidence that the dinuclear copper center of tyrosinase is flexible during catalysis. J Biol Chem 281:8981–8990

Oetting WS, Fryer JP, Shriram S, King RA (2003) Oculocutaneous albinism type 1: the last 100 years. Pigment Cell Res 16:307–311

Plonka PM, Passeron T, Brenner M, Tobin DJ, Shibahara S, Thomas A, Slominski A, Kadekaro AL, Hershkovitz D, Peters E, Nordlund JJ, Abdel-Malek Z, Takeda K, Paus R, Ortonne JP, Hearing VJ, Schallreuter KU (2009) What are melanocytes really doing all day long...? Exp Dermatol 18:799–819

Ramsey MR, Sharpless NE (2006) ROS as a tumour suppressor? Nat Cell Biol 8:1213–1215

Sardi B (2004) The anti-aging-pill, is the 125-year life span about to become common? 1st edn. Here and Now Books, San Dimas, CA. ISBN 097742720X / 0-9774272-0-X

Scandalios JG (2002) The rise of ROS. Trends Biochem Sci 27:483–486

Scherz-Shouval R, Elazar Z (2010) Regulation of autophagy by ROS: physiology and pathology. Trends Biochem Sci. 2010 Aug 19. [Epub ahead of print]

Schweikardt T, Olivares C, Solano F, Jaenicke E, García-Borrón JC, Decker H (2007) A three-dimensional model of mammalian tyrosinase active site accounting for loss of function mutations. Pigment Cell Res 20:394–401

Seiler A, Schneider M, Förster H, Roth S, Wirth EK, Culmsee C, Plesnila N, Kremmer E, Rådmark O, Wurst W, Bornkamm GW, Schweizer U, Conrad M (2008) Glutathione peroxidase 4 senses and translates oxidative stress into 12/15-lipoxygenase dependent- and AIF-mediated cell death. Cell Metab 8:237–248

Semenza GL (2008) Tumor metabolism: cancer cells give and take lactate. J Clin Invest 118:3835–3837

Sen CK, Khanna S, Roy S (2006) Tocotrienols: vitamin E beyond tocopherols. Life Sci 78:2088–2098

Shigeoka S, Ishikawa T, Tamoi M, Miyagawa Y, Takeda T, Yabuta Y, Yoshimura K (2002) Regulation and function of ascorbate peroxidase isoenzymes. J Exp Bot 53:1305–1319

Stadtman E (1992) Protein oxidation and aging. Science 257:1220–1224

Szurmant H, Ordal GW (2004) Diversity in chemotaxis mechanisms among the bacteria and Archaea. Microbiol Mol Biol Rev 68:301–319

Unden G, Becker S, Bongaerts J, Holighaus G, Schirawski J, Six S (1995) O_2-sensing and O_2-dependent gene regulation in facultatively anaerobic bacteria. Arch Microbiol 164:81–90

van Kuijk FJ, Dratz EA (1987) Detection of phospholipid peroxides in biological samples. Free Radic Biol Med 3:349–354

Vinogradov SN, Hoogewijs D, Bailly X, Mizuguchi K, Dewilde S, Moens L, Vanfleteren JR (2007) A model of globin evolution. Gene 398:132–142

Wang X, Quinn PJ (1999) Vitamin E and its function in membranes. Prog Lipid Res 38:309–336

Wells WW, Xu DP, Yang YF, Rocque PA (1990) Mammalian thioltransferase (glutaredoxin) and protein disulfide isomerase have dehydroascorbate reductase activity. J Biol Chem 265:15361–15364

Wood JM, Decker H, Hartmann X, Chavan B, Rokos H, Spencer JD, Hasse S, Thornton J, Shalbaf M, Paus R, Schallreuter KU (2009) Senile hair greying: H_2O_2-mediated oxidative stress affects human hair colour by blunting methionine sulfoxide repair. FASEB J 23:2065–2075

Chapter 4
Aerobic Metabolism: Benefits from an Oxygenated World

Contents

4.1   The Advantage to Being Aerobic

In the preceding chapter, we have emphasized the dangers that the advent of dioxygen presented to the existing anaerobic organisms, and the ways they evolved to deal with the problems. However, this is only part of the story and were it to have ended here, we and the world we know would not exist. What happened instead was quite remarkable; for life seized upon an opportunity presented by the presence of free dioxygen to become many-fold more efficient in extracting energy from foodstuffs. As we shall see, this aerobic, oxidative metabolism opened in turn a multitude of new opportunities for growth and diversification.

Skipping over a lot of complicated biochemistry, we can explain the advantage of oxidative metabolism over anaerobic metabolism in a very simple way. What counts in metabolic energy generation is how much free energy you can extract by processing a single foodstuff-molecule. The effective energy production is more directly measured in terms of how many molecules of ADP are converted to ATP per nutrient molecule, because ATP is the universal energy currency, used to drive

every kind of cellular process. Now consider an anaerobic fermentation process using glucose ($C_6H_{12}O_6$) as a foodstuff (see Fig. 2.5). A typical overall reaction can be written as:

glucose + 2 ADP + 2 phosphate → 2 lactic acid + 2 ATP
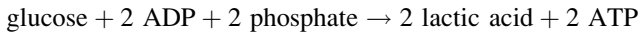
In brief, two ATP are gained per glucose fermented. Now consider the same glucose molecule, passing through the oxidative metabolism of a typical aerobic cell: the overall reaction (actually involving many intermediate steps), can be summarized as

glucose + 6 O_2 + 38 phosphate + 38 ADP → 6 CO_2 + 38 ATP + 6 H_2O
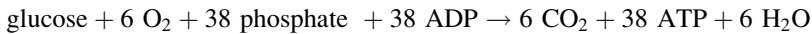
The energy (ATP) gained in completely oxidizing glucose to CO_2 and water is 19-fold greater than in the anaerobic process! Organisms could not reap this reward of energy until free dioxygen was available on Earth. Thus, the advent of dioxygen was, in fact, an enormous bonanza for life.

The actual process of extracting this energy from aerobic metabolism is by no means as simple and direct as the equation we have written would imply. A complex molecular apparatus and an elaborate series of reactions are required. We are not interested here in all the details, but must illustrate some of the mechanism, in order to ask – how could this new metabolism have evolved?

4.2 Evolution of an Aerobic Metabolism

4.2.1 Special Mechanisms Needed for Aerobic Metabolism

The entry of dioxygen into metabolism, in energy-releasing processes, requires that there be something to be oxidized – that is, a molecule or molecules that can transfer electrons to oxygen to reduce it to H_2O. The substance that plays the major role as a reducing agent is reduced nicotinamide adenine dinucleotide, abbreviated NADH; its oxidation and reduction occur as shown in Fig. 4.1. NADH and its oxidized form, NAD^+, may have been very ancient participants in metabolism as electron carriers. They are also essential in fermentation (see Fig. 2.5a), and many even have been produced abiotically at the time of the origin of life (Cleaves and Miller 2001).

Aerobic metabolism of a substance like glucose proceeds first through the citric acid cycle (Fig. 4.2) in which the carbohydrate is eventually completely oxidized to CO_2. Electrons are transferred to a large number of NAD^+ molecules, reducing them to NADH. These then pass through the electron transport chain (Fig. 4.3) where they eventually give up electrons to oxygen, producing water. The electron transport chain is found in aerobic bacteria and in *mitochondria*, special cellular
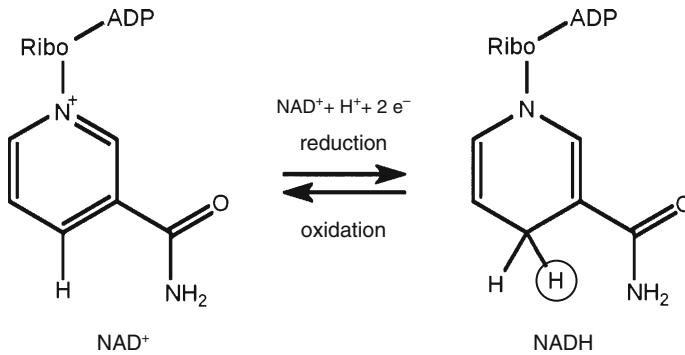
Fig. 4.1 Reduction of NAD$^+$ to NADH. The position of the added hydrogen is marked by a *circle*

organelles of higher organisms (see below). A schematic of this complex apparatus is shown in Fig. 4.3. The working principle is remarkable; electron transport through the system drives the pumping of protons across the mitochondrial membrane. This creates a pH gradient and a strong driving force to bring protons back. But the protons can only reenter the mitochondrial matrix via an "ATP generator" – it is here that most of the ATP is produced. This way of generating ATP from ADP is called *oxidative phosphorylation*.

4.2.2 When and How Did Aerobes Arise?

How did such a complex and seemingly bizarre mechanism evolve? We certainly cannot yet recreate the path but there are certain hints. Recall that there must have existed previously another complex energy-transducing system – the photosystems that carried out oxygenic photosynthesis. These must have been fully functional and well evolved before aerobic ATP-generation could be developed; for there would have been no point in having an oxygen-utilizing system before the environment was provided with dioxygen.

Could the mitochondrial system (in its most primitive bacterial form) have evolved from the photosynthetic apparatus? They are certainly very different, but some startling similarities appear if we compare the electron flows in mitochondria and chloroplasts (Fig. 4.4). If we imagine the photosystem with electron flow reversed, the similarity becomes intriguing. Could there have been a period in evolution when some kind of microbe could fix carbon and release oxygen when in sunlight using photosynthesis, and then "reverse its motor" to burn carbohydrate with dioxygen in periods of darkness? Today, of course, most plants have both chloroplasts and mitochondria, and use these for the separate processes.

An alternative (or complementary?) explanation lies in the similarity of "proton pumps" that are used to remove, from the cell, protons that accumulate in anaerobic
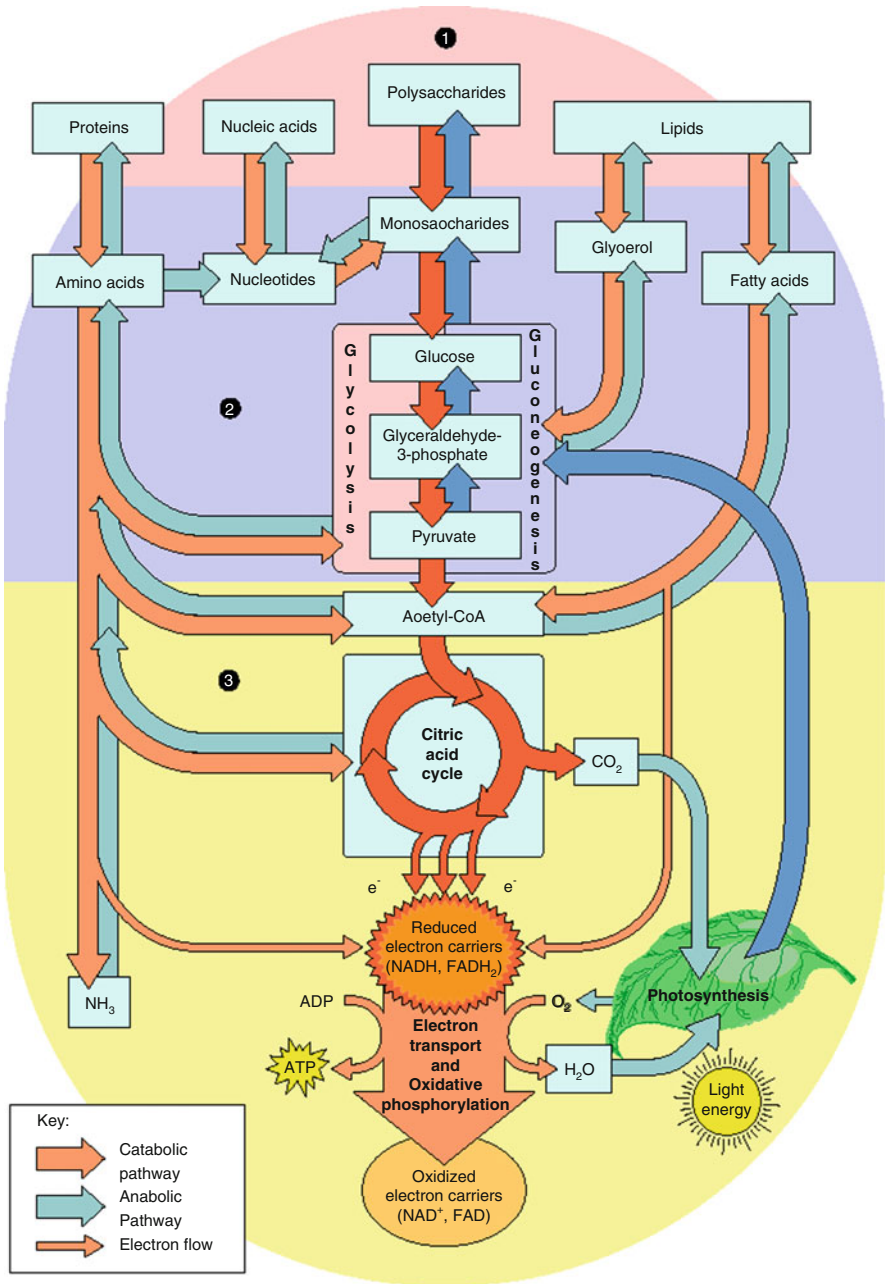
Fig. 4.2 An overview of oxidative metabolism. Catabolic and anabolic pathways and electron flow (Mathews et al. 2000)
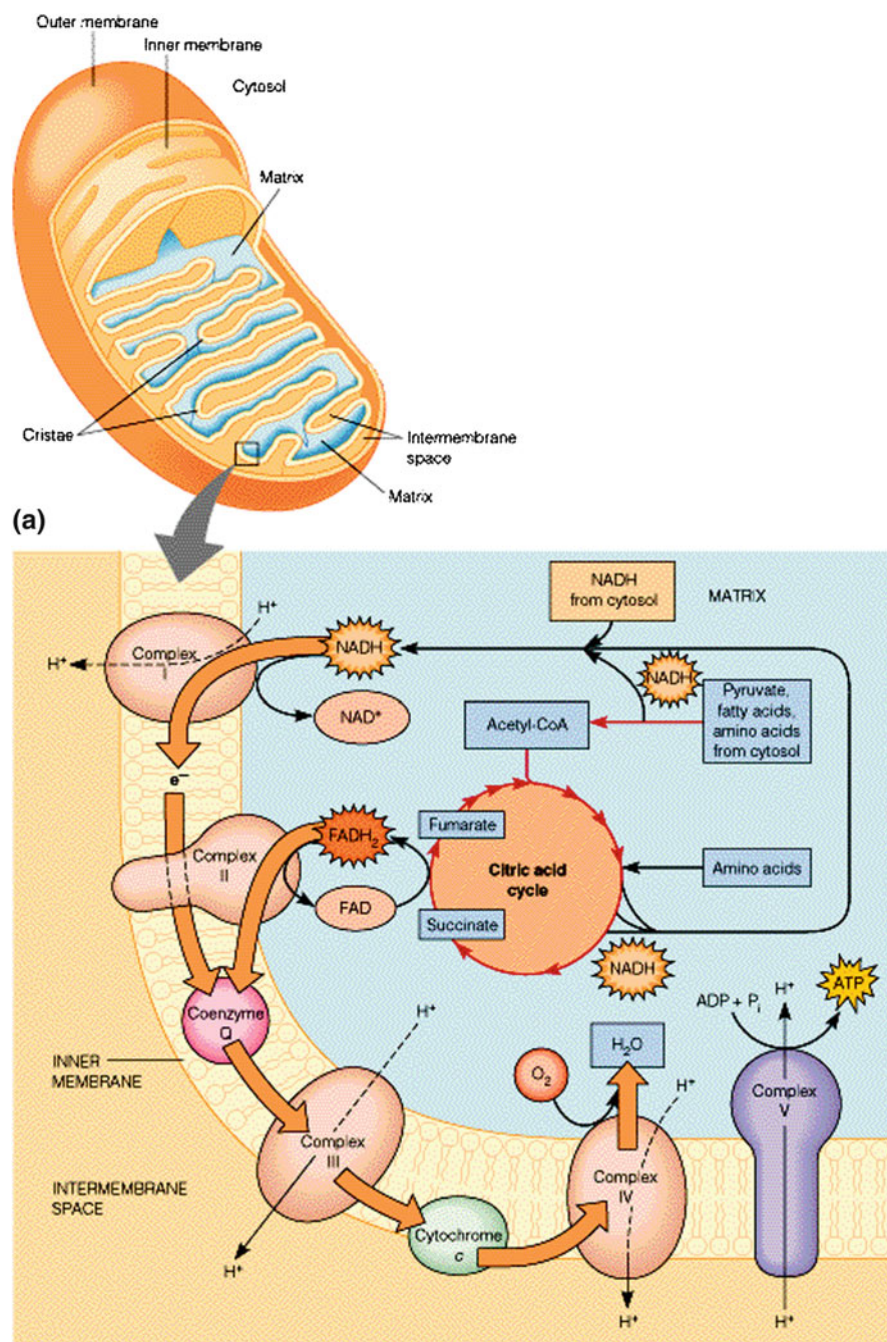
Fig. 4.3 Localization of respiratory processes in the mitochondrion (Mathews et al. 2000)
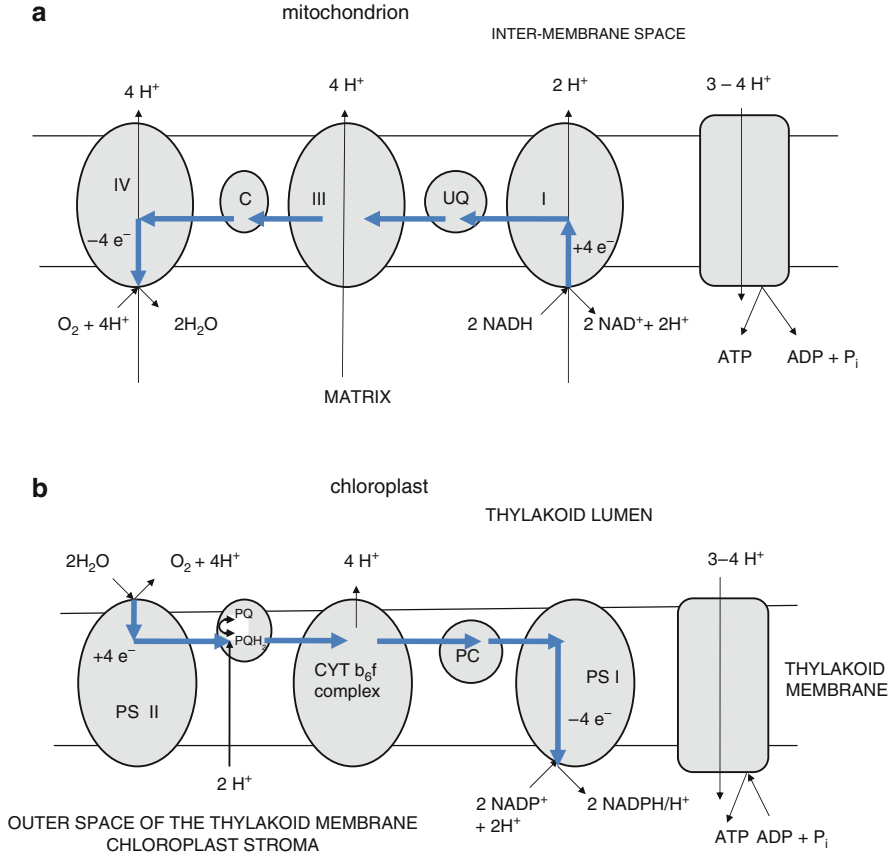
Fig. 4.4 Schematic comparison of electron transport and overall reaction in (**a**) mitochondria and (**b**) chloroplasts. Note the remarkable similarity of functional units, and the opposite direction of electron flow

metabolism to the "ATP generator" that functions in the opposite direction during oxidative phosphorylation.

If the depiction of dioxygen rise, shown in Fig. 2.6, is correct, it would seem reasonable that aerobic metabolism, expressed first in bacteria, would not likely predate 2.2 BYA. Most probably, the first were facultative aerobes, which would fall back on anaerobic metabolism in the frequent situations when dioxygen was absent. Such organism would have had use for the oxygen sensors described in Chap. 3. It should be noted that even aerobes like us can carry out anaerobic ATP generation in muscles that have been depleted of oxygen by strenuous exercises.

There are those, however, who emphasize molecular genetic evidence for the antiquity of certain enzymes of the respiratory chain. They argue that oxidative metabolism predated the rise of atmospheric dioxygen (see for example Castresana and Saraste 1995). This is not, however, the generally accepted view, for the

selective advantage in possessing an oxidative metabolism in the virtual absence of dioxygen is hard to see. However, one must remember that the first appearance of oxygenic photosynthesis may have long predated the "Great Oxidation Event". Microbes that evolved dioxygen may also have learned to use it at a very early stage, extracting it from their local oxygen-rich environment.

It may be, indeed, that it is impossible to even define the "first aerobe". The acquisition of the complex set of metabolic pathways that we consider characteristic of existing aerobes (the citric acid cycle, oxidative phosphorylation, etc.) are not likely to have evolved in a single neat package. Rather, there must have existed intermediate states of evolution, including long extinct organisms which had partial or limited aerobic functionality. The implication of data on protein evolution is of a gradual increase in metabolic complexity. But there was plenty of time. It is at least a half a billion years from the great oxidation event to the emergence of even the earliest ancestor of modern organisms (see Fig. 2.6). Although it is impossible to set a precise date or even period for the first aerobes it certainly must predate the evolution of eukaryotes which is dated somewhere between 1.8 and 0.8 BYA (see below). A thoughtful discussion of possible models of evolution of aerobic metabolism is provided by Albert et al. (1994).

4.3 Eukaryotes: The Next Step in Evolution

4.3.1 Distinction Between Prokaryotes and Eukaryotes

It seems probable that the first photosynthesizers and the first aerobes were both prokaryotes-members of the kingdom now known as Eubacteria. These ancient lineages are characterized by the lack of internal organelles. For example, they do not have a nucleus in which their DNA is separated from the cytoplasm. An eukaryotic cell, such as found in all protozoa, algae, higher plants and animals, is characterized by the presence of a nucleus, separated from the cytoplasm by a nuclear membrane, plus (usually) mitochondria, and in some cases (plants and algae) chloroplasts as well. Some aerobic prokaryotes look like free mitochondria; some photosynthesizing bacteria look like individual chloroplasts.

4.3.2 The Symbiotic Hypothesis

If we examine chloroplasts and mitochondria more closely, we find strong evidence for their prokaryotic ancestry. For example, these organelles each have their own DNA, which codes for some but not all of their proteins. The structural organization of this DNA is much more like that of bacteria than like the nuclear DNA found in the same cells. The protein-synthesizing machinery of chloroplasts

and mitochondria resembles closely that found in prokaryotes and is different from that in the cytoplasm of the eukaryotic cell. In addition, some mitochondrial enzymes seem related to those in prokaryotes, rather than their homologues in the eukaryotic cytoplasm. All of these similarities are consistent with the proposals made over 100 years ago (Mereschowsky 1905) and elaborated and refined by Lynn Margulis (1970) – that eukaryotic cells arose by symbiotic engulfment of aerobic bacteria by a "host" cell. In plants, a similar symbiosis with photosynthetic bacteria by primitive nucleated cells is postulated (Fig. 4.5). These proposals are now widely accepted, although there have been questions raised in recent years (see, for example, Martin and Müller 1998; Doolittle 1998). The process may be continuing, with symbiosis of sulfide-metabolizing bacteria into molluscs and annelids near hydro-thermal vents (Lallier 2006). Even more suggestive is the report that sea slugs harboring photosynthetic algae have experienced nuclear gene transfer of photosynthesis genes (Rumpho et al. 2008). It should be noted that there are discussions about eukaryotes without mitochondria, but having a nucleus, being the basis of the phyla of eukaryotes (König et al. 2007).

The evolution of eukaryotes was a major advance, for a cell with a nucleus, capable in principle of mitosis and sexual reproduction, and also provided by its mitochondria with an internal source of abundant ATP, possessed enormous competitive advantages. Dates are uncertain. The earliest fossils that might be considered protozoan eukaryotes date to 1.5–1.8 BYA (Knoll et al. 2006). However, some
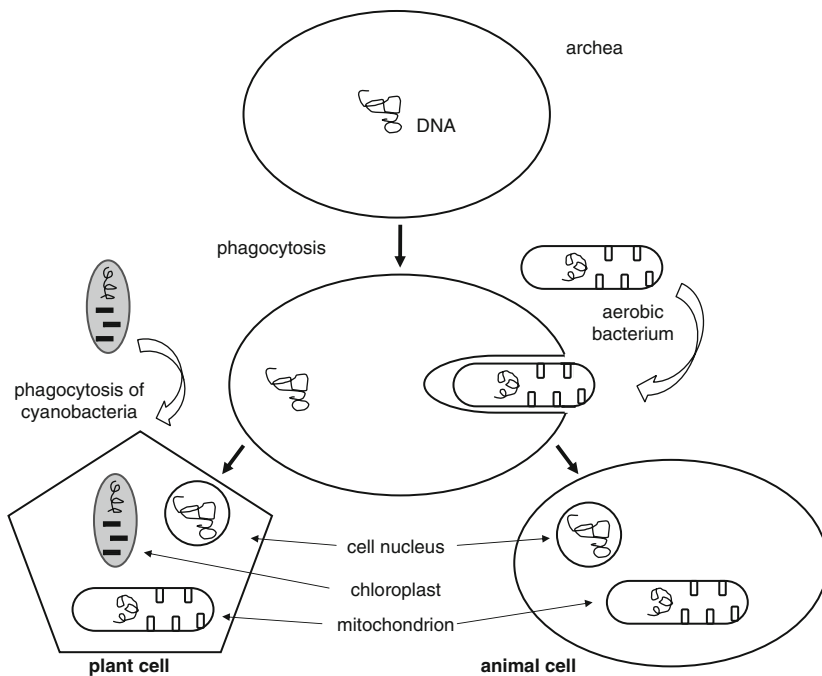
Fig. 4.5 Evolution of eukaryotes by symbiosis

argue that these were more likely to have been advanced prokaryotes, and that origin of eukaryotes lies nearer 1.0–0.9 BYA (see Cavalier-Smith et al. 2006). In apparent contrast is the 1.8 MYA fossil *Grypania*, a half-meter long spiral organism tentatively identified as an alga (see Runnegar 1994 for discussion). If *Grypania* has been correctly identified, the origin of Eukaryotes must lie near the earlier date proposed by Knoll et al. (2006). The problem, of course, is in recognizing sufficient internal detail in ancient fossils to draw a definite conclusion. To complicate the question, molecular clock estimates argue for about 2 BYA (see, for example, Hedges et al. 2004). There remains much discussion, but probably most would agree that by about 1.5 BYA true eukaryotes existed.

4.4 The Last Great Leap: Multicellular Organisms, "Metazoans"

4.4.1 When, Why, and How?

Sometime before 1.0 BYA, the first multicellular organisms (metazoa) emerged. The earliest evidence is of leaf-like forms that may represent multicellular algae that are found as early as 1.7 BYA (see Shixing and Huineng 1995; Knoll et al. 2006). More definitive algae are observed by about 1.2 BYA (Weiguo 1994). From this date forward, there exist multiple small "fossils," generally of unidentified phylogeny. Definite evidence for primitive sponges is found well before 600 MYA (Müller et al. 2001); these probably represent the first multicellular organisms that can be clearly related to living species. However, Müller et al. (2001), argue that the first metazoa must have evolved much earlier. The emergence of such multicellular organisms with discrete macrostructure, implying development of specialized cell types, represents a major advance in animal evolution. It requires, in particular, a mechanism for the specific expression (or nonexpression) of some genes in differentiated cells. Because each cell in a given organism contains the same DNA, and a complete set of genes, specific gene regulation was the key to multicellularity.

The existence of multiple, specialized cell types, as well as larger size, gives metazoa advantages over protozoa in both obtaining food and resisting predation. What level of predation did exist 1 BYA is completely unknown; it must at least have existed among protists. Therefore, a multicellular creature gained an intrinsic defense, and at the same time, distinct advantages in seeking protozoan and prokaryotic prey. We have mentioned in Chap. 3 that melanization, which might have occurred on the dermal cells of primitive metazoans, could offer protection from ROS; it might also have formed a defensive cuticle against predation. Once the principle of cell specialization was established, the way was open for almost infinite pathways of development of special organs, functions and animal diversity.

How did multicellularity originate? Although we have virtually no evidence, there is no dearth of speculation (see Reiger 1994, for an excellent discussion). Three major models exist:

(a) Division colonies:
    Formed by multiple cellular division without cellular separation.
(b) Multinucleate cells:
    Formed by nuclear division without cell division. Such cells are known even today, for example, in *Drosophila* larvae. A later stage of segregation would be needed to produce true multicellularity.
(c) Cell association:
    This is seen today as certain slime molds which exist in both unicellular and multicellular stages.

There is really no evidence on which to choose among those, or other models. But it is obvious that multiple possibilities exist. For the development of cell specialization, a mechanism was needed to provide differential control of gene expression among cells with identical genomes. This is accomplished in all metazoans today via the complexation of genomic DNA with histones and nonhistone proteins to form chromatin. This complex facilitates selective expression of specific regions of the genome. Thus, it is a reasonable hypothesis that the development of differentiated metazoans depended on the evolution of the eukaryotic nucleus, and its modern ultrastructure. At some point in the transition from prokaryotes to metazoa, a new mechanism of gene regulation involving chromatin structure and its dynamics must have arisen. But we have very little idea as to when this happened.

4.4.2   Collagen and Cholesterin

What holds the cells in a metazoan together? In those we see today, this is accomplished by an extracellular matrix composed of a variety of fibrous proteins. Chief among these is collagen, formed from extremely long polypeptide chains rich in glycine, proline, and hydroxyproline. The latter amino acid is unusual; it is not coded for in the genome, but is formed from proline residues in the collagen fibers by unusual reaction (Fig. 4.6).

Dioxygen is absolutely required for this step in the formation of mature collagen fibrils. Remarkably, ascorbic acid is also essential – and this is one of the few coenzymatic functions known for the ascorbic acid in animals, while in plants especially in chloroplasts there are many. The whole process looks very much like the kind of enzyme/scavenger coupled process we described in Chap. 3. Remarkably dioxygen is also needed for another step in the maturation of collagen – the formation of crosslinks between fibrils. Thus, collagen in its functional form in the matrix could not have existed in an anoxic world, and the advent of dioxygen
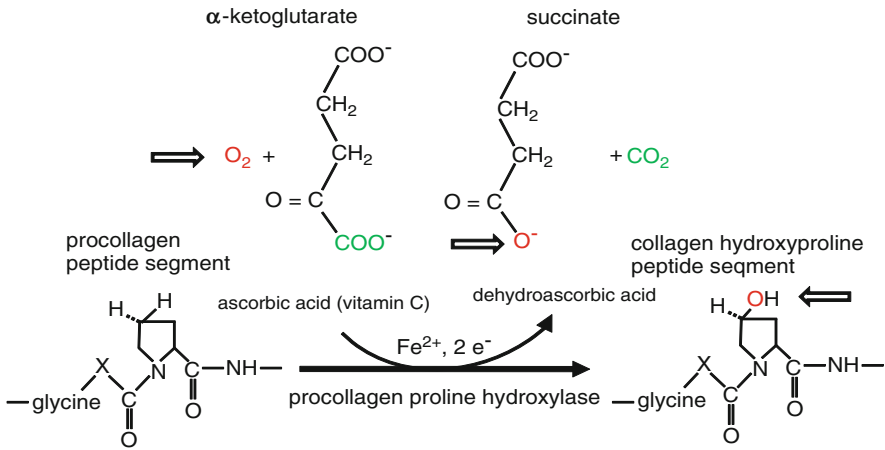
Fig. 4.6 Oxidation of proline (in procollagen) to hydroxyproline enzymatic hydroxylation of procollagen proline residues in the synthesis of collagen. *Arrows* point to the transferred oxygen (after Mathews et al. 2000)

would seem to be necessary for the development of metazoa, at least as we know them.

There is another important biomolecule which needs dioxygen for synthesis: cholesterin. This is needed to make the cell membranes more rigid. However, cholesterin can occur only after dioxygen was available to convert the biogen squalen to cholesterin.

In summary, the increase in dioxygen that occurred during the middle Proterozoic allowed first the development of aerobic metabolism in certain bacteria. These, probably by symbiosis with other bacteria, lead to the first eukaryotes: mitochondria-bearing protists and algae. This, in turn, as dioxygen levels increased, permitted the assembly of multicellular structures, the metazoa.

4.4.3 Half a Billion Years of Stasis?

By about 1 BYA, the stage would seem to have been set for an explosion of large animal development. It did not occur at this point. In fact, there is no evidence for large creatures until over 400 million years later. There are at least three possible reasons that can be suggested:

(a) Rather than being static, increases in animal complexity and diversity were continuing during the period, but the organisms involved were both small and soft-bodied, and so escape much of the fossil record.
(b) According to some (see, for example Holland 2006, and Fig. 2.6), the Earth's dioxygen concentration did not rise to near present levels until nearly 0.6 BYA. As we shall see in Chap. 5, large, complex animals have special problems in

using dioxygen, especially at low concentrations, and their appearance may have awaited the evolution of oxygen transport protein.

(c) There occurred, first at about 0.8 BYA and later at about 0.6 BYA, periods of intense, worldwide glaciations ("snowball Earth," see Fig. 2.6). These, by covering most of Earth's oceans with ice, must have provided a major stress on the whole ecosystem. Perhaps, only after those glaciations were over, could multiple niches for new speciation become available. But the first major group to emerge was peculiar, to say the least.

4.4.4 Emergence and Extinction of the Ediacaran Fauna

Toward the end of the Proterozoic, about 575 MYA, we find the evidence for a group of large, soft-bodied animals that cannot be clearly related to any creatures existing before or after. These are the Ediacaran fauna, named after the location in Australia where such fossils were first found. An example is shown in Fig. 4.7. General overviews are given by Fedonkin (1994) and Grazhdankin (2004). The Ediacaran were almost universally thin (ca 1–2 mm), flattened creatures that either reclined on sediments or projected from them like modern sea fans. Some could be quite large – specimens of *Dickinsonia* (Fig. 4.7) have approached 1 m in maximum dimension.

There is no clear evidence of a mouth, anus, digestive system, or circulatory system, implying that the Ediacarans may have ingested nutrients and dioxygen directly through their surfaces – perhaps this is why they were so thin (see Chap. 5 for further discussion of such matters).

The Ediacarans vanished almost as suddenly as they appeared, about 550 MYA, in the very end of the pre-Cambrian. This was the period of a great explosion of shell-bearing and carapace-protected species (see below). This suggests the advent

Fig. 4.7 Fossil of *Dickinsonia costata* (http://de.academic.ru/dic.nsf/dewiki/370936)

Fig. 4.8 Early cambrium predator *Anomalocaris canadensis*, the top predator from the middle Cambrian Burgess Shale of British Columbia. It is shown from the *side*, in swimming position (Janice Lilien, *Natural History* magazine December1985)
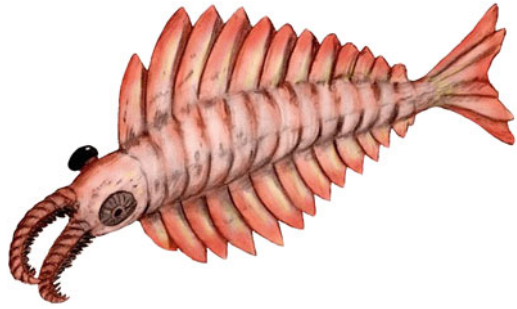
of mega-predation and there is indeed evidence for large predators early in the Cambrian (see Fig. 4.8). The thin, soft-bodied, nonmotile Ediacarans would have been sitting ducks for any predator, no matter how inefficient. Thus, their rapid demise is probably not surprising. It is, however, surprising that no clear evidence for predation on Ediacarans has emerged. Perhaps their extinction was very rapid, once predators appeared. McMenamin (2003) has estimated that predation, once introduced, could encompass the world in a few million years.

4.4.5 The Bilateral Body Plan

A characteristic that is shared by almost all higher animals – including all vertebrates, molluscs, and arthropods – is a bilateral body plan. This refers to the pseudosymmetry of even numbers of legs, eyes, etc. across a plane of reflection. We say pseudosymmetry, because it is not exact – we have only one heart, for example, and it is on one side.

It is now recognized that the key to bilaterity lies in the regulated expression of certain genes during embryonic development (see Peterson and Davidson 2000, for details). How early did the bilateral plan evolve? It is clearly present in many of the phyla that are recognized early in the Cambrian, but it is not obvious that the soft-bodied Ediacarans were true bilaterians (see Fedonkin 2003).

For this reason, the discovery in 580–600 MYA strata in China, of a very small, but clearly bilaterian fossil was of major importance (see Bottjer 2005, for background). The fossil called *Vernamimalcula* has a complex ultrastructure, showing evidence of a sophisticated evolutionary history. Although probably contemporary with or slightly predating the Ediacarans, *Vernamimalcula*, is clearly unrelated to them, and more likely in the line of descent that led to one of the existing major phyla. By 555 MYA, still Precambrian, there existed some large bilaterian animals having some rigid parts such as *Kimberella quadrata* found in Australia and in Russia (see Fig. 4.9; Fedonkin and Waggoner 1997).
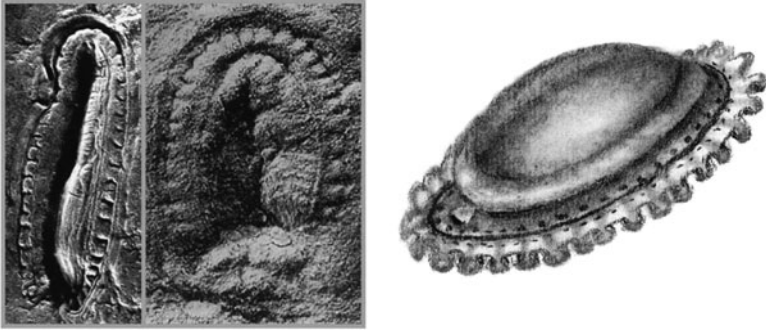
Fig. 4.9 Photo and a reconstruction of the fossil *Kimberella* (UC Berkeley museum of paleontology and drawn by Alexandra Cowen; Images courtesy of Waggoner and Fedonkin)

4.4.6   The "Cambrian Explosion": Fact or Artifact?

A long view over the fossil record reveals a veritable discontinuity at about 545 MYA the beginning of the Cambrian. Before this date, fossils are scarce, and fossils with evidence for shells, skeletons, or other hard parts are almost nonexistent. Before this date, it is hard to recognize organisms that can be clearly identified as ancestor of any living species. Aside from the sponges, mentioned above, the sole exception may be the mollusk-like *Kimberella* (see Fig. 4.9 and Fedonkin and Waggoner 1997). On the other hand, shortly after the beginning of the Cambrian, most extant phyla are represented.

This sudden apparent diversification of species has been termed the "Cambrian explosion," and it was a puzzle even to Darwin (see Conway-Morris 2006). After many years of debate, scientific opinion is still divided between those who take the explosion literally, and believe that all of the major phyla came into being in a relatively short period about 550 MYA, and those who believe that specific antecedents of most phyla existed much earlier. The latter have to explain the lack of fossils. There are two common attempted explanations:

(a) Perhaps the ancestor of mollusks, arthropods, annelids, etc. were soft-bodied, and did not preserve well. However the preservation of the Ediacaran fauna would seem to counter this argument.

(b) Perhaps all the predecessors were very small and thus, their fossils have escaped detection. This is also hard to accept, considering the very large number of small or even microscopic fossils of protists and prokaryotes that have been recognized, even back to the Archaean. The argument that major eukaryotic phyla diverged at about the Cambrian/pre-Cambrian boundary has been strongly presented by Conway-Morris (2000, 2006).

In contrast to these arguments, the majority of "molecular clock" data from analysis of RNA or protein sequences point to much earlier dates (800–1,200 MYA) for

the divergence of the major eukaryotic phyla (see, for example Hedges et al. 2004; Wang et al. 1999). There is another logic to this view. Each of the extant phyla possesses particular and complex structures. It is hard to see how all of these could have simultaneously sprung forth from some single generalized ancestor in the brief period at the Cambrian boundary. Rather, one would expect that the evolution of their unique characteristics must have had its roots in earlier periods. For example, molluscs and arthropods developed completely different chemistries for constructing shells and carapaces, for the same protective function. These organisms must have previously developed their different oxygen transport proteins (OTPs) to permit the existence of such structure (see Chap. 5). As we shall see later, the OTPs of molluscs and arthropods are different and apparently evolved from different proteins with different functions. This argues that the common ancestors existed much earlier.

Despite strong claims to the contrary (Conway-Morris 2000, 2006), it does not seem to us that the controversy concerning the "Cambrian Explosion" is yet reconciled, or will be in the near future. The two extreme models are depicted on Fig. 4.10.

One feature that clearly characterizes the early Cambrian is the rather sudden appearance of skeletonization, characterized especially by the shells of molluscs, the exoskeletons of arthropods, and the cuticle of annelids. By several different means, animals were suddenly armoring themselves. These creatures are in sharp contrast to the soft-bodied Ediacaran, soon marked for destruction. The simplest explanation for this is defense against predators. Indeed, there is evidence that macropredation became significant at about this time. The primitive arthropod
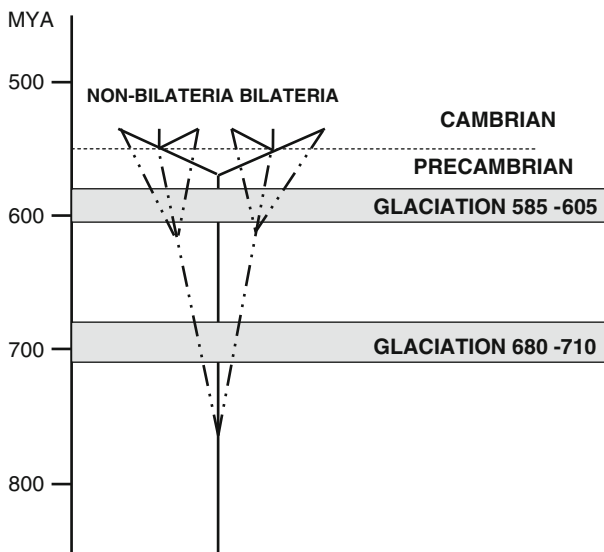

Fig. 4.10 "Deep" and "shallow" models for animal evolution. The shallow model (*solid lines*) assumes no major divergence of phyla until around the Precambrian–Cambrian boundary. The "deep" model assumes divergences began much earlier

Anomalocaris (up to 2 ft in length, Fig. 4.8) occurring about 540–500 MYA must have been a formidable predator even to sizeable prey.

But this raises another question: Why did creatures suddenly become large? We may suggest the following: only at about this time did dioxygen levels finally increase to near present value (see Fig. 2.6). As we shall see in the next chapter, low dioxygen levels causes difficulties for large animals, and restrict size increase unless special provisions are made to deliver dioxygen to tissues. This is especially true for the organism armored against predators. We shall see in the next chapter how such provisions can be made. Dioxygen may have, for one more time, played a critical role in evolution.

References

Albert BD, Lewis I, Raff M, Robert K, Watson IW (1994) Molecular biology of the cell, 3rd edn. Garland, New York, pp 699–763

Bottjer DJ (2005) The early earth evolution of animals. Sci Am 293:42–47

Castresana J, Saraste M (1995) Evolution of energetic metabolism: the respiration-early hypothesis. Trend Biochem Sci 20:443–448 http://www.ncbi.nlm.nih.gov/pubmed/8578586?ordinal-pos=2&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum

Cavalier-Smith T, Brazier M, Embley TM (2006) Introduction: how and when did microbes change the world. Philos Trans R Soc Lond B Biol Sci 361:845–850

Cleaves HJ, Miller SL (2001) The nicotinamide biosynthetic pathway is a by-product of the RNA world. J Mol Evol 52:73–77

Conway-Morris S (2000) Evolution: bringing molecules into the fold. Cell 100:1–11

Conway-Morris S (2006) Darwin's dilemma: the realities of the Cambrian 'explosion'. Proc R Soc Lond B Biol Sci 361:1069–1083

Doolittle WF (1998) A paradigm gets shifty. Nature 392:15–16

Fedonkin MA (1994) Vendian body forms and traces. In: Nengston S (ed) Early life on earth. Columbia University Press, New York, pp 370–388

Fedonkin MA (2003) The origin of the metazoa in the light of the proterozoic fossil record. Paleontol Res 7:9–41

Fedonkin MA, Waggoner BM (1997) The Late Precambrian fossil *Kimberella* is a mollusc-like bilaterian organism. Nature 388:868–871

Grazhdankin D (2004) Patterns of distribution in the Ediacaran biotas: facies versus biogeography and evolution. Paleobiology 30:203–221

Hedges SB, Blair JF, Venturi M, Shoe J (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. BMC Evol Biol 4:2–10

Holland HD (2006) The oxygenation of the atmosphere and oceans. Philos Trans R Soc Lond B Biol Sci 361:903–915

Knoll AH, Javaux EJ, Hewitt D, Cohn P (2006) Eukaryotic organisms in Proterozoic oceans. Philos Trans R Soc Lond B Biol Sci 361:1033–1038

König H, Fröhlich J, Lil L, Wenzel M, Berchtold M, Dröge S, Breunig A, Pfeiffer P, Radek R, Brugerolle G (2007) The flagellates of the Australian termite *Mastotermes darwiniensis*: identification of their symbiotic bacteria and cellulases. Symbiosis 44:51–65

Lallier FH (2006) Thioautotrophic symbiosis: towards a new step in eukaryotic evolution. Cah Biol Mar 47:391–396

Margulis L (1970) Origin of eukaryotic cell. Yale University Press, New Haven, CT

Martin W, Müller M (1998) The hydrogen hypothesis for the first eukaryote. Nature 392:37–41

Mathews CK, van Holde KE, Ahern K (2000) Biochemistry, 3rd edn. Addison-Wesley-Longman, San Francisco, CA

McMenamin MAS (2003) Origin and early predators. In: Kelley P, Kowalewski M, Hansen T (eds) Predator-prey interactions in the fossil record. Kluwer/Academic, New York, pp 379–400

Mereschowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Zent Bl 25:595–596

Müller WEG, Li J, Schröder HC, Qiao L, Wang X (2001) The unique skeleton of siliceous sponges (Porifera; Hexactinellida and Demospongiae) that evolved first from the Urmetazoa during the Proterozoic: a review. Biogeosciences 4:219–232

Peterson KJ, Davidson EH (2000) Regulatory evolution and the origin of the bilaterians. Proc Natl Acad Sci USA 97:4430–4433

Reiger RM (1994) Evolution of the lower metazoan. In: Bengstron S (ed) Early life on earth. Colombia University Press, New York, pp 475–488

Rumpho ME, Worful JM, Lee J, Kannan K, Tyler MS, Bhattacharya D, Moustafa A, Manhart JR (2008) Horizontal gene transfer of the algal nuclear gene psbO to the photosynthetic sea slug *Elysia chlorotica*. Proc Natl Acad Sci USA 105:17867–17871

Runnegar B (1994) Proterozoic eukaryotes: evidence from biology and geology. In: Bengstron S (ed) Early life on earth. Colombia University Press, New York, pp 287–297

Shixing Z, Huineng C (1995) Megascopic multicellular organism from the 1700 million-years-old Tuanshanzi formation in the Jixian area, North China. Science 279:620–622

Wang DYC, Kumar S, Hedges SB (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. Proc R Soc B Lond Biol Sci 266:163–171

Weiguo S (1994) Early multicellular fossils. In: Bengstron S (ed) Early life on earth. Colombia University Press, New York, pp 358–369

Chapter 5
Facilitated Oxygen Transport

Contents

5.1   How to Deliver Dioxygen to Animal Tissues?

The amount of dioxygen an organism needs for aerobic metabolism depends on many factors, size and activity being the most important. However, as an approximate figure, we may say that a typical higher eukaryote will utilize about 3.5 ml dioxygen kg^{-1} body weight per minute. This must reach the tissues where active metabolism is occurring and be maintained there at a steady-state pressure of approximately 2 Torr. This will assure a sufficient rate of delivery to mitochondria

and allow continued utilization therein for oxidative reactions (see Chap. 4). The problem faced by the organism is how to assure sufficient delivery to all the tissues, even those buried deep in the body, sometimes while inhabiting an oxygen-poor environment. There are a number of possible strategies to solve this problem. All have been used at one time or another in the course of evolution.

5.2 Modes of Delivery

5.2.1 Diffusion from the Surface

The simplest (but rarely effective) way in which dioxygen might be delivered to respiring tissues is by simple diffusion from the surface of the organism. The practicality of this solution depends on a number of factors: the surrounding dioxygen concentration, the permeability of the integument, and the distance over which diffusion must occur. The latter factor is often critical and precludes the use of simple diffusion by "compact" animals, unless they are very small (a few mm in diameter, see below). There are essentially two ways that a larger organism can diminish the diffusion path: first, by providing invaginations such as insect trachea and second, by being very thin and sheet like. A very thin organism can, in principle, respire effectively even if it is quite large in other dimensions. How thin is essential? A classic calculation by Raff and Raff (1970) has given approximate answers (see also Runnegar 1991). They used an approximate solution for the differential equations for diffusion into a respiring thin slab, and asked what thickness would permit a critical level of O_2 to be maintained at its center. If the ambient pressure were 160 mm Hg (present dioxygen level [PAL]), a body thickness of 1–4 mm could support a reasonable metabolic activity. However, if the O_2 content were only 10% of this (a reasonable pre-Cambrian level, or one found in many hypoxic environments today), a thickness greater than 1 mm could not be allowed (Runnegar 1991).

This bears interesting implications for some Precambrian organisms, especially the Ediacaran fauna. An example may be found in the sheet-like creature (*Dickinsonia*) shown in Fig. 4.8. Although up to 1 m in diameter, *Dickensonia* was only a few mm thick. Whether or not *Dickensonia* could have metabolized on the basis of dioxygen diffusion is still unsure because we do not know the ambient dioxygen pressure in late Precambrian times or the activity of the organism (see Runnegar 1991). However, it is intriguing that most Ediacaran organisms were sheet like and none showed evidence of a circulatory system. It is possible that they grazed on or lived in proximity to photosynthetic bacterial mats which would have provided a locally higher O_2 concentration. There are a number of living organisms that depend on diffusion (flatworms, protista, coelenterates), but they are invariably thin; often only a few cell layers in thickness.

Of course, all those animals that carry shells, other impermeable body coverings, or are larger than a few millimeters, simply cannot depend on diffusion of dioxygen

from the outer surface. They must use some kind of a circulatory system to bring dioxygen from lungs or gills to tissues. As we shall see later, there was likely evolutionary pressure in the late Precambrian that demanded a drastic solution to the oxygen delivery problem. Thinness may have been only one attempt, and a largely unsuccessful one.

5.2.2 Transport via Blood as a Dissolved Gas

If blood (or as it is often referred to in invertebrates, hemolymph) devoid of an oxygen transport protein (OTP) is circulated through a "gas exchanger" (lung or gill) and then to the tissues, it can transport a soluble gas. This can work quite well in the reverse sense for the elimination of CO_2, a very soluble product of aerobic metabolism, but is much less successful for O_2 delivery. Dioxygen is just not very soluble in water. For example, in human blood, the dissolved dioxygen represents only about 2% of that which can be carried by the hemoglobin present (see Fig. 5.1). Despite the inefficiency of this method of transport, there exist organisms that utilize it. Many have low metabolic rates. On the other hand, organisms living in cold water will encounter more dioxygen and may live just on the physically dissolved dioxygen as do some Antarctic fish. It should be noted that dioxygen is about 60% more soluble in sea water at 0°C than at 20°C. An interesting example is found in the Antarctic "ice fish" of the family Channichthyidae. These fish are unique among vertebrates in having no hemoglobin at all; they even lack essential globin genes (Sidell and O'Brien 2006). Nor do they use any of the other dioxygen
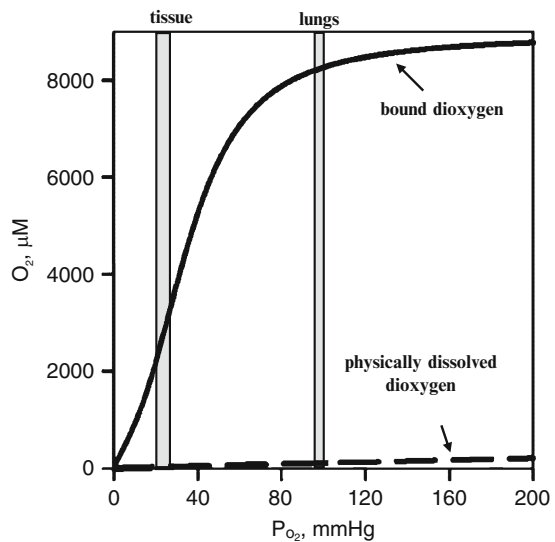
Fig. 5.1 Oxygen transport capacity of human blood. The sigmoid oxygen-binding curve increases the oxygen-binding capacity in the lungs by up to two orders of magnitude in comparison to the physically solved dioxygen, which increases linearly with oxygen partial pressure

carriers we shall discuss later. Their metabolism is not excessively slow, so there must be compensating factors. Some may be found in an unusually high circulation rate and an increase in tissue vascularization, which increases the total contact area between the blood vessels and the tissues (Sidell and O'Brien 2006). There is presumably some compensation to the animal in not having to synthesize hemoglobin. Nevertheless, this is unusual; all other vertebrates carry hemoglobin, even those living in polar waters. For active, large vertebrates like humans, the dissolved dioxygen is completely insufficient. The following example gives an idea, why: given the solubility of dioxygen is 0.00003 ml O_2/ml blood/Torr. In the alveolus, the oxygen partial pressure is about 100 Torr. Thus, arterial blood contains 0.003 ml O_2/ml blood. A 75-kg man needs at least 0.075 ml O_2/ml blood assuming a cardiac output of 5 l blood/min and a consumption of 5 ml per kg body mass and per minute. At a concentration of only 0.003 ml dissolved O_2/ml blood at 100 Torr, the 75-kg man would need to pump the blood with a 25-fold higher rate in order to provide the required amount of dioxygen. This would correspond to a pump rate of 125 l blood/min, which is not possible: in order to reach this pump rate, the heart would need to beat with a 25-fold higher frequency, meaning a rate of 60×25 Hz = 1,500 Hz! Thus, OTPs are essential to provide the necessary addition amount of oxygen.

5.2.3   Oxygen Transport Proteins: What They Must Do?

Given the ineffectiveness of diffusion or plasma solubility in delivering oxygen, it is not surprising that almost all higher organisms have evolved proteins dedicated to dioxygen transport in the blood. We shall abbreviate these *oxygen transport proteins* as OTPs. In some cases, these proteins are simply dissolved in the plasma; in others, they are carried in circulating cells. In any event, the advantage in having such a protein can be enormous, as Fig. 5.1 has shown.

Another, more subtle advantage in using an OTP can be seen by a close examination of Fig. 5.1. Solubility has an almost linear dependence on pO_2, whereas the curve for hemoglobin has a sigmoidal shape. This curve shape is typical of *cooperative binding*, which proteins can facilitate and which can make for much more efficient loading and unloading of the dioxygen. We shall have much more to say about cooperative binding later. For now, we should also note that cooperative binding has a corollary feature: it allows for *allosteric control* of binding. This means that other substances involved in metabolism can act to fine-tune the oxygen delivery process. Cooperative O_2-binding and allosteric controls are features of most OTPs.

For a substance to serve as an oxygen transporter in the blood, it must be able to accept dioxygen from a surface in contact with external air or water (lung, gill), be circulated to the tissues, and there deliver the dioxygen. The carrier must not itself be irreversibly oxidized by dioxygen. It seems, from evolutionary
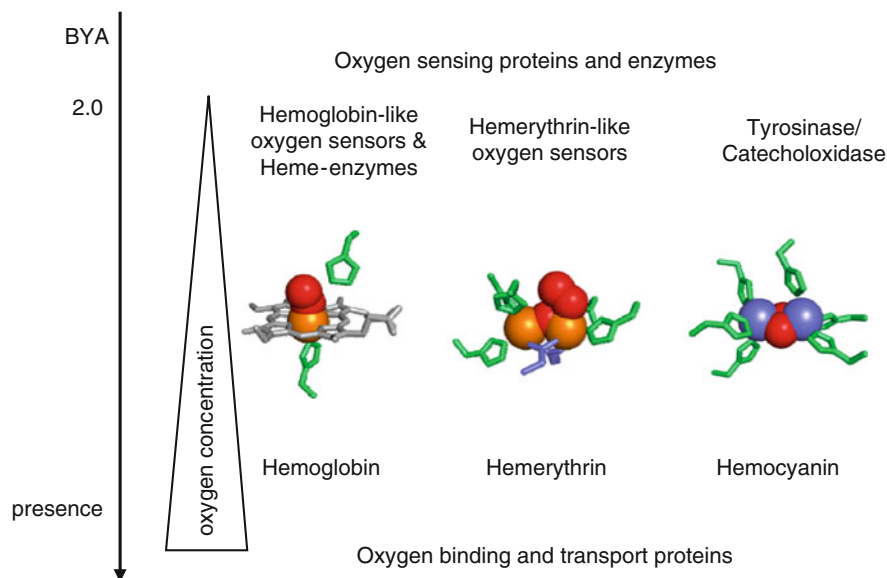
Fig. 5.2 Binding sites for dioxygen in different OTPs. (**a**) Hemoglobin: dioxygen is bound between the iron atom in the center of the porphyrin ring and a histidine residue. (**b**) Hemerythrin: The dioxygen is bound in a complex way between two iron atoms. (**c**) Hemocyanin: dioxygen is bound between two copper atoms. Color code: iron (*orange*), copper (*blue*), oxygen (*red*). The today's oxygen transport proteins evolved from oxygen sensor proteins and enzymes with increasing oxygen concentration in the atmosphere. Courtesy of Elmar Jaenicke

history, that transition metals in their lower oxidation states, particularly iron(2) and copper(1), are especially well suited to this task. However, they will not so function as free ions, for these would be quickly oxidized. Therefore, they are invariably carried in proteins, which provide, in some still not fully understood manner, protection against irreversible oxidation. At present, all known OTPs contain either Fe(2) or Cu(1) at their dioxygen-binding sites, and we know of no evidence for any other. The sites found in the four recognized classes – hemoglobins, hemerythrins, and two types of hemocyanin – are shown in Fig. 5.2, and general characteristics of these protein classes are summarized in Table 5.1. In all four classes, there are changes in either electron spin distribution or electronic charge of the metal ions upon oxygen binding and these accompany changes in the local protein geometry. These changes are also reflected in spectroscopic changes, which in some cases (notably hemocyanin) produce visible color changes. In hemocyanin, the copper is actually oxidized by the bound dioxygen resulting in a blue color, but this is reversed upon oxygen release. Active site geometries for each class have now been determined by high-resolution X-ray diffraction studies (Perutz 1970; Hazes et al. 1993; Magnus et al. 1994; Cuff et al. 1998; Farmer et al. 2001).

Table 5.1 Comparison of different OTP classes

Protein type	Phylogenetic distribution (sources)	Metal	Number of O_2-binding sites	Localization occurrence
Hemoglobin	*Many phyla*	Fe heme		
	Vertebrates		4	Intracellular
	Invertebrates (annelids, some arthropods, molluscs, minor phyla)		2–144	Some intracellular, some extracellular
Hemerythrin	*Rare, in 3 phyla*	Fe–Fe		
	Sipunculids, brachiopods, occasionally annelids		3–8	Intracellular
Hemocyanin	*2 phyla*			
	Only molluscs	Cu–Cu	70–160	Extracellular
	Only arthropods	Cu–Cu	6–48	Extracellular

5.3 Modes of Dioxygen Binding to Oxygen Transport Proteins

5.3.1 Cooperative and Noncooperative Binding

There exist two distinct ways in which a protein can bind a small molecule-like dioxygen (often called the ligand): the binding may be *noncooperative* or *cooperative*. The molecular requirement for the two is different, and their functionality is very different. In cooperative binding, the occupation of some binding sites modifies the affinity of other sites for the ligand. This requires some kind of communication, direct or indirect, between binding sites. In noncooperative binding, the sites are completely independent of one another: any given site has no "knowledge" of the state of any other site. As we shall see, cooperative binding is most useful in oxygen transport, noncooperative binding is usually found in proteins used for dioxygen storage, such as myoglobin.

A simple analysis of binding behavior can easily distinguish between noncooperative and cooperative binding. This is expressed as a binding curve (Fig. 5.3). In such curves, we graph the fraction of filled sites in the sample versus the ligand activity, which can be expressed in case of gases as dissolved gas concentration or the partial pressure of the gas. For dioxygen, we designate this as pO_2. A typical result for noncooperative binding is a hyperbolic curve, only slowly approaching a limit as pO_2 becomes large (curve A). In contrast, a cooperative binder shows the very different sigmoidal curve B. We shall discuss possible mechanisms of cooperative binding below. But for now, a qualitative picture will suffice. The cooperative curve acts as if it begins at low pO_2 with weak binding, which becomes stronger as more and more dioxygens are added. This gives the sigmoidal shape. Its advantage in dioxygen transport is the following: a cooperative binder can nearly completely load at moderate pO_2, and yet almost completely
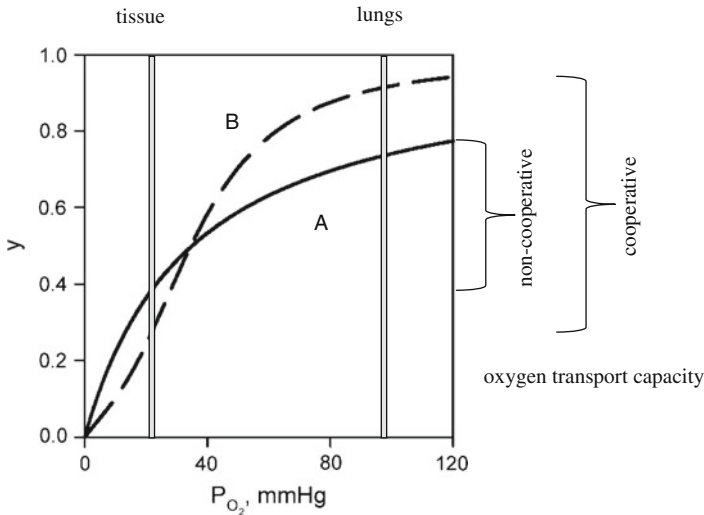
Fig. 5.3 Comparison of noncooperative (A) and cooperative (B) binding of oxygen to an OTP. The two bands represent approximately oxygen pressures in lungs and tissues, respectively. Note how much more effective B will be in delivering oxygen. The abscissa (Y) represents fraction of saturation

unload at the lower oxygen concentration in the tissues (see Figs. 5.1 and 5.3). As cooperativity becomes stronger and stronger, the system begins to act like a "molecular switch" giving almost no binding below a critical value and nearly complete saturation above. Cooperative binding is one form of allostery, the phenomenon in which binding of one ligand can influence the binding of others. Since in this case it is dioxygen binding influencing further dioxygen binding, we call it homotropic allostery.

What is required for cooperative binding? The essence of cooperativity is that one binding site is somehow informed of the status of other binding sites: are they occupied or empty? Only in this way can the molecules respond by changing the affinity of site X in response to the occupancy of site Y. This requires that there be some connections, covalent or noncovalent, between binding sites. In nature, this has been achieved in two ways: (a) The molecular unit containing the sites which be duplicated in a tandem multidomain protein structure, as a consequence of tandem gene duplication (Fig. 5.4a). (b) More often, individual protein's oxygen-binding units may associate to form noncovalent oligomers (Fig. 5.4b). In either case, individual functional units, each with one binding site, must communicate. We shall see that these methods of generating potentially cooperative-binding structures have been used over and over in evolution. Various kinds of organisms have sought to satisfy the need for cooperative binding in a variety of ways, but invariably multiple binding units are combined, covalently or noncovalently, to form the functional molecule.

Subunit with one binding site

(a) Tandem gene duplication

(b) Evolution of capacity for association

This protein can be cooperative
if two conformations
with different affinities exist

If two conformational states
have different affinities
will get cooperativity
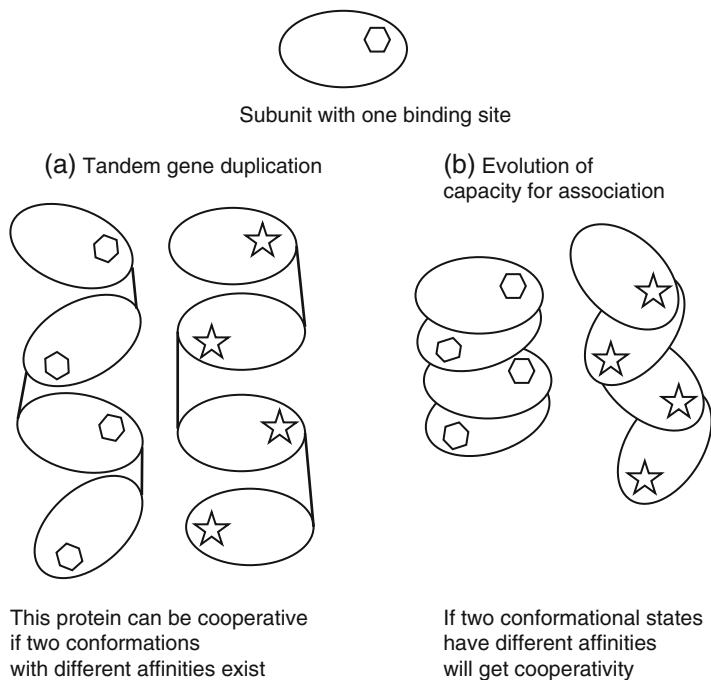
Fig. 5.4 Two ways in which cooperative binding could evolve. (**a**) By tandem gene duplication, one protein carries several oxygen-binding sites in different domains and functional units. (**b**) More often, individual proteins carrying one oxygen-binding site associated to noncovalent oligomers. They can adopt two different conformations characterized by different affinities indicated by the different symbols

5.3.2 How Does Cooperativity Work?: Models for Allostery

Cooperativity can only be ensured when oxygen-binding sites are in communication with each other. There is still a discussion whether cooperativity can be established by single functional units undergoing transition from weak to strong binding state one at a time (Fig. 5.5a) or if the observed cooperativity is the result of the possibility of the entire multisite OTP molecules having two overall conformations, which differ in oxygen affinity (see Fig. 5.5b). The first possibility requires that some kind of conformational change be propagated from each functional unit to another one as dioxygen is bound. This kind of model was first proposed by Koshland et al. (1966). The second type of model calls for concerted changes in a multisubunit structure. Here, whole multisubunit OTP molecules can occur in either a high (R) or a low affinity state (T) characterized by different affinities of the subunits to oxygen. The model requires that all subunits have the same oxygen affinity and switch synchronously from one conformation to the other one.
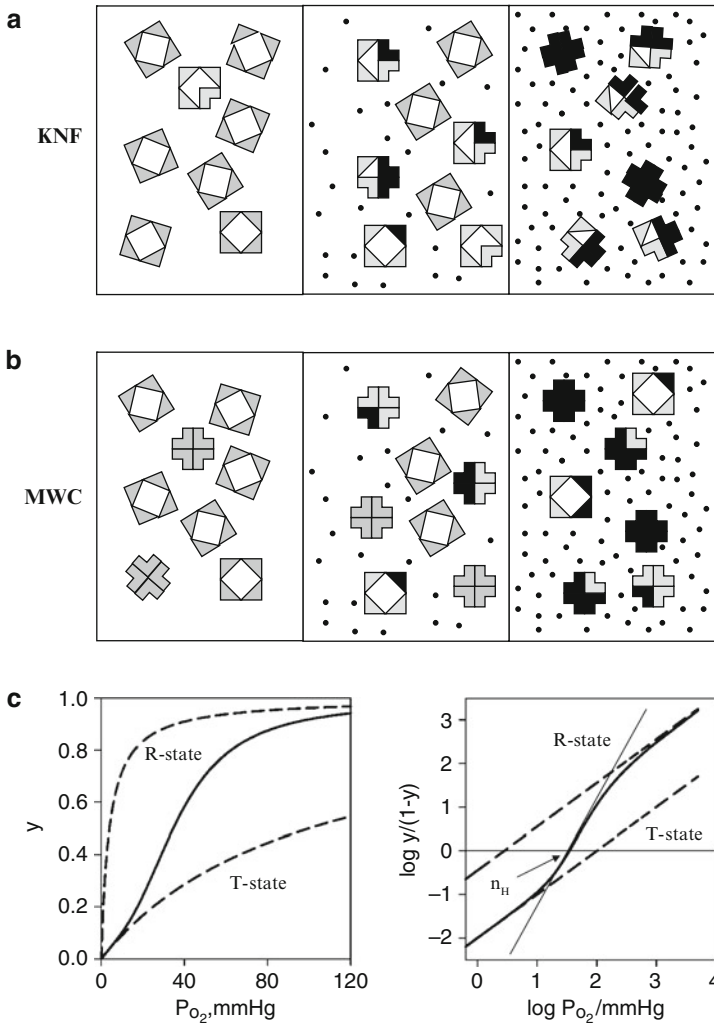
Fig. 5.5 Two major models for cooperative binding. (**a**) Binding of dioxygen to one subunit modifies the affinity of the neighboring subunits (sequential KNF model). (**b**) An assembly (the allosteric unit) can adopt two different conformations characterized by different oxygen affinities. Within each conformation, after binding of dioxygen does not influence the other binding sites (concerted MWC model) (Mathews et al. 2000). The *dark dots* represent oxygen molecules, whose concentration increases from *left* to *right*. (**c**) Here a cooperative binding curve is shown for a mixture of allosteric units within the two states (R- and T-states). The subunits bind dioxygen noncooperatively in the R- or T-state with different affinities. But the population of R-state will increase as binding progresses. The data can be represented as a Hill plot (*right*). The slope of the cooperative-binding curve at half-saturation (Y = 0.5) gives the Hill coefficient n_H, a measure for the degree of cooperativity. Behavior of molecules that were constrained to remain in R- or T-states is shown by *broken lines* with slope of unity

The oxygen-binding curve of blood is a superposition of some OTP molecules being in the R- and some in the T-state. The equilibrium can be shifted in either direction by stabilizing the one or the other conformation (Monod et al. 1965).

5.3.3 Self-Assembly and Nesting

As we shall see, most OTPs are multisubunit proteins, formed from subunits, each of which carries one binding site. This seems to be the simplest way to accomplish the necessary function. These multisubunit proteins are formed via a self-assembly process establishing various symmetry axes. Self-assembly has many advantages. At the in vivo concentration, all proteins will be associated as oligomers. The same principle can be used to build up larger complexes. Oligomeric blocks such as the hexameric hemocyanins (see below) have established all functional properties to behave as an OTP. However, in the hemolymph of many arthropods we find molecular structures that are integer multiples (2, 4, 6, or 8) of hexamers depending on the species (Markl and Decker 1992; van Holde and Miller 1995). This higher level of organization is also formed by self-assembly. This arrangement allows different effectors to modulate the function at different structural levels along the various symmetry axes (see below). We call a model possessing these features a nested model (see Fig. 5.6 and Robert et al. 1987). Here hierarchies of regulation mechanisms are assigned to corresponding hierarchies in the structure and function (Decker and Sterner 1990; Decker and Hellmann 2010; in preparation). Self-assembly has a number of advantages. It allows complex structures to develop without any additional scaffold protein and also results in a more efficient use of gene expression; for mis-folded subunits would simply not take part in the assembly process. This is important, because oligomers containing mis-folded units might not function cooperatively. Instead, the mis-folded proteins will be rejected in assembly and degraded. Thus, it seems that building up cooperative structures by a self-assembly of monomers is more economical than to build a cooperative device in one piece from a tandemly reiterated gene. In addition, genetic material is more economically used, in self-assembly.

Fig. 5.6 Nesting model. Large protein complexes often consist of assembled allosteric units. This offers many functional possibilities/properties. The embedded/nested allosteric units can be modulated by cooperative interactions between them. Thus, this nesting model respects the obvious hierarchies in the complex protein structure

$$L_o = \frac{[T_o]}{[R_o]}$$

$$l_T = \frac{[tT_o]}{[rT_o]} \qquad l_R = \frac{[tR_o]}{[rR_o]}$$
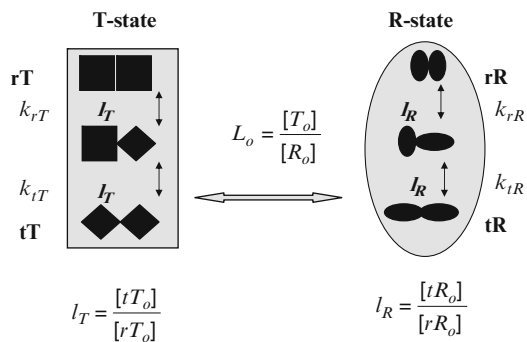
5.3.4   Why Complex Multisubunit Oxygen Transport Proteins?

The molecules that function as OTPs – especially these found free in the blood of invertebrates – are often remarkable huge molecules. Why are they so large? One obvious reason is to produce a high cooperativity; the more sites there are in one connected assembly, the higher the cooperativity can be. However, this cannot be the whole answer, for often the observed cooperativity is far less than that possible. There are large OTPs with very low cooperativity. Thus, other reasons must be sought. One could be that oligomerization stabilizes the complex. When OTPs are freely exposed in the hemolymph, they are surrounded by an environment which contains enzymes, bacteria, etc., which may lead to molecular damage. Oligomerization also protects against natural degradation. Often, oligomers assembled from subunits can hide possible cleavage position on their surfaces and the often-flexible C- and N-terminal parts in the interior of the assembly. But this will only be the case when the assembly is perfect.

There are also simple physical reasons to favor large size in dissolved OTP molecules. To obtain high oxygen-carrying capacity, the concentration of the protein must be large (sometimes as high as 100 mg/ml). This would create a very high colloid osmotic pressure if the protein were present as free subunits, for the osmotic pressure depends only on the number of particles per unit volume. Furthermore, the viscosity of a hemolymph containing such a concentration as dissociated units would be high. Evolution has solved these problems in two ways. Either:

(a) Make the multisubunit aggregates very large, if they must circulate free in the blood. This is the path taken by many invertebrates.
(b) Alternatively, pack the OTP molecules into circulatory cells, such as the erythrocytes in mammals and other vertebrates.

Either mode prevents loss of the protein due to ultrafiltration in the kidney. Thus, the same effects are achieved either way. Why one or the other evolved in different phyla is something of a mystery. Like the fact that different phyla have evolved different OTPs, this may be simply another example of randomness of evolutionary paths.

5.4   Modulation of Dioxygen Delivery by Oxygen Transport Proteins: Heteroallostery

The optimal functioning of OTPs in animal bodies requires that they deliver the appropriate level of dioxygen to the tissues. This requirement becomes particularly acute under special circumstances, as when, under strenuous exercise, aerobic metabolism is forced to the limits and anaerobic pathways are put to use. To respond to such situations, effector molecules binding elsewhere except at the

active site would help. We have already pointed out that cooperative binding (homeotropic allostery) increases efficiency of dioxygen unloading. But heterotrophic allostery, in which other molecules (allosteric effectors) or ions can modify the transport efficiency, is of equal importance. We give some examples that are quite general over the whole animal kingdom.

5.4.1 Modulation by the Products of Anaerobic Metabolism: the Bohr Effect

In many animals, aerobic metabolism becomes partially replaced by anaerobic metabolism under exercise stress. A common product of the latter is lactic acid (see Chap. 4), which is secreted into the blood. Although lactate itself results in a higher affinity of OTP for dioxygen, the protons released compensate for this effect to make the dioxygen unloading by the OTP more efficient by stabilizing the low affinity state, thus countering the O_2 deficiency. This dependence of the oxygen affinity on the protein concentration ($\log_p 50/\log [H^+]$) is known as the *Bohr Effect* (Bohr et al. 1904). Another contribution to the Bohr Effect comes from the high level of CO_2 transported from tissues to blood when exercise is pronounced. The concentration of protons increases due to the increase of released carbon dioxide in the hemolymph to form H_2CO_3 according to:

$$CO_2 + H_2O \leftrightarrow H_2CO_3$$
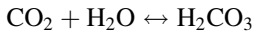
where H_2CO_3 is quickly dissociated by an enzyme, carbonic anhydrase present within erythrocytes, to yield

$$H_2CO_3 \leftrightarrow HCO_3^- + H^+$$

Thus, CO_2 forms HCO_3^- ions and protons (see Sect. 1.8). These protons decrease the pH of the blood or hemolymph, shifting an oxygen-binding curve to the right resulting in a release of dioxygen bound to OTP. This effect happens at around the physiological pH range between 7 and 7.5. The molecular mechanism has been known for decades for hemoglobin (Bohr et al. 1904; Perutz 1970). Of course, if the demand for oxygen continues to exceed the supply, the tissue becomes steadily more acidic until eventually it is unable to function.

5.4.2 The Haldane Effect

We have seen that proton-binding facilitates the release of dioxygen from many OTPs (the Bohr Effect). In converse, the binding of dioxygen (as in the lungs)

releases protons from OTP molecules and makes the blood more acid. This favors the release of CO_2 from bicarbonate, followed by diffusion through the membrane of the erythrocytes and permits it to be exhaled from the lungs

$$HbH^+ + O_2 \leftrightarrow HbO_2 + H^+$$

$$H^+ + HCO_3^- \leftrightarrow H_2O + CO_2$$
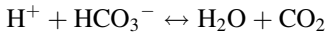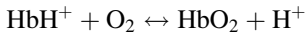
This *Haldane Effect* explains how CO_2 is transferred out of the erythrocytes. It can be considered as the inverse function of the Bohr Effect. Thus, deoxygenation of the blood results in an increase of dissolved carbon dioxide (Haldane and Priestley 1935). Vice versa, oxygenated blood has a reduced capacity for carbon dioxide. This is a consequence of the fact that reduced (deoxygenated) OTP is a proton acceptor but not the oxygenated form. This simple molecular accommodation facilitates both the release of O_2 to the tissues and the release of CO_2 in the lungs, a remarkable case of biochemical efficiency.

5.4.3 The Root Effect

OTPs are not only used for oxygen supply but also to meet other physiological needs. One that might not be expected is the floatation of fishes. Water is an excellent environment for life, but swimming is energy consuming. It is more relaxing for an organism to keep its body buoyancy neutral (that is, having the same overall density as the surrounding water). Then the organism does not have to work to keep its body from floating to the top or sinking to the bottom. A solution to maintain neutral buoyancy was realized in a large number of bony fish by evolving a gas containing bladder as a hydrostatic organ, which can be filled or emptied to achieve the desired density for the whole fish. This effect can be achieved by isoforms of hemoglobin. Here the regulation of the release of dioxygen from blood into the bladder is also controlled by increasing the proton concentration to very low pH values, i.e., 6.5. This function is locally limited to a small region referred to the rete mirabilis (Latin for wondrous network) with specialized epithelial cells (the gas giant cells) sitting on the surface of the gas bladder (Berenbrink et al. 2005; Berenbrink 2007). Specific metabolic pathways are involved, which produce lactic acid via the glycolytic pathway and CO_2 via the pentose phosphate cycle. The pH drop in the *rete mirabilis* releases a large fraction of the dioxygen carried by the specialized hemoglobin isoforms into the bladder.

This remarkable ability, called the *Root Effect* (Root 1931), enables hemoglobins to pump dioxygen from the blood vessel against a high O_2 gradient into the gas bladder. For this, fish have evolved isoforms of hemoglobin with special properties. At low pH such as <6.5, these isoforms of hemoglobins lose the ability to bind dioxygen cooperatively, instead binding dioxygen noncooperatively with a very

low affinity. Thus, dioxygen is released into the blood as physically dissolved dioxygen resulting in a tremendous increase of the pO_2, reaching values of over 700 Torr in the rete mirabilis which are much higher than those in the gas bladder. Thus, a steep dioxygen gradient is created between the blood and the gas bladder, which forces dioxygen to diffuse into the gas bladder. However, the total oxygen-binding capacity of the blood is only slightly reduced. In order not to lose functional dioxygen transport completely under these conditions, the blood must contain at least two different isoforms (Mylvaganam et al. 1996; Weber and Vinogradov 2001; Decker and Hellmann 2007; Verde et al. 2007). One type provides the Root Effect, but the other types possess the normal Bohr Effect being responsible for oxygen supply to tissues. Thus, the Root Effect opens a new evolutionary way using the appropriate physics to overcome the environmental conditions: floating within a liquid.

The Root Effect is also found in organs other than the gas bladder, such has the eyes of some fishes. They have a *choroid rete* (the choroid is the vascular space between the retina and the schlera) (Root 1931; Pelster and Decker 2004). As in the swim bladder, the blood flowing into the choroid is made more acidic, so bound dioxygen is released. The dioxygen diffuses into the eye to support the higher metabolic needs of the retinal cells and associated neurons. The physiological advantage may be for deep-sea fish to maintain high levels of oxygenation in the retinal tissues. The rete isolates both acidity and high dioxygen concentration to the circulation around the eye.

5.4.4   Temperature Dependence

Many organisms, especially invertebrates, adapt their body temperature to the environmental temperature. Changes in the temperature will also change the pH of the blood, and may thereby influence oxygen transport. This, however, is an often neglected topic. Increase in temperature results in a decrease in pH of an aqueous solution. The result is to shift the oxygen-binding curve to the right resulting in a lower affinity. Thus, more dioxygen will be released from OTP which is needed due to higher activity at higher temperatures. But depending on the environmental condition, the optimum dioxygen-binding function may also change. Thus, for each organism, there should be a particular pH–T dependence with respect to cooperativity and affinity. The higher the cooperativity, the more dioxygen will be released from the OTPs within a small change in the pO_2. On the other hand, a broad pH–T landscape guarantees that enough dioxygen will still be provided by OTPs under nonideal conditions. At the extremes, however, not enough O_2 will be released to the tissues even though the OTPs may even be fully loaded with dioxygen. An example is *Octopus*, which will asphyxiate in cold polar seas, when low temperatures result in high affinity, not enough dioxygen will be released to the tissues by the hemocyanin even though much dioxygen is available within those organisms.

High temperature also causes another problem for terrestrial organisms, if accompanied by a decrease in the humidity. Water vapor is important for respiration, since it covers the respiratory areas such as the epithelia cell layer forming the alveoli deep in the lungs. This prevents them from sticking together, so that dioxygen cannot pass. Some spiders live in long narrow tubes at the end of which the temperature may be below 30°C, under very moist conditions, while at the entrance the sun heats up the atmosphere to more than 70°C with a very low humidity. The tube provides a refuge from dehydration.

5.4.5 Evolutionary Aspects of Regulation

The Bohr Effect in fish appears to have been evolved at least 500 MYA when fish first appeared in the oceans, while the Root Effect evolved along with the gas bladder about 350 MYA followed by the choroid rete at about 250 MYA. However, as revealed in a recent study (Berenbrink et al. 2005), evolution of a swim bladder rete occurred independently at least four times among fishes. During the last 100 million years, the swim bladder rete, the choroid rete, and, surprisingly, even the Bohr Effect have been lost several times within various groups of fish and sometimes regained again. However, it seems that the Root Effect is always present when the animal possesses a *rete mirabilis*. It seems that when no *rete mirabilis* has been evolved, then hemoglobins lack the Root Effect. This is observed for hemoglobins from nonteleost fish and from tetrapods and more complex terrestrial vertebrates where you would not expect a *rete mirabilis*. While the structural basis of the Bohr Effect is clear, the basis for the Root Effect is still controversial, since no unique molecular structure to account for it has yet been found (Mylvaganam et al. 1996; Verde et al. 2007).

5.5 Diversity of Oxygen Transport Proteins

As Table 5.1 indicates, there exist four classes of OTPs. The four classes are distributed in a very peculiar fashion among the animal phyla (Fig. 5.7). Hemoglobins are almost ubiquitous, having examples in some species in almost every phylum (Weber and Vinogradov 2001; Brunori and Vallone 2007). On the other hand, hemerythrins are the major OTPs in only two minor phyla, and also occur, but rarely, in annelids. Arthropod and Molluscan hemocyanins are found only in their respective phyla, but some arthropods and molluscs have hemoglobin instead of hemocyanin or have both. Despite their common names, and some common features in the binding site with its copper atoms, the arthropod and molluscan hemocyanins are quite distinct proteins (see Fig. 5.8, and van Holde and Miller 1995; van Holde et al. 2001; Burmester 2002) with no clear evidence of common ancestry (see below). Its ubiquitous distribution indicates that the hemoglobin gene family is very ancient, predating the divergence of most animal phyla. The molluscs
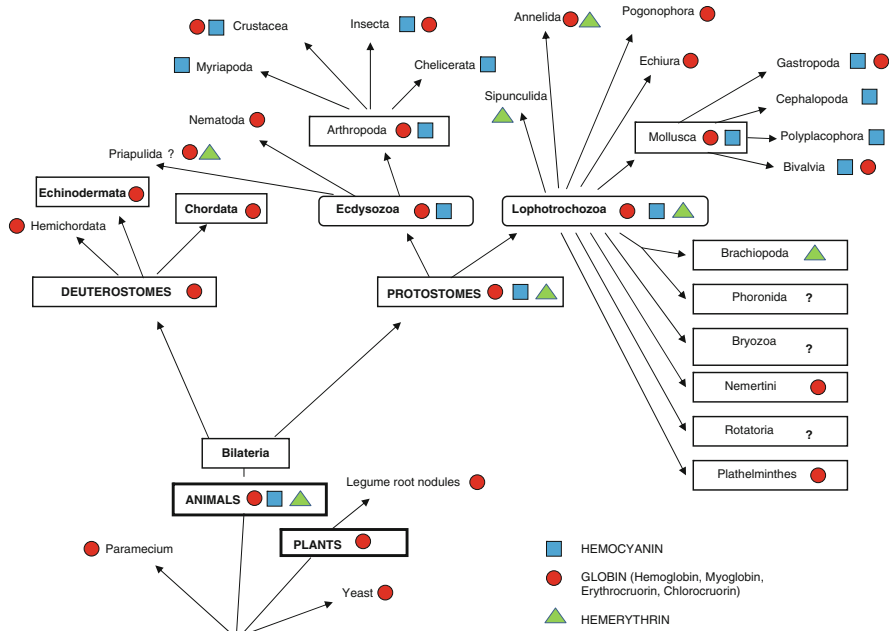
Fig. 5.7 Phylogenetic distribution of oxygen transport proteins. The schematic distribution of the three principal different OTPs (globins, hemocyanins, and hemerythrins) among phyla is shown

and arthropod hemocyanin, on the other hand, must have arisen only after their specific phyla diverged, for there are no examples of "crossing over." We now examine some of those proteins more deeply.

5.5.1 Hemoglobins

The word hemoglobin connotes an enormous class of proteins, of wildly varied structures, but almost surely a common evolutionary ancestry. The name has even been applied to a number of monomeric intracellular heme proteins, which, like myoglobin, seem to store or transmit dioxygen within cells or function as oxygen sensors. These are not, however, OTPs in the sense we use here.

The transport hemoglobins are all formed from myoglobin-like subunits or tandem copies of the globin gene. The number of subunits ranges from 2 (the smallest to allow cooperative binding) to 144. Although the hemoglobins of most organisms are carried in circulatory cells, some of the very large ones are free in the blood. Many molecular structures have been determined. All vertebrate animals (including ourselves) have the kind of tetrameric hemoglobin depicted in Fig. 5.9 (Perutz 1970). However, some invertebrates have much more elaborate oxygen transporters. The hemoglobin of the common earthworm *Lumbricus terrestris* has the enormous, elegant structure shown in Fig. 5.7. It contains 144 globin chains,
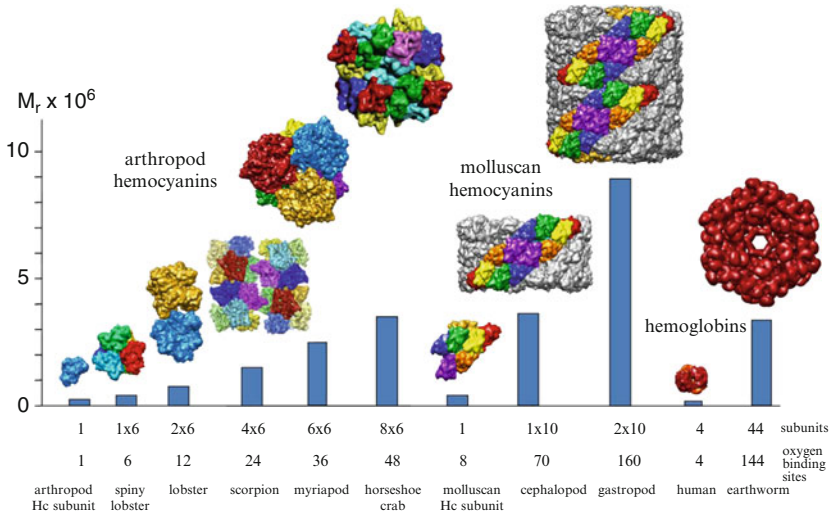
Fig. 5.8 Schematic survey of the huge extracellular hemoglobin and hemocyanin molecules. The numbers of the subunits and the oxygen-binding sites are given. The *blue bars* represent the approximate molecular masses in each class. The tetrameric human hemoglobin occurs within the erythrocytes (Courtesy of Jürgen Markl)
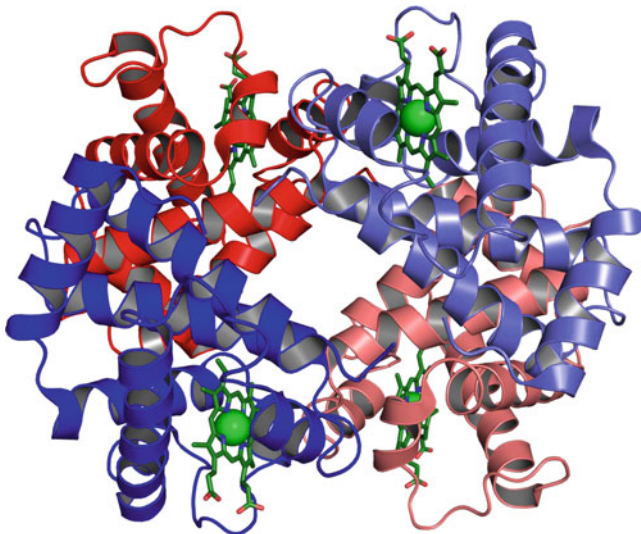
Fig. 5.9 Human hemoglobin structure. The tetrameric human hemoglobin is composed of two different subunit types: two α (*blue*) and two β (*red*). Each one binds one dioxygen by an iron atom (*green ball*) in the center of a porphyrin ring (*green*) being fixed in a pocket of the globin

each with a heme oxygen site. In addition, there are 36 "linker" chains, which form a collar-like structure holding the whole ensemble together (Royer et al. 2006). A quite different example of hemoglobin complexity is that found in the brine shrimp, *Artemia* (Coleman et al. 2001; Chyou et al. 2007). Here, the multisubunit structure is formed not by the self-association of globin units, but by a repeated gene duplication that has led to the formation of tandem arrays of nine globin gene units, which translate into monomeric protein domains. There actually are two varieties of arrays, which combine to form an 18-unit "necklace-like" structure. Recently, hemoglobin was also found in the insects where it was thought for a long time to be unnecessary since the trachea system is very effective (Burmester and Hankeln 2007).

We have emphasized these examples of an enormous diversity to stress the following point: hemoglobin function, as an oxygen transporter, seems to have been invented again and again in different phyla. However, this by no means describes the limits of evolutionary invention for OTP function. Consider the other completely different classes of OTPs.

5.5.2 Hemerythrins

The hemerythrins, found only in a few invertebrate phyla (Fig. 5.7), represent a class of iron-containing OTPs completely distinct from the hemoglobins (French et al. 2008; Vanin et al. 2006; Meyer and Lieb 2010). The oxygen-binding site does not contain a heme, but a pair of Fe(2) atoms (see Fig. 5.2, Wirstam et al. 2003). The molecular structure of the basic hemerythrins subunit is entirely unlike that of the globins. The protein folds in the three dimensions to a common motif, a four-helix bundle, which holds the residues that coordinate the iron. Hemerythrins exist in either trimeric or octomeric structures and there is a monomeric myohemerythrin analogous to myoglobin. Recent evidence indicates that present hemerythrins have evolved from such monomeric precursors (Vanin et al. 2006). Hemerythrin may be evolutionary ancient (considering the phyla in which it is found), but it seems to have represented a rather unsuccessful evolutionary experiment.

5.5.3 Hemocyanins

Not only iron is used for binding oxygen by OTP, two types of proteins use copper. The mode of binding is surprising, since the two copper atoms at the active site occur as Cu(I) with a fully saturated d_{10} shell. Normally, such an atom would be inert; however, dioxygen does bind reversibly as peroxide, via electron transfer, forming a Cu(II)-O_2^{2}-Cu(II) center (see Fig. 5.2). There exist two copper-based OTPs, the arthropod and molluscan hemocyanins. Each binds the dioxygen in "side on" coordination (Fig. 5.2). The change in electronic states of the copper ions in oxygenated hemocyanin determines another feature: they turn deep blue upon

oxygenation. However, molluscan and arthropod hemocyanins have little else in common. They are wholly different in their molecular structures. Although the similarity at the dioxygen-binding site suggests some distant relation between the proteins, some authorities consider them to be unrelated (Burmester 2002; Burmester et al. 2001). However, recently, a comparison of the folding motifs revealed that hemocyanins from both arthropods and molluscs possess a cupredoxin-like domain most likely having served for copper loading at the active site at early times. During evolution, adaption of this domain results in different functions such as connecting elements or allosteric force transfer (Jaenicke et al. 2010).

5.5.3.1 Molluscan Hemocyanin

A unique class of hemocyanins is found only in molluscs. However, these are not the only OTPs in the phylum. Some molluscs have hemoglobins instead. Some snails have both, with hemocyanin in the hemolymph and myoglobin as an oxygen storage protein in the radula. Molluscan hemocyanins show a common structural pattern (Fig. 5.10). Each is made up of 10 subunits which may dimerize to a 20-subunit hemocyanin. Each subunit in turn represents tandem arrays of seven to eight "functional units" (FUs) in one enormous polypeptide chain. Each FU contains one oxygen-binding site: a copper pair as shown in Fig. 5.2. Thus, an intact molluscan hemocyanin may contain up to 160 FUs. This means that these are very large molecules since each FU has a molecular mass of about 50,000 Da (Cuff et al. 1998; Perbandt et al. 2003). The 10- or 20-mers, which are the functional molecules with molecular masses up to 8 million Dalton, are cylindrical structures with fivefold symmetry. A representative structure is shown in Fig. 5.10, along with the higher resolution structure of one FU that has been determined (Gatsogiannis et al. 2009). Molluscan hemocyanins are always found dissolved in the blood, often in very high concentrations.
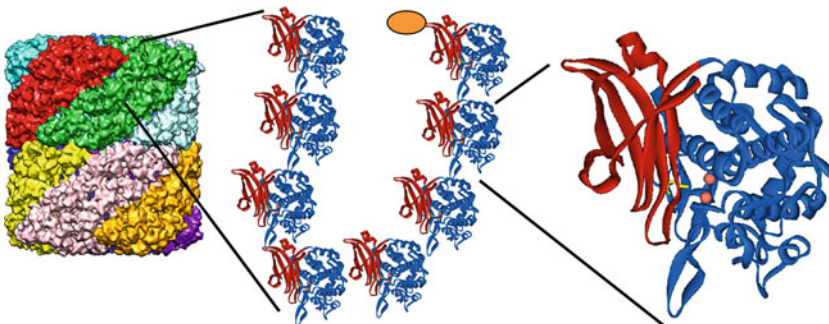
Fig. 5.10 Molluscan hemocyanin molecule. The largest hemocyanin with 160 oxygen-binding sites (from the keyhole limpet hemocyanin KLH1) is composed of subunits (shown in different colors) each with eight functional units (FUs). Each FU folds into two different domains, the *blue* one carries the active site with copper pair shown as *orange* balls (right picture). The C-terminal cupredoxin domain is shown as orange loop (middle picture) (Courtesy of Jürgen Markl)

5.5.3.2 Arthropod Hemocyanins

Among arthropods, hemocyanins are mostly found in crustaceans such as lobsters and crabs and chelicerates (such as spiders, scorpions and horseshoe crabs) but also in a few insects and isopods (Hagner-Holler et al. 2004; van Holde and Brenowitz 1981). While in the latter two the function is still uncertain; all other arthropod hemocyanins serve as OTPs. They represent the major fraction of proteins dissolved in the hemolymph of crustaceans and chelicerates. The smallest natively occurring hemocyanins consist of six subunits (Markl and Decker 1992). Each subunit is folded into three domains characterized by different folding motifs. The first consists of five to six α-helices, the second is a "four α-helix bundle" carrying the active site and the third fold in another motif, a β-barrel. The active site domain bears a slight, but intriguing similarity to hemerythrins (see above). Three of these kidney-shaped subunits associate back to back as a trimer, establishing a threefold rotational axis (Fig. 5.11). Two of these trimers dimerize in a sandwich fashion along the threefold axes, being rotated by 60° on this axes with respect to one another. Depending on the species, these hexamers or integer multiples of them are found (1 × 6, 2 × 6, 4 × 6, 6 × 6, and 8 × 6; Fig. 5.11). These structures provide the basis for cooperative oxygen binding, including examples of the strongest cooperativity observed in nature. In addition, only assembled hemocyanins possess broad functional properties (Fig. 5.11). Interestingly, only even multiples of hemocyanins hexamers have been found in vivo, but never intermediate such as dimers or

Fig. 5.11 Arthropod hemocyanin structures. The basic structure is a hexamer having a threefold symmetry. Multiples of hexamers (1 × 6, 2 × 6, 4 × 6, 6 × 6, 8 × 6) are found depending on the species. Each subunit consists of three domains with different folding motifs represented by different colors (*green*, *red*, *blue*). The middle domain (*red* contains the active site). In the enlargement of the second domain (to *right*), the coppers are represented by *orange* balls, dioxygen in *blue* (Courtesy of Barbara Markl and Elmar Jaenicke)

subunit

active site within a four α helix bundle
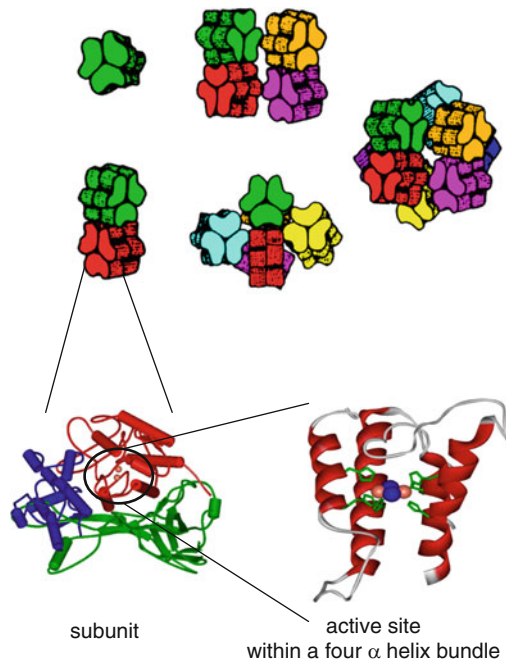
trimers. This requires sophisticated control mechanisms. The control operates through the self-assembly process; mis-folded subunits will not be incorporated.

A new technique, the cryoelectron microscopy, allowed to resolve the structure of these very large structures (Gatsogiannis et al. 2007; Cong et al. 2009; Martin et al. 2007; Markl et al. 2009; Lieb et al. 2010; Gatsogiannis and Markl 2009; see also Figs. 5.7 and 5.10).

5.6   Evolution of Oxygen Transport Proteins

As we have argued in Sect. 5.1, OTPs become essential only for macroscopic animals utilizing aerobic metabolism. Very small or very thin animals do not need them, and they could not have been advantageous until dioxygen levels in the sea had raised above the *Pasteur point* or thereabout, a concentration of dioxygen at which microorganisms (aerobic or facultative anaerobes) adapted aerobic respiration. At this point, mechanisms were evolved to integrate dioxygen in the metabolism to gain much more energy (see Chaps. 3 and 4). The Pasteur point at about a concentration of 1% of the present atmospheric concentration corresponds to a partial oxygen pressure at about 1.5–2 Torr which is at least necessary to enable any diffusion of dioxygen into the mitochondria for generating ATP molecules.

OTPs became especially important when predation began to occur, for they allowed the development of shells and protective carapaces. In fact, it can be argued that macroscopic animals could not have developed such protection without OTPs. Thus, it is clear why OTPs have evolved and become established in all higher phyla; there are many selective advantages. The fact that several different kinds of OTPs have evolved independently implies that several of the animal phyla had differentiated before the Cambrian explosion, at which point body armor is first observed. The questions as to how and exactly when the proteins evolved are more difficult to answer. As we have seen above, it is likely that many of the OTPs arise from monomeric oxygen-binding proteins (OBP) – a myoglobin-like protein for hemoglobin, myohemerythrin for hemerythrins, and perhaps a monomeric tyrosinase for the molluscan hemocyanins and a phenoloxidase for arthropod hemocyanin (Terwilliger and Decker 2000; van Holde et al. 2001; Lieb et al. 2001).

In any event, the critical step in going from a simple OBP to a complex OTP is establishing a protein structure capable of cooperative binding. As Fig. 5.3 shows cooperative binding is important for efficient oxygen delivery, and it is a feature of almost all OTP. The simplest structure to accomplish this is one with two binding sites. Such a protein can be evolved in at least two ways: The protein product of a tandemly duplicated gene may be folded into a tandem polypeptide with two or more identical sequences. If these protein domains are capable of interacting in two ways, with different binding strengths, the prerequisites for allostery and cooperative binding are satisfied. Alternatively, a copy of the monomeric genes may undergo a mutation, so as to lead to noncovalent dimerization (or higher

oligomerization). Again, if there exist conformations for such a complex with different oxygen affinities, cooperative binding is possible. Both pathways have been used, but B seems to be much more common. Most OTPs are noncovalently linked oligomers.

When OTPs evolved is a much more difficult question to answer. We can, however, put some logical constraints on the problem:

(a) An oxygen transport system is a prerequisite for the development of shells or hard carapaces. It is not possible to breathe through such, so transport from gills or lungs is required. This either requires or is facilitated by an OTP. This means, in turn, that the development of OTPs, at least, in molluscs and arthropods, must have occurred *before* the appearance of shelly fossils and hard-bodied arthropods in the early Cambrian.

(b) Because different phyla use quite different OTPs, which in many cases have no common protein ancestry, we conclude that many or most OTPs evolved *after* the divergence of major phyla.

(c) The independent emergence and presence of different OTPs in different phyla at about the same time argues that there was a common evolutionary challenge faced by all organisms.

(d) To form OTPs, nature has evolved strategies to bind dioxygen without reacting it chemically. For this, two metals seem to be advantageous – iron and copper. Iron was readily available in the earliest oceans, whereas copper became accessible only after the great oxidation event. This implies an earlier origin for the iron-based OTPs.

These arguments do not take us very far in assigning dates but together are rather inconsistent with the traditionalist view that the divergence of major animal phyla occurred at the beginning of the Cambrian, the period often referred to as the "Cambrian explosion." We would argue that the appearance of fossils of both arthropods and molluscs at this time is a consequence of the development of hard carapaces (and increase in size) by phyla that had already independently existed for some time. The possession of a shell requires the *prior* existence of an oxygen transport system. It just does not work the other way around. But that also means that a kind of pump (a heart) and at least a rudimentary circulation system had to have been evolved already. Likely protein precursors to OTPs seem to have developed early. For example, tyrosinase have the same active site as hemocyanins (van Holde et al. 2001; Decker et al. 2007) and are present already in sponges which arose about 800 MYA (Müller et al. 2004). If the OTPs are of Precambrian origin, the question is how early? There is a growing body of molecular biological evidence.

Recall that the molluscan hemocyanin molecule is constructed from subunits, which are in turn tandem arrays of "functional units." These units are all clearly derived from a common ancestral gene, and we can follow the evolution of a given unit through the well-documented evolution of molluscs. The times of divergence of molluscan classes are quite well known from the fossil record. This provides a "time scale" for evolution of the molluscan hemocyanin. On this basis, we find that the initial divergence of the functional units predates any known molluscan fossil.
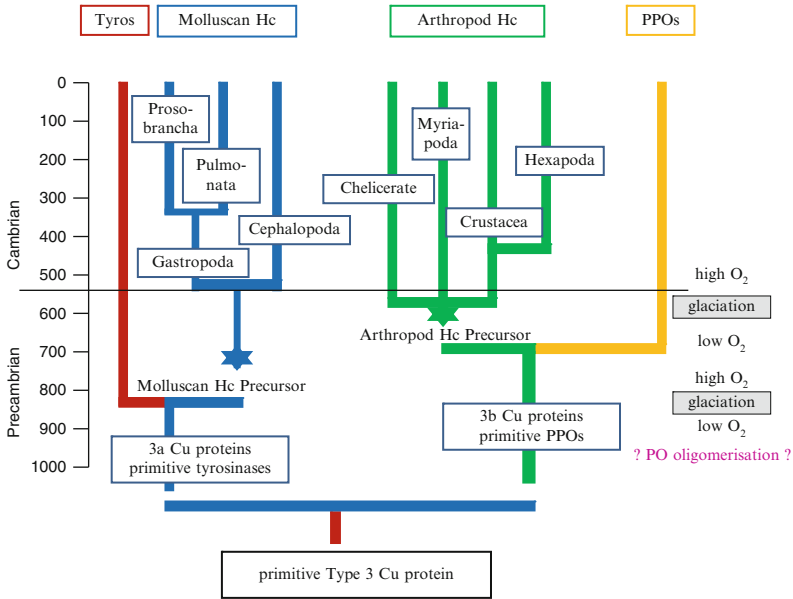
Fig. 5.12 Postulated evolution of hemocyanin. Primitive type 3 copper proteins may have evolved to tyrosinases (Tyros) or related polyphenoloxidases (PPO). They can be considered as the ancient precursors of hemocyanins (adapted from van Holde et al. 2001)

In fact, a date of approximately 700 MYA is indicated (Fig. 5.12). This corresponds to or predates the great Proterozoic glaciation, but postdates the evolution of sponges. A very similar date can be ascribed to the initial divergence of the globin units of the hemoglobin of branchiopod arthropods. The origin of arthropod hemocyanins has likewise been estimated between 600 and 700 MYA (Fig. 5.12). We believe that during this period, all or most of the existing classes of OTPs evolved from monomeric ancestors. The problem with this hypothesis is the paucity of the fossil record in this period which indicates that the organisms in which these evolved must have been small and soft bodied. If so, why did they evolve OTPs? We realize that there is a strong scientific opposition to this view (see for example: Conway-Morris 2006). Nonetheless, the idea that complex organisms with functioning circulatory systems sprang into full-blown existence from nothing at the "Cambrian explosion" also seems unreasonable.

5.7 Was Snowball Earth a Possible Trigger for OPT Evolution?

We have argued above that protein-mediated oxygen transport must have evolved prior to 550 MYA, but certainly not earlier than about 800 MYA, the age of the

oldest fossil sponges. An upper limit of 700 MYA seems to be credible. Interestingly, this span corresponds closely to two great worldwide glaciations – the Sturtian (710–680 MYA) and the Varanjer-Marinean (605–585 MYA) (see Chap. 6 for details). It has been suggested (see for example, Fedonkin 2003) that the "snowball Earth" episodes, which must have been ecological disasters for many species, may have in fact played a role in triggering the evolutionary bursts that followed. We could suggest that rather than "triggers" the glaciation acted as "screens," sorting out the fit from the unfit.

The covering of much of Earth's surface and seas with ice would have greatly inhibited photosynthesis with consequent decrease in oxygen levels in the oceans (Fedonkin 2003). This then required that aerobes to survive must remain very small, very thin, or develop oxygen transport systems to more efficiently extract the little dioxygen available. It is clear that all Precambrian animals remained small (except Ediacarans, which, however, were thin), but we suggest that the most successful were those that also developed dioxygen transport. These then survived into the Cambrian and were readily able to develop protective armor and other hard parts against the predation that obliterated the soft-bodied Ediacarans. One need not postulate a mysterious evolutionary "explosion," rather the opportunity for a number of kinds of creatures to survive and occupy new niches that arose after the glaciers left. There might have been a good chance that without the extinction of many species by two glaciations, Earth would still be populated by primitive organisms.

5.8   From What Proteins Did Oxygen Transport Proteins Evolve?

Because OTPs would have no function until fairly complex aerobes had evolved, we must assume that the proteins themselves are derived from some previously existing proteins; but which ones? For the hemoglobins, Vinogradov et al. (2007) have carried out an extensive analysis of many eukaryotic, bacterial, and archeal proteins. Their conclusion is that all existing hemoglobin may have arisen from a very ancient oxygen sensor protein, in which a heme is functionally coupled to a flavin moiety which transmits the signal. At first, these may have signaled oxygen warning, but as aerobes appeared, led to oxygen seeking. On the other hand, hemocyanins may have developed as a result of modification of the enzymes phenoloxidases, especially tyrosinases. They were needed to neutralize increasing concentrations of dioxygen by forming diphenols and quinones which are also necessary for hardening the cuticula of arthropods. Phenoloxidases share with hemocyanins the same active site and activation mechanism (Decker et al. 2007; Decker and Tuczek 2000). Thus, in all cases, it is at least possible to assign reasonable precursors, which were carrying out useful functions before the need for oxygen transport arose.

5.9   Oxygen Transport Proteins and "Intelligent Design"

There is a branch of creationist thought that contends that Darwinism cannot, in principle, explain the existing diversity and complexity of organisms and their metabolism. Rather, it is proposed that evolution has been guided by an "intelligent designer" who has helped solve the problems encountered in the development of complex structures and processes. It is illuminating to consider the evolution of OTPs in this regard. Recall that there exist several classes of such proteins, which bear essentially no structural similarity to one another. Also note that the hemoglobin gene is very ancient and appears to be carried by all higher metazoans including in particular the arthropods and molluscs. But almost all species of arthropods and molluscs that employ a transport protein use hemocyanins for this purpose. One would suppose that the "designer," were he/she/it intelligent, would have simply used the already available globin gene to fashion transport proteins in all phyla, rather than having to invent new solutions for the arthropods and molluscs (and hemerythrins for some minor phyla). That hemoglobin could have been employed is evidenced by the fact that a few arthropods and molluscs do use unique hemoglobins. It is not that the hemocyanins possess some intrinsic advantage over hemoglobin. In fact, in terms of efficiency as oxygen carriers (mass of protein needed to transport one oxygen molecule), the hemoglobins are three to four times more efficient than the hemocyanins. Everything that hemocyanins can do hemoglobins can do at less cost to the organism. The additional mass is believed to be necessary for building up the contact areas between the subunits establishing cooperative interactions.

One must then conclude that the "designer," if she/he/it exists, is either sloppy or not very intelligent. The idea of a "dumb designer" should be unattractive to even the most ardent creationist. On the other hand, the peculiar kind of phylogenetic distribution of OTPs we observe is exactly what would have been expected from Darwinian natural selection. When circumstances gave advantage to those phyla which had already evolved OTPs, those organisms survived. But there are many ways to develop an OTP from preexisting proteins by simple mutation. Several different ones happened in separate phyla and their carriers were subsequently selected for.

References

Berenbrink M (2007) Historical reconstructions of evolving physiological complexity: O_2 secretion in the eye and swimbladder of fishes. J Exp Biol 210:1641–1652

Berenbrink M, Koldkjaer P, Kepp O, Cossins AR (2005) Evolution of oxygen secretion in fishes and the emergence of a complex physiological system. Science 307:1752–1757

Bohr C, Hasselbach KA, Hrogh A (1904) Über einen in biologischer Beziehung wichtigen Einfluß, den die Kohlensauerstoffspannung des Blutes auf dessen Sauerstoffbindung übt. Skand Arch Physiol 15:401–412

Burmester T (2001) Molecular evolution of the arthropod hemocyanin superfamily. Mol Biol Evol. 18:184–195.

Brunori M, Vallone B (2007) Neuroglobin, seven years after. Cell Mol Life Sci 64:1259–1268

Burmester T (2002) Origin and evolution of arthropod hemocyanins and related proteins. J Comp Physiol B 172:95–107

Burmester T, Hankeln T (2007) The respiratory proteins of insects. J Insect Physiol 53:285–294

Chyou DT, Rawle VL, Trotman CN (2007) Quaternary structure of *Artemia* haemoglobin II: analysis of T and C polymer alignment and interpolymer interface. BMC Struct Biol 18:7–26

Coleman M, Matthews CM, Trotman CN (2001) Multimeric hemoglobin of the Australian brine shrimp P*arartemia*. Mol Biol Evol 18:570–576

Cong Y, Zhang Q, Woolford D, Schweikardt T, Khant H, Dougherty M, Ludtke SJ, Chiu W, Decker H (2009) Structural mechanism of SDS-induced enzyme activity of scorpion hemocyanin revealed by electron cryomicroscopy. Structure 17:749–758

Conway-Morris S (2006) Darwin's dilemma: the realities of the Cambrian "explosion". Philos Trans Roy Soc B 361:1069–1083

Cuff ME, Miller KI, van Holde KE, Hendrickson WA (1998) Crystal structure of a functional unit from *Octopus* hemocyanin. J Mol Biol 278:855–870

Decker H, Hellmann N (2007) Negative cooperativity in root-effect hemoglobin: role of heterogeneity. Integr Comp Biol 47:656–661

Decker H, Tuczek F (2000) Phenoloxidase activity of hemocyanins: activation, substrate orientation and molecular mechanism. Trends Biochem Sci 25:392–397

Decker H, Sterner R (1990) Hierarchien in der Struktur und Funktion von sauerstoffbindenden Proteinen. Naturwissenschaften 77:561–568

Decker H, Schweikardt T, Nillius D, Salzbrunn U, Jaenicke E, Tuczek F (2007) Similar activation process and catalysis in hemocyanins and tyrosinases. Gene 398:183–191

Farmer CS, Kurtz DM, Jr LZJ, Wang BC, Rose J, Ai J, Sanders-Loehr J (2001) The crystal structures of *Phascolopsis gouldii* wild type and L98Y methemerythrins: structural and functional alterations of the O_2 binding pocket. J Biol Inorg Chem 6:418–429

Fedonkin MA (2003) The origin of the Metazoa in the light of the Proterozoic fossil record. Paleontol Res 1:9–41

French CE, Bell JM, Ward FB (2008) Diversity and distribution of hemerythrin-like proteins in prokaryotes. FEMS Microbiol Lett 279:131–145

Gatsogiannis C, Markl J (2009) Keyhole limpet hemocyanin: 9-A CryoEM structure and molecular model of the KLH1 didecamer reveal the interfaces and intricate topology of the 160 functional units. J Mol Biol 385:963–983

Gatsogiannis C, Moeller A, Depoix F, Meissner U, Markl J (2007) *Nautilus pompilius* hemocyanin: 9 Å cryo-EM structure and molecular model reveal the subunit pathway and the interfaces between the 70 functional units. J Mol Biol 374:465–486

Hagner-Holler S, Schoen A, Erker W, Marden JH, Rupprecht R, Decker H, Burmester T (2004) A respiratory hemocyanin from an insect. Proc Natl Acad Sci USA 101:871–874

Haldane JS, Priestley SJ (1935) Respiration. Oxford Press, Oxford, UK

Hazes B, Magnus KA, Bonaventura C, Bonaventura J, Dauter Z, Kalk KH, Hol WG (1993) Crystal structure of deoxygenated *Limulus polyphemus* subunit II hemocyanin at 2.18 A resolution: clues for a mechanism for allosteric regulation. Protein Sci 2:597–619

Jaenicke E, Büchler K, Markl J, Decker H, Barends TRM (2010) Cupredoxin-like domain in haemocyanins. Biochem J 426(3):373–378

Koshland DE, Nemethy G, Filmer D (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 5:365–385

Lieb B, Altenhein B, Markl J, Vincent A, van Olden E, van Holde KE, Miller KI (2001) Structures of two molluscan hemocyanin genes: significance for gene evolution. Proc Natl Acad Sci USA 98:4546–4551

Lieb B, Gebauer W, Gatsogiannis C, Depoix F, Hellmann N, Harasewych MG, Strong EE, Markl J (2010) Molluscan mega-hemocyanin: an ancient oxygen carrier tuned by a ~550 kDa polypeptide. Front Zool 7:14

Magnus KA, Hazes B, Ton-That H, Bonaventura C, Bonaventura J, Hol WG (1994) Crystallographic analysis of oxygenated and deoxygenated states of arthropod hemocyanin shows unusual differences. Proteins 19:302–309

Markl J, Decker H (1992) Molecular structure of arthropodan hemocyanins. Adv Comp Environ Physiol 13:325–376

Markl J, Moeller A, Martin AG, Rheinbay J, Gebauer W, Depoix F (2009) 10-A cryoEM structure and molecular model of the Myriapod (Scutigera) 6x6mer hemocyanin: understanding a giant oxygen transport protein. J Mol Biol 392:362–380

Martin AG, Depoix F, Stohr M, Meissner U, Hagner-Holler S, Hammouti K, Burmester T, Heyd J, Wriggers W, Markl J (2007) *Limulus polyphemus* hemocyanin: 10 Å cryo-EM structure, sequence analysis, molecular modelling and rigid-body fitting reveal the interfaces between the eight hexamers. J Mol Biol 366:1332–1350

Meyer A, Lieb B (2010) Respiratory proteins in *Sipunculus nudus* – implications for phylogeny and evolution of the hemerythrin family. Comp Biochem Physiol B 155:171–177

Monod J, Wyman J, Changeaux JP (1965) On the nature of allosteric transitions: a plausible model. J Mol Biol 12:88–118

Müller WE, Grebenjuk VA, Thakur NL, Thakur AN, Batel R, Krasko A, Müller IM, Breter HJ (2004) Oxygen-controlled bacterial growth in the sponge *Suberites domuncula*: toward a molecular understanding of the symbiotic relationships between sponge and bacteria. Appl Environ Microbiol 70:2332–2341

Mylvaganam SE, Bonaventura C, Bonaventura J, Getzoff ED (1996) Structural basis for the root effect in haemoglobin. Nat Struct Biol 3:275–283

Pelster B, Decker H (2004) New physiological insights in the root effect. Micron 35:73–74

Perbandt M, Guthöhrlein EW, Rypniewski W, Idakieva K, Stoeva S, Voelter W, Genov N, Betzel C (2003) The structure of a functional unit from the wall of a gastropod hemocyanin offers a possible mechanism for cooperativity. Biochemistry 42:6341–6346

Perutz MF (1970) Stereochemistry of cooperative effects in haemoglobin. Nature 228:726–739

Raff RA, Raff EC (1970) Respiratory mechanisms and the metazoan fossil record. Nature 228:1003–1005

Robert CH, Decker H, Richey B, Gill SJ, Wyman J (1987) Nesting: hierarchies of allosteric interactions. Proc Natl Acad Sci USA 84:1891–1895

Root RW (1931) The respiratory function of the blood of marine fishes. Biol Bull 61:427–456

Royer WE Jr, Sharma H, Strand K, Knapp JE, Bhyravbhatia B (2006) *Lumbricus* erythrocruorin at 3.5 A resolution: architecture of a megadalton respiratory complex. Structure 14:116–1177

Runnegar B (1991) Nucleic acid and protein clocks. Philos Trans R Soc Lond B Biol Sci 333:391–397

Sidell BD, O'Brien KM (2006) When bad things happen to good fish: the loss of hemoglobin and myoglobin expression in Antarctic icefishes. J Exp Biol 209:1791–1802

Terwilliger N, Decker H (2000) COPs and robbers: evolution of copper oxygen proteins. J Exp Biol 203:1777–1782

van Holde KE, Brenowitz M (1981) Subunit structure and physical properties of the hemocyanin of the giant isopod *Bathynomus giganteus*. Biochemistry 20:5232–5239

van Holde KE, Miller KI (1995) Hemocyanins. Adv Protein Chem 47:1–81

van Holde K, Miller KI, Decker H (2001) Hemocyanins and invertebrate evolution. J Biol Chem 276:15563–15566

Vanin S, Negrisolo E, Bailly X, Bubacco L, Beltramini M, Salvato B (2006) Molecular evolution and phylogeny of sipunculan hemerythrins. J Mol Evol 62:32–41

Verde C, Warage S, Guordano D, Mazzarella D, di Prosco G (2007) The root-effect – structural and evolutionary perspective. Antarct Sci 19:271–278

Vinogradov SN, Hoogewijs D, Bailly X, Mizuguchi K, Dewilde S, Moens L, Vanfleteren JR (2007) A model of globin evolution. Gene 398:132–142

Weber RE, Vinogradov SN (2001) Nonvertebrate hemoglobins: functions and molecular adaptations. Physiol Rev 81:569–628

Wirstam M, Lippard SJ, Friesner RA (2003) Reversible dioxygen binding to hemerythrin. J Am Chem Soc 125:3980–3987

Chapter 6
Climate Over the Ages; Is the Environment Stable?

Contents

As described in Chaps. 3 and 4, the advent of oxygenic photosynthesis triggered worldwide environmental changes. A world that had been reductive passed over into a state in which free dioxygen was available in the oceans and the atmosphere. We have already described the likely catastrophic effects on an anaerobic biota, but the changes were much broader than that. Dioxygen in the seas led to major changes in seawater chemistry. Iron, which had previously been soluble as ferrous salts, was precipitated in the ferric form. Copper, which had been insoluble in the anaerobic ocean as cuprous sulphide (Cu^+-state), now became moderately soluble in the cupric form (Cu^{++}-state). Other changes in metal behavior that must have occurred are described in Fig. 2.3 and summarized in Table 6.1. In turn, life forms changed during this period depending on changes in available resources.

Not only did the altering environment influence evolution, but in turn, the evolution of new forms and new metabolisms helped reshape the environment. Indeed, over the past 3 billion years, there has been continual interplay between life and the environment that has often been manifested in abrupt and drastic changes in the Earth's climate. It is the aim of this chapter to provide an overview of those changes, and to raise the questions: has a stable equilibrium been reached? Can a stable equilibrium be achieved? In which way is dioxygen involved?

One way to introduce these questions is to use as markers the series of massive glaciations that have interrupted the past of the Earth (Fig. 6.1). We shall see that these glacial episodes are likely linked to CO_2 levels in the atmospheres, and that those, in turn, correlate with changes in dioxygen content. Here, only the most

Table 6.1 State of element in reducing and oxidizing environment (see Frausto da Silva and Williams 2001)

Element	Reducing environment	Oxidizing environment
Iron	Fe (II) (high)	Fe (III) (low)
Copper	Cu(I) (low)	Cu(II) (moderate)
Sulphur	HS (high)	SO_4^2 (high)
Molybdenum	$(MoO_nS_{4n})^2$, MoS_2 (low)	MoO_4^2 (moderate)
Vanadium	V^{3+}, V(IV) sulphides (low)	VO_4^3 (moderate)

Note that each metal develops different chemical properties depending on the environment with respect to the presence or absence of oxygen
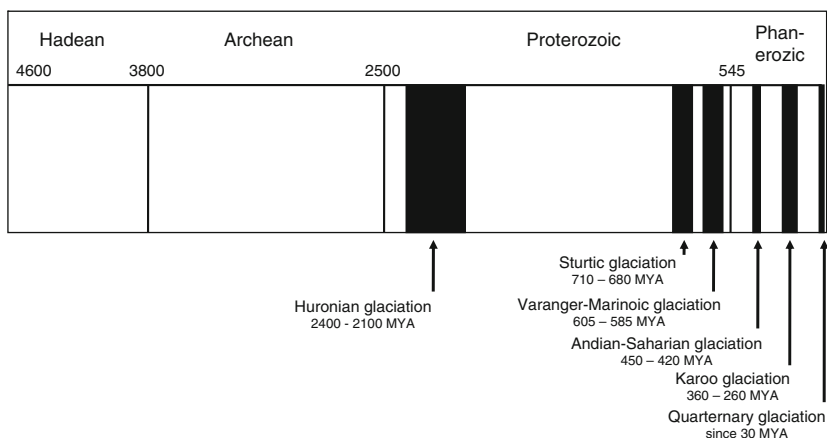

Fig. 6.1 Periods of glaciation. The major glaciations are indicated in *black*. The periods and names are given

drastic glacial episodes are emphasized, in which a major portion of the Earth's surface was covered by ice or snow, the so-called "Snowball Earth" events. It should not be surprising that this can happen; for glaciations have a positive feedback feature. The more surface that is covered by ice or snow, the more sunlight is reflected, thus leading to more cooling. It is obvious that counteracting processes must also be at play, or the Earth and all life would have long ago frozen. Section 6.5 will describe these contradictory processes in more detail.

6.1 Climate and Glaciations in Earth's History

6.1.1 The First Massive Glaciations; the Huronion Event: A Role for Methane?

Although there is some evidence for glaciations as early as 2.9 BYA (see Kasting and Ono 2006), the first clearly identifiable major glacial period occurred about 2.45–2.2 BYA (the Huronion Event). It appears that just previous to that time, the

Earth was very warm although probably not as warm as it had been around 4.2–3.0 BYA (60–80°C). Between 2.7 and 2.5 BYA, it was only a bit cooler, probably about 60°C (Kasting and Ono 2006). Such temperatures are understandable, for there must have existed at that time a massive greenhouse effect. Not only was carbon dioxide abundant in the atmosphere, but an even more effective greenhouse gas, methane, would have been present under the reducing conditions preceding this date (Kasting 2001, 2004; Pavlov et al. 2000, 2003; Fig. 6.2). A CO_2 concentration with about 50 times the present value would have been necessary to keep the Earth at such a temperature. But relatively small amounts of methane would have remarkably reduced the CO_2 requirement. It has been argued that the methane levels in the Archaen atmosphere were many times those found today. A very schematic view of these changes in atmospheric composition is shown in Fig. 6.2.

Today the concentration of these greenhouse gases are low: CO_2 (380 ppm) is roughly 220 times more concentrated than methane (1.7 ppm) and water vapor about 6,000 times less (Kasting 2004). 380 ppm means 380 molecules of CO_2 among one million gas molecules. Although this number seems insignificant, it has a significant effect on the climate.

Methane is a product not only of volcanoes (to a very small extent) but also by about 100-fold more from methanogenic bacteria. Methane-producing microbes make up nearly half of all known surviving Archaebacteria and are thought to have been among the first organisms to have evolved. The recently discovered marine bacterium *Methanosarcina acetivoraies* may be representative of that primitive anaerobic population (Moran et al. 2005; Ferry and House 2006). It utilizes carbon monoxide (probably also abundant in the Archaean) to produce methane. Such organisms were probably very abundant in the anoxic Archean ocean.
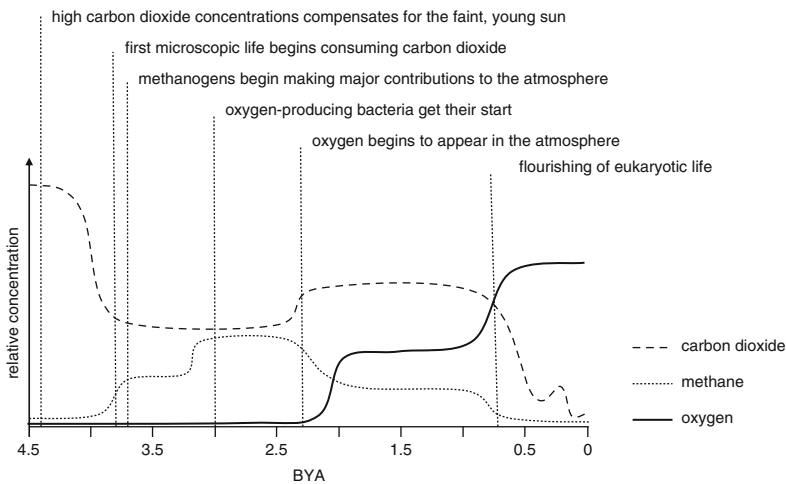

Fig. 6.2 An approximate representation of the relative concentrations of atmospheric gases over the Earth's history. Only long-term trends are shown. For more details in Phaenrozoic, see Fig. 6.3 (Kasting 2004)

It is perhaps not surprising that the Huronion Event coincides quite closely with the "Great Oxidation Event" at about 2.2 BYA, at which point free dioxygen first appeared in significant quantity in the atmosphere (Holland 2006). This would have quickly diminished methane (perhaps within 100,000 years) and many methanogens. Newly produced methane would survive only 10 years in the present oxygen atmosphere, since methane will be destroyed both by photolysis and by the hydroxyl radical OH itself, a product of atmospheric dioxygen.

Thus, the strong greenhouse gas methane disappeared and the bloom of oxygenic photosynthesizers would have drawn down carbon dioxide from the atmosphere. At this time, there was probably no efficient aerobic metabolism to replace carbon dioxide. Thus, the greenhouse warming probably collapsed totally, and with it the Earth's temperature, perhaps to as low as $-50°C$. Ice formed even at the equator, of the order of several hundreds of meters thick (Kirschvink et al. 2000). Many organisms must have died, although some found ways to survive (see Sect. 6.2).

It is interesting to imagine what might have happened had oxygenic photosynthesizers never evolved. In such a case, if CO_2 and methane had continued to rise, it is possible that a "runaway greenhouse" could have occurred with ultimate destruction of all life. We might have gone the way of Venus.

But how was the glaciation reversed? How was a "runaway ice house" averted? Again, it is believed that life itself came to the rescue. We must consider that some organisms survived this and other glaciations (see below), even though their numbers were greatly reduced. This provides a negative feedback, for dioxygen generation and carbon dioxide draw down were diminished as photosynthesizers died. On the other hand, volcanic production of green house gases proceeded unabated. Another factor may have been the rise of aerobic metabolism at about that time, which by producing carbon dioxide from oxidation of carbon-containing substrates added to the greenhouse gases. It is believed that by the late Proterozoic, seawater temperatures may have again risen to levels at $50-60°C$.

Thus, a mild greenhouse effect could be reestablished. In fact, with both positive and negative factors at work, a fragile quasiequilibrium could be established. Remarkably, this persisted for a billion years, throughout much of the Proterozoic until evolutionary changes again upset the balance.

6.1.2 Later Proterozoic Glaciations

For a long time, the balance held. The next glaciation at about 950 MYA (the Algonkic Ice Age) was of limited extent. Then, two massive glaciations (together termed the Cryogenian Event) occurred in the Neo-Proterozoic at 710–680 MYA (Sturtic glaciations) and 605–585 MYA (Varanger-Marinoic glaciations) (see Fig. 6.1). Evidence has been found in many parts of the world. The two glaciations were separated by a short warm period of about 75 million years. It has been suggested that both were "Snowball Earth Events," in which ice nearly completely covered the Earth (Hoffman et al. 1998). Although the exact extent is disputed, these glaciations were clearly severe. One might expect that only simple cells could

survive, but the rapid emergence of complex organisms (the Ediacaran fauna, for example) closely after the Varanger-Marinoic glaciation suggests that the situation was more complex and that some metazoans persisted through the glaciations (see also Sect. 6.2). Thus, there is evidence for sponges and higher algae already before the Neo-Proterozoic glaciations (Dawkins 2004).

If there were a plant/lichen colonization of the land at ~700 MYA, it could have contributed to the Neo-Proterozoic glaciations. The consequence of such colonization would have been a burst of photosynthesizers, pouring even more dioxygen into the atmosphere and drawing down the previously high carbon dioxide levels. Given the weakness of the Proterozoic sun, the inevitable consequence is extensive glaciations. It seems likely that such factors may have played a major role in triggering the widespread and long-lived periods of Proterozoic glaciations (see Hedges 2004). If we do not accept land plants at 700 MYA, the fact that higher algae developed at about this time may alternatively account for the glaciations.

In fact, the end of this cold period is marked by the rise of the Ediacaran fauna, followed by the Cambrian explosion at 545 BYA. This was a time of rapid diversification of multicellular life (Rieu et al. 2007). One important basis may have been that collagen was evolved. It allows developing complex multicellular organisms since it enables cells to stick together and favors proliferation. For its synthesis, however, dioxygen is necessary (see Chap. 4), which was present in reasonable amounts at that time but not before. By about 545 BYA, the warm planet had no restriction to life. CO_2 and dioxygen were present in the atmosphere in large amounts and the lack of climatic stress and perhaps the concordant separation of the continents apparently resulted in a rapid evolution.

6.1.3 Phanerozoic Climate and Glaciations

The Phanerozoic (the era postdating the Cambrian/Pre-Cambrian boundary when diverse hard-shelled animals first appeared) has also exhibited some remarkable variations in atmospheric composition and climate (see Graham et al. 1995). By about 500–600 MYA, the dioxygen sinks of reduced metals in the ocean and on land surfaces had been saturated resulting in a quick increase of dioxygen up to about 12–15%, based on photosynthesis. This remained constant until about 380–400 MYA, when it rose rapidly to a maximum above 30% in the Carboniferous period (Berner et al. 2007; see Fig. 6.3, Sect. 6.3). While the CO_2 slightly increased in the Cambrian to about 0.5% shortly before 500 MYA (Fig. 6.3), a steep drop to a minimum followed at about 350 MYA, corresponding roughly to the period of maximum dioxygen (Fig. 6.3). Since then, CO_2 increased until about 200 MYA and has decreased since. To summarize what we know of Phanerozoic climate: Until about 380 MYA (the Devonian/Carboniferous boundary), dioxygen was lower than at present and carbon dioxide much higher and the Earth must have been warm. Then, a remarkable change occurred: carbon dioxide precipitously dropped and dioxygen peaked to unusual levels. What could have triggered such a remarkable change? Many believe it must have been the overwhelming flowering of the land
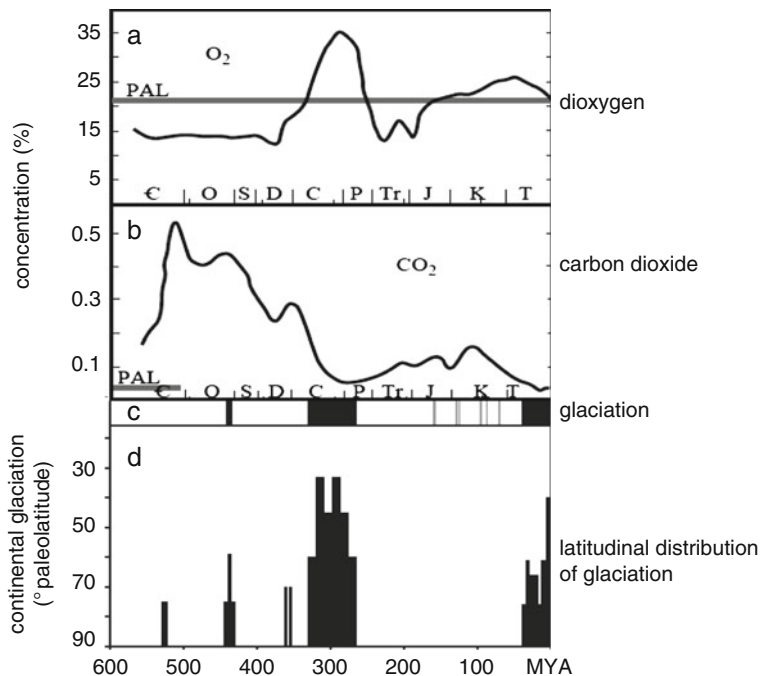
Fig. 6.3 Approximate variation of dioxygen (**a**) and carbon dioxide (**b**) in the Earth's atmosphere over Phanerozoic time and extension of ice formation (**c, d**). Estimates (on a linear percentage scale) are shown over the past 600 million years. PAL = present atmospheric level. (**c**) Intervals of glacial (*black*) or cool climates (*gray*). (**d**) Latitudinal distribution of direct glacial evidence (as evidenced by tillites, striated bedrock, etc.) throughout the Phanerozoic. This demonstrates nicely that not all the Earth's surface was covered with snow [(**a, b**) reproduced with permission of the *Journal of Experimental Biology*; taken from Dudley (1998); (**c, d**) after Royer et al. (2004); see also Crowley (1998)]

surface with broad-leaved, large vascular plants conquering the third dimension and seed plants which could readily spread over large areas. In itself, a large increment in photosynthesis would tend to bring down carbon dioxide and raise dioxygen. But the situation in the presence of aerobes is a bit more complicated. Consider the organic carbon cycle, expressed in a simple form. All plant material is designated C (Fig. 6.4). Some carbon can be buried and may or may not reenter the organic carbon cycle.

If all plant carbon is simply reoxidized to carbon dioxide, the balance is unchanged. It is the extent to which some carbon is physically buried (as coal, petroleum, carbonates, etc.) that will lead to a net loss of carbon dioxide from the atmosphere and a net gain of oxygen. This is precisely what must have happened in the Carboniferous period (~290–370 MYA). The first vascular terrestrial plants appear about 420 MYA (*Cooksania*) and were followed by seed plants (about 350 MYA). Vast forests and swamps and the absence of large grazing creatures led to the burial of much of the carbon and hydrocarbon we mine today. High

Fig. 6.4 Scheme of the organic carbon cycle. All plant material designated with C is obtained as a consequence of the photosynthesis. A part of C can be buried and will not reenter the organic carbon cycle for being used in the aerobic metabolism
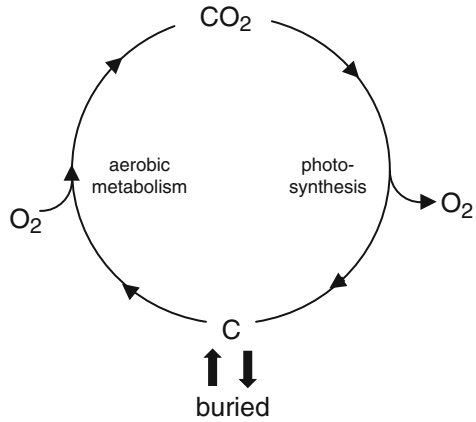
dioxygen levels also favored the formation of carbon-based polymers such as lignin which were important for wood structure. Lignin and cellulose are relatively difficult to metabolize. Thus, CO_2 was concentrated in flourishing forests and then buried as fossil carbon. This effect withdrew much of the atmospheric CO_2 to the very low minimum shown in Fig. 6.3.

More difficult to explain is the decline of dioxygen in the Permian. Part may be accounted for by burning of the Carboniferous forests, but there is not a sufficient corresponding rise in CO_2 as would be expected. Perhaps inorganic carbonate burial (see Sect. 6.4) may have been involved.

The drop in carbon dioxide in the Carboniferous period can probably account for the most extensive glaciation in Phanerozoic times. This occurred at roughly 280–320 MYA, at the end of the carbon dioxide minimum and dioxygen maximum (Fig. 6.5). Figure 6.3 demonstrates that ice occurred down to the latitude of 30° (Royer et al. 2004). Despite the severity, there is no evidence that the glaciations led to massive extinction.

However, later, at 251 MYA, after the supercontinent Pangaea has been formed (Fig. 6.5), a major extinction event (sometimes informally called the Great Dying) occurred at the boundary between the Permian and Triassic geologic periods (Yin et al. 2000). It was the Earth's most severe extinction event, with up to 96% of all marine species (Benton 2005) and 70% of terrestrial vertebrate species became extinct (Table 6.2). The reason for this extinction remains unclear. Suggestions include effects of plate tectonics, an impact event, a supernova of a far sun, large-scale volcanism, or some combination of factors. The discovery of a crater hidden under the East Antarctic Ice Sheet and indication of another to be north of Australia has given rise to a new theory that meteors caused the mass extinction and possibly began the breakup of the Gondwana supercontinent by creating the tectonic shift that pushed Australia northward. A recent proposal (see Kerr 2008) cites evidence for major dioxygen depletion in the ocean at about this period. It is noteworthy that this coincides quite nicely with the minimum in atmospheric dioxygen shown in Fig. 6.3, and Berner et al. (2007).
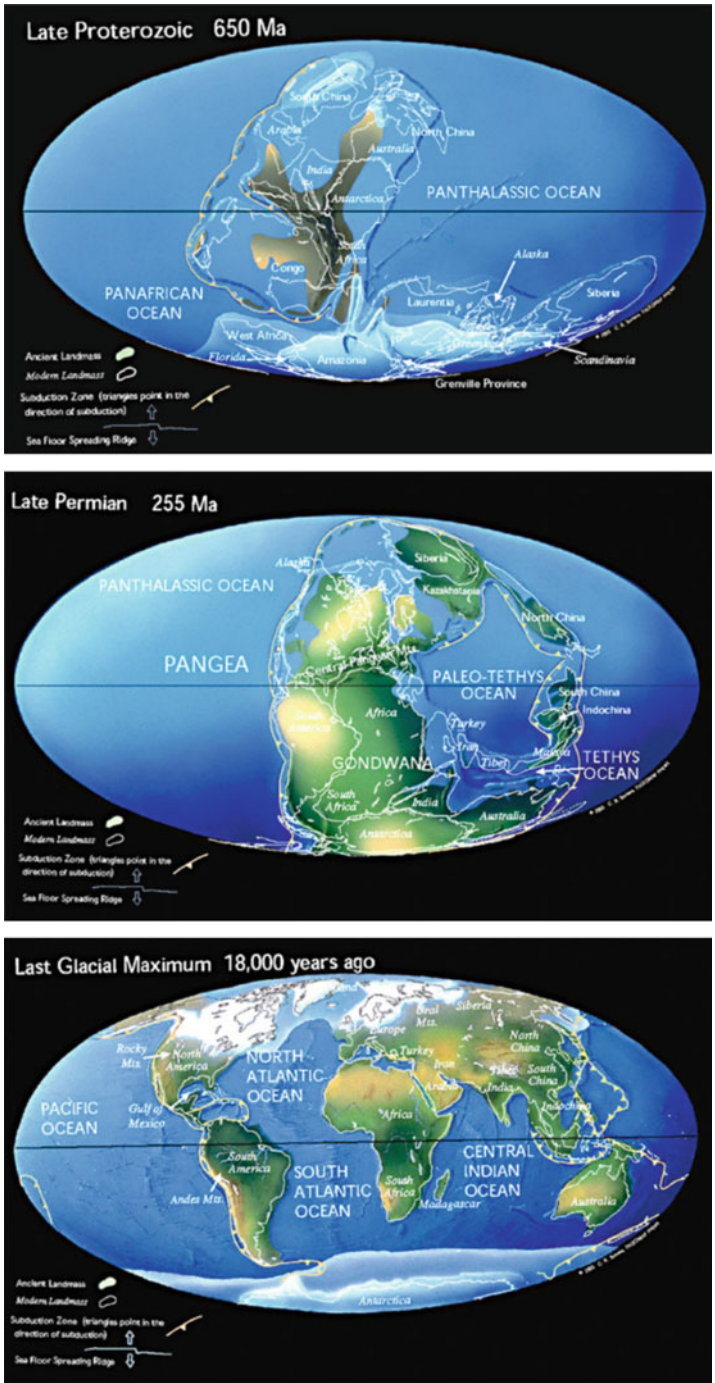
Fig. 6.5 Continental drifts during the last 650 million years (Scotese 2001, 2009)

Table 6.2 The Great Extinction Events

MYA	Time	Examples of types of organisms becoming extinct
445	Late Ordovician	Trilobites, brachiopods, echinoderms, corals
365	Late Devonian	Many marine families, corals, brachiopods, bivalves, sponges
250	End Permian	Most marine families, molluscs and brachiopods, trilobites, vertebrates
200	End Triassic	Mollusca, sponges, marine vertebrates, large amphibians
65	End Cretaceous	Dinosaurs, pterosaurs, marine plankton, brachiopods, bivalves, echinoderms

After http://blog.everythingdinosaur.co.uk/mass_extinction_table.jpg

Since about 250 MYA, the Earth's atmosphere has shown two definite trends (Fig. 6.3). Dioxygen levels have slowly recovered to a value of about 20–25%. Carbon dioxide has been decreasing from about 0.4% to less than 0.1% (Fig. 6.3). In fact, recent CO_2 levels are about the same as they were during the major Permian glaciations (see Fig. 6.3).

This would imply that we are in a major ice age, and all evidence supports this. The last million years have shown a definite cooling trend, with increasing glaciations. The evidence for such a current "glacial epoch" is strong. Several glaciations have been recorded for the northern part of the Earth during the last 650,000 years. During the recent glaciations, about 32% of the land was covered with ice starting at the poles, compared to about 10% today. The times, duration, and the area of glaciations are still strongly debated, including even the names of the glaciations, which reflect the areas that were covered: Guenz (about 650,000–540,000), Mindel, Yarmouth (about 460,000–250,000, being divided in several periods, about 460,000–420,000, 370,000–330,000, 290,000–250,000), Riss, Illinoian (200,000–130,000). The last one (between about 100,000 and 12,000) occurred at different regions in different strength and names are attached according to the places: Wurm or Weichsel ice period in continental Europe, Devonian in England, Wisconsin in the US, and Waldai in Russia (Kukla 2005; Bond et al. 2001).

Each cold period may be divided into subperiods such as the Würm ice periods (~60,000, 40,000, and 18,000 years ago). The periodical decreases in temperature by up to 5°C would explain why much more ice was formed compared to today. This also resulted in a decrease of the sea level: for example, up to 135 m about 18,000 years ago. Thus, large parts of the North Sea bottom were passable by land and no Gulf Stream could have warmed Europe. Estimated temperatures of about −20°C in January would have been much lower than today. Nevertheless, a great variety of organisms found areas where they could survive the cold or adapt to the cold, such as the wooly mammoth or even the Neanderthals. One reason may be the current high level of available dioxygen, which has favored heat generation in mammals and birds.

Although the incoming energy from the sun has averaged more or less the same over recent periods of millions of years, fluctuations of the orbit of Earth influence the distribution of the seasons with respect to the latitudes (Kerr 1987; Karner and Muller 2000; Bond et al. 1997, 2001). Three different cycles modulate the heating and cooling of Earth:

(a) The eccentricity of Earth's orbit, with a period of 100,000 years
(b) The ecliptic (angle between orbit and the rotational axis of Earth) which cycles between 21.6° and 24.5° with a period of 41,000 years
(c) The rotational axis of Earth completes a precession with a period of 28,000 years

Beside these three cycles, there appears to be an additional one of 400,000 years (Ruddiman 2000). Despite the overall cooling trend, the cyclic periodicity is such that we presently live in a relatively warm interglacial period, which is expected when the eccentricity of Earth's orbit is at a minimum. The last time that was about 400,000 years ago. Thus, the next severe glaciation is expected in a few thousands of years (Berger and Loutre 2002; Paillard 2001). It is possible that man-made global warming by increasing CO_2 might delay the next glaciations. The new period in which man significantly influences climate has been called "Anthropocea" (Crutzen and Steffen 2003; Rahmstorf and Schellnhuber 2006; Ruddiman 2000). However, there is still some controversy. In addition, whether the length of a day has any influence is still strongly discussed. During the Cambrian, the length of a day was about 21 h compared to 24 h today. Thus, the climate was much more balanced with respect to temperature than today.

6.2   How Did Life Survive Glaciations?

The major Proterozoic and Phanerozoic glaciations covered much of Earth in ice, sometimes kilometers thick. What happened to life during such glaciations? Why was life not extinguished, as one might assume, when Earth and seas were largely covered? The answers may well be different for different glaciation periods. Before the great oxidation event (2.2 BYA), the seas and metabolism were largely anoxic. When the ocean is completely covered with ice, it acts as a cover to isolate the conditions of seawater below the glaciers. This may be the case today for some extraterrestrial planets or moons (see Chap. 9). Special chemistry must have happened under sea, when volcanoes released reduced material including metals such as iron and manganese, along with H_2S and CO for anaerobe metabolism. Anaerobes could have actually thrived in this environment, although primitive photosynthesizers would probably not (see below). However, immediately after being freed from ice cover, carbon dioxide, metals (iron, manganese, magnesium, etc.), and other nutrients were available to allow the development of extensive photosynthesis. This resulted in an enormous increase in free dioxygen, mostly made by phytoplankton. This is evidenced by deposition of oxidized iron and manganese compounds shortly after the first large glaciations.

In later glaciations, when life had begun to depend more and more on photosynthesis, the situation became very different. The availability of light below a glacial covering is certainly decreased. Yet photosynthesis must have continued to some extent. This is the focus of recent research in the Arctic and other ice-covered regions. It has been discovered that life can thrive even among the ice crystals

themselves. In sea ice, a high-salt and viscous solution called brine is trapped between the ice crystals. Within this environment, organisms/cells will not be destroyed by ice formations (Fig. 6.6). It is believed that between 10 and 30% of
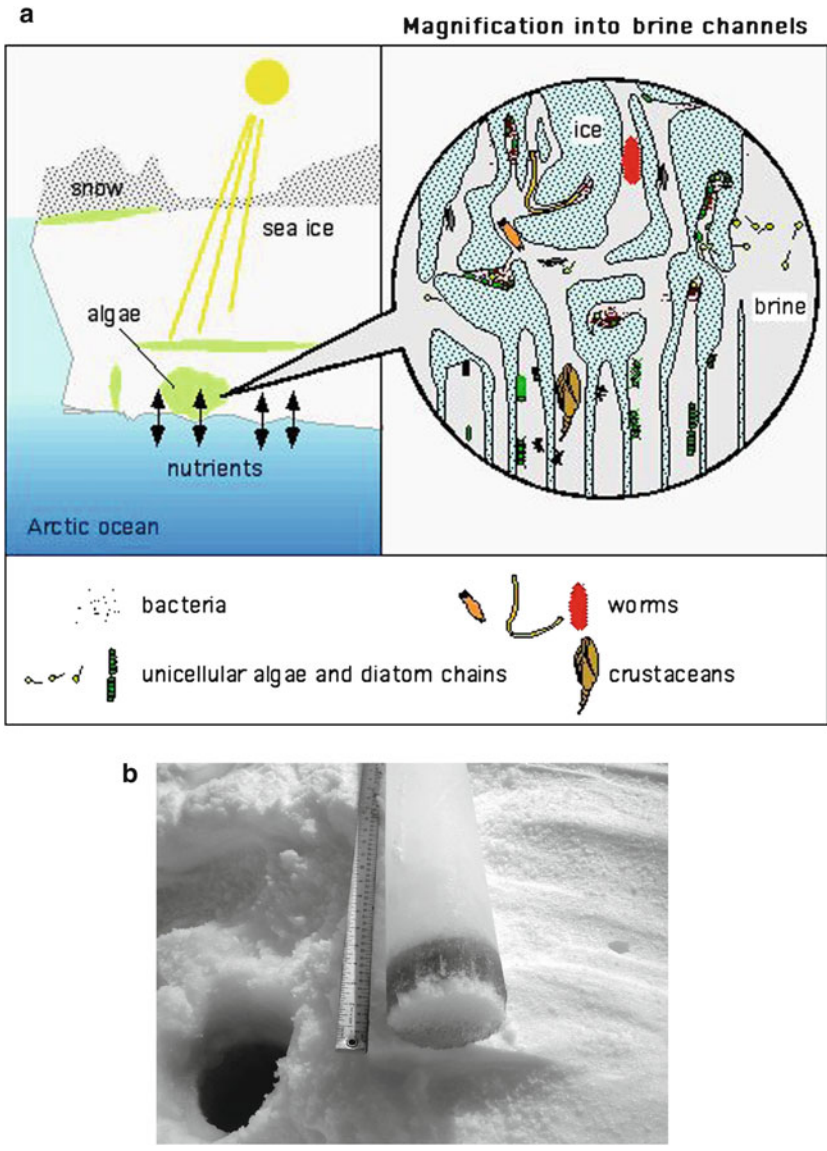

Fig. 6.6 Ice algae. (**a**) Schematic view of the flourishing life within the briny habitat of sea ice; magnified view of the brine channels. The ice-specific ecosystem includes bacteria, viruses, unicellular algae, diatom chains, worms, and crustaceans; (**b**) Ice core showing algae band (adapted from Christopher Krembs and Jody Deming, University of Washington; http://www.arctic.noaa.gov/essay_krembsdeming.html)

the sea ice volume is filled with brine. Life within the brines exhibit great variety: diatoms, bacteria, algae, and even small arthropods are found (Krembs et al. 2000; Krembs and Deming 2006). But the size of the brine-pockets of a few millimeters limits the size of the organisms. Although only about 10% of the sunlight compared to the ice surface is available at the bottom of sea ice, single-celled plant species such as algae and diatoms are still able to perform photosynthesis. Some even seem to be specialized to low light intensity. Similar conditions could have existed during early Earth and on other planets.

The size of brine-pockets and their contact through channels depend on the temperature. The necessary nutrition and organic molecules are present in the sea everywhere. Normally, brine channels are found at the bottom of ice at the interface with water. Even on terrestrial glaciers, there seems to be a small water underlay due to the high pressure of ice. There creatures can survive, since during cold winter months, the temperature varies from $-2°C$ at the bottom with the contact of seawater to $-35°C$ or lower at the wind-chilled surface (Cox and Weeks 1983). These ice algae are a very important source of the marine food web.

This environment is of advantage to algae-eating zooplankton who are in turn protected by the ice from their predators (Gosselin et al. 1997; Horner et al. 1992; see Fig. 6.6). The algae growth turns the underside of the ice pack into a lush pasture that is grazed by a surprising array of animals. Amphipod, crustaceans, flatworms, and other small organisms are found. It is believed that fish like the Arctic cod ultimately owe their existence to the plants living among the ice crystals. Under these conditions, at this cold temperature, dioxygen and carbon dioxide are available in large amounts; they are dissolved at the highest concentration (see Chap. 5). We may understand therefore that during the Proterozoic and Phanerozoic glaciations, life persisted even if "Snowball Earth" conditions existed. This could also be a scenario which holds for other ice-covered planets or moons such as Europa, a satellite of Jupiter (see Chap. 9 for further discussion).

Even in deeper water layers, where we would expect O_2 to be markedly decreased during a long "snowball" event, life may have adapted. It is now known that a variety of aerobic organisms survive today in nearly anoxic water layers, depending on a high efficiency in collecting dioxygen (Childress and Seibel 1998).

6.3 Milestones of Life in the Phanerozoic

Possibly by around 900 MYA, true multicellularity had evolved in animals living in water (see Chap. 4). At first, it probably somewhat resembled today's sponges, where all cells were totipotent and a disrupted organism could reassemble itself completely. Thus, a variety of cell types had been evolved at that time as well as genetic information for assembling metazoan organisms (Dawkins 2004).

We have discussed in earlier chapters the evolution of life up to the Cambrian explosion (Conway-Morris 2006; Fortey 1999). At that point, a wide variety of organisms were evolved and evolutionarily tested. Some organism depended only on light (plants and algae). Others became animals. In the earliest Phanerozoic, many were protected and given forms by an exoskeleton, later an endoskeleton avoided the necessity of molting. The evolution of the exoskeleton seems to have begun very early: about at the beginning of the Cambrian for mollusks and arthropods and even earlier for annelids (Vermelic 1989). In some rare cases, oxygen could be absorbed by skin as in contemporary frogs. But many must have already developed oxygen transport system (see Chap. 5).

Now we briefly summarize the milestones of evolution in the remaining time until the present. In the *Silurian* period, somewhat over 450 MYA, macroscopic plants began to occupy the land in large numbers. Three major requirements were needed for this development:

(a) Structural strength is needed, because plants living on the Earth's surface no longer have the buoyant support of surrounding water.
(b) A controlled fluid transport system has to be established being based on turgor.
(c) New leaf structures have to be extended for exchanging gases with atmosphere instead of water.

Molecular evidence suggests that the fungi may have colonized the land as early as more than 1,000 MYA and some plants in form of green algae by 700 MYA (Bhattacharya and Medlin 1998). However, these dates seem doubtful, for the oldest recognized fossils of land fungi and plants date to 480–460 MYA (Conway-Morris 2006). Perhaps this colonization started at the edge of the water before proceeding entirely to the land.

After glaciations, the land was empty, ready to be filled with descendents of those organisms that had survived. There was no limitation for nutrition, energy, and the two important types of gas – dioxygen and carbon dioxide – were abundant. Mutation and variations resulted in further colonization of this new environment. The time at which the first animals left the oceans is not precisely known: the oldest clear evidence shows arthropods on land around 450 MYA or even (unconfirmed) 530 MYA (Pisani et al. 2004). By 480 MYA, these may have thrived due to vast food source provided by the terrestrial plants which were already established on the surface. How arthropods could have survived on land at the earlier date (530 MYA) is unclear. Probably, the first ones spent only short periods ashore. Even today, coconut crabs (*Birgus latro*) climb up palm trees to harvest coconuts, and crabs such as *Potamon* in the Mediterranean Sea leave the sea to steal and eat grapes. Crustaceans have the advantage of being protected by a carapace against radiation, dehydration, and against predation. Behind the carapace, the gills are hidden, being kept humid. For these needs, phenoloxidases such as tyrosinase and/or catecholoxidase had already been evolved to build up the carapace and cuticula and hemocyanins to provide a reliable oxygen supply.

Around 400 MYA, the first tetrapods evolved from fish (Niedzwiedski et al. 2010). It is thought that they evolved limbs allowing them to get from one pond to

another. In this, they had to lift their heads out of the water into the atmosphere requiring them to breathe air. They may have later ventured on land for brief periods. Eventually, some of them became so well adapted to terrestrial life that they spent their adult lives on land, although they still hatched in the water and returned to lay their eggs. It may also be significant that the first vertebrates to move out of the waters and onto the land did so in the late-Devonian/early-Carboniferous, just at the time when atmospheric dioxygen was rising (see Fig. 6.3). This was the origin of amphibians shortly before the next glaciations (Karoo Event). Perhaps, this event forced them to adapt to low temperature of about −4°C. Today, frogs survive these temperature covered by snow due to antifreezing proteins in their blood.

About 340 MYA, the amniotic egg evolved, which could be laid on land, giving a survival advantage to tetrapod embryos. This resulted in the divergence of amniotes from amphibians. Shortly after at about 310 MYA saw the divergence of synapsids (including mammals) from the sauropsids (including birds and non-avian reptiles).

The increase in the atmospheric dioxygen content to 35% in the Carboniferous period (Graham et al. 1995) probably explains the existence during this period of giant arthropods as shown in Fig. 6.7. The dragonfly with a wing span of 75 cm occurred as well as a giant sea scorpion (*Jaekelopterus rhenariae*) with a length of 2.5 m. The latter was recently found in Germany and apparently lived about 390 MYA (Fig. 6.7). We have mentioned (Chap. 5) that the insect tracheal system is efficient for dioxygen uptake only over small distances (at present oxygen levels). When the percentage of dioxygen is nearly doubled, tracheal efficiency will increase markedly. And the length of trachea and size of the insects can correspondingly increase (Fig. 6.7). The great increase in dioxygen content in the Carboniferous period also increased the density of the atmosphere by about 20%, which supported the evolution of flying insects. Later during the second oxygen peak at 140–200 MYA, some flying vertebrates, in particular the birds, evolved (Ward 2006).

The most severe of all extinction events occurred 250 MYA, at the boundary of the Permian and Triassic periods (see Sect. 6.2; Table 6.2). Fewer than 10% of living sea species and about 30% of land species survived (Erwin 1993). The recovery of life on Earth took significantly longer than after other extinction events. This event has been described as the "mother of all mass extinctions" (Erwin 1993). But life continued, and around 230 MYA, dinosaurs split off from their reptilian ancestors. By 180 MYA, the vast continent of Pangaea broke up into Laurasia and Gondwana. *Archaeopteryx*, traditionally considered one of the first birds, lived around 150 MYA. Dinosaurs were probably already in decline for various reasons when, 65 MYA, a meteorite of about 10-km diameter struck Earth just off the Yucatan Peninsula, ejecting vast quantities of particulate matter and vapor into the air that occluded sunlight, inhibiting photosynthesis (Table 6.2). Most large animals became extinct. Hereafter, the small mammals rapidly diversified, grew larger, and became the dominant vertebrates. Perhaps a couple of million years later (around 63 MYA), the last common ancestor of primates lived.
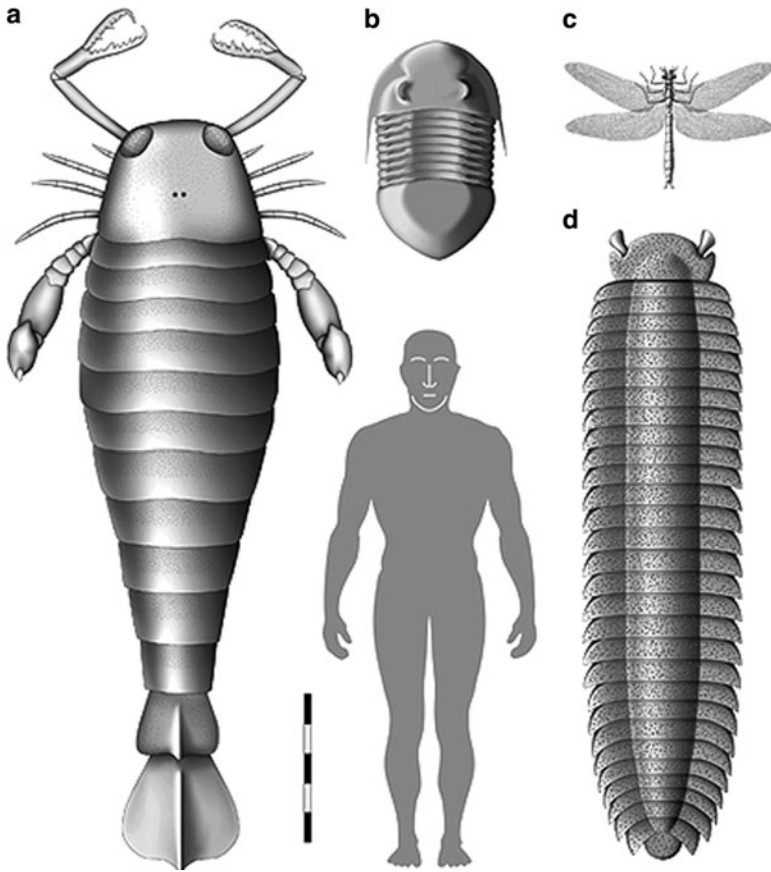
Fig. 6.7 Comparison of various giant ancient arthropods with a human male. (**a**) Ancient sea scorpion *Jaekelopterus rhenaniae* (this 390-million-year-old giant sea scorpion occurred in the Carboniferous period. It is proposed that this animal with its 46-cm fossil claw represents the world's largest known bug at 2.5 m), (**b**) the trilobite *Isotelus rex*, (**c**) the dragonfly *Meganeura monyi*, (**d**) the millipede *Arthropleura armata* (Braddy 2008)

6.4 Inorganic Cycling of Carbon Dioxide

In previous sections, we have seen how atmospheric carbon dioxide and dioxygen are mutually regulated through the organic carbon cycle. But there are also wholly inorganic processes that must be considered in elucidating the overall carbon balance of the Earth. First, silicates from terrestrial rocks may either react with atmospheric carbon dioxide or be washed into the ocean and there react with dissolved carbon dioxide. In either case, the basic reaction (for calcium silicate as an example) is

$$CaSiO_3 + CO_2 \rightarrow CaCO_3 + SiO_2$$
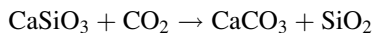
The products, calcium carbonate ($CaCO_3$) and quartz (SiO_2), are both insoluble and will accumulate on the sea floor or remain on the land surface. In either case, carbon dioxide is (at least temporarily) buried. However, the burial can to some extent be reversed. If the calcium carbonate is then transported to the subduction zones by movement of continental plates, it will be melted and decomposed in the hot mantle of the Earth. A part will become magma, which will be exposed to the atmosphere through volcanoes thereby releasing carbon dioxide:

$$CaCO_3 + SiO_2 \rightarrow CaSiO_3 + CO_2$$
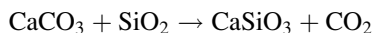
This carbon dioxide along with water vapor is responsible for much of the explosive eruption of volcanoes. Again without tectonic movement, which gives rise to volcanoes, there could be no life on Earth for carbon dioxide would be continually irreversibly buried. The final result would be a frozen planet. Because carbon dioxide is much less soluble in water at high temperature, the formation of carbonate is temperature dependent. Therefore, the inorganic carbon dioxide cycle can be considered as a feedback process for regulating the climate. Consider the following reactions:

$$CO_2 + H_2O \rightarrow H_2CO_3 \rightarrow H^+ + HCO_3^- \rightarrow 2H^+ + CO_3^{2-}$$

$$H_2CO_3 + CaCO_3 \rightarrow Ca(HCO_3)_2$$
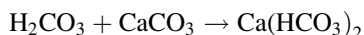
In contrast to calcium carbonate ($CaCO_3$), calcium hydrogen carbonate, $Ca(HCO_3)_2$, is very soluble in water. At low temperatures, this solublizing effect is enhanced since more carbon dioxide can be dissolved. When the temperature increases, carbonate will be left on the bottom while the carbon dioxide leaves the water due to the decrease of solubility of gas in water. The equilibrium between hydrogen ions and carbonate will be adjusted slowly. Thus, carbonate will be formed faster, extracting the dissolved carbon dioxide, which will be dissolved rapidly. Since carbon dioxide only slowly leaves the water and atmospheric carbon dioxide is dissolved fast in water, carbon dioxide in the atmosphere is decreased resulting in a cooling of Earth due to the lack of the greenhouse gas. Thus, the carbon dioxide cycle provides a feedback protection against runaway greenhouse. The time of a cycle is about several thousands of years. This is the reason why it is dangerous when the carbon dioxide increases in a much shorter interval of time since the compensating feedback by the carbon dioxide cycle is too slow.

6.5   Is Our Environment Stable?

With minor exceptions, Earth's climate, in particular the temperature range, has remained remarkably stable over the past 3 billion years. Average seawater temperatures have probably rarely exceeded the limits of 0–50°C in that time. This is in
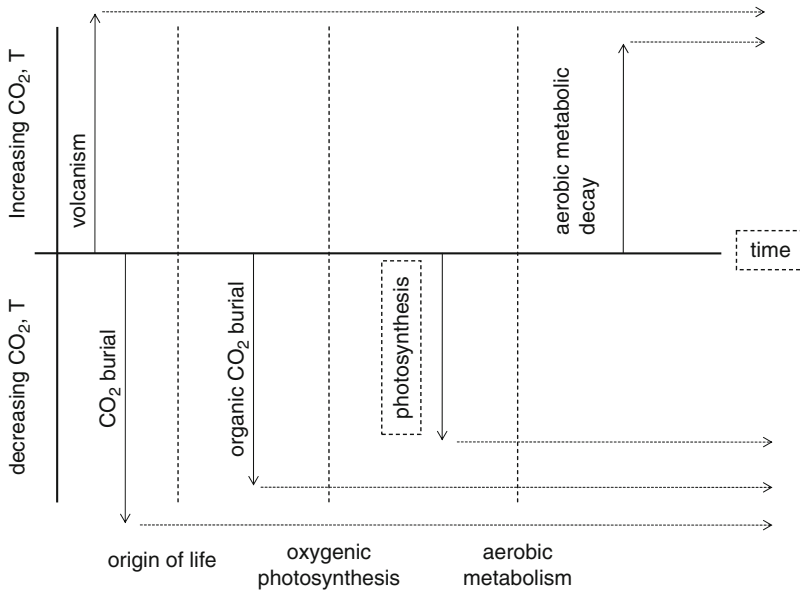
Fig. 6.8 A schematic depiction of the major positive and negative effects on atmospheric CO_2 that have arisen during the Earth's history. *Arrows above the line* indicate contributions to atmospheric CO_2, *arrows pointed down* denote draw down

marked contrast to our sister planets, Venus and Mars, one of which is excessively hot and the other frozen. One might ascribe this in part to Earth's intermediate distance from the sun, but this cannot be the whole story, for the sun's energy output has increased by 20–30% over the past 3 billion years, and yet this has been compensated in some fashion. Somehow, total runaway greenhouse and runaway icehouse have been avoided. It is tempting to ascribe this stability to the presence of life, which has provided a number of feedback mechanisms for climate stability. Figure 6.8 depicts, in a very schematic way, some of the opposing controls on climate on a world either with or without life, with or without photosynthesis or aerobes. In a planet close to the sun, lacking life, there is likely to be a steady increase of carbon dioxide in the atmosphere (from volcanism), until the boiling point of water is exceeded, in which case a runaway greenhouse is likely, because water vapor is also a strong greenhouse gas. The evolution of photosynthetic life provides a carbon dioxide sink: In the absences of aerobic metabolism, there will be little return of carbon dioxide to the atmosphere. This condition can favor glaciations (such as the great Proterozoic glaciations), but even here life exerts a compensatory effect, because of the die-off associated with massive glaciations, carbon dioxide draw down decreases. Furthermore, the advent of aerobic metabolism provided a mechanism to restore CO_2 to the atmosphere. Thus, the evolution of life on Earth has continually added more positive and negative controls to climate. Although we may yet suffer climatic changes that could be devastating from the

human point of view, the possibility of either a true runaway greenhouse (as in Venus) or a new "Snowball Earth" event seems remote.

To take the very long view, consider the expanded range in Fig. 6.3, neglecting small changes. Dioxygen seems to exhibit an unregular increase toward a limit. But carbon dioxide appears to be in continual decrease! Over the very long range, we can only expect the Earth to grow colder (assuming that anthropogenic global warming is controlled). Eventually, release of buried CO_2 from volcanic eruptions may reach equilibrium with its burial, but this may be a cold distant future.

The ultimate test, from our point of view, is whether the environment will stay within limits within which life can survive. It has done so for 3 billion years, so there is room for optimum.

6.6   Recent Global Warming

We are currently in an interglacial episode of what can be considered overall a glacial period (see above). However, a recent and alarming perturbation is the rise of carbon dioxide levels and terrestrial temperatures during the past 200 years (Masklin 2004). There is much discussion as to whether this rise, unprecedentedly rapid on a geological time scale, can or cannot be compensated by the planet's normal responses. It is believed to be at least in part the consequence of human operations, involving the burning of fossil fuels. It is the object of considerable concern which will be discussed in the next chapter. A frequently overlooked fact is that much of this combustion is of buried carbon. For the first time in millions of years, this carbon is reentering the atmosphere as carbon dioxide. This represents an unusual perturbation in the long-term carbon balance. If we were simply burning large amounts of vegetative material of contemporary origin, there would be, in the long run, no imbalance in the carbon amount – most of it would have aerobically decayed anyway. But we are dipping into the "carbon bank." The overall consequences remain to be seen.

We have seen that every major shift in dioxygen and carbon dioxide levels has resulted from a new perturbation introduced by life: photosynthesis, carbon burial, aerobic metabolism have all served as "triggers" to some extent of environmental change. On the other hand, life's evolution appears in many instances to be in response to major climatic changes.

References

Benton MJ (2005) When life nearly died: the greatest mass extinction of all time. Thames and Hudson, London

Berner RA, Kothavala Z (2001) GEOCARB III: a revised model of atmospheric CO_2 over Phanerozoic time. Am J Sci 301:182–204

Berner RA, Vandenbrooks JM, Ward PD (2007) Oxygen and evolution. Science 316:557–558

Berger A, Loutre MF (2002) An exceptionally long interglacial ahead? Science 297:1287–1288

Bhattacharya D, Medlin L (1998) Algal phylogeny and the origin of land plants. Plant Physiol 116:9–15

Bond G, Showers W, Cheseby M, Lotti R, Almasi P, de Menocal P, Priore P, Cullen H, Hajdas I, Bonani G (1997) A pervasive millennial-scale cycle in North Atlantic Holocene and glacial climates. Science 278:1257–1266

Bond G, Kromer B, Beer J, Muscheler R, Evans MN, Showers W, Hoffmann S, Lotti-Bond R, Hajdas I, Bonani G (2001) Persistent solar influence on North Atlantic climate during the Holocene. Science 294:2130–2136

Braddy SJ, Poschmann M, Tetlie OE (2008) Giant claw reveals the largest ever arthropod. Biol Lett 4:106–109

Childress JJ, Seibel BA (1998) Life at stable low oxygen levels: adaptation of animals to oceanic oxygen minimum layers. J Exp Biol 201:1223–1232

Conway-Morris S (2006) Darwin's dilemma: the realities of the Cambrian explosion. Philos Trans R Soc Lond B Biol Sci 361:1069–1083

Cox GFN, Weeks WF (1983) Equation for determining the gas and brine volumes in sea-ice samples. J Glaciol 29:306–316

Crowley TJ (1998) Significance of tectonic boundary conditions for paleoclimate simulations. In: Crowley TJ, Burke K (eds) Tectonic boundary conditions for climate reconstruction. Oxford University Press, New York, pp 3–17

Crutzen PJ, Steffen W (2003) How long have we been in the Antropocene era? Climate Change 61:251–257

Dawkins R (2004) The ancestor's tale: a pilgrimage to the dawn of life, 1st edn. Mariner Books, New York

Dudley R (1998) Atmospheric oxygen, giant Paleozoic insects and the evolution of aerial locomotor reformance. J Exp Biol 201:1043–1050

Erwin DH (1993) The Great Paleozoic crisis; life and death in the Permian. Columbia University Press, New York

Ferry JG, House CH (2006) The stepwise evolution of early life is driven by energy conservation. Mol Biol Evol 23:1286–1292

Fortey R (1999) Life: a natural history of the first four billion years of life on Earth. Vintage Books, New York. ISBN 0-375-70261-X

Frausto da Silva JJR, Williams RJP (2001) The biological chemistry of the elements: the inorganic chemistry of life. Oxford University Press, New York

Gosselin M, Levasseur M, Wheeler PA, Horner RA, Booth BC (1997) New measurements of phytoplankton and ice algal production in the Arctic Ocean. Deep Sea Res II 44:1623–1644

Graham JB, Dudley R, Agullar N, Gans C (1995) Implications of the late Palaeozoic oxygen pulse from physiology and evolution. Nature 375:117–120

Hedges SB (2004) Molecular clocks and a biological trigger for neoproterozoic snowball Earth events and the Cambrian explosion. In: Doneghue PCJ, Smiti MP (eds) Telling the evolutionary time. Molecular clock and the fossil record. CRC, Boca Raton, pp 27–39

Holland HD (2006) The oxygenation of the atmosphere and oceans. Philos Trans R Soc Lond B Biol Sci 361:903–915

Hoffman PF, Kaufman AJ, Halverson GP, Schrag DP (1998) A Neoproterozoic snowball Earth. Science 281:1342–1346

Horner RA, Ackley SF, Dieckmann GS, Gulliksen B, Hoshiai T, Legendre L, Melnikov IA, Reeburgh WS, Spindler M, Sullivan CW (1992) Ecology of sea ice biota. 1. Habitat, terminology, and methodology. Polar Biol 12:417–427

Kasting JF (2001) Earth history. The rise of atmospheric oxygen. Science 293:819–820

Kasting JF (2004) When methane made climate. Sci Am 291:79–85

Kasting JF, Ono S (2006) Palaeoclimates: the first two billion years. Philos Trans R Soc Lond B Biol Sci 361:917–929

Kerr RA (2008) Mass extinction. Most devastating mass extinction followed long bout of sea sickness. Science 222:359

Kerr RA (1987) Milankovitch climate cycles through the ages: Earth's orbital variations that bring on ice ages have been modulating climate for hundreds of millions of years. Science 235: 973–974

Karner DB, Muller RA (2000) Paleoclimate: a causality problem for Milankovitch. Science 288: 2143–2144

Kirschvink IL, Gaidos ES, Bertani LE, Benkas NJ, Gatzmer J, Maepa L, Sternberger RE (2000) Paleoproterozoic snowball Earth: extreme climatic and geochemical global changes and its biological consequences. Proc Natl Acad Sci USA 97:1400–1405

Krembs C, Deming J (2006) Sea ice: a refuge for life in polar seas? http://www.arctic.noaa.gov/essay_krembsdeming.html#ref7

Krembs C, Gradinger R, Spindler M (2000) Implications of brine channel geometry and surface area for the interaction of sympagic organisms in Arctic sea ice. J Exp Mar Ecol 243:55–80

Kukla G (2005) Saalian supercycle, Mindel/Riss interglacial and Milankovitch's dating. Quatern Sci Rev 24:1573–1583

Masklin M (2004) Global warming: a very short introduction. Oxford University Press, Oxford, UK

Moran JJ, House CH, Freeman KH, Ferry JG (2005) Trace methane oxidation studied in several Euryarchaeota under diverse conditions. Archaea 1:303–309

Niedzwiedski G, Szrek P, Narkiewicz K, Narkiewicz M, Ahlberg PE (2010) Tetrapod trackways from the early Middle Devonian period of Poland. Nature 463:43–48

Paillard D (2001) Glacial cycles: towards a new paradigma. Rev Geophys 39:325–346

Pavlov AA, Kasting JF, Brown LL, Rages KA, Freedman R (2000) Greenhouse warming by CH_4 in the atmosphere of early Earth. J Geophys Res Planets 105:11981–11990

Pavlov AA, Hurtgen MT, Kasting JF, Arthur MA (2003) Methane-rich proterozoic atmosphere? Geology 31:87–90

Pisani D, Poling LL, Lyons-Weiler M, Hedges SB (2004) The colonization of land by animals; molecular phylogeny and divergence times among arthropods. BMC Biol 2:1

Rahmstorf S, Schellnhuber HJ (2006) Der Klimawandel – Diagnose, Prognose, Therapie. C.H. Beck, München. ISBN 3-406-50866-9

Rieu R, Allen PA, Plötze M, Pettke T (2007) Climatic cycles during a Neoproterozoic "snowball" glacial epoch. Geology 35:299–302

Royer DL, Berner RA, Montanez IP, Tabor NJ, Beerling DJ (2004) CO_2 as a primary driver of Phaneroerzoiv climate. GSA Today 14:4–10

Ruddiman WE (2000) Earths climate: past and future. Freeman, New York

Scotese CR (2001) Atlas of Earth History, Volume 1, Paleogeography, PALEOMAP Project, Arlington, Texas, 52 pp

Scotese CR (2009) Plate Tectonic and Paleogeographic maps and animations, PALEOMAP Project (www.scotese.com)

Vermelic GJ (1989) The origin of skeletons. Paleos 4:585–589

Yin YG, Wang Y, Wang W, Shang QH, Cao CQ, Erwin DH (2000) Pattern of marine mass extinction near the Permian–Triassic boundary in South China. Science 289:432–436

Ward PD (ed) (2006) Out of Thin Air: Dinosaurs, Birds, and Earth's Ancient Atmosphere. The National Academic Press

Chapter 7
Global Warming: Human Intervention in World Climate

Contents

7.1 Recent Climate Changes

In the preceding chapter, we described climate changes that have occurred over very long geological periods. We concluded that Earth is currently in an interglacial interval within a rather long period of glaciations. Indeed, average carbon dioxide concentrations in the atmosphere have been slowly decreasing over the past 600,000 years, with accompanying cooling (Fig. 6.3). There have been, of course, many periodic changes in the CO_2 concentrations and average temperature over this period (see Fig. 7.1). However, very recently, something quite unique and startling has occurred. As Fig. 7.1 shows, there has been a remarkable increase in CO_2 levels, actually during the past 200 years, from 288 to 385 ppm today. Mean world temperature has shown an accompanying rise. This period is but an instant on any geological time scale, and it is very difficult to imagine any natural process that could account for it. The amount of CO_2 in the atmosphere
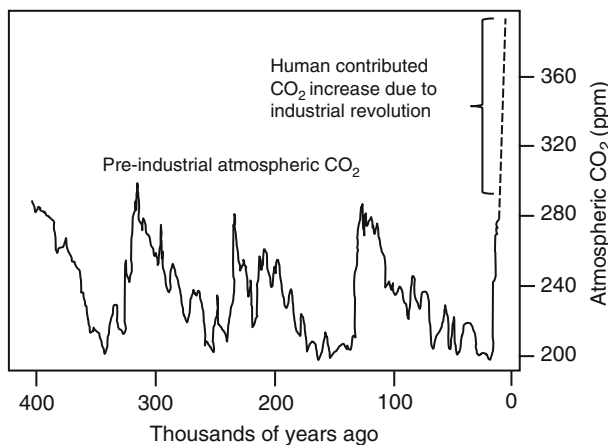
Fig. 7.1 Trends in atmospheric CO_2 (adapted from Zedillo 2008). Note that this figure shows that even though there have been significant fluctuations of CO_2 in the past 400,000 years, the recent jump is unprecedented

has increased by 30% in this period to reach a level not attained in the past million years. Nor do CO_2 and temperature show any signs of immediately leveling off. Is this only a brief fluctuation or may this even disturb the balance built up in the past? What is happening, and why? This "global warming" has been the object of intense international inquiry (see IPCC 2007, or for a compact summary see Kerr 2007).

The peculiar nature of this change, and its almost exact correspondence to the development of an industrial society burning fossil fuels, has led most scientist to conclude that it is anthropogenic. A parallel increase in atmospheric methane from 350 ppb (at last glaciations) to 1750 ppb (in 2005) can also be correlated with human activity. In this case, it is likely due to the great increase in cultivation of livestock and plants such as rice. We are, it would seem, introducing new parameters into the normal variations of atmospheric composition and climate. The same factors are involved, O_2, CO_2, water vapor, and methane, but mixing is now influenced by man. So far, the changes in carbon dioxide and temperature are miniscule compared to changes that have occurred naturally over eons (compare the scale of Figs. 6.3 and 7.1). However, the rate of change is remarkable, and the limits of these excursions are not known. They may depend very much on the course of action humanity employs to deal with them. Furthermore, it must be noted that there is something unique about these anthropogenic contributions to climate change: We are burning fossil fuels; carbon that has been buried for ages is now put into the atmosphere, which would not have been expected for ages in the normal course of events. Thus, we are perturbing the organic carbon cycle itself. It is amazing and disturbing that human activity can have a measurable effect on so vast a venue as the whole Earth. We briefly describe these changes below, and their possible consequences for humanity.

7.2 Physical Consequences of Global Warming

7.2.1 Shrinking Ice and Glaciers

As Earth warms, it is inevitable that glaciers and regions of sea ice will melt, so we should expect to observe such (Overpeck et al. 2006; Petit et al. 1999; Rignot and Kanagaratnam 2006; Thomas et al. 2006). Indeed, there are many such recent observations, although not always consistent. The strongest evidence concerns the ice covering of the Arctic Ocean and Antarctica. NASA satellite photographs of the North Polar Region clearly show a reduction of 20% in sea ice cover in the 8-year period from 1979 to 2003 (Fig. 7.2). Similarly, there have been substantial releases from the ice shelves of the Antarctic (Velicogna and Wahr 2006). The changes are of considerable interest because of the positive feedback effect they have. Ocean areas covered by ice reflect about 95% of the radiation they receive (an albedo of 0.95), whereas open ocean reflects only about 10% (albedo of 0.1). Thus, whenever warming melts ice, even more warming is induced. Exact calculations are difficult because of different effects of varying land surfaces, but the positive feedback is worrisome.

Despite many assertions to the contrary, the data with respect to glaciers are somewhat ambiguous. The Intergovernmental Panel on Climate Change (IPCC)
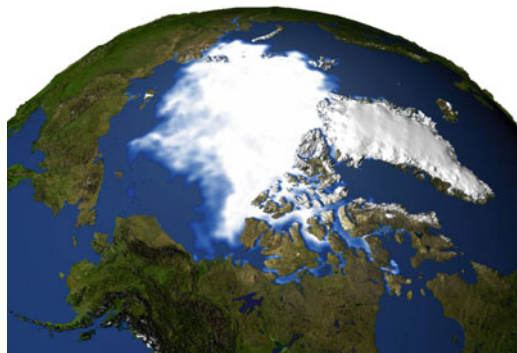

Fig. 7.2 Extent of arctic summer ice in 1979 (*top* satellite image) and in 2003 (*lower* satellite image). A decline in Arctic sea ice of 8.5% per decade is calculated. Further decreases have been observed in the past 3 years (http://www.nasa.gov/ centers/goddard/news/ topstory/2003/1023esuice. html)

monitors some 20 major glaciers and notes a general shrinkage over recent decades (Houghton et al. 2001). On the other hand, a more extensive study of 246 glaciers between 1946 and 1995 gave a less clear answer (Braithwaite 2002). Glaciers in many parts of the world were indeed shrinking, but others were static and some even growing. It seems likely that the explanation may lie in the fact that climate in different parts of the world is responding differently to the overall process, and that different aspects of weather (i.e., snowfall) are being influenced in complicated ways. Such may also explain equivocal results concerning the Greenland ice cap (Alley et al. 2005; Johannessen et al. 2005) and the uneven melting of the polar ice cap evident in Fig. 7.2.

We must also keep in mind that even within recorded history, there have been substantial, apparently natural changes in the sea ice cover and glaciations. Notable are the warm period around 800–1200 AD and the "little ice age" that followed it. In the former period, navigation in Arctic seas was much easier than at any time since, and Greenland had sufficient arable land for settlement. In the subsequent cold period, access to Greenland and even Iceland was restricted by sea ice, and the Greenland colony vanished. The causes for the warm period and the "little ice age" are unknown, but might have involved ocean current shifts (see below). Major volcanic eruptions may also contribute to changes even at the global level. It is thought, for example, that the 1991 eruption of Mount Pinatubo in the Philippines temporarily lowered mean global temperature by about 0.5°C. It is incidents like this that make the use of local climatic data unreliable in assessing global warming. Detailed data over large regions are essential and are still under strong discussion (Boehm 2008).

7.2.2   Sea Level Changes

As the ocean warms, and land ice melts, the sea level around the world will rise (Church and White 2006; Meehl et al. 2005; Nerem et al. 2006; Rahmstorf 2007; Siddell et al. 2003). The available evidence indicates that over the past 100 years the rise has been of the order of 10–20 cm (Fig. 7.3). This is a rough measurement, for the actual level a century ago is difficult to determine, and calculations must involve not only the melting of glaciers (sea ice melting does not count since this is already floating and displacing its own mass of water) but also the expansion of the sea with increasing temperature. This in turn is complex because of the nonuniform ocean temperature. In any event, the effects up to now seem to be small. On the other hand, should a substantial fraction of the Antarctic and/or Greenland ice sheets melt or simply move into the sea, very substantial rise could be expected. It has been estimated that the former change alone could produce a sea level increase of up to 60 m (Stoddart 2008)! The total sea level rise since the last glacial episode (~20,000 years) was about 140 m. Total melting of these great ice sheets is very unlikely in the foreseeable future, but even sliding of portions of these sheets from land to sea would have dramatic (and perhaps sudden) effects. The IPCC 2007
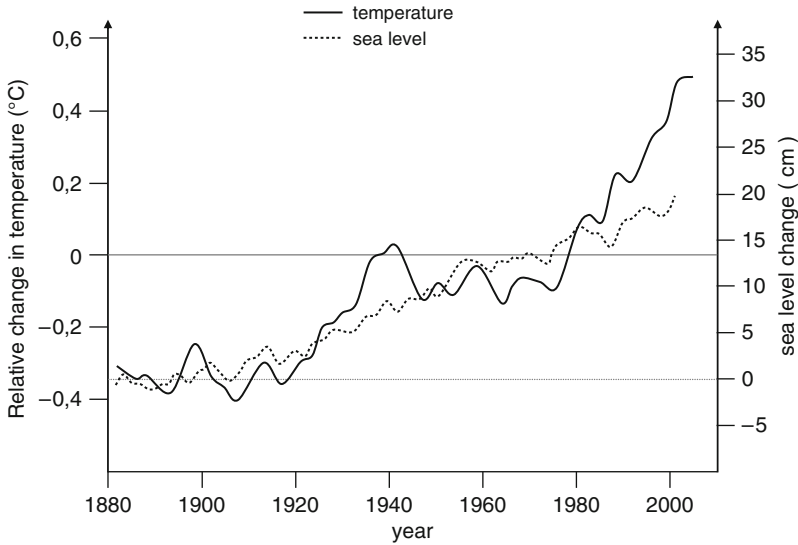
Fig. 7.3 Recent sea level rise and global surface temperature trends between 1880 and 2005. Due to a common practice of the IPCC, the zero on this figure is the mean temperature from 1961 to 1990. The temperature is a 5-year average, the sea level a 3-year average (adapted from Hansen et al. 2006)

report suggests that a possible partial melting in the next century could produce a 4–6 m rise. On the other hand, complete melting of all the glaciers in the Alps would yield a rise of only half a meter. As we shall see in a later section, even such changes that would seem small on a geological scale might have major human consequences.

7.2.3   Changes in Ocean Currents

Another potentially large climate change may be due to the modification of Earth's complex system of ocean currents by global warming (see Fig. 7.4). To take one example, consider the Gulf Stream. The driving force for this current, which warms much of western Europe, is believed to be the sinking of cold dense water in the North Atlantic (Broecker 1991). This water is preferentially dense because of high salt content. This in turn is a consequence of the freezing of surface water in the subarctic regions. Salt is excluded from the ice and makes the underlying water denser. Thus, the "motor" of the Gulf Stream is the cooling of North Atlantic waters. It is at least conceivable that global warming could block the Gulf Stream, thereby leading to the seemingly paradoxical result of cooling portions of Europe. The interconnections of the several major ocean current systems around the globe are still not fully understood nor are the likely effects of global warming on these,
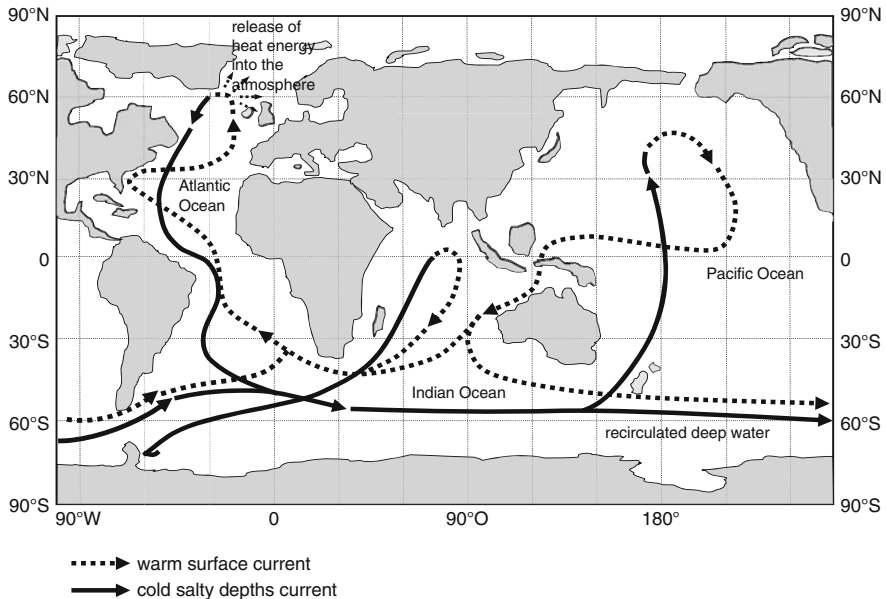
Fig. 7.4 Scheme of the major ocean currents. Surface currents are shown as *dotted lines*, deeper countercurrents are shown as *solid lines* (redrawn after Norbert Noreiks, Max-Planck-Institute for Meteorology based on Broecker (1991) and changes by Hansen et al. (2004))

and consequently on local climates. It is entirely possible that even moderate warming may make currently wet areas arid, and vice versa.

7.2.4   Local Climate and Weather

In addition to the possibilities mentioned above, global warming may more directly influence weather patterns. For example, there appears to have been an increase in violent tropical storms in recent decades, and some attribute this to warming. While difficult to prove at this point, such a connection is not unreasonable. It has long been known that the intensity of such cyclones increases or decreases in response to the temperature of the water over which they travel, so a warming of the tropical ocean might well be expected to have a significant effect. There have been, until now, no rapid perturbations in the terrestrial environment, so we are ill-equipped to guess at all the possible climate consequences of rapid global warming. It is a well-known principle in physics and chemistry that systems that have attained an equilibrium or steady state may be perturbed drastically by even small changes, if rapid. The problem is so complex that it is impossible at this time to predict with accuracy how climate and weather will change in different localities around the world. We may expect surprises.

7.2.5 The Danger of Methane Releases

Methane is a greenhouse gas. It is present in the atmosphere at about 2 ppm, much less than carbon dioxide (currently about 380 ppm). However, because it is about 30 times as effective as CO_2 in blocking infrared radiation from leaving the Earth, it makes a significant contribution to the greenhouse effect. The methane content of the atmosphere has been increasing for the last century or so, probably as a consequence of increased animal husbandry. However, this trend seems unlikely to present a major problem. The real potential danger from methane comes from the very large amounts that are sequestered as hydrate complexes deep in the oceans or in permafrost in Arctic regions (Buechel 2007). These complexes are stabilized by high pressure and/or low temperatures. If global warming should melt a significant portion of the perma frosts in Siberia or northern Canada, very large amounts of methane could be liberated in a short time. It is to be expected, however, that dioxygen in the atmosphere will oxidize the methane within only a few decades (eventually to CO_2), but this response might not be fast enough to prevent a rapid "spike" in world temperatures, which might have other unexpected effects. There is precedent for such events. According to Sluijs et al. (2007), the concentration of CO_2 increased in the atmosphere about 55 MYA probably due to intense volcanic activity. The warming of Earth resulted in the rapid decomposition of frozen methane hydrate deposits found in the deep ocean near continental margins and in the Arctic. Large amounts of methane were released into the atmosphere which reacted with dioxygen to produce additionally huge amounts of carbon dioxide. This chain reaction resulted in an increase by about 6°C within a few thousands of years.

7.2.6 Greenhouse to Icehouse and Vice Versa?

There also exists the possibility that the effects of changes in atmospheric CO_2 concentration may not be gradual, but involve abrupt shifts from one global state to another. Recent analysis (DeConto et al. 2008; Pekar 2008) indicate that about 35 MYA, Earth may have shifted rather abruptly from a "greenhouse world" in which glaciation was essentially nonexistent, to its present, rather heavily glaciated state – an "icehouse world." The transition appears to have been triggered when CO_2 levels fell below about 1,000 ppm (see Fig. 7.5). The disturbing thing about this analysis lies in the unstated question: If CO_2 once again rose to over 1,000 ppm (which might occur within a century), could the reverse transition abruptly happen? The disappearance of all glaciations in a short period would indeed be a catastrophe. Although the analysis is still somewhat speculative, it points up the fact that there is much we still do not know about climate changes.
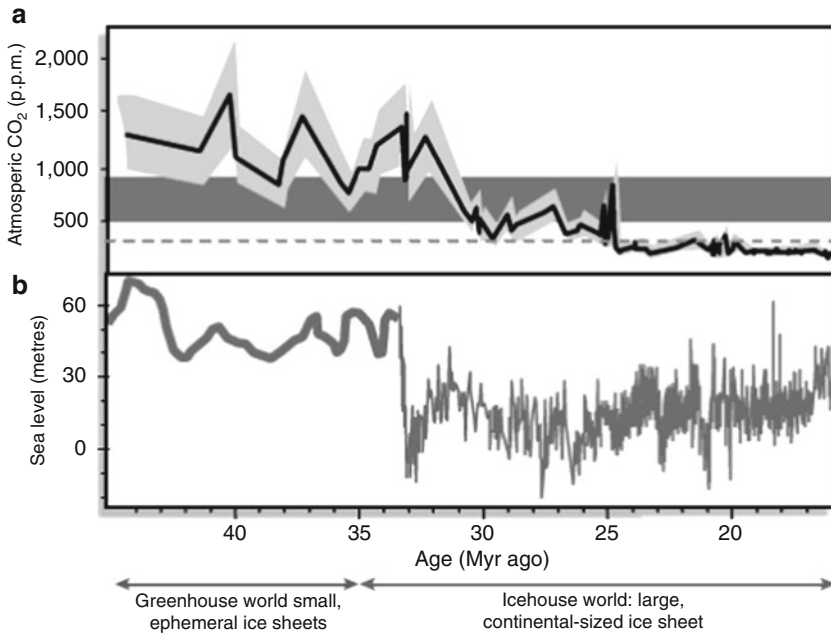
Fig. 7.5 Greenhouse to icehouse. (**a**) Shows estimated CO_2 levels over the last 50 MYA. The *dotted line* shows current level, the *shaded* region estimates for values to be reached in next century. (**b**) Shows estimated global sea levels relative to a zero at which Antarctica is fully glaciated. Estimates from 45 to 34 MYA and 34 to 0 MYA are from different sources (adapted from Pekar 2008; DeConto et al. 2008)

7.3 Human Consequences of Global Warming

In considering the possible effects climate changes may have on human society, it is essential to understand that society is much more fragile than the physical world, or even the biosphere as a whole. Environmental changes that might be trivial in the physical sense or to plants and animals as a whole can be devastating to a complex, interdependent social order. Thus, perceived or predicted consequences of global warming may be infinitesimal in comparison to changes that have occurred naturally in the distant past, yet of great concern to humanity. The more global and interdependent our society has become, the more vulnerable it is. In this section, we must consider global warming from this perspective.

7.3.1 Direct Consequences of CO_2 and Temperature Increase

As can be seen from Table 7.1, there is no likelihood that CO_2 levels per se will increase to values directly dangerous to human health. The expected scenario of

Table 7.1 Physiological impact of CO_2 (vol%) to man

Carbon dioxide		
ppm	Vol%	Comment
380	0.038	Concentration in the atmosphere at present
15,000	1.5	Increase in respiration volume by more then 40%
40,000	4	Concentration of exhaled air
50,000	5	Headache, dizzy
80,000	8	Loss of consciousness, exposure for 30–60 min leads to death

global temperature increase over the next century depends a good deal on the measures that humanity elects to pursue to contain the problem. The IPCC Report in 2007 predicts between 1°C and 4°C increase in the global average by 2100. This will, however, be far from uniform over the globe, with some areas expecting significantly greater changes. But this should not be, even at its extreme, a generally deadly change for humans, and while some regions become less habitable, others will become more attractive. As the change will probably be gradual on a human time scale, it will likely result in a gradual shift in populations. Energy demands will shift in complicated ways, as more artificial cooling is required in hot areas, but less heating is needed in formerly cool, now warmer regions. However inconvenient as it may be to some, the temperature change *itself* does not seem likely to constitute a major threat to humanity.

7.3.2 Sea Level Rise

As mentioned above (Fig. 7.2), the estimated sea level rise from anthropogenic global warming that has occurred to date is of the order of 20 cm. A prediction based mainly on the projected melting of continental glaciers is about another 50 cm in the next century. This will have modest effects, mainly in low-lying coastal regions (Florida, the Netherlands, Bangladesh, Venice, etc.). Even in such localities, the fact that such rise will be gradual allows the possibility of construction of dikes and barriers. Nonetheless, the likelihood of extensive flooding from violent storms will increase (see below).

Much more serious, even catastrophic, would be widespread melting (or moving onto the sea) of the great ice caps of Antarctica and/or Greenland. As described above, this could result in tens of meters of sea level rise, enough to drown many of the world's great coastal cities, and make many lowland areas uninhabitable. Such an event would be particularly dangerous if it happened suddenly, even on a scale of a few decades. This is in contrast to melting of Arctic sea ice shields, which will not change the sea level (see above).

The fact that very large changes in sea level have occurred during glacial and interglacial episodes in the past should not mean such can be dismissed as inconsequential. There did not then exist an evolved civilization which had to respond to

these earlier fluctuations. There is a great deal of uncertainty as to the likelihood of catastrophic sea rise, but it remains a possibility that must not be forgotten.

7.3.3 Extreme Weather

There is evidence that the number and ferocity of violent storms has increased in many parts of the world in recent years. Many believe that this may be accounted for by recent global warming, but the evidence is far from being complete. If this is true, it marks another very serious consequence of warming, at least from the human point of view. Hurricanes like Katrina would be unimportant events on a world without humans; they become catastrophes when they strike great metropolitan areas like New Orleans. In the future, the effects of such storms on coastal areas will be further compounded by sea level rise; thus, warming may have a twofold effect. Again, this is a possible consequence that must be watched with caution. Although the press reports the increased severeness of tropical storms, it has also to be considered that a storm today will cause a relatively higher damage than an equivalent one hundred years ago. Tropical storms mostly occur in the same area. In the past, not so much valuable property was present in these areas that could be destroyed. Thus, after correcting for these effects, an increase of destruction by tropical storms cannot be observed. Fig. 7.6 demonstrates for the area of the
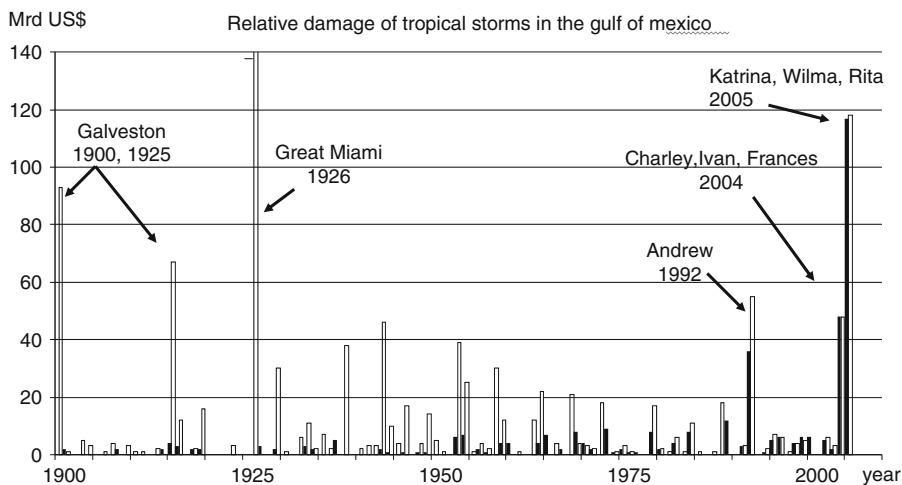
Fig. 7.6 Relative damage of tropical storms in the Gulf of Mexico. Scheme of economical damages (of more than 20 billion dollars) by hurricanes at the Atlantic and Gulf coast of the USA between 1900 and 2005. Given in dollar and taking into account the inflation occurring by 2005. Additionally, taking into account the development of population and their properties like expensive homes. *Black columns*: Inflation corrected; *open columns*: corrected for inflation and increase in properties (after Boehm 2008 and Pielke et al. 2008)

Gulf of Mexico that the storms of the early 1900s would have been as devastating as Katrina, if occurring with present population and evaluated at current dollar value.

7.3.4 Effects on Agriculture

The effects on agriculture (including forest productivity) are likely to be both considerable and varied in impact. It has been predicted that some presently productive areas, particularly in midlatitude regions, will become more arid, and lose productivity. On the other hand, northerly regions such as Canada and Siberia may become more suitable for food crops. Overall, some increase in both food and forest productivity should be expected because of the higher CO_2 concentration (Canadell et al. 2007).

It has been predicted that CO_2 levels will rise to near 1,000 ppm within a century; interestingly, this is near the maximum productivity value for many crops. Greenhouse operators today often adjust the CO_2 level to about this value for maximum efficiency. So there is at least one probably positive aspect to the current trend. There may be quite different effects on various kinds of plants. Some plant species will rapidly spread and others will equally rapidly go extinct. This is bound to constitute a drastic disturbance to intricate ecosystems, hard to cope with by evolutionary adaptation because the changes may occur so fast. The repercussion of these changes in plant populations (and the resulting impact on entire ecosystems) on our planet's climate is hard to predict. There exist two major classes of plants, termed C3 plants and C4 plants, distinguished by certain metabolic pathways. C4 plants evolved from C3 plant. The key enzyme is an ancient one, Rubisco (ribulose-1,5-bisphosphate-carboxylase/-oxygenase). At that time, plenty of CO_2 was available and much less O_2 was present. The younger C4 plants, including maize and sugar cane (being members of the family of grasses) as major examples, are better adapted to high temperatures, and thus may be expected to be favored by global warming. They can keep their stomata partially closed and therefore conserve water. Stomata (singular: stoma) are openings in the leaves for the exchange of O_2, CO_2, and water vapor. In cooler climates, C3 plants are favored because it takes less energy for them to fix carbon dioxide. In hot and dry climates on the other hand, C4 plants can more effectively use CO_2 and water than C3 plants which need two to three times as much water than C4 plants. The optimal temperature range is between 30 and 45°C for C4 and 15 and 25°C for C3 plants. This may explain why C4 grasses dominate the savannah of West Africa.

Human agriculturalists are adaptable. For example, Oregon vintners, who have hitherto favored white grapes, because of the cool climate, are already anticipating longer, hotter summers by planting red grape varieties long favored in California. A similar observation was reported for England during the warm time from 800 to 1200 AD when wine was produced in what is normally a wine-poor climate. Agricultural practices will surely respond to climate change. A strong indication

of a warming effect is the shift of harvesting time of grapes in Middle Europe from October to September by about 2 weeks during the last few hundreds of years.

7.4 Control of Global Warming

7.4.1 Positive and Negative Natural Feedback Mechanism

We must first consider: what built-in features of the Earth's ecosystem exist to effect global warming, now that it has begun? Two major classes can be distinguished: positive feedback, which tends to accelerate any change, and negative feedback, which tends to damp out changes. Positive feedback is characteristic of runaway processes, such as the runaway greenhouse on Venus. Negative feedback tends to make a process self-restricting, leading to equilibrium. A system subject to positive feedback, on the other hand, tends to be unstable. We have discussed feedback with respect to long-range climate control in Chap. 6. Here, we ask whether such factors could mitigate/regulate anthropogenic global warming. Of course, complex natural processes like the global climate usually contain both kinds of effects, with the result that the overall behavior is complicated and hard to predict. Let us consider a few of the factors affecting global warming.

Positive feedback A classic example here is the albedo effect, which can accelerate either warming or freezing, depending on circumstances. When Earth is cooling, ice cover grows, leading to more reflection of sunlight and increased cooling. On the other hand, while the Earth is warming, ice melts, and more solar energy can be absorbed. Another example arises from the fact that water vapor is a quite efficient greenhouse gas. As the oceans warm, the amount of water vapor in the atmosphere rises, leading to more warming. It is believed that this is what may have happened on Venus, where the temperature finally increased enough to completely vaporize the ocean.

A more subtle form of positive feedback may result from effects of dissolving excess CO_2 in the oceans. It has been calculated that if CO_2 concentration in the atmosphere rises to 1,000 ppm, the pH of surface water may drop from 8.2 to about 7.7 (see Sect. 6.4). This might be enough to negatively effect CO_2-consuming phytoplankton, and thereby decreasing their drawdown of CO_2 from the atmosphere.

Negative Feedback The albedo effect provides another example. As water is vaporized, more clouds are formed, which reflect sunlight, producing cooling. Which effect of vaporization will dominate is difficult to predict, and perhaps there is an unstable balance. The rise in CO_2 in the atmosphere will stimulate the growth of plants that remove this compound. If a portion of the carbon is "buried," this will reduce the greenhouse effect. If areas of Earth that were formerly covered by vegetation become desert, they will reflect more sunlight and again reduce the heat input.

    In brief: there are many opposing positive and negative natural influences on the global warming process, so as to make prediction quite difficult. Earth's ecosystem is resilient, however, and we may take comfort in the observation that over nearly 4 billion years, the Earth has never experienced a truly runaway greenhouse or icehouse effect as observed on other planets of the solar system.

7.4.2   Human Efforts to Control Global Warming

Perhaps the most interesting (and effective) feedback mechanism comes from the recognition by humanity of the fact of global warming, and the decision to do something about it. Indeed, intelligence is the most flexible and potentially efficient of all feedbacks; it is what keeps us alive. Especially during the past decade, there has come recognition of the potential magnitude of the problem, and international planning to deal with it. Establishment of the IPCC is a first step, and it seems likely that widespread protocols will be developed in the near future to decrease CO_2 emissions. Dealing with this problem is within our grasp. On the other hand, no panic should be created since as we have indicated above, there exist various mechanisms to compensate even major man-made fluctuations.

7.4.3   The Long View

How long will Earth be habitable by humankind? We have argued above that a runaway greenhouse effect is unlikely, for there are both natural (and now human) feedback mechanisms to regulate extreme climate changes. It may well be that we will experience new ice age episodes in the next hundreds of thousands of years, and some very warm interglacial's. But these will develop slowly in comparison to human perceptions, and can be accommodated. On the other hand, the overall slow decrease in CO_2 mentioned in Chap. 6 may point to an inevitable series of unrelenting ice ages in a distant future. Major changes in the dioxygen content of the atmosphere are unlikely, given the balance of photosynthetic and aerobic metabolism on this planet. The only serious near-term danger is from asteroid or major meteorite impacts. Calculations tell us that an asteroid of 250–400 km diameter would vaporize the oceans, and raise surface temperatures to about 1,000°C. Unfortunately, while we can imagine developing technology to deflect smaller bodies, such giants would seem beyond our control.

    If we survive this threat, which seems very unlikely to occur the next decades, life on Earth is still ultimately doomed. Within a few billion years, our sun will expand into a red giant, enveloping the orbit of the Earth. But it is our conviction that by this far distant time, humankind will itself have expanded throughout the galaxy.

References

Alley RB, Clark PU, Huybrechts P, Joughin I (2005) Ice-sheet and sea-level changes. Science 310:456–460

Boehm R (2008) Heisse Luft. Bene, Wien-Klosterneuburg. ISBN 978-3-85167-213-8

Braithwaite RJ (2002) Glacier mass balance: the first 50 years of international monitoring. Prog Phys Geogr 26:76–95

Broecker WS (1991) The great ocean conveyor. Oceanography 4:79–91

Buechel KG (2007) Der Klimaschwindel. C. Bertelsmann Verlag, München. ISBN 978-3-570-01010-5

Canadell JG, Le Quéré C, Raupach MR, Field CB, Buitenhuis ET, Ciais P, Conway TJ, Gillett NP, Houghton RA, Marland G (2007) Contributions to accelerating atmospheric CO_2 growth from economic activity, carbon intensity, and efficiency of natural sinks. Proc Natl Acad Sci USA 104:18866–18870

Church JA, White NJ (2006) A 20th century acceleration in global sea-level rise. Geophys Res Lett 33:L01602

DeConto R, Pollard B, Wilson PA, Pälcke H, Lear CH, Pagini M (2008) Thresholds for cenozoic bipolar glaciations. Nature 455:652–654

Hansen B, Østerhus S, Quadfasel D, Turrell W (2004) Already the day after tomorrow? Science 305:953–954

Hansen J, Mki S, Ruedy R, Lo K, Lea DW, Medina-Elizade M (2006) Global temperature change. Proc Natl Acad Sci USA 103:14288–14293

Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) (2001) Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. Cambridge University Press, New York

IPCC (2007) Climate change 2007: the physical science basis. Summary for policymakers for working group I of the intergovernmental panel on climate change (available from http://www.ipcc.ch)

Johannessen OM, Khvorostovsky K, Miles MW, Bobylev LP (2005) Recent ice-sheet growth in the interior of Greenland. Science 310:1013–1016

Kerr RA (2007) Global warming is changing the world. Science 316:188–190

Meehl GA, Washington WM, Collins WD, Arblaster JM, Hu A, Buja LE, Strand WG, Teng H (2005) How much more global warming and sea level rise? Science 307:1769–1772

Nerem RS, Leuliette E, Cazenave A (2006) Present-day sea-level change: a review. CR Geosci 338:1077–1083

Overpeck JT, Otto-Bliesner BL, Miller GH, Muhs DR, Alley RB, Kiehl JT (2006) Paleoclimatic evidence for future ice-sheet instability and rapid sea-level rise. Science 311:1747–1750

Pekar SF (2008) When did the icehouse come? Nature 455:602–603

Petit JR, Jouzel J, Raynaud D, Barkov NI, Barnola JM, Basile I, Bender M, Chappellaz J, Davis M, Delaygue G, Delmotte M, Kotlyakov VM, Legrand M, Lipenkov VY, Lorius C, Pepin L, Ritz C, Saltzman E, Stievenard M (1999) Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature 399:429–436

Pielke RA Jr, Gratz J, Landsea CW, Collins D, Saunders MA, Musulin R (2008) Normalized hurricane damage in the United States: 1900–2005. Nat Hazard Rev 9:29–42

Rahmstorf S (2007) A semi-empirical approach to projecting future sea-level rise. Science 315:368–370

Rignot E, Kanagaratnam P (2006) Changes in the velocity structure of the Greenland ice sheet. Science 311:986–990

Siddell M, Rohling EJ, Almogi-Labin A, Hemleben Ch, Meischner D, Schmelzer I, Smeed DA (2003) Sea-level fluctuations during the last glacial cycle. Nature 423:853–858

Sluijs A, Brinkhuis H, Schouten S, Bohaty SM, John CM, Zachos JC, Reichart G-J, Damsté JSS, Crouch EM, Dickens GR (2007) Environmental precursors to rapid light carbon injection at the Palaeocene/Eocene boundary. Nature 450:1218–1221

Stoddart DM (ed) (2008) Australia's contribution to Antarctic climate science. Australian Antarctic Division. http://www.aad.gov.au/default.asp?casid=34583

Thomas R, Frederick E, Krabill W, Manizade S, Martin C (2006) Progressive increase in ice loss from Greenland. Geophys Res Lett 33:L10503

Velicogna I, Wahr J (2006) Measurements of time-variable gravity show mass loss in Antarctica. Science 311:1754–1756

Zedillo E (2008) Global warming. In: Rahmsdorf S (ed) Anthropogenic climate change: revisiting the facts. Brooklings University Press, Washington, pp 34–53

Chapter 8
Oxygen in Medicine

Contents

The implications of oxygen for medicine are basically of two kinds. First, there is the continued need by human tissues for an adequate supply of dioxygen; if that is not met, a condition called hypoxia may arise, with serious medical consequences. There are a wide number of causes for hypoxia, and a variety of medical responses.

On the other hand, as we have seen in Chap. 3, the normal metabolism of oxygen in the body can give rise to reactive oxygen species (ROS). These in turn can have serious medical consequences. Alternatively, the proper functioning of human metabolism also depends on certain ROS serving as signal molecules. We shall briefly consider these two aspects of a vast field.

8.1 Hypoxia

In general, any situation in which an organism or any tissues therein, may be deprived of an optimal dioxygen supply is termed hypoxia. Briefly, we can consider the major reasons for whole-body deprivation in humans to be either environmental, particularly high altitudes, where oxygen pressure is insufficient to allow normal tissue supply, or from medical condition, which limits oxygen distribution to tissues.

8.1.1 High-Altitude Hypoxia

Everybody realizes that dioxygen supply becomes short at high altitudes. The tolerance of different animals varies widely. While at sea level the atmospheric oxygen partial pressure is about 160 Torr, it will decrease to about 95 Torr at 4,000 m (where lamas will be found) or to 45 Torr at about 9,000 m. Here, flying bar-headed geese have been encountered in the Himalayan area. Even at 11,300 m at Africa's Ivory Coast, Rüppel's griffon have been observed flying in an atmosphere with only about 33 Torr oxygen partial pressure. But for humans, at about 7,000 m, the pO_2 in the alveoli in our lungs will be close to 30 Torr, the consequential low level of dioxygen saturation in the blood of most humans under these circumstances is not sufficient to transfer dioxygen to the mitochondria. This problem will be even more extreme at the top of Mount Everest at about 8,840 m.

The pressure and therefore concentration of dioxygen steadily decrease as we climb to higher altitudes. In order to keep the available dioxygen at the tissues (about 20 mmHg) constant, physiological adaptation is needed. The simplest response is hyperventilation. This, however, has a bad consequence since CO_2 will also be released and therefore its buffering effect in the blood will be disturbed (see Table 7.1). The dizziness caused by fast breathing is due to this alkalosis. A more gradual adaptation is the increase in 2,3-bisphosphoglycerate (2,3 BPG), which results in a right shift of the oxygen-binding curve, which releases more dioxygen at low pO_2 (see Chap. 5 and Benesch et al. 1968). An even slower adaptation is to induce the increase in the hemoglobin mass in the blood. This is regulated by erythropoietin in the kidney as a reaction to the local dioxygen concentration. Erythropoietin is also well-known as a doping drug. Bicyclists have used it, hoping to gain more dioxygen by increasing the red cell mass. Most athletes like to exercise at high altitude before any sports event to increase the concentration of 2,3 BPG and to increase the production of hemoglobin. The latter will only have its onset after exposure for at least 4 weeks. For immediate compensation for high-altitude hypoxia, it is necessary to provide supplemental dioxygen in a breathing apparatus of some kind, as is common with high-altitude climbers, military aviators, and air passengers.

Although not strictly a medical question, it is interesting to ask how birds handle the problem of low oxygen concentration considering that flying needs a lot of energy. Mountain climbers struggling to breathe on Mt. Everest report that birds are flying above them in an altitude of about 11,000 m without any problem. There are two major reasons: The avian respiratory system is more than 30% more efficient than any mammalian lung since it allows a unidirectional airflow while the blood flow is oriented in a countercurrent way. This allows extracting most of the dioxygen, much more than our sac-lungs can provide. Another way is through an increase of the high affinity myoglobin molecules in the tissue. At the molecular level, a left shift of the sigmoid oxygen-binding curve of hemoglobin is required resulting in a higher affinity of hemoglobin, thereby allowing a complete saturation with dioxygen even in an environment with low dioxygen concentration. This may be negative for the

unloading of dioxygen at the tissue to load the myoglobin in the cells. But an increase in capillarization may compensate for that (Weber 2007). Another special way of molecular adaptation is the use of inositol phosphate (ITP) by birds instead of 2,3 DPG by humans (Chap. 5). Due to the less negative charges, ITP does not stabilize the T state as strongly as 2,3 DPG. Thus, a higher oxygen affinity at low pO_2 will be obtained. However, most important seems to be the fact that hemoglobin from high flying birds has a few mutations resulting in a higher affinity (Weber 2007).

8.1.2 Hypoxia Arising from Medical Conditions

Many instances of hypoxia in humans arise from medical conditions that restrict the oxygenation of blood, the flow of blood in specific tissues, or the capacity of the blood to carry dioxygen. They include diverse causes such as pulmonary disease, ischemic restriction of blood flow in the brain, carbon monoxide poisoning. We must include the numerous varieties of anemia that decrease hemoglobin and/or red cell content of the blood. Although there are many specific treatments for these hypoxic states, a general method for acute hypoxia, applicable to many hypoxias, is hyperbaric oxygen therapy (Neubauer and Walker 1998).

The idea is relatively simple: the patient is placed in a chamber ventilated with pure dioxygen, at pressures up to twice ambient. Since air at one atmosphere is about 20% dioxygen, pure dioxygen at two atmospheres corresponds to a dioxygen concentration about ten times normal. Recall from Chap. 5 that the concentration of dioxygen *dissolved* in blood is linearly related to the oxygen partial pressure; we find from Fig. 5.1 that at a pressure tenfold higher than normal, the blood plasma will dissolve more dioxygen than normally carried by the hemoglobin. In fact, the amount carried by the hemoglobin is not significantly increased as it is normally close to saturation in the lungs. Under such hyperbaric conditions, dioxygen can be absorbed directly into tissues to aid in wound healing, etc.

Prolonged exposure to such high levels of dioxygen can have deleterious effects. One major class is observed in accentuated production of ROS (see Chap. 3). In fact, problems with ROS are common in humans and are frequently encountered, in a situation referred to as oxidative stress.

8.2 Oxidative Stress

8.2.1 Nature of Oxidative Stress

A major problem that higher organisms have with dioxygen arises from ROS. In Chap. 3, we saw how many organisms cope with ROS. Here, we consider a few examples of how humans deal with and even utilize ROS. Man is very dependent on

dioxygen and utilizes a large amount of it without damage. However, disturbances in the normal redox state can cause oxidative damage through the production of peroxides and free radicals (Ketteridge 2000) that can damage all components of the cell, including proteins, lipids, and DNA (Chap. 1 and 3). It is estimated that the DNA in a typical cell will undergo about many thousands of oxidative hits per day, resulting in at least a dozen DNA lesions per day. Most of them cause somatic mutations. In humans, oxidative stress is believed to be involved in many diseases, although we often do not know the precise mechanisms. A number of examples that have been suggested are related to aging such as atherosclerosis, Parkinson's and Alzheimer's diseases (Hogg 1998; Fang 2004; Wang et al. 2006).

ROS can originate from exogenous sources such as ultraviolet radiation and ionizing radiations or from a number of normal and essential intracellular processes. A free radical theory of aging suggests that aging results from the accumulated damage generated by ROS (Gerschman et al. 1954; Harman 1956; Szilard 1959; Beckman and Ames 1998). According to this theory, aging is correlated with the amount of dioxygen consumed during life. In fact, long-living organisms seem to have lower energy consumption than the short-living ones. For example, consider birds, which do not live as long as humans presumably because they have a many times higher metabolic rate. On the other hand, the slowly moving and acting turtles can live more than 100 years. A nice comparison can be made between mice and cows. Both consume similar total amounts of energy per unit mass during their lives but the active mice live only 3 years, while the sedentary, lethargic cows live about 30 years. It is also proposed that a low calorie diet extends median and maximum lifespan in many animals. This effect may involve a reduction in oxidative stress since less energy has to be produced based on a lower energy metabolism. The role of oxidative stress in aging seems to be well understood in model organisms such as *Drosophila melanogaster* and *Caenorhabditis elegans*. In mammals, however, the mechanism is less clear. Diets with a high antioxidative capacity as found in fruit and vegetables have been suggested to promote health and reduce the effects of aging. However, substances, which neutralize oxidants (called antioxidants), do not explain the differences in longevity between mice and rats (Andziak et al. 2005). Similarly, several studies indicate that antioxidant vitamin supplementation seems to have no detectable effect on aging processes. This may be a hint that the positive effects of fruit and vegetables depend not only on their antioxidant contents as is so often assumed but on other factors as well. To summarize, we can manage oxidative stress, but with age, it seems that a continued creation of radicals will occur with deleterious effects.

8.2.2 Special Examples of Medical Consequences of Oxidative Stress

A few examples will help clarify the role of oxidative stress in human health. The role of oxidative stress in cardiovascular disease is based on oxidation of low-density

lipoprotein (LDL) which appears to trigger the process of arthrosclerosis, and finally ends in cardiovascular disease (Wang et al. 2006; Bhakdi et al. 2004; Schwarz et al. 2008). Especially in the presence of homocysteine, which is an indicator for "risk" patients, production of ROS may result in homocystinuria as well as atherosclerosis and stroke.

8.2.2.1 DNA

Another target for ROS is DNA. According to the "DNA Damage Theory," accumulation of damaged DNA may be a major factor in aging. Antioxidative enzymes such as SOD, catalase, and glutathion peroxidase may play a role in delaying aging. Interestingly, when transgenic fruit flies overexpress these antioxidants, they show a more than 30% increase in average longevity and a delayed aging process. Similar effects were observed for the nematode worm *C. elegans*, but not in much more complex mammals (Andziak et al. 2005).

Among the four nucleotides from which DNA is built, one – guanosine – can easily be affected by ROS by forming a keto group yielding 7, 8-dihydro-8-oxo-deoxy-guanosine (8-oxo-dG) from deoxy-guanosine (Helbock et al. 1998). During cell division, the latter will not be read correctly by DNA polymerase but will be read as deoxy-thymidin (dT). This will cause a mutation, possibly resulting in instability, misfolding, or inactivation of the protein that is coded for by the mutated genes. These mutations occur more often with increasing age. Of course, cells have developed repair mechanisms which cut out the damaged part of the DNA and replace it by a correct sequence but these are never 100% efficient.

It is the DNA of mitochondria that is especially affected by oxidative stress probably because of the production of ROS by the oxidative processes in mitochondria (see Chap. 4). Although there exist many DNA repair mechanisms not all function in mitochondria. In contrast to chromosomal DNA in the cell nucleus, mitochondrial DNA is not associated into a complex with histone proteins (called chromatin) which provides some protection of the DNA against ROS.

During recent years, the role of oxidative stress in neurologic diseases including Alzheimer's disease, Parkinson's disease, multiple sclerosis and AIDS dementia, etc., has been extensively studied (Cherubini et al. 2005). Still, although several such diseases appear to be related to failure of DNA repair mechanism, it is not known whether these changes are effects or causes of aging. In a disease such as Alzheimer's, oxidative stress/oxidative damage is felt to play a key role in the loss of neurons and the progression to dementia (Rice-Evans and Gopinathan 1995; Cherubini et al. 2005), but the exact mechanism is uncertain.

8.2.2.2 Proteins

Proteins may be directly damaged by ROS lowering their stability and/or eliminating their function (Petropoulos and Friguet 2006). Several amino acid residues

are susceptible to oxidation including cysteine, tryptophan, tyrosine, and histidine. Many ROS especially react with methionine residues in proteins to form methionine sulfoxide, thus scavenging the reactive species. However, the more polar sulfoxide side chain may result in disrupting the protein structure or in a loss of binding metal ions at the active site. During evolution, cells have learned to recover these oxidized methionines. Almost all cells contain methionine sulfoxide reductases. They catalyze a thioredoxin-dependent reduction of methionine sulfoxide back to methionine (see Chap. 3). Thus, methionine residues may act as catalytic antioxidants (Luo and Levine 2009). On the other hand, all cells make a tiny bit of hydrogen peroxide (Chap. 4), but with increasing age, this little bit becomes a lot. This is the reason why our hair turns gray and white with increasing age (Wood et al. 2009). Gray or white hair develops with advancing age in an entirely natural aging process which results in the generation of less and less color pigments, mainly melanin. Free oxygen radicals are significantly involved in the loss of hair color. The originator of the entire graying process is hydrogen peroxide. With advancing age, hydrogen peroxide builds up in larger amounts in the hair follicle and ultimately inhibits the synthesis of the color pigment melanin. Hydrogen peroxide is a by-product of metabolism in all mitochondria in each cell (see Chaps. 3 and 4), and as such it is generated in small amounts throughout the human body, consequently also in hair follicles. With increasing age, the quantity builds up, because the human body can no longer continue neutralizing the hydrogen peroxide using the enzyme catalase (Chaps. 1 and 3). In aging cells, this enzyme is still present but in very limited concentration. The rise in H_2O_2 has dramatic consequences. Among other proteins, hydrogen peroxide attacks the enzyme tyrosinase by oxidizing methionine, at the active site. As a consequence, this key enzyme, which normally starts the synthesizing pathway of the coloring pigment melanin, no longer functions.

Molecular dynamic studies revealed the underlying process based on the structure of human tyrosinase (Matoba et al. 2006; Decker et al. 2006; Schweikardt et al. 2007): The oxidized methionine interacts with another amino acid in the protein, a histidine, which then cannot orientate the substrate precisely to begin melanin formation (Wood et al. 2009). It should be pointed out that this effect has nothing to do with artificial bleaching of hair with peroxide. The applied hydrogen peroxide is low in concentration to avoid irritation of the skin of our head and will therefore not diffuse into the hair follicles but only bleach the hair. After several days, colored growing hair will be observed again. Oxidation by in vivo-generated hydrogen peroxide not only interferes with the production of melanin but also inhibits other enzymes that are needed for the repair of damaged proteins. As a result, a cascade of events is set off, at the end of which stands the gradual loss of pigments in the entire hair from its root to its tip. This understanding may also be the basis for future therapy of vitiligo, a skin pigment disorder also induced by hydrogen peroxide. Here, melanin is not synthesized for the same reason, causing white spots on the skin.

8.3   Treatment of Oxidative Stress

As implied by our discussion of ROS in Chap. 3, organisms have evolved modes of dealing with oxidative stress.

Our major interest here, however, is how medical and pharmacological treatments may limit oxidations. As we have discussed in Chap. 3, one strategy is to fight oxidative stress with antioxidants such as the well-known vitamins C and E, which neutralize the radicals.

How effectively vitamins can be used for treatment is not entirely clear yet. It is thought that oxidation of LDL in the blood contributes to heart disease, and initial studies found that people taking vitamin E supplements had a lower risk of developing heart disease (Sies and Stahl 1995). Consequently, a number of clinical trials were conducted to test the effects of antioxidant supplement with vitamin E, in doses ranging from 50 to 600 mg per day. However, the topic is still controversial. None of these later trials found a statistically significant effect of vitamin E on overall number of deaths or deaths due to heart disease.

The brain is of special interest due to its high metabolic rate and high concentration of polyunsaturated lipids that are targets for peroxidation. Antioxidants would seem to be ideal candidates for protection of the brain and nervous system from oxidative stress/damage. They are well absorbed from the gastrointestinal tract and then effectively enter the nervous system, to improve the GSH level which regulates the oxidative level in a cell. Alpha lipoic acid and probably another cofactor CoQ 10 (a part of the respiratory chain) fulfill such important characteristics. Thus, antioxidants are used in the treatment/prevention of Alzheimer's disease, Parkinson's disease, and similar neurological disorders. Attesting to the importance of antioxidants in the brain is the fact that there are observed the highest concentrations of two enzymes that oppose ROS: superoxide dismutase and catalase.

It is also believed that certain trace minerals (e.g., selenium and zinc) are necessary for certain antioxidants to function properly. A modern strategy is to synthesize antioxidant compounds that meet the necessary requirement for uptake and minimal side effects.

Physical exercise is well-known for maintaining good health, but reasons are not entirely clear. Under exercise, production of oxidants is increased and results in damage that contributes to muscular fatigue during and after exercise. This is due to the fact that dioxygen consumption can increase by a factor of more than 10. It seems that the *inflammatory response* that occurs after extended exercise is also associated with oxidative stress. On the other hand, the immune system responds by production of free radicals through *neutrophils* to remove damaged tissue. The better health of those who undertake regular exercise is believed to be based on a strengthening of the body's antioxidant defenses, particularly the glutathione system, to deal with the increased oxidative stress. This seems a bit strange, but we must recognize that ROS also play many useful and even contradictory roles in human physiology.

8.4 Beneficial Roles of ROS

As we have seen in earlier chapters, ROS can also be beneficial. During evolution, organisms have learned to use the normally dangerous oxidants for important functions. Thus, ROS are thought to be involved in cell proliferation, programmed cell death, expression of genes, and modulation of signaling pathways (Butler and Nicholson 2003). A major defense line against microbes involves neutrophils and other phagocytic white blood cells (Nathan and Shiloh 2000). The killing agents produced by these cells include superoxide, nitric oxide ($^{•}$NO), and their particularly reactive product, peroxynitrite ($OONO^{-}$) (see Chap. 1). Another defense against bacteria is based on NADH oxidase (Fang 2004), a membrane-bound multisubunit complex in neutrophils. The complex reduces molecular oxygen to superoxide (consuming NADPH) and releases it into a phagosome, a cellular compartment in which pathogenic microorganisms can be killed and digested (Jones et al. 2000; Ketteridge 2000).

Organisms also use various other enzymes, producing superoxide or hydrogen peroxide, which occurs in phagocytes. These cells also kill pathogen host cells. Thus, a broad spectrum of response is available, with the consequence that a pathogen cannot easily escape, by mutation, from being attacked. In plants, another mechanism seems to be successful. Here, ROS induces a programmed cell death, which hinders dispersal of the pathogens (Hancock et al. 2001; Gupta et al. 2005). Even respiratory proteins of vertebrates (hemoglobin) and invertebrates (hemocyanin) are involved in defending from pathogenic organisms. They can generate highly reactive oxygen intermediates, the release of which can be activated by microbial compounds (Bogdan 2007; Jiang et al. 2007).

8.4.1 SCN and Primary Immune Response

Thiocyanate, SCN, is a compound formed in mitochondria where it detoxifies cyanide (CN), which is extremely toxic for the respiration chain. After certain plants are eaten, CN is released from its glycosylated form which occurs in their vacuoles. It is considered to be a part of the plant defense system. At low concentration, most free CN is converted to thiocyanate (SCN) by *rhodanese*. Rhodanese is an ancient enzyme, reminding us that CN was present in the early history of earth. SCN is considered to be neutral and is released in urine – through the skin, through perspiration, and via body openings such as mouth and eyes (Weuffen and Decker 2004). In presence of peroxidase and traces of H_2O_2, SCN is converted to a very potent radical OSCN. This compound is responsible for the killing of most of the 10^8 bacteria and fungi invading through the mouth per day (Conner et al. 2002; Wijkstrom-Frei et al. 2003; Ashby 2008). This kind of disinfection is also used by many animals when licking open wounds.

8.4.2 Nitric Oxide

Nitric oxide, NO, also known as the "*endothelium-derived relaxing factor*" (EDRF), is one of the few gaseous signaling molecules (Butler and Nicholson 2003; Van Faassen and Vanin 2007; Moncada et al. 1991; Pacher et al. 2007). Within vertebrates, it plays a role in many biological processes. Most prominent is the usage of NO for *vasodilation* and increasing blood flow. Its long lifetime of a few seconds enables NO to diffuse about 200 μm even across a membrane, so as to act as a signal molecule between adjacent cells and within cells. It reacts only with a few biomolecules, although it has an unpaired electron and is therefore very similar to superoxide. In particular, it reacts with other free radicals (i.e., peroxyl and alkyl radicals). Thus, it can be considered as a radical scavanger (Pacher et al. 2007).

Within the body, NO is synthesized by converting L-arginine to L-citrulline plus NO, catalyzed by a complicated protein complex nitric oxide synthase (NOS) (Fig. 8.1). There are different isoforms, eNOS, iNOS, nNOS, and mtNOS, which can be distinguished by the occurrence and function: eNOS reduces the tension of the blood vessels via NO. The activity within the endothelia of blood vessels is dependent on both calcium and mechanical forces. nNOS occurs in brain neurons, acting as a second messenger to induce the increase of cGMP synthesis. cGMP itself is also a most important second messenger in almost all organisms, responsible for signal transduction within a cell. It is involved in the process of vision in the retina cells, the relaxation of smooth muscles and regulation of the insulin level. mtNOS, possibly a splicing variant of nNOS, modulates the metabolism in mitochondria for synthesis, proliferation, apoptosis, and regulation of oxygen consumption.

Endotoxines from bacteria (lipopolysaccharides) or cytokines (TNF-alpha, IFN-gamma) induce the isoform iNOS within macrophages, cardiomyocytes, and smooth muscle cells. The resulting high concentration of NO kills bacteria by diffusing into the bacteria or even cancer cells, binding to iron and copper proteins (Anggard 1994) and resulting in a release of the iron and copper from the active site. In addition, both metals activate dioxygen producing toxic ROS. Thus, NO can be considered as an important part of the immune system.
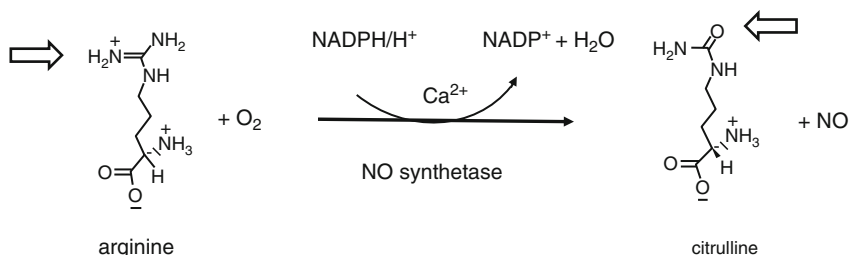
Fig. 8.1 Scheme of synthesis of NO. L-arginine is catalyzed to L-citrulline by a complex enzyme (NO synthetase) producing one molecule of NO. For this reaction, the oxygen and NADPH/H$^+$ are necessary

However, there is always a balance in nature; during evolution, many bacterial pathogens developed mechanisms for nitric oxide resistance (Janeway et al. 2005; Nathan and Shiloh 2000).

The dangerous effect of very high dioxygen levels may be based on the effect that higher dioxygen level may also result in higher amounts of superoxide and increasing iNOS activity. This superoxide reacts with *NO to peroxonitrite $ONOO^-$ which damages proteins especially in membranes when their tyrosines are modified. This is thought to result in cell death.

Early in the human history, Sumerians used NO as an antimicrobial agent without knowing the mechanism when they conserved meat with nitrite salt (NO_2^-). After some time, NO will be formed killing such dangerous bacteria as the well-known *Botulinus clostridium*. In addition, meat so treated will keep the red
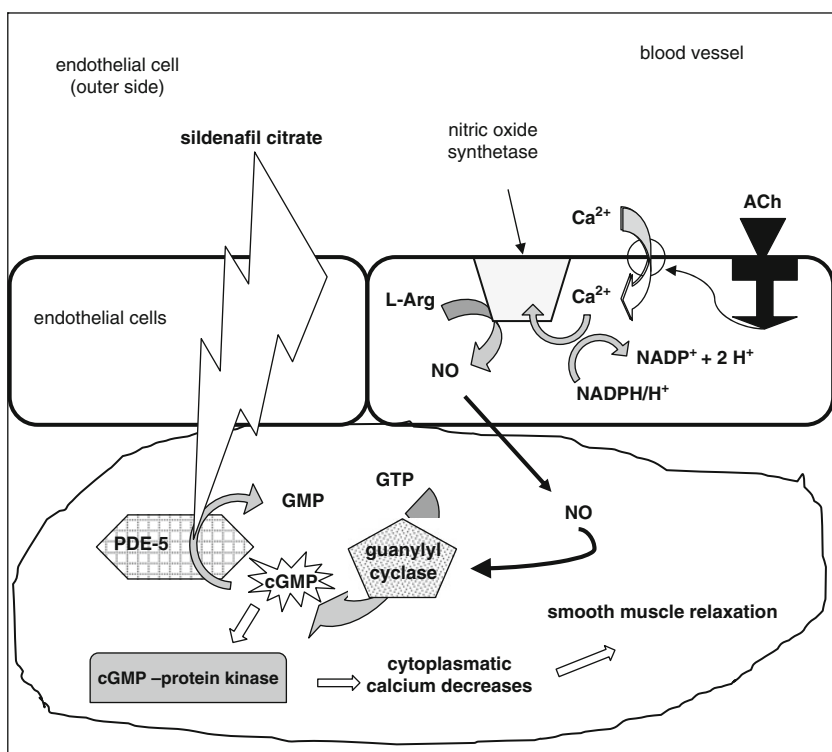
Fig. 8.2 NO stimulation of penile erection. Acetylcholine (ACh) released by cholinergic nerves favors the import of calcium ions into endothelial cells which activate the nitric oxide synthetase. These physiologically produced NO induce accumulation of cGMP in smooth muscle cells binding to an enzyme, the guanylyl cyclase. As a result, the muscle cells relax and induce a vasodilatation of vessels. Thus, blood will flow into the penis allowing an erection. Under normal conditions, cGMP is removed by the action of cGMP phosphodiesterase (PDE-5), which converts it to GMP. This neutralization of cGMP can be blocked by drugs like a citrate salt of sildenafil allowing a prolongation of erection

color due to the binding of NO by myoglobin occurring freely as residual amounts in the muscle.

Another function of NO is observed for people from Tibet adapted to high altitude (about 4,000 m). They have a tenfold higher NO concentration in blood compared to people living at sea level. Here, NO increases the blood circulation by vasodilation and therefore delivers more oxygen to the tissues (Erzurum et al. 2007).

A popular application of NO is the stimulation of the penis erection by vasodilation of the blood vessel. In a complicated cascade of biochemical interactions, NO is produced (Fig. 8.2). There are two sources for the necessary NO. First, it is liberated by the cavernous nerve which innervates smooth muscle surrounding penile arteries. These muscles normally hinder blood to flow into the *corpus cavernosum* resulting in a flacid state of the penis. The second source is the endothelial cells lining the deep arteries of the penis (For a good overview see Boyer 2002). Whatever be the source, endothelial cells or cavernous nerve, NO diffuses into the muscle cell membrane and binds to guanylyl cyclase. This enzyme catalyzes the synthesis of cyclic GMP, thereby relaxing the muscle cells resulting in penile erection due to the inflow of blood in the *corpus cavernosum*. For relaxation, cGMP has to be degraded with an enzyme PDE-5 (phosphodiesterase-5). This, however, can be blocked by drugs such as sildenafil which therefore help maintain penile erection.

References

Andziak B, O'Connor TP, Buffenstein R (2005) Antioxidants do not explain the disparate longevity between mice and the longest-living rodent, the naked mole-rat. Mech Ageing Dev 126:1206–1212

Anggard E (1994) Nitric oxide: mediator, murderer, and medicine. Lancet 343:1199–1206

Ashby MT (2008) Inorganic chemistry of defensive peroxidases in the human oral cavity. J Dent Res 87:900–914

Beckman KB, Ames BN (1998) Mitochondrial aging: open questions. Ann NY Acad Sci 854:118–127

Benesch R, Benesch RE, Enoki Y (1968) The interaction of hemoglobin and its subunits with 2, 3-diphosphoglycerate. Proc Natl Acad Sci USA 61:1102–1106

Bhakdi S, Lackner KJ, Han SR, Torzewski M, Husmann M (2004) Beyond cholesterol: the enigma of atherosclerosis revisited. Thromb Haemost 91:639–645

Bogdan C (2007) Oxidative burst without phagocytes: the role of respiratory proteins. Nat Immunol 8:1029–1031

Boyer R (2002) Interactive concepts in biochemistry. Wiley, New Jersey

Butler A, Nicholson R (2003) Life, death and NO. Cambridge Press, Cambridge. ISBN 13: 978-0-85404-686-7

Cherubini A, Vigna G, Zuliani G, Ruggiero C, Senin U, Fellin R (2005) Dual effects of antioxidants in neurodegeneration: direct neuroprotection against oxidative stress and indirect protection via suppression of glia-mediated inflammation. Curr Pharm Des 12:3521–3533

Conner GE, Salathe M, Forteza R (2002) Lactoperoxidase and hydrogen peroxide metabolism in the airway. Am J Respir Crit Care Med 166:S57–S61

Decker H, Schweikardt T, Tuczek F (2006) The first crystal structure of tyrosinase: all questions answered? Angew Chem Int Ed Engl 45:4546–4550

Erzurum SC, Ghosh S, Janocha AJ, Xu W, Bauer S, Bryan NS, Tejero J, Hemann C, Hille R, Stuehr DJ, Feelisch M, Beall CM (2007) Higher blood flow and circulating NO products offset high-altitude hypoxia among Tibetans. Proc Natl Acad Sci USA 104:17593–17598

Fang FC (2004) Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. Nat Rev Microbiol 2:820–832

Gerschman R, Gilbert DL, Nye SW, Dwyer P, Fenn WO (1954) Oxygen poisoning and x-irradiation: a mechanism in common. Science 119:623–626

Gupta KJ, Stoimenova M, Kaiser WM (2005) In higher plants, only root mitochondria, but not leaf mitochondria reduce nitrite to NO, *in vitro* and in situ. J Exp Bot 56:2601–2609

Hancock JT, Desikan R, Neill SJ (2001) Role of reactive oxygen species in cell signalling pathways. Biochem Soc Trans 29:345–350

Harman D (1956) Aging: a theory based on free radical and radiation chemistry. J Gerontol 11:298–300

Helbock HJ, Beckman KB, Shigenaga MK, Walter PB, Woodall AA, Yeo HC, Ames BN (1998) DNA oxidation matters: the HPLC-electrochemical detection assay of 8-oxo-deoxyguanosine and 8-oxo-guanine. Proc Natl Acad Sci USA 95:288–293

Hogg N (1998) Free radicals in disease. Semin Reprod Endocrinol 16:241–248

Janeway CA, Travers P, Walport M, Schlomchik M (2005) Immunobiology: the immune system in health and disease, 6th edn. Garland, New York

Jiang N, Tan NS, Ho B, Ding JL (2007) Respiratory protein-generated reactive oxygen species as an antimicrobial strategy. Nat Immunol 8:1114–1122

Jones RD, Hancock JT, Morice AH (2000) NADPH oxidase: a universal oxygen sensor? Free radic Biol Méd 29:416–424

Ketteridge DJ (2000) What is oxidative stress? Metabolism 49:3–8

Luo S, Levine RL (2009) Methionine in proteins defends against oxidative stress. FASEB J 23:464–472

Matoba Y, Kumagai T, Yamamoto A, Yoshitsu H, Sugiyama M (2006) Crystallographic evidence that the dinuclear copper center of tyrosinase is flexible during catalysis. J Biol Chem 281:8981–8990

Moncada S, Palmer RMJ, Higgs EA (1991) Nitric oxide: physiology, pathophysiology, and pharmacology. Pharmacol Rev 43:109–142

Nathan C, Shiloh M (2000) Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. Proc Natl Acad Sci USA 97:8841–8848

Neubauer RA, Walker M (1998) Hyperbaric oxygen therapy. Avery, New York

Pacher P, Beckman JS, Liaudet L (2007) Nitric Oxide and peroxynitrite in health and disease. Physiol Rev 87:315–424

Petropoulos I, Friguet B (2006) Maintenance of proteins and aging: the role of oxidized protein repair. Free Radic Res 40:1269–1276

Rice-Evans CA, Gopinathan V (1995) Oxygen toxicity, free radicals and antioxidants in human disease: biochemical implications in atherosclerosis and the problems of premature neonates. Essays Biochem 29:39–63

Schwarz M, Spath L, Lux CA, Paprotka K, Torzewski M, Dersch K, Koch-Brandt C, Husmann M, Bhakdi S (2008) Potential protective role of apoprotein J (clusterin) in atherogenesis: binding to enzymatically modified low-density lipoprotein reduces fatty acid-mediated cytotoxicity. Thromb Haemost 100:110–118

Schweikardt T, Olivares C, Solano F, Jaenicke E, Garcia-Borron JC, Decker H (2007) A three-dimensional model of mammalian tyrosinase active site accounting for loss of function mutations. Pigment Cell Res 20:394–401

Sies H, Stahl W (1995) Vitamins E and C, beta-carotene, and other carotenoids as antioxidants. Am J Clin Nutr 62(6 Suppl):1315S–1321S

Szilard L (1959) On the nature of the aging process. Proc Natl Acad Sci USA 45:30–45

van Faassen EE, Vanin AF (eds) (2007) Radicals for life: the various forms of nitric oxide. Elsevier, Amsterdam. ISBN 13: 978-0-444-52236-8

Wang J, Wen L, Huang Y, Chen Y, Ku M (2006) Role of antioxidants in atherosclerosis: epidemiological and clinical update. Curr Pharm Des 11:2017–2032

Weber RE (2007) High-altitude adaptation in invertebrate hemoglobins. Respir Physiol Neurobiol 158:132–142

Weuffen W, Decker H (eds) (2004) Thiocyanat – ein bioaktives Ion mit orthomolekularem Charakter. I.S.M.H. Verlag, Sarow. ISBN 3-934043-06-2

Wijkstrom-Frei C, El-Chemaly S, Ali-Rachedi R, Gerson C, Cobas MA, Forteza RS, Salathe M, Conner GE (2003) Lactoperoxidase and human airway host defense. Am J Respir Cell Mol Biol 29:206–212

Wood JM, Decker H, Hartmann H, Chavan B, Rokos H, Spencer JD, Hasse S, Thornton J, Shalbaf M, Paus R, Schallreuter KU (2009) Senile hair greying: H_2O_2-mediated oxidative stress affects human hair colour by blunting methionine sulfoxide repair. FASEB J 23:2065–2075

Chapter 9
Oxygen and the Exploration of the Universe

Contents

Humankind has begun, in a tentative way, the immense project of exploring, and perhaps colonizing, other worlds. The grand enterprise has hardly begun and will certainly suffer many defeats and reversals, but it seems destined to go forward. In the course of this, both in seeking life in extraterrestrial environments and voyaging into them, we shall encounter a number of problems concerning the existence or provision of oxygen. The basis for this has been described in previous chapters. First, we would like to summarize arguments as to why life could have evolved on other planets. We need to know what to expect.

9.1 What Is Essential for the Development of Life as We Know It?

Most discussions have made certain common suggestions of factors favoring the development of life on Earth-like planets:

(a) A large planet such as Jupiter or Saturn located farther away from the star will help protect the Earth-like planet from asteroids and comets.
(b) The distance to the star has to be appropriate, so that the planet obtains enough heat but not too much. The phrase "habitable zone" is frequently used.

(c) A moon may be advantageous to slow down the period of day and night as well as the precession of the planet around the star to obtain moderate climate.

(d) An initial atmosphere with at least some (but not too much) greenhouse gases is needed to moderate climate.

(e) The planet should, in its earlier years, provide enough radioactive heating to offset rapid cooling.

(f) Tectonic movement is necessary for releasing greenhouse gases to the atmosphere during periods of glaciations.

(g) For life to persist at or near the surface, a UV shield such as ozone appears essential, but this will not be a factor in the origin of life.

However, as we shall see below, recent discoveries have made questionable many of these assumptions. For example, the conditions now recognized on some of the moons of giant planets may displace the whole concept of a "habitable zone" (see below). In any event, life is unlikely on a planet without either an atmosphere or hydrosphere. On the other hand, a mixture of nitrogen, oxygen, and water vapor as found in our atmosphere is also very unlikely to be encountered on a dead planet.

9.2   What Makes O_2 Necessary for Complex Life on Habitable Planets?

The reduction of dioxygen to water provides the largest possible free energy release per electron transfer, except for the reduction of fluorine and chlorine (which do not occur free in nature). While oxygen is abundant in the universe (see Chap. 2), it does not occur as free dioxygen in most circumstances. This means that dioxygen-dependent metabolism must wait upon oxygenic photosynthesis. However, the bonding of O_2 ensures that dioxygen is sufficiently stable to accumulate in a planetary atmosphere, whereas the more weakly bonded halogen gases are far too reactive ever to achieve significant abundance. Consequently, an atmosphere rich in O_2 provides the largest feasible energy source. This universal uniqueness suggests that abundant O_2 is necessary for the high-energy demands of complex life (see Chap. 4). Thus, the oxygenation time is likely to be a key rate-limiting step in the evolution of life on any habitable planet.

9.3   Seeking Evidence for Extraterrestrial Life

One facet of the great program of space exploration is to ask whether life, in particular metazoan life, already exists elsewhere. The scope of this investigation will include the solar system but extend beyond, for we have already abundant evidence for at least dozends of extrasolar planets (Schneider 2010), including

some that might be Earth like. An example is the planet orbiting the red dwarf star *Gliese 581*. The gravity on the surface is estimated to be 1.25 times as strong as Earth if it is an icy and/or watery planet or 2.3 times if it has a large iron core (http://en.wikipedia.org/wiki/Gliese_581_c). It seems almost certain that many more Earth-like planets will be discovered. Thus, it is probable that the conditions that gave rise to microbial life on Earth have existed elsewhere. However, we have also seen (Chap. 4) that the progression from anaerobic microorganisms to metazoan life forms with oxidative metabolism was fraught with obstacles and pitfalls. There are many possibilities (runaway greenhouse, runaway glaciations, asteroid impact, or simple starvation, for example) that could have extinguished any primitive ecology. To proceed to a true metazoan level, life must take two major and difficult steps: the evolution of oxygenic photosynthesis and aerobic metabolism. Probably both are essential for a stable, higher level ecology. As we have seen in Chap. 2, the evolution of photosynthesis alone, if not followed by aerobic metabolism, could lead to catastrophic glaciations by drastically decreasing CO_2 and CH_4 levels.

If oxygenic photosynthesis has evolved and thrived on a planet, there should exist a significant amount of dioxygen in the atmosphere. More important for our observations, this is the only temperate situation we know in which such an atmosphere can form. The simple observation of the existence of significant levels of O_2 in the atmosphere of an Earth-like planet is compelling evidence for the presence of photosynthesizing life forms. This means that we have a powerful potential tool in the search for advanced life elsewhere. The lack of such evidence in the solar planets means, as we suggest, that photosynthesis does not occur in our planetary system, other than on Earth (see Sect. 9.2). This does not mean that other bodies in the solar system do not, or did not, harbor a more primitive, anaerobic life or did not at some earlier time have photosynthetic organisms. But if so, they are now gone, and the dioxygen they produced has been abiotically consumed.

To look for specific gases in the atmosphere of a planet orbiting a distant star will not be easy, although the technical ability appears to be rapidly approaching. Highly sensitive methods will be needed, using either UV or infrared spectroscopy. Dioxygen does absorb in the UV, but the bands overlap with water absorption. Dioxygen and nitrogen would not be detectable in the infrared but carbon dioxide, ozone, methane, and water vapor all absorb strongly within the infrared spectrum (Fig. 9.1). Infrared light will be emitted from the planetary surface, if it is illuminated by light within the visible spectrum. Within the emitted infrared spectrum, the above gases absorb light at specific wavelengths around 7 μm (water vapor), 8 μm (methane), 10 μm (ozone), and 15 μm (CO_2). This is why these gases form the basis of the greenhouse effect.

Already there exist examples of the detection of water vapor, carbon, and dioxygen in the atmosphere of "hot Jupiters" (Rojo et al. 2004; Tinetti et al. 2007). These data should not be taken as evidence for the existence of life, because these giant planets, orbiting close to their suns, are subject to intense radiation and represent extremely non-Earth environments. But these first results show feasibility. The IR bands listed above have been observed in Earth's atmosphere by orbital satellites (Fig. 9.1, and see Traub and Jucks 2002).
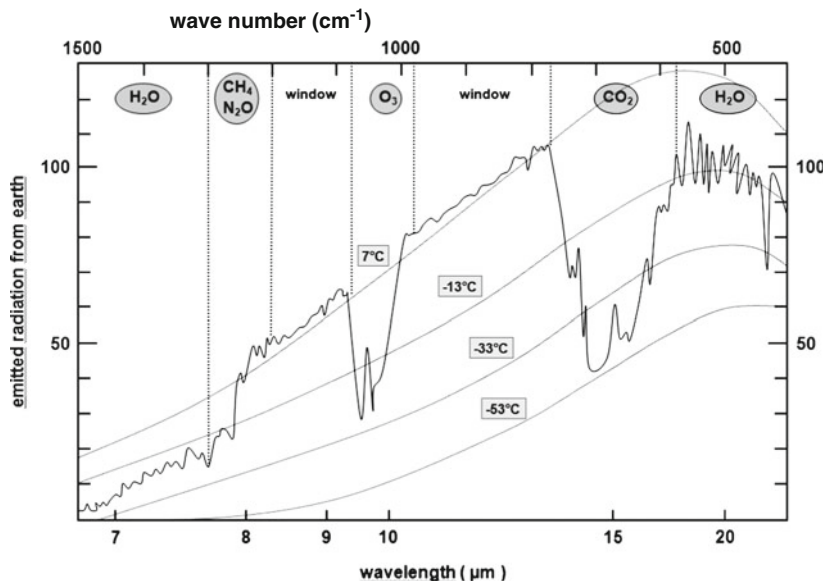
Fig. 9.1 Detection of O_3, CO_2, CH_4, and H_2O in the Earth's atmosphere. Observation of thermal radiation from space using an orbiting satellite with infrared detection. The deep "valleys" compose to absorption bands of the molecules indicated. The *smooth curves* correspond to black body radiation at the temperatures indicated

Because of the difficulties in observing O_2 itself, it would seem that the best way to look for dioxygen would be to look for ozone, instead. Ozone has strong absorption bands in the near UV, at about 250 nm as well as strong absorption in the infrared. The persistent presence of ozone in an atmosphere is a strong indication of dioxygen, for the only important reactions to produce ozone begin with dioxygen. Ozone has another consequence important to life: the formation of ozone in the atmosphere by UV radiation produces (Chap. 1) a UV shield permitting life in the ocean surface waters and on and above land surfaces. The detection of ozone in an extrasolar planetary atmosphere would be de facto evidence for the presence of a significant concentration of free dioxygen. Needless to say, the detection of ozone in the atmosphere of a small Earth-like planet orbiting a distant star will require technology beyond our present means. Space-based interferometric telescopes may be required (Bada 2001). Nevertheless, the way is clear and the technology is advancing very rapidly. We may expect such studies in the not too distant future. A positive result will be exciting, for we will then ask if life has proceeded to oxygenic photosynthesis, has it also approached sentience?

There exists a potential alternative method to seek extraterrestrial photosynthetic life. One might search for spectroscopic features typical of photosynthetic pigments in the reflectance spectrum of a planet. This approach is described in a number of laboratories (see Kiang et al. 2007). They consider not only the spectral characteristics we know to be associated with photosynthesis on Earth but also hypothetical

pigments that might have evolved on planets orbiting stars with very different spectra than the sun. This approach, although certainly more direct than searching for dioxygen via ozone, will also be technically more demanding, for detection will depend upon subtle features of the spectrum. It would seem most appropriate as a follow-up technique after initial ozone screening. The potential for detection of higher life by methods not far beyond our present capabilities raises a most interesting question: If we can detect, why cannot "others" detect us? The Earth must stand out as clearly as a beacon of life to any advanced civilization. Have we been seen?

9.4 Life in the Solar System?

9.4.1 Terrestrial Planets

Despite extensive observations, there has as yet been no detection of evidence for life, at least at the photosynthetic level, on any of the planets or moons of the solar system. Traces of dioxygen and ozone detected on the surface ice of Jupiter's moon Europa are most probably photochemical in origin. We cannot eliminate the possibility of anaerobic microorganisms below the surface of Mars, or in possible buried oceans of satellites of the gas giants (see below). Neither can we exclude the possibility that life may once have existed on the terrestrial planets, Venus and Mars. If it did, it is clearly extinct in the former case and has at best not far developed in the latter. It is of interest to examine these possibilities in more detail. To do so, we must consider the probable early atmospheric evolution of the terrestrial planets (Hunten 1993). Most current discussions make certain assumptions:

(a) The terrestrial planets are formed by accretion of rocky planitesimals and had roughly equal chemical compositions, including substantial quantities of water.
(b) Gas molecules are lost from a planet if their velocity exceeds that required for escape. This in turn depends on temperature, molecular mass of the gas particles, and the gravitational mass of the planet. Heavier gases can also be carried along in the escape of light gases like hydrogen (Hunten 1993). Small planets will rapidly lose lighter gases like H_2 and He.
(c) Content of hydrogen and light gases was originally much higher than today.
(d) The sun's intensity at the time of terrestrial planet formation (about 4.5 BYA) was only about 70% of the current value.
(e) The presence of the large planet Jupiter, farther from the sun, substantially protected Earth from bombardment from asteroids and comets.

These and other conditions must have been very similar for all four terrestrial planets shortly after their formation. Thus, one might expect that the very early planetary atmospheres of Mercury, Venus, Earth, and Mars would have been relatively similar. In contrast, we find the remarkably disparate current results as

shown in Table 2.2. What processes might have occurred to account for these very different atmospheric outcomes, and could life have been involved? As we have seen both water and essential elements for life are present all over the universe (Chap. 2). We shall consider only Venus, Earth, and Mars, for Mercury is too small and too close to the sun to have held an appreciable atmosphere for any significant time.

9.4.1.1 Venus

Venus appears to be the classic example of a runaway greenhouse. It has been argued that in its early stages, Venus might well have been cool enough (given the lower solar radiation) to possess a hot ocean. However, Venus has always received about twice the solar flux as Earth and the early Venusian atmosphere would have been quite humid. Therefore, a positive feedback would have come into play: water vapor is a powerful greenhouse gas and as temperature rises, the water vapor content of the atmosphere above an ocean will rise. It has been proposed that there is a crucial point at which the oceans will completely evaporate (see Fig. 9.2). Beyond this point, the temperature will continue to rise. Photodissociation of water vapor in the atmosphere will release hydrogen, which will escape, and dioxygen, which will combine with reduced metals and carbon. The latter will yield more carbon dioxide, to further augment the greenhouse effect and produce the exceedingly hot, CO_2-rich atmosphere we see in Venus today. It is hard to see how life, even if it had arisen, could influence such a massive, irreversible process, or even survive it.
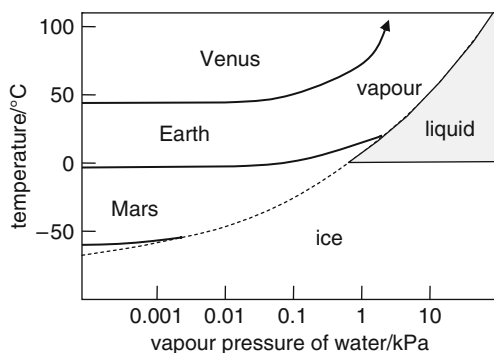
Fig. 9.2 A schematic diagram of the proposed course of evolution of the atmosphere of Venus, Earth, and Mars. The postulated initial temperatures of the three planets are about +50, 0, and −50°C, respectively. The conditions under which water vapor, liquid water, and ice are stable are indicated. At the lines separating these regions, two phases can be in equilibrium. As the diagram indicated, the Mars atmosphere came to equilibrium with ice at a low temperature, whereas the Earth's atmosphere came to equilibrium with water. On Venus, no such point could be reached, and all of the water became heated vapor, a "runaway greenhouse" (Houghton 2004)

9.4.1.2 Mars

Mars represents a very different and perplexing scenario. There now seems to be good evidence that liquid water existed on the surface at various periods billions of years ago. Given the greater distance of Mars from the sun than Earth or Venus, this must be explained by initial warmth from accretion and radioactivity, an early greenhouse atmosphere, or all three. It may simply be that earlier planetary cooling of a small planet rather far from the sun decreased the atmospheric water vapor, which in turn favored more cooling – a "runaway icebox" if you will. While we cannot rule out early life (and contested evidence from a meteorite argues for it), there is no obvious signature of life today on Mars. There is virtually no free dioxygen in the Martian atmosphere. Whatever was produced by water photolysis or by extinct life is probably tied up in the evidently oxidized crust. When we have a better analysis of that crust, it would be interesting to see if all of that oxygen can be accounted for by inorganic processes.

The evolution of Earth's atmosphere has been described in great detail in earlier chapters. In view of the pitfalls exemplified by Venus and Mars, the Earth seems to have steered a risky course between the Scylla of roasting and the Chyrabis of freezing. Life has provided negative feedback on CO_2-level fluctuations that may account for this.

9.4.2 Icy Moons

It has come to our realization, in recent years, that there exists a second class of bodies in the solar system that could conceivably harbor life. These are some of the moons of the gas-giant planets Jupiter, Saturn, and perhaps Neptune (see for overviews: Pappalardo 2007; Chyba and Phillips 2001, 2002). Many of these moons are ice covered, and recent evidence indicates that in some cases a water or water-ammonia sea may lie beneath a thick frozen surface.

The most carefully studied example is the Jovian satellite, *Europa* (see Fig. 9.3). The visible surface of Europa is a highly fractured ice layer whose constantly changing surface must reflect dynamic process below. Europa is especially interesting because its ocean may be quite shallow. This means that the liquid layer will be in contact with the rocky core of the moon, and not be underlain by dense forms of ice that exist under high pressure as may be the case with other icy moons (Pappalardo 2007). Thus, heat and gases from the core should vent directly into the ocean. The entire aqueous environment seems remarkably similar to that proposed for early anaerobic life forms on Earth (see Chap. 2). Whether life can *arise* in such an environment is another question, although many would say yes. It seems likely that once arisen, life could persist there, at least at the anaerobic level.

A question may arise as to the source of the heat to maintain liquid oceans on satellites so far from the sun. Two sources can be envisioned:
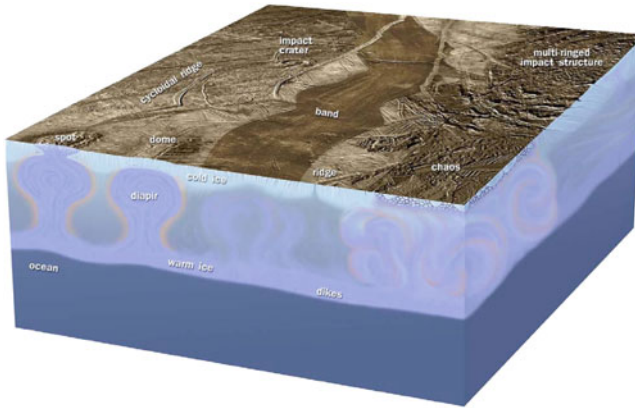
Fig. 9.3 Europa's bizarre surface features suggest a churning, icy interior. Warm ice "blobs" rise from the warmed under surface of the ice to fracture and continually rearrange the surface. The scarcity of impact craters indicates that the surface is continuously renewed (modified after Pappalardo 2007; from http://sse.jpl.nasa.gov/scitech/display.cfm?ST_ID=1790) and http://solarsystem.nasa.gov/multimedia/display.cfm?IM_ID=5543)

(a) Heat from radioactive decay in the rocky core (which must be steadily diminishing).
(b) Heat from gravitational distortion of the satellite in the immense gravitational field of its principal. The latter seems more likely to be important and constant.

Consideration of these icy moons, so far removed from sun, basically changes our thinking about life in the cosmos. The idea of "habitable zone" no longer seems so important, and the number of possible life-bearing worlds seems much greater.

9.5   Oxygen Supply Problems in Extraterrestrial Voyages

It seems likely that within the next few decades, extended human extraterrestrial voyages, to the moon or even Mars, may be attempted. Indeed, the European Space Agency is already planning for such. Essential for any human extraterrestrial voyaging is the provision of a reliable, adequate source of dioxygen. Each adult human consumes about 0.25 L of dioxygen per minute, or 360 L/day, which corresponds to 0.25 kg/day. For short voyages, or near-Earth orbiting satellites, where resupply is possible, the use of bottled dioxygen is adequate and simple. However, when longer voyages are contemplated such as a 500-day round trip expedition to Mars, with a crew of six astronauts, bottled dioxygen becomes prohibitably heavy. In this example, which is probably the minimal interplanetary

voyage, about 800 kg of O_2 must be carried. This does not include the mass of storage tanks, or the necessary equipment and chemicals for scrubbing CO_2 from the atmosphere. It is essential that CO_2 will be removed efficiently, for even fairly low concentrations are quite toxic to humans (see Table 7.1). Finally, one must consider the fact that such expeditions are essentially not rescuable. Were delays to prevent a return voyage at the optimal time, it might be necessary to wait months or even years for a return, while dioxygen reserves ran out. Thus, an extensive backup supply must be assured.

This compels us to think that some means of continually producing dioxygen (and consuming carbon dioxide) would be highly advantageous for long voyages. The potentially most interesting possibility is to make use of the same system that has provided, for ages, a more or less stable ecosystem on Earth: Combine plant photosynthesis with animal oxidative metabolism. Writing these two processes in the most elementary form, we have:

$$Photosynthesis: CO_2 + H_2O + light\ energy \rightarrow CH_2O + O_2$$
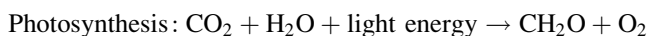
Here, CH_2O represents carbohydrate in a shorthand form.

$$Metabolism\ CH_2O + O_2 \rightarrow CO_2 + H_2O + metabolic\ energy$$

At least in principle, this forms a closed cycle. The beauty of the coupled system is that, at least ideally, there is no net mass requirement. This is not exactly true, of course, and there will be some mass needed for cultivation of plants or algae, and there will be losses. Water as urine and perspiration must be recovered and recycled as efficiently as possible. Some stored dioxygen would be advisable for emergencies. But in principle, such a system has no time limit. On the spaceship Earth it has functioned for ages, and we rely on it without question.

Nevertheless, to grow plants in a limited, closed environment can be tricky. If it were possible to produce a synthetic enzyme system that could carry out the basic reactions of oxygenic photosynthesis, a more reliable and finely controllable process could be imagined. But this is by no means simple: one needs not only the water-splitting function but also the CO_2-fixing function for the cycle to be complete. Both the light and dark reactions of photosynthesis are required, a formidable task, but perhaps within the competence of future molecular biologists. Efforts have already begun (see Sanderson 2008).

What has been said here concerning extraterrestrial voyages applies with even greater force when we consider even temporary extraterrestrial settlement. However, it is possible that in such cases, local material supplies can relax the stringency of the requirements. For an outpost on Mars, for example, local water might be in sufficient supply to allow the necessary O_2 production by electrolysis, given sufficient energy. Note that this is not useful during a voyage, for the extra mass of water that would have to be carried would exceed the mass of oxygen produced. Likewise, available mineral supplies might facilitate CO_2 scrubbing. In the long

run, however, the kind of human/plant "symbiosis" described above is probably the ideal.

9.6 Problems Facing Extended Extraterrestrial Settlement or Colonization

Looking still further into the future, we may imagine human efforts to form permanent settlements on other worlds. These might involve either solar planets (Mars) or conceivably, in the more distant future, extrasolar planets. These latter would be considered only if the atmosphere and other conditions closely approximated those of Earth. This may cause problems, for the prior existence of a dioxygen-rich atmosphere would imply that we would not have the planet to ourselves. Furthermore, the much extended, perhaps multigenerational travel required would necessitate a long-term spaceship "settlement." Likewise, establishing any colony on Mars, the moon, or an asteroid would require an enclosed environment with a constant supply of dioxygen, and means for CO_2 removal. The best long-term solution would be agriculture, which could also guarantee a food supply.

Although long-term extraterrestrial settlements in closed, environmentally controlled enclaves are potentially viable, they are intrinsically unsatisfactory. Even if elaborate backup systems are provided, they will always remain vulnerable to catastrophic failure. Furthermore, except for research stations, which are not meant to be permanent habitats, there is something repugnant to humans about living in a box. If we are really going to colonize another world, we should be free to live fully in it. That would certainly not be possible on any other planet we recognize today. Even extrasolar "Earth-like" planets would not likely have an atmospheric composition that would be usable by humans. The science fiction picture of humans stepping out of a spaceship, sniffing the air, and doffing their helmets is just that. Thus, permanent settlement of other worlds represents an extremely challenging problem. Two solutions, each far beyond current technology, come to mind.

9.6.1 Adjusting the Planetary Environment: Terraforming

Most extraterrestrial rocky planets that have not experienced photosynthesis may be expected to have atmospheres that are not breathable by humans, but are not excessively toxic. Gases like the halogens are chemically unlikely. If a planet has undergone a runaway greenhouse, like Venus, it would seem to be intrinsically uninhabitable. But the atmosphere of Mars, although currently unusable by humans, is not infinitely deficient (see Table 2.2). Basically, it is very low on dioxygen and somewhat high on carbon dioxide. Conceivably, the latter could be decreased and

the former increased by extensive growth of cold-resistant plants (probably genetically engineered) over a long period of time. The problem here is that the decrease in CO_2 would diminish what little greenhouse effect Mars has today, and make it even colder. If ways could be found to evaporate a portion of the water held in the polar caps (and perhaps elsewhere), a positive feedback could be established, for water vapor itself is a greenhouse gas. Basically, such a project would be an attempt to recapitulate, in a period of hundreds or thousands of years, what happened on Earth on a billion-year scale. A critical determinant will be whether or not the planet in question has already a well-oxidized surface. We certainly cannot wait for worldwide red beds to form. Fortunately, in the case of Mars, there is some evidence that surface soils are already oxidized. Such a project seems outlandishly difficult by present standards. We must, however, remember two things: First, current technology would have seemed absurd two centuries ago. Second, if plants were the instruments, the project would be largely autonomous.

9.6.2 Adjusting the Organism: Bioforming

An extreme alternative to adjusting the environment to fit humans would be to genetically alter humans to survive in an alien environment. To even consider this may seem absurd, or even repugnant to some, but in some situations it might be easier than terraforming. Human populations can already genetically adjust to considerable differences in environment: consider those adapted to live at high altitudes under low oxygen pressure. Further adjustments, either anatomic (lung capacity, musculature) or physiological (respiration rate, altered effector control of hemoglobin function) can be visualized. There is the huge ethical problem concerning making genetic changes in the human germ plasma; this would probably be more difficult to resolve than the technology.

The problems become even more acute if more radical modifications of a portion of the species than those suggested above are contemplated at some far future date. Will we be willing to create humans with the diving and swimming abilities of sea mammals to inhabit a watery world? Can we contemplate a "human" with skin cells containing chloroplasts, allowing energy gathering from sunlight? Although this may seem utterly fantastic, it has recently been shown that certain sea slugs that incorporate algae into their epithelium for photosynthesis have actually experienced horizontal gene transfer of photosynthesis genes into the animals' genome (Rumpho et al. 2008). But will the "green man" be human? In contrast to plants, however, this seems not likely since the surface of a human including the hair is too small for collecting enough sun energy to keep the body at the high temperature of 37°C.

Life has come a very long way in 4 billion years, on a path interlaced in multitudinous ways with the oxygenation of our planet. Our plans for the future seem a continuation of this same theme.

References

Bada JL (2001) State of the art instruments for detecting extraterrestrial life. Proc Natl Acad Sci USA 98:797–800

Chyba CF, Phillips CB (2001) Possible ecosystems and the search for life on Europa. Proc Natl Acad Sci USA 98:801–804

Chyba CF, Phillips CB (2002) Europa as an abode of life. Orig Life Evol Biosph 32:47–68

Houghton J (2004) Global warming, 3rd edn. Cambridge University Press, Cambridge

Hunten DM (1993) Atmospheric evolution of the terrestrial planets. Science 259:915–919

Kiang NY, Segura A, Tinetti G, Govindjee G, Blankenship RE, Siefert J, Cohen M, Crisp D, Meadows VS (2007) Spectral signatures of photosynthesis II: coevolution with other stars and the atmosphere on extrasolar worlds. Astrobiology 7:252–279

Pappalardo R (2007) Oceans in the outer solar systems. Planetary Report 27 No.4, pp 12–17

Rojo P, Harrington J, Zeehandelaar D, Dermody J, Drake D, Stegert D, Richardson LT, Wiedemann G (2004) Transit spectroscopy of the extrasolar planet HD 2094586: the search for water. In: Holt SS, Deming D (eds) The search for other Worlds: Fourteenth Astrophysics Conference, American Institute of Physics

Rumpho ME, Worful JM, Lee J, Kannan K, Tyler MS, Bhattacharya D, Moustafa A, Manhart JR (2008) Horizontal gene transfer of the algal nuclear gene psbO to the photosynthetic sea slug *Elysia chlorotica*. Proc Natl Acad Sci USA 105:17867–17871

Sanderson K (2008) The Photon trap. Nature 452:400–402

Schneider J (2010) Interactive extra-solar planets catalog. The Extrasolar Planets Encyclopedia http://exoplanet.eu/

Tinetti G, Vidal-Madjar A, Liang M-C, Beauöieu J-P, Yung Y, Carey S, Barber RJ, Tennyson J, Ribas I, Allard N, Ballester GE, Sing D, Salsip F (2007) Water vapor in the atmosphere of a transiting planet. Nature 448:169–171

Traub WA, Jucks KW (2002) A possible aeronomy of extrasolar terrestrial planets. In: Mendilla M, Nagy A, Watts J (eds) Atmospheres in the solar system. Am Geophys Union, Washington, DC, pp 370–380

Index