THE INORGANIC CHEMISTRY OF BIOLOGICAL PROCESSES

M. N. Hughes

THE
INORGANIC CHEMISTRY
OF
BIOLOGICAL PROCESSES

M. N. Hughes

Lecturer in Chemistry, Queen Elizabeth College, University of London

A Wiley–Interscience Publication

JOHN WILEY & SONS

London · New York · Sydney · Toronto

Library of Congress Catalog Card No. 72–5717

ISBN 0 471 42020 4

Printed in Great Britain by J. W. Arrowsmith Ltd., Winterstoke Road, Bristol BS3 2NT

PREFACE

The overlap region between inorganic chemistry and the biological sciences is one where exciting and significant developments are taking place. It is now appreciated that metal ions control a vast range of processes in biology; that life is really as dependent upon inorganic chemistry as organic chemistry. New developments in instrumental techniques have further accelerated the growth of 'inorganic biochemistry' so that it is now probably true to say that this subject involves one of the most rapidly expanding areas in the chemical and biochemical sciences.

This book is intended to present an introduction to this most important field. It has its origins in third year undergraduate courses at Queen Elizabeth College and is written primarily for chemists, particularly inorganic chemists. The material presented includes a survey of the occurrence and role of the metal ions of biological importance and shows how the function of these ions may be studied experimentally. While most topics of current interest are discussed, the coverage is not intended to be exhaustive. The book does not depend upon a prior knowledge of biological subjects, some relevant material is summarized in Chapter 1. It is also hoped that this book may be of interest to workers in the biological sciences, and so, primarily for this purpose, a brief survey of the relevant properties of transition metal complexes is presented in Chapter 2, together with an account of the mechanisms of their reactions in solution.

I am happy to acknowledge the assistance of a number of colleagues and friends; in particular Dr. K. J. Rutt for his helpful comments on the early chapters, Dr. C. W. Bird for his encouragement throughout all stages of writing this book and Miss Jane Cooper for her excellent typing of the manuscript. I am grateful to Professors J. Brachet, J. Coleman and S. Lindskog for permission to reproduce Figures 1.1, 4.3, 4.4 and 4.5 respectively and also to the appropriate Editors as specified in the text.

CONTENTS

CHAPTER ONE

INTRODUCTION

Metal ions play a vital role in a vast number of widely differing biological processes. Increasing knowledge will almost certainly serve to demonstrate this fact more effectively. Some of these processes are quite specific in their metal ion requirements in that only certain metal ions, in specified oxidation states, can fulfil the necessary catalytic or structural requirement, while other processes are much less specific, and it is possible to replace one metal ion by another, although the activity may be reduced.

Metal ion dependent processes are found throughout the Life Sciences and vary tremendously in their function and complexity. Three examples, from biochemistry, physiology and cytochemistry respectively, are given to illustrate this point. Thus the metal[1] ions potassium, magnesium, manganese, iron, cobalt, copper, molybdenum and zinc are all important catalysts of a variety of enzyme reactions such as, for example, group transfer, redox or hydrolytic processes. Not only, however, are these metal ions involved in such processes, but, in certain cases, there are other protein systems involved in storing and controlling the concentration of the metal ion, and then in transporting it to the appropriate site for incorporation into the necessary enzyme system. Sodium, potassium and calcium, on the other hand, are heavily involved in certain physiological control and trigger mechanisms, while potassium, calcium and magnesium ions are all important in maintaining the structure and controlling the function of cell walls. The metals cited in these examples are not the only ones involved in biological processes, others although quantitatively less important, also have biological functions. Of the cited metal ions, Na^+, K^+, Mg^{2+} and Ca^{2+} are present in much greater amounts than the heavy metal ions. Thus, in the human body these four cations constitute some 99% of the total metal ion content.

The physiological and biochemical function of the metal ion in all these processes is obviously a matter of fundamental importance, but the difficulties involved in attempting to clarify their role should not be minimized. Such a study presents many difficult problems, the solutions of which often require an overlap of disciplines. For the Inorganic chemist, without doubt, the field which holds out most hope of effective exploitation is that of the role of metal ions in enzyme and similar systems. It should be stressed that here there are often additional advantages associated with the

presence of the metal ion which contribute very markedly to an understanding of the system. This is particularly true of transition metal ions. Thus, as a result of the electronic properties of the metal, a variety of powerful instrumental techniques may be brought into play. Again the presence of the metal ion provides an extra guide in elucidating the mechanism of the enzyme action, if only in providing an extra check on its correctness in terms of correlating the specificity of the system for that ion in the light of modern awareness of the preferred environment, stereochemistry and electronic properties of the ion.

Recent developments in inorganic and organometallic chemistry have resulted in a very significantly increased understanding of the bonding, structure and reactivity of coordination compounds. This has been practically reflected in certain areas of inorganic chemistry, for example that of the activity of small molecules such as CO, H_2 and olefins on coordination to a metal, and this is leading to an increased understanding of certain important catalytic processes. Equally, however, the border-area between inorganic chemistry and the Life Sciences should present a challenge to the inorganic chemist to apply his increased understanding in the design of model systems that throw light on the behaviour of metal ions in biological processes, and ultimately to look more closely at these processes themselves. These developments in inorganic chemistry have been matched by developments in biochemistry in that it is now possible from the biochemical point of view to consider processes at the molecular level. Certain progress in the field of metal ion activated enzymes has been made, particularly in the case of metallo-enzymes where the metal is firmly bound to protein. Most metal cations in living organisms will in fact be associated with proteins and so the subject of metal-protein binding is a most fundamental one in the overall context of this book.

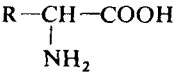
BACKGROUND MATERIAL

It is necessary at this stage to introduce some background material in order to explain the terms and concepts used in later chapters. Much of this material is necessarily presented at an elementary level.

Amino acids, peptides and proteins

The proteins are macromolecules of great biological importance. They are made up of α-amino acids of L configuration, linked together *via* peptide bonds —CONH—. By suitable treatment they can be degraded to smaller peptides and finally to the constituent amino acids. Some twenty of these amino acids are found in nature, together with the α-imino acids

TABLE 1.1 Naturally occurring amino acids



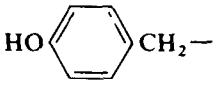
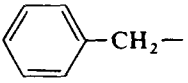
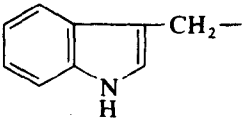
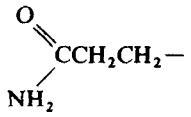
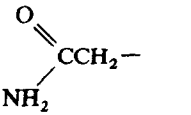
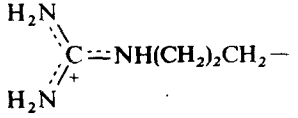
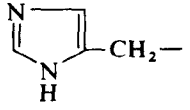
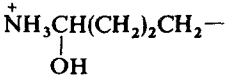
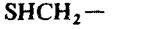
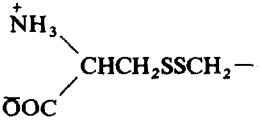
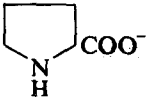
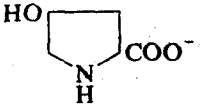
R	
H—	Glycine
CH_3—	Alanine
$(CH_3)_2CH$—	Valine
$(CH_3)_2CHCH_2$—	Leucine
$CH_3CH_2CH(CH_3)$—	Isoleucine
$HOCH_2$—	Serine
$CH_3CH(OH)$—	Threonine
	Tyrosine
	Phenylalanine
	Tryptophan
$^-OOCCH_2$—	Aspartic acid
$^-OOCCH_2CH_2$—	Glutamic acid
	Glutamine
	Asparagine
$^+NH_3(CH_2)_3CH_2$—	Lysine

TABLE 1.1 (cont.)

R	
	Arginine
	Histidine
	Hydroxylysine
	Cysteine
	Methionine
	Cystine
	Proline (imino acid)
	Hydroxyproline (imino acid)

proline and hydroxyproline. These are listed in Table 1.1. Each naturally occurring polypeptide or protein involves a specific sequence of amino acid residues, which may be determined by chemical and biochemical methods. The sequence of amino acid residues will determine the physical and chemical properties of the protein in terms of the chemical and physical interactions occurring between the side chains R in $NH_2CH(R)COOH$. The important side chains in this connection are those involving aromatic

groups, sulphur-containing groups, and $-NH_2$, $-OH$ and $-COOH$ groups. The nature of the residues may generate hydrophobic or hydrophilic environments in certain regions of the protein chain.

Each amino acid has at least two ionizable groups, the amino and carboxyl groups. The $-NH_3^+$ group is less acidic than the carboxyl group and so, between pH 4–9, the amino acid exists as a zwitterion, $H_3N^+CH(R)CO\bar{O}$. The α-carboxyl and amino functions are involved in the formation of the peptide link, and so each peptide has terminal $-NH_2$ and $-COOH$ groups, together with peptide links and side chains. These are all possible metal binding sites. Certain low molecular weight peptides are biologically important molecules.

Proteins

The molecular weights of proteins are in the range of 10^4–10^6 g. The determination of the chemical structure and the spatial configuration of proteins is a very important matter, as this is bound up with the biological function of the proteins. At the present time an increasing number of protein structures have been determined to a high degree of resolution by X-ray diffraction techniques. This is the only method for determining the complete structure.

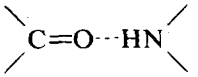
The structure of proteins is discussed in terms of primary, secondary, tertiary and quaternary structure.

Primary structure

This is the sequence of amino acid residues in the chain. For a particular enzyme some residues are less important than others and may, in fact, in enzymes from different species, be replaced by other residues.

Secondary structure

This is concerned with the configuration of the protein chain that results largely from hydrogen bonding between peptide links. Pauling and Corey



began their classic study of this problem by determining the structures of a range of simpler compounds, including amides. This has demonstrated that for a stable secondary structure the peptide link is always planar; that the carbon atoms each side of the peptide link are *trans* to each other, so lowering repulsive forces; and finally that there is a maximum amount of hydrogen bonding between carbonyl oxygen and amide nitrogen atoms.

This hydrogen bonding may either be intramolecular or intermolecular. In the first case, it gives rise to an α-helical structure and in the latter case to a pleated sheet structure. X-ray studies have confirmed the α-helix structure and have shown that there are 3·7 residues per turn of the helix. Helical structures occur in globular and fibrous protein. Pleated sheet structures may involve peptide chains having all N-termini at one end or with every other N-terminus at one end. These are the parallel and anti-parallel forms respectively.

Secondary structure is also dependent upon the nature of the side chains in that interactions between these may, for example, lower the stability of the α-helix through repulsion.

Tertiary structure

This is concerned with the way in which the protein chain (with its secondary structure) folds upon itself, and the resulting shape of the molecule. Thus it may be globular or rod-like. This results from the interactions between side chains, and may reflect the formation of S—S bonds between cysteine residues, hydrogen bonding between side chains, so called 'salt' or ionic linkages between oppositely charged groups such as $-\overset{+}{N}H_3$ and $-COO^-$ and hydrophobic interactions between aromatic residues. It appears that the last named of these is most important.

The overall shape of the protein molecule may be studied in a number of ways. The most powerful technique is of course that of X-ray diffraction, but other techniques include the measurement of the intrinsic viscosity and the light scattering properties of the macromolecule. X-ray diffraction will also confirm or determine the primary structure.

A schematic representation of the structure of myoglobin is given in Figure 1.1. This globular protein is made up of eight helical segments

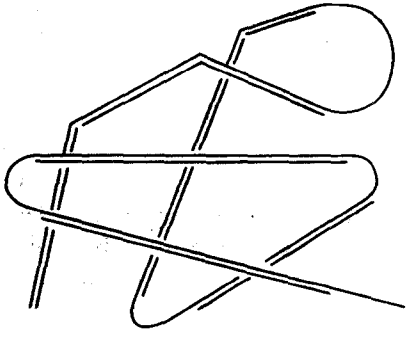


Figure 1.1 Schematic representation of the structure of myoglobin (the double lines indicate α-helical segments).

separated by regions of random coil. It has a high α-helix content of 77 per cent, and has no S–S bridges. It is rather atypical in this.

Quaternary structure
Many proteins are made up of linked subunits held together by other than covalent bonds. The quaternary structure reflects the way in which these subunits come together. Hemoglobin is made up of four such units, and substantial changes in interaction between these occur when oxygen is taken up or lost.

Proteins in solution, denaturation and other topics

One important question is the extent to which protein structures, determined in the solid state, change in solution. Only if there is good reason for assuming that little change has occurred can we extrapolate the results of X-ray structure determinations to the problems of enzyme mechanisms in solution. It appears, however, that solution and crystal structures are similar. Thus, the alkylation of side chains in myoglobin proceeds readily for those residues which the X-ray determinations show to be exposed and does not proceed for those deep inside the molecule.

It is clear from the preceding paragraphs that the secondary and tertiary structure of the protein is dependent upon a variety of interactions between various groups in the protein molecule. On heating or change of pH some of these interactions may be affected. The addition of solvents may break down some hydrogen bonding, while treatment with reducing agents may break S–S linkages. Clearly, therefore, proteins are readily subject to changes which convert an ordered structure to a disordered one. This is termed denaturation. Sometimes these changes are reversible and sometimes they are irreversible. Denaturation results in a number of changes in the chemical and physical properties of a protein. The sensitivity of proteins to denaturation is a major difficulty in experimental work.

Much work has been carried out on synthetic polypeptides such as poly-L-alanine and this has provided a basis for the study of naturally occurring proteins; for example, the value of certain physical techniques in measuring the α-helical content of proteins.

The determination of the helical content of proteins is an important measurement. Optical rotatory dispersion (O.R.D.) studies have been widely used, as the helix itself contributes to this in addition to the constituent amino acids. Other approaches have involved the use of infrared dichroism studies using plane polarized infrared light and hydrogen-deuterium exchange. Standard texts should be consulted for a full account of these techniques.

The use of models and the study of synthetic polypeptides with only one type of amino acid residue have also contributed to an understanding of the secondary structure of proteins. Thus it has been demonstrated that a mixture of D and L amino acid residues is not compatible with the α-helix. Similarly the requirements of the α-helix structure cannot be met by proline residues and so these must terminate the α-helix configuration. van der Waals' interactions between substituents such as in valine and isoleucine will also lower the stability of the helix.

One important property of proteins is the fact that they have a number of ionizable groups that may be involved in acid–base equilibria. These are the side chains of aspartic and glutamic acids, arginine, cysteine, histidine, lysine and tyrosine. As a result of the large number of titratable groups present in each protein the interpretation of protein titration curves is complex. In addition, the environment of the protein (Hydrogen bonding, medium effects) may well affect the pK_a of the amino acid residue so that it is greater or less than the value for the free acid by up to one pK unit. Certain residues may also be buried within the interior of the protein molecule and so not be accessible during the titration. By carrying out a back titration it is possible to discover if denaturation has occurred, resulting in inaccessible groups now being available. However, in favoured cases the analysis of titration curves has allowed the estimation of the number and type of the protonation sites available.

Enzymes

These are proteins which by reason of their particular three-dimensional structure are able to act as highly specific biochemical catalysts. The catalytic effect is considerable. Thus sometimes rate constants for enzyme catalysed reactions and model reactions differ by as much as a factor of 10^{12}. A number of explanations have been suggested for this, and we shall consider some later. Because the enzymes are proteins the general problems of protein stability hold here, and therefore pH and temperature must be carefully controlled.

In general the cell requires a different enzyme for each of its reactions, although a limited number will catalyse reactions of a general type. The esterases will thus catalyse ester hydrolysis. The need for such high specificity can readily be seen. Life processes usually consist of a series of interrelated complex reactions. Often the product of one reaction becomes the starting material for another. The reactions must be very specific in order to avoid complications from other simultaneously occurring systems.

The need for extra factors in an enzyme reaction can often be shown by the process of dialysis in which the enzyme solution, in a cellophane

container, is suspended in distilled water. The cellophane pores allow the low molecular weight components to diffuse through into the surrounding water, leaving the protein molecules in the cellophane container. If the enzyme depended upon any of the low molecular weight substances it will fail to function until they are added back.

The enzyme activators can be metal ions or complex organic molecules such as nucleotides or certain B vitamins. These are termed co-enzymes and are bound to the enzyme protein, being removed on prolonged dialysis. Sometimes a co-enzyme is bound so firmly that it is not removed by dialysis, in which case it is termed a prosthetic group.

The mechanism of enzyme action

An old established analogy is that of the lock and key. This represents the complex three-dimensional relationship between an enzyme and the substrate on which it acts; a relationship which may require the incorporation of certain cofactors or activators before it is complete. The result of this is to activate the substrate so that it is able to react in the required way. This analogy is still useful in that X-ray studies on enzymes have shown, in all cases, the presence of a cleft at the active site into which the substrate must fit. The lock and key analogy is, however, inadequate in that it does not indicate the other effects which occur; for example, the substantial conformational changes that result when the substrate is bound to carboxypeptidase. The active site itself, that portion of the protein chain involved in the interaction with substrate, will only involve a few residues, although of course they may be from well separated parts of the protein chain as a result of protein folding. It is the configuration of the protein around this portion that provides the cleft into which the substrate must fit to become activated. This explains the specificity of the enzyme. Why, for example, an enzyme may only be able to attack certain optical isomers. Thus malate dehydrogenase oxidizes L-malate exclusively to oxaloacetate in the presence of co-enzyme 1 (NAD), while D-malate is unaffected, thus providing an effective method for obtaining the D-isomer from the DL form. The role of inhibitors can also be understood in a general sense, in terms of their modifying the overall molecular shape at the active site so that the substrate may not fit into it, or by their competing with the substrate for the active centre of the enzyme.

However, the portion of the enzyme known as the active site may not only act by generating a certain steric specificity. The nature of the side chains is also important in terms of their hydrophobic or hydrophilic properties. It appears that the active site is surrounded always by non-polar residues, thus providing an environment at the active site which is of

lower dielectric constant than that of the aqueous solution in which the enzyme is found. This means that a number of interactions are different from what might have been expected; pK_a values are affected, ionic interactions will be stronger. All of these will be necessary for the correct functioning of the enzyme. Again, the active site provides the directional hydrogen bonding and van der Waals interactions to bind the substrate to the enzyme. In a number of cases X-ray studies have suggested that the initial interaction between substrate and enzyme induces further interactions that cause other parts of the enzyme to close in upon the substrate.

The actual source of catalytic power has been ascribed to a number of causes such as proximity and orientation effects and electron push-pull effects. It is, however, difficult to put these on a quantitative basis. One recent suggestion which has aroused some controversy is that of orbital steering, i.e. an orientation effect involving the orientation of orbitals in the reacting atoms of enzyme and substrate.

The cell

The complexity of the cell is illustrated by the fact that not only are there many celled organisms, but there are also one-celled species, such as bacteria, viruses and moulds whose behaviour must be entirely accounted for by the activities occurring within that single cell. The following comments apply, in general terms, to animal cells and unicellular plants (green plants will also contain chloroplasts, associated with photosynthesis) but not to bacterial cells which differ in many respects, e.g. they do not possess a nucleus and therefore belong to the prokaryotic class of cells.

Cells, in general, are encased by a membrane, the function of which is extremely important and which is dependent on metal ions. This membrane is selectively permeable to different metal ions (and other species) and this is readily associated with the function and distribution of, for example, the *s* block metal cations. Thus Mg^{2+} and K^+ are concentrated in the cell by the action of the membrane, while Na^+ and Ca^{2+} are rejected by it. This is associated with the utilization of Ca^{2+} as a structural factor in teeth, bones and shells, and as an activator of extracellular enzymes, while Mg^{2+} and K^+ are associated with intracellular processes, both as structural stabilizers and also as enzyme activators. The function of the membrane will be discussed at length in a later chapter. Plant cells are surrounded by an additional wall of cellulose giving extra rigidity.

The complex internal structure of the cell has been demonstrated by electron microscopy. Figure 1.2 shows a typical cell. In the cytoplasm of the cell are a number of structures, whose existence must be correlated with the various degradative and synthetic pathways involved in the

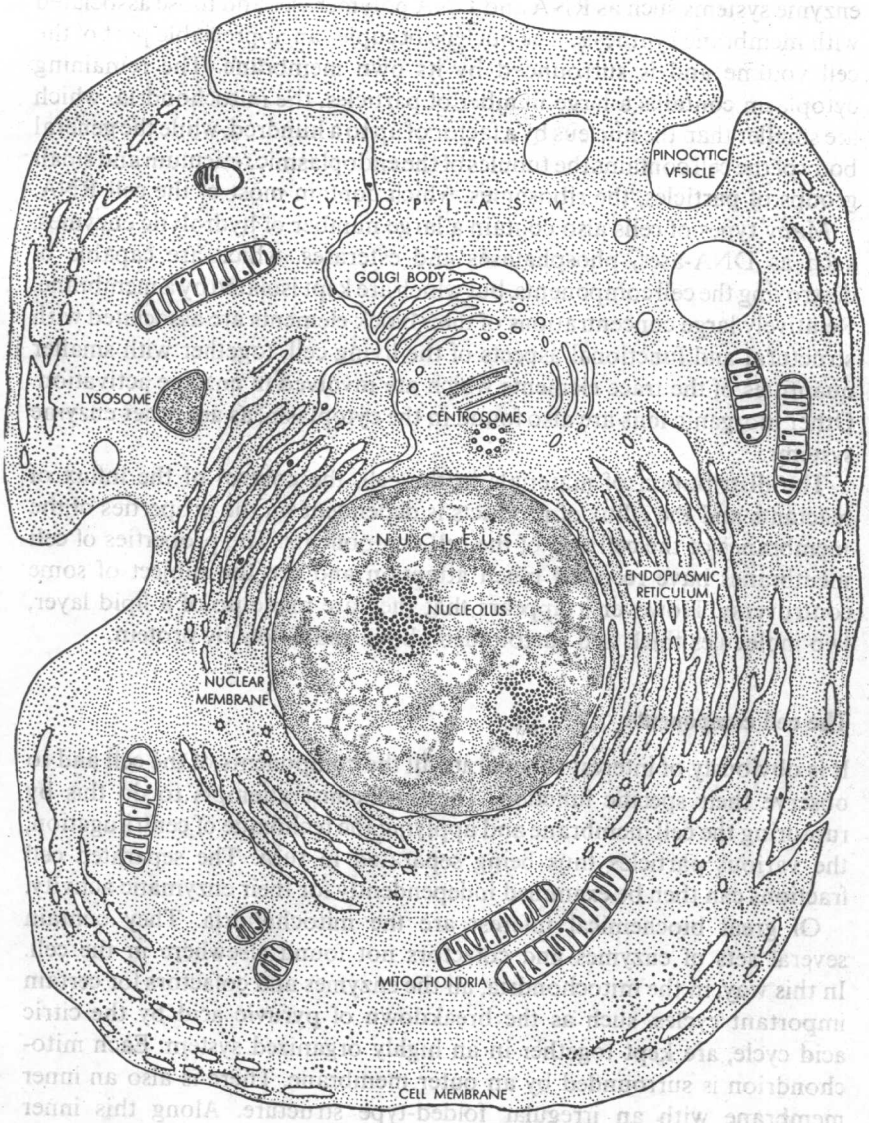


Figure 1.2 A typical cell (from *The Living Cell*, Brachet. Copyright © 1961 by Scientific American, Inc. All rights reserved).

working of the cell. All cells have a nucleus which contains the cell DNA, a little RNA and certain associated proteins, together with relevant enzyme systems such as RNA and DNA polymerases and those associated with membrane synthesis. The nucleus occupies an appreciable part of the cell volume and is surrounded by its own membrane. The remaining cytoplasm contains a large number of particles, the mitochondria, which are smaller than the nucleus by a factor of up to a hundred, while the general body of the cell contains the tubules of the endoplasmic reticulum and small groups of particles, the ribosomes, which contain most of the cell RNA.

Many types of cells also contain a large number of hydrolytic enzymes, such as DNA-ases, phosphatases and esterases, which are capable of destroying the cell components, but which are kept isolated by a membrane. Relatively large concentrations of the *s* block elements are associated with structural and functional aspects of the living cell, together with smaller quantities of the trace elements which are involved in enzyme activation. Other inorganic ions are also found in the cytoplasm, as are other enzyme systems.

The distribution of materials in the cell, the control of the influx of reactants and the efflux of products is dependent on the properties of the membranes of the cell and of the cell components. The properties of cell membranes have received much attention and are the subject of some controversy. However it appears that they are made up of a lipid layer, two molecules thick and surrounded above and below by protein.

The cell components

It is obviously of interest to separate all the components of the cell and to observe their specific function. The biochemist attempts to do this by rupturing the cell membrane and applying the technique of centrifugation, the various particles thus being separated in turn. The separated cell fractions can then be examined independently for their enzymatic activity.

Of great biochemical interest are the mitochondria. They contain several sets of enzymes and cofactors not found elsewhere in the cell. In this way, in the mitochondria, all the enzymes and cofactors for certain important cycles, such as the breakdown of pyruvic acid by the citric acid cycle, are kept together in an highly organized system. Each mitochondrion is surrounded by an outer membrane. There is also an inner membrane with an irregular folded-type structure. Along this inner membrane is distributed a complex system involved in oxidation and phosphorylation. The complexity and interlinking of various biochemical processes certainly implies, in this case, a highly organized system of enzymes, co-enzymes and electron carriers. This is provided by attaching