Thesis for the degree
of Candidatus Scientiarum
Espen Harbitz

Spectroscopic and Sequence Studies of Cytochrome c-554 from *Methylosinus trichosporium* OB3b

DEPARTMENT OF BIOCHEMISTRY
FACULTY OF MATHEMATICS
AND NATURAL SCIENCES
UNIVERSITY OF OSLO 2003

Espen Harbitz

Spectroscopic and Sequence Studies of Cytochrome c-554 from *Methylosinus trichosporium* OB3b

DEPARTMENT OF BIOCHEMISTRY
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
UNIVERSITY OF OSLO MMIII

Acknowledgements

The work presented in this thesis was carried out under the supervision of Professor K. Kristoffer Anderssons at the Department of Biochemistry, University of Oslo.

Professor K. Kristoffer Andersson has provided valuable supervision and given me the opportunity to develop my interest for the research field of inorganic biochemistry. He has also provided financial support, making it possible for me to visit inspiring conferences.

I will thank Dr. Bettina Katterle for supervision during my first half year, and getting me started with the project. I will also thank Giorgio Zoppellaro who worked in our laboratory in 2001 and performed the EPR simulations. Discussing the project with him gave me valuable insight in the EPR properties of our systems. Solveig Karlsen made a valuable contribution by preparing the structures in Figure 5.2. Hamid Shergarfi and Innocent Chukwu helped growing the bacteria.

Dr. Matthias Kollberg has read through my thesis. I am grateful for the corrections they gave and the discussions regarding the scientific presentation of my work.

Finally I will thank my research group, all my fellow students and the employees at the Department of Biochemistry for making a great social environment.

Oslo, 04.03.2003

Espen Harbitz

1 Abstract 1
2 Introduction 3
2.1 Metals in Biology 3
2.2 Iron in Biology 4
2.3 The Haem Group 5
2.4 Haemproteins 7
2.5 Cytochromes 7
2.6 Cytochrome c 8
2.7 The Highly Axial Low-Spin EPR Signal 8
2.8 Aim of the Thesis 9
3 Methods 11
3.1 Growth of Methylosinus trichosporium OB3b 11
3.1.1 Culture Maintenance 11
3.1.2 Large Scale Growth 11
3.1.3 Cell Harvesting 12
3.2 Purification of Cytochrome c-554 from *Methylosinus trichosporium* OB3b 13
3.2.1 Cell Disruption and Preparation of Cell Free Extract 13
3.2.2 Ammonium Sulphate Precipitation 13
3.2.3 Reversed Phase Chromatography 14
3.2.4 Cation-Exchange Chromatography 14
3.2.5 Gel Filtration 15
3.2.6 SDS Polyacrylamide Gel Electrophoresis 15
3.3 Reduction / Oxidation of the Haem Centre 17
3.4 Haem Cleavage 18
3.5 Protein Sequencing 20
3.6 Protein Cleavage 21
3.7 Sequence Alignment using pBLAST 22
3.8 MALDI-TOF Mass Spectroscopy 23
3.9 Optical Spectroscopy 24
3.9.1 Pyridine Haemochrome 25
3.10 Circular Dichroism (CD) 26
3.11 Electron Paramagnetic Resonance (EPR) 27
3.11.1 Introduction to EPR Theory 27
3.11.1.1 The Zeeman Effect 27
3.11.1.2 The g-Value 28
3.11.1.3 The Boltzmann Distribution, Relaxation and Saturation 31
3.11.2 EPR states observed in haems 32
3.11.3 EPR sample preparation 32
3.11.4 EPR instrumentation 33
4 Results 35
4.1 Cytochrome c-554 from *Methylosinus trichosporium* OB3b 35
4.2 Protein Purification 36
4.3 Optical spectroscopy 37
4.3.1 The Alkaline Transition 39
4.4 Circular Dichroism (CD) 40
4.5 Haem Cleavage 43
4.6 Mass Spectroscopy 44
4.7 EPR studies 46

4.7.1 The Highly Axial Low-Spin (HALS) EPR Signal....46
4.7.2 pH Dependence of the HALS EPR Signal....48
4.7.3 Horse Heart Cytochrome c and the Alkaline Transition....48
4.7.3 Horse Heart Cytochrome c and the Alkaline Transition....49
4.7.4 Microwave Power Saturation....50
4.7.5 Temperature dependence....52
4.8 Amino acid sequencing....54
4.8.1 Sequencing of Cytochrome c from *Methylococcus capsulatus* Bath...55
4.9 Sequence Alignment....56
4.9.1 Protein Sequence Database....56
4.9.2 Structural Database....57
4.10 Mass determination of a cyanogen bromide fragment of cytochrome c-554....58
4.11 Protein Modelling....60

5 Discussion....61

5.1 Optical Properties of Cytochrome c-554 from *Methylosinus trichosporium* OB3b....62
5.2 Molecular Mass and Haem Content....62
5.3 The Highly Axial Low-Spin EPR Signal....63
5.4 Amino Acid Sequence Studies....65
5.5 Structural Modelling....66
5.6 Conclusion....68
5.7 Further Studies....68

Appendix....69

Growth Medium for Methylosinus trichosporium OB3b:....69
Materials....70
Terms and Abbreviations....73

Reference List....75

1 Abstract

Methane oxidation to methanol at ambient temperature and pressure is a difficult chemical problem. Methylotrophic bacteria, like *Methylosinus trichosporium* OB3b, can carry out this reaction using two electrons from an electron donor and molecular oxygen.

Cytochrome c-554 from *Methylosinus trichosporium* OB3b is a redox protein with unknown biological function, that might have an important electron-transfer role.

Cytochrome c-554 from *Methylosinus trichosporium* OB3b exhibits a ferric HALS (highly axial low-spin) signal when subjected to low-temperature electron paramagnetic resonance (EPR) spectroscopy. Spectropotentiometric measurements indicate haem heterogeneity.

The EPR spectrum has a typical HALS lineshape with a g_{max} value of 3.41 at pH 7.0 and pH 8.2. The EPR spectrum changes slightly with pH. The EPR spectrum probably consists of two different low-spin species. The presence of two distinct ferric low-spin species can explain the observed haem heterogeneity.

The molecular weight of cytochrome c-554 from *Methylosinus trichosporium* OB3b has been determined to 12 230 Da by mass spectrometry. The cytochrome is found to contain only one haem group.

Optical spectroscopy has determined one axial haem ligand to be methionine.

The sequence of the thirty-nine N-terminal amino acids has been determined. This fragment reveals the characteristic haem-c binding motif -CXXCH-, where the cysteines covalently attach the haem group to the protein and the histidine provides one of the axial haem ligands. The amino acid sequence of another twenty-eight residue fragment is also determined. 62% of the amino acid sequence of cytochrome c-554 has been determined.

The partial amino acid sequence reveals that this cytochrome is homologous to cytochrome c's from a diverse range of organisms. An identity of more than 50% is found with two *Rhodopseudomonas* species. Sequence alignments indicate that the smaller amino acid fragment is located at the C-terminal end of the protein.

A putative 3D structure model has been made based on the amino acid sequence and structures of homologous proteins.

Cytochrome c-554 was compared with other cytochrome c's that exhibit HALS EPR signals and have histidine and methionine ligation of the haem iron. No correlation between the geometry of the histidine and methionine ligands and the HALS EPR signal has been established.

2 Introduction

2.1 Metals in Biology

The importance of trace elements in biology has been known long before the advent of inorganic biochemistry as a distinct discipline in the late 1960s. Inorganic biochemistry is at the meeting point between different branches of science, and requires input from physics, chemistry, and molecular biology. The topic's broad scope of dealing with inorganic compounds causes it to impact medicine, pharmacology, agriculture and environmental sciences (Wilkins and Wilkins, 1997). The importance of this field can be recognised by the fact that one third of all known enzymes require the presence of metal atoms for full catalytic activity.

Of the 90 elements that have been present throughout evolution (every element up to 93 except Tc, Po and At), about 30 are recognized as essential to some form of life (Emsley, 2001). Figure 2.1 displays the biologically important elements. Of the metal ions, calcium, magnesium, iron, copper, zinc, manganese, cobalt, and chromium are found in most organisms. Nickel, vanadium, molybdenum, and tungsten are found in more specialized bacterial systems (Wilkins and Wilkins, 1997).

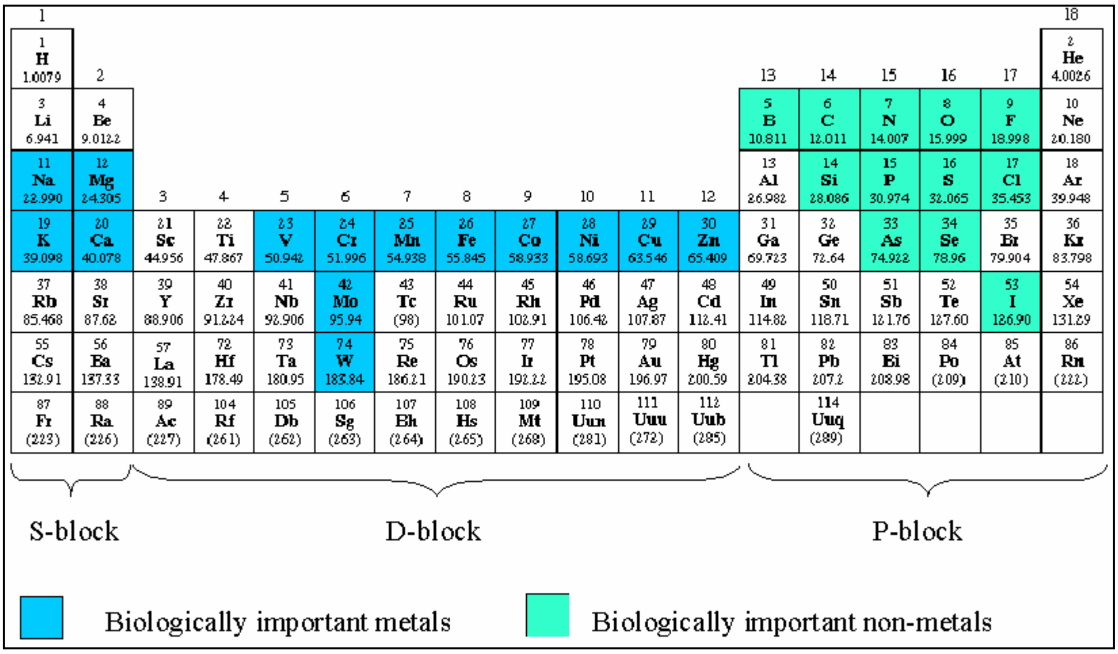

S-block
D-block
P-block
Biologically important metals
Biologically important non-metals

Figure 2.1 The periodic table and the biologically important elements

The S-block elements (Na, K, Mg and Ca) are the most abundant metal ions in biological systems. They participate in a wide range of processes from the triggering of biochemical reactions by activating enzymes, to electrostatic stabilization of biomolecules. Calcium also has an important role in the skeleton. The S-block metals in biological systems occur mostly as ions in solution. One important exception is chlorophyll, the primary photoreceptor in photosynthesis (Voet and Voet, 1995), where magnesium is covalently bound in a cyclic tetrapyrrole. This structure resembles the redox cofactor haem in which iron is the central metal ion. The principal difference between these two cofactors arises from the fact that iron is redox active, whereas magnesium only exists as Mg^{2+} in biological systems. Figure 2.2 shows the molecular structure of Chlorophyll *a*.

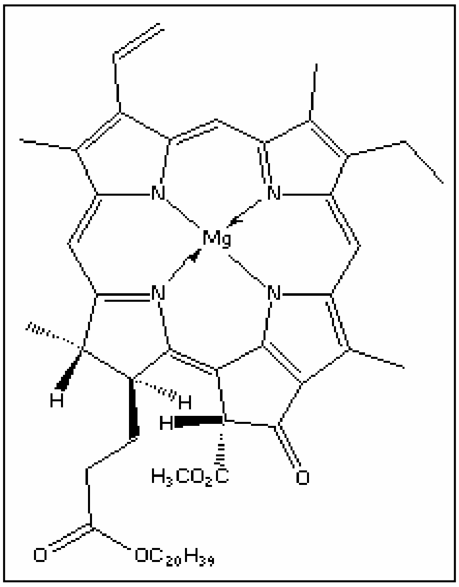


Figure 2.2 The molecular structure of Chlorophyll *a*

The P-block elements are the main constituents of all living organisms, but most of these elements are classified as non-metals. Some of the metals in this group have been suggested to have a biological function but no specific role has been found. Rats fed on a tin-free diet fails to grow normally but recover when given a tin supplement. No evidence is found for humans lacking or needing tin (Emsley, 2001).

The D-block elements are the major concern of inorganic biochemists. The metals in the first row of transition elements occur 1000 times more concentrated in blood plasma than in seawater (Wilkins and Wilkins, 1997), and are of vital importance to all organisms. They are essential in a great diversity of functions such as structural stabilization, electron transfer, enzymatic actions, transport and storage of O_2, and control of gene activity. Furthermore, processes important for the existence of life, and undeniably the formation of life, like photosynthesis, nitrogen fixation and the respiratory chain are all utterly dependent on transition metals.

2.2 Iron in Biology

Iron is the most predominant of all trace metals, and has indeed been known to be essential since the seventeenth century (Wilkins and Wilkins, 1997). The reason for this early discovery is the shear amount of iron needed in the organism. A grown up person normally contains roughly 4 grams of iron (Emsley, 2001), and iron is essential in substantial amounts for all cellular organisms from bacteria to mammals (Smith, 1997). Iron plays a key role in biological electron transfer, enzymatic actions, and transport and storage of O_2. Although vitally necessary, iron is also the major cause of oxidative stress. Presence of Fe^{2+} ions together with endogenously produced hydrogen peroxide triggers the formation of reactive oxygen species as superoxide anions and hydroxyl radicals (Wardman and Candeias, 1996). Complex mechanisms exist for transporting and storing iron to minimize the development of these oxygen radicals. Furthermore, enzymes such as superoxide dismutase and catalase have evolved to remove these reactive oxygen species from the organism. Superoxide dismutase disproportionates superoxide to O_2 and H_2O_2, and the haem containing enzyme catalase completes the oxygen radical detoxification by disproportionating H_2O_2 to H_2O and O_2 (Voet and Voet, 1995).

The usefulness of iron is mainly due to its ability to exist in several oxidation states. Fe^{2+}, Fe^{3+}, and Fe^{4+} are commonly occurring states n biology. In a reaction catalysed by an iron-containing enzyme, the iron often undergoes changes in oxidation state during the enzymatic cycle.

2.3 The Haem Group

Haem is the name of any prosthetic group consisting of an iron-porphyrin coordination complex. The porphyrin-ring, shown in Figure 2.3, is a tetrapyrrole commonly found in prosthetic groups because of its ability to chelate metal ions. The structure is aromatic and it absorbs electromagnetic radiation in the visible light range. This causes the porphyrin-based prosthetic groups to have distinct colours and absorption spectra. Haem, in which an iron ion is coordinated to the four nitrogen atoms in the porphyrin-ring, have distinct optical features which varies with the oxidation state of the iron, the spin state, and the binding of small iron ligands (e.g. O_2 or CO). Absorption bands around 550 nm are the origin of the strong red colour of blood and muscles. This is caused by the haemproteins haemoglobin and myoglobin that are abundant in red blood cells and muscle cells, respectively. There are several different haem groups found in proteins (Degtyarenko *et al.*, 1999). Haem *a,* haem *b,* haem *c* and haem *d*, are the most widely found haem groups in nature. The haem groups differ in the side chains on the porphyrin skeleton (Figure 2.4). This influences the spectroscopic properties and the way the haem group is attached to the protein. Haem *b* is by far the most prevalent haem group found in the animal kingdom, and hence this group is often associated by the word haem. Haem *b* contains two vinyl side chains, which affect the spectroscopic behaviour. In other haem groups, namely haem *c* and haem P460, cysteine residues react with these two vinyl groups to form thioether linkage between the haem group and the protein backbone. Strictly speaking, these are not prosthetic groups as they are covalently attached to the proteins in which they are found. In haem *a* and haem *o* a farnesyl group is bound to one of the vinyl group thus removing the aromatic character of the vinyl group, and adding an aliphatic side chain that promotes binding to the hydrophobic protein core. Haem *a* also contains a formyl side chain which add to the aromatic character of that haem group. Another important factor for the spectroscopic behaviour is the degree of aromaticity in the porphyrin ring. Most haem molecules are fully aromatic. Exceptions are haem *d*, haem d_1 and sirohaem, which have one or two non-aromatic pyrrole rings.

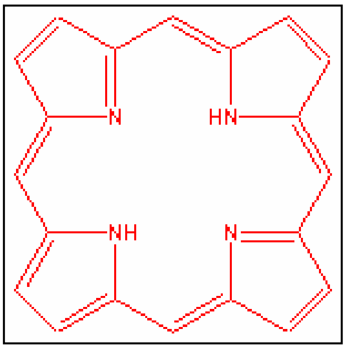
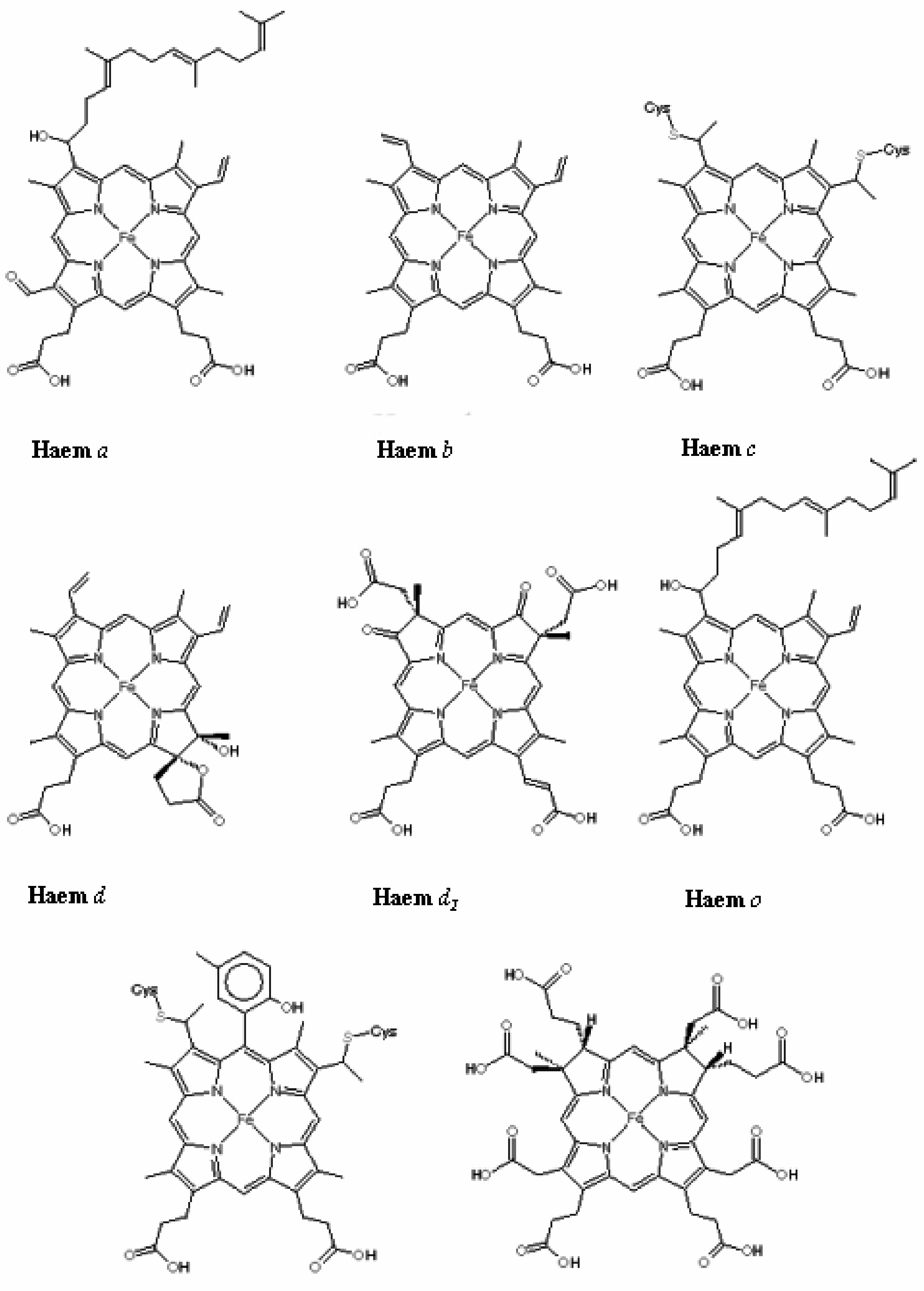


Figure 2.3 The porphyrin skeleton


HO
OH
O
N
Fe
S
Cys
H

Haem *a*

Haem *b*

Haem *c*

Haem *d*

Haem d_1

Haem *o*

Haem P460

Sirohaem

Figure 2.4 Some of the haem groups found in proteins

2.4 Haemproteins

Haemproteins is probably the most extensively studied group of proteins. Due to the amount in which they occur, and the straightforward detection due to their colour, haemproteins were among the first purified and studied. This has also led to the use of simple haemproteins as model systems both for general protein studies, the study of larger haem proteins and the development of new biochemical and biophysical methods.

Haemproteins perform a great variety of tasks in biological systems. The redox active iron in the haem is suitable for catalysing redox reactions and for electron transfer. The main functions performed by haemproteins are those of catalysis, electron transfer, O_2 transport and storage and NO transport.

Oxygen transport and storage is performed by the globin superfamily. A pocket on one side of the haem group gives room for a dioxygen molecule, which binds strongly to the haem iron. This protein family includes vertebrate haemoglobins (Hb), vertebrate myoglobins (Mb), invertebrate globins, plant globins, and bacterial and fungal flavohaemoproteins (Degtyarenko *et al.*, 1999). Another haemprotein that participates in small molecule storage is nitrophorin, which stores NO in the saliva of blood-sucking insects. The insect uses nitrophorins to deliver NO to host tissues during feeding where NO, released at neutral pH, induces vasodilation and inhibition of platelet aggregation (Weichsel *et al.*, 2000). There are also examples where NO binding to a haem group induces conformational changes in the protein leading to an increased catalytic activity, e.g. guanylate cyclase (Voet and Voet, 1995).

The variety of reactions catalysed by haemproteins is vast, but falls mainly into the categories of oxidases, reductases, peroxidases and catalases. These are important in reduction of O_2 in aerobic organisms, detoxification processes of organic substrates, removal of H_2O_2 and other reactive oxygen species, and generation of the signalling molecule NO.

Electron transfer performed by haemproteins falls in to two categories; proteins that transfer electrons between different donors and acceptors, and proteins that use haem-groups for electron transport to and from their active site. One example is hydroxylamine oxidoreductase in which one haem group is at the catalytic site and seven other haems move electrons away from the active site (Prince and George, 1997). Haemproteins that solely transfer electrons belong to the family of cytochromes.

2.5 Cytochromes

Cytochromes are the family of haemproteins that transfer electrons. The first cytochromes were purified and described in the 1880's (MacMunn, 1886) and in 1925 the name "cytochrome", which simply means cellular pigment, was introduced (Keilin, 1925). The name was used to describe a group of intracellular haemproteins that undergo oxidation and reduction, and in the reduced form exhibit intense absorption bands between 510 and 615 nm. As currently used, the name appears to include all intracellular haemproteins with the exception of globins, the peroxidases, catalase, tryptophan 2,3-dioxygenase, some of the haem-thiolate proteins (P450) and the nitrite and sulphite reductases. Consequently, proteins of markedly different function are found in this family.

Thus a number of enzymes are also commonly referred to as cytochromes. These include cytochrome-*c* oxidase, cytochrome P450 (a family of haem-thiolate proteins), and nitric oxide synthase.

There are four major subclasses of cytochromes established named cytochrome *a, b, c* and *d* according to the haem group they contain (Palmer and Reedijk, 1991). Cytochromes are essential components in all organisms that have a respiratory chain or a photosynthetic system. This includes all aerobe organisms.

2.6 Cytochrome c

Cytochrome c is probably the most thoroughly studied example of any class of redox protein (Cowan, 1993). Cytochrome c is commonly associated with the soluble mitochondrial matrix protein that transfers electrons between complexes III and IV of the mitochondrial respiratory chain. The class of cytochrome c's does however comprise a large variety of proteins. Cytochrome c's can be defined as electron - transfer proteins having one or several haem c groups, bound to the protein by one or, more commonly, two thioether bonds involving the sulphydryl groups of cysteine residues. The fifth haem iron ligand is always provided by a histidine residue. Cytochrome c possesses a wide range of properties and functions in a large number of different redox processes (Pettigrew and Moore, 1987;Moore and Pettigrew, 1990). There are currently four types of totally unrelated cytochrome c's, which are classified by sequence and 3D structure.

Type I includes the low-spin soluble cytochrome c of mitochondria and bacteria, with the haem attachment site towards the N-terminus, and the sixth ligand provided by a methionine residue about 40-60 residues further on towards the C-terminus. The proteins contain three conserved "core" helices, which form a basket around the haem group with one haem edge exposed to the solvent.

Type II includes the high-spin cytochrome c' and a number of low-spin cytochromes. The haem attachment site is close to the C-terminus. The protein fold comprises a four α-helix bundle (Moore, 1991).

Type III comprises the low redox potential multiple haem cytochromes: cytochrome c_7 (tri-haem), c_3 (tetra-haem), and high molecular weight cytochrome c. The haem *c* groups, all bishistidine coordinated, are structurally and functionally non-equivalent and present different redox potentials in the range 0 to -400 mV (Coutinho and Xavier, 1994).

Type IV was established to comprise the complex proteins that have other prosthetic groups as well as haem *c*, e.g. flavocytochrome *c* and cytochromes *cd*.

2.7 The Highly Axial Low-Spin EPR Signal

Electron paramagnetic resonance is a spectroscopic technique that probes unpaired electrons in radicals and transition metal ions (see Chapter 3.11). EPR signals from transition metals are dependent on the oxidation state and the metal ligands. This makes EPR spectroscopy a useful technique in studying the behaviour of metal containing proteins.

EPR has been crucial in the study of many important systems like cytochrome c oxidase of the mitochondrial respiratory chain, photosystem I and II, and ribonucleotide reductase.

Highly axial low-spin (HALS) EPR spectra have been observed in a small number of very diverse cytochromes. These include several membrane bound b-type cytochromes, for example cytochrome b_{558} from *Bacillus subtilis* (Hederstedt and Andersson, 1986;Friden *et al.*, 1990), cytochrome b_{562} and b_{566} in Complex III of the bovine mitochondrial respiratory chain (complex bc_1) (Tsai and Palmer, 1982;Tsai and Palmer, 1983), and the multi-subunit cytochrome *bf* complex of the photosynthetic electron transport in spinach chloroplasts (Schunemann *et al.*, 1999).

C-type cytochromes that exhibit this kind of EPR signal include cytochrome f from the cytochrome *bf* complex (Schunemann *et al.*, 1999), cytochrome c peroxidase from *Nitrosomonas europaea* (Arciero and Hooper, 1994), cytochrome c-553 from *Bacillus pasteurii*, and cytochrome c-552 from *Nitrosomonas europaea* (Arciero *et al.*, 1994). The b-type cytochromes that exhibit HALS EPR signals generally have bishistidine ligation of the haem iron, whereas the c-type cytochromes generally have histidine and methionine ligation of the haem iron. One exception is the c-type cytochrome in cytochrome f from the cytochrome *bf* complex, were the haem iron is ligated by a histidine and the α-amino group of the N-terminal tyrosine (Schunemann *et al.*, 1999).

Model compounds of low-spin ferric haem with high-basicity pyridines or imidazoles as the axial ligands have been shown to exhibit EPR signals that are dependent on the geometry of the ligands (Walker, 1999;Ogura *et al.*, 2001). HALS EPR signals are observed for complexes where the two planar axial ligands are aligned perpendicular to each other. Complexes that exhibit rhombic EPR spectra are found to have parallel or near-parallel alignment of the planar axial ligands. The electronic ground state is however $(d_{xy})^2(d_{xz},d_{yz})^3$ in both these cases (Walker, 1999).

In the b-type cytochromes that exhibit HALS EPR signals the two axial histidine ligands are aligned perpendicular or near-perpendicular to each other. Perpendicular histidine ligands are also found in myoglobin $\text{His}^{64}\rightarrow\text{Val}$ / $\text{Val}^{68}\rightarrow\text{His}$ double mutants which exhibit a HALS EPR signal (Dou *et al.*, 1995).

A similar correlation between the HALS EPR signal and structure has not been established for other combinations of axial ligands.

2.8 Aim of the Thesis

Cytochrome c-554 from *Methylosinus trichosporium* OB3b has been found to exhibit a HALS EPR signal. The aim of these studies has been to characterise this protein and the behaviour of the HALS EPR signal exhibited by this protein. Cytochrome c's with similar histidine and methionine ligation of the haem iron can exhibit different EPR signals. The HALS signal could indicate a difference in the electron transport pathway to and from the haem iron.

Better understanding of this signal in histidine and methionine ligated haem systems will be helpful in the future study of proteins exhibiting similar EPR signals.

3 Methods

3.1 Growth of Methylosinus trichosporium OB3b

Methylosinus trichosporium OB3b is a Type II methanotroph (a bacterium able to utilize methane as the sole source of both carbon and energy (Morris, 1992)), classified by the content of paired membranes extending throughout the cytoplasm, and the serine pathway utilized for formaldehyde assimilation (Hanson *et al.*, 1992). This bacterium can produce two different forms of the methane monooxygenase enzyme (MMO) depending on the copper concentration. Low copper content in the medium will lead to predominantly production of the soluble sMMO while at copper concentrations over 5 μM the membrane bound form of the enzyme, pMMO, will dominate (Stanley *et al.*, 1983). In this project *Methylosinus trichosporium* OB3b was grown in a medium containing 10 μM copper. The growth medium used for both the agar plates and the large-scale growth is described in Appendix A.

3.1.1 Culture Maintenance

The *Methylosinus trichosporium* OB3b cell culture was continuously maintained on agar plates in a gas-tight container at room temperature. The atmosphere in the container consisted of 20% methane and 80% air. The atmosphere was exchanged weekly through a 0.2 μm sterile filter.

3.1.2 Large Scale Growth

Larger particles like bacterial cells will scatter light with wavelengths from around half their diameter and longer. This phenomenon is known as Mie Scattering (Meyer-Arendt, 1972). This causes bacterial cells to scatter light in the upper half of the visible spectrum. The optical density of a bacterial culture at wavelengths higher than 500 nm is therefore a direct measure of the relative number of bacteria in the solution. OD_{540} (optical density at 540 nm), which is measured as light absorption at 540 nm, was used to monitor the cell density throughout the growth process.

Cultivation of *Methylosinus trichosporium* OB3b was started by inoculation of a single colony from an agar plate in 50 ml medium. The medium is described in Appendix A. These pre-cultures was grown in sterile 250 ml Erlenmeyer flasks sealed with a sterile rubber stopper and a 0.2 μm sterile filter. The atmosphere of 20 % methane and 80 % air was added through the 0.2 μm sterile filter and exchanged daily. The flasks were kept at 30°C and at constant shaking at 200 rpm.

When the pre-culture reached an OD_{540} of approximately 0.5 (after about a week) it was transferred to 2000 ml Erlenmeyer flasks and diluted 10 times (to a total volume of 500 ml). The 2000 ml Erlenmeyer flasks were sealed with a sterile rubber stopper and a 0.2 μm sterile filter through which the 20% methane and 80 % air atmosphere was added. The cell culture was cultivated under the same conditions as the pre-culture until it reached an OD_{540} of 1-2. This cell density was reached after 5 to 7 days.

The culture obtained from this cultivation was used to inoculate a large-scale culture in a 10 l Biostat fermentor. The culture was gradually diluted with medium (one litre each time) until the total volume reached 8 litres. Filling the fermentor completely would make occasional foaming clog the exhaust filter and block the airflow.

The Biostat fermentor continuously monitors pH and temperature, and controls stirring and gas-flow through the culture. The temperature was kept at 30°C and the stirring at 200 – 300 revolutions per minute. Sulphuric acid was added when needed to keep the pH in the range of 6.8 – 7.0. The fermentation was stopped after 10 to 14 days, when the OD_{540} had reached 3 – 5.

3.1.3 Cell Harvesting

The cells were harvested by first pouring the bacterial culture onto crushed ice. Using crushed ice made from tap water is normally not a good idea when working with metalloproteins, as it is contaminated with metal ions. In this case this was not a problem as the contaminants could be removed during the purification of cytochrome c-554 without harming the cytochrome itself.

The cooled cell culture was filtered through a Millipore Masterflex Cross-flow filtration unit with a cut-off of 0.45 μm, until the volume of the cell culture was around 1 litre. The concentrated cells were then centrifuged at 7 000 x g_n for 10 minutes, washed in 50 mM Pipes buffer pH 7.0 and centrifuged once more. The cell pellet was transferred to small plastic bags in portions of 30 g and stored at – 80 °C for later protein purification.

3.2 Purification of Cytochrome c-554 from *Methylosinus trichosporium* OB3b

The procedure for purification of cytochrome c-554 from *Methylosinus trichosporium* OB3b was developed by my predecessor Siv Fauchald. The method described here is a slightly modified version of the one described in her *Cand. Scient.* thesis from 2000 (Fauchald, 2000).

3.2.1 Cell Disruption and Preparation of Cell Free Extract

Principle:

When squeezed through a small opening under high pressure the membranes of frozen cells will disrupt. This is the principle for the **X-press**® method (Magnusson and Edebo, 1976).

Procedure:

1. The **X-press**® was cooled in a -20 °C freezer over night.
2. 30 g of cells were taken directly from the -80 °C freezer into the cold **X-press**® (AB **BIOX**, Gothenburg) and crushed three times. The **X-press**® equipment were kept on dry ice when not in use in order to keep it cold.
3. The broken cells were dissolved in 120 ml 10 mM Tris HCl buffer at pH 8.2.
4. 40 μl of DNase I solution (2000 Kunitz units/ml) was added in order to reduce viscosity. (One Kunitz unit will produce a ΔA_{260} of 0.001 per min per ml at pH 5.0 at 25 °C, using DNA (Type I or III) as substrate and 4.2 mM Mg^{2+} (Kunitz, 1950).)
5. The cell suspension containing DNase I was stirred for 60 minutes at 4 °C.
6. The cell suspension was homogenized using a 50 ml Teflon hand homogeniser.
7. The resulting solution was centrifuged at 12 000 x g_n for 20 minutes.
8. The pellet was discarded and the supernatant centrifuged at 75 000 x g_n for another three hours. The supernatant from this centrifugation was termed cell free extract.

3.2.2 Ammonium Sulphate Precipitation

Principle:

Water-soluble proteins have polar surfaces and interact with water mainly through hydrogen bonds and other polar interactions. The hydration sphere of proteins packs a large amount of water molecules in ordered positions around the protein. When large amounts of salts are dissolved into the solution water molecules coordinate around the ions and stabilize them. When the total amount of ions in the solution increase the energetically most favourable interactions changes from being those between water and the proteins to being interactions between the proteins themselves. Differences in surface properties cause different proteins to precipitate at different concentrations of ammonium sulphate (Green, 1932).

Procedure:

1. Ammonium sulphate was dissolved gradually into the "cell free extract" until the solution was 75 % saturated. This corresponds to adding 47.6 g ammonium sulphate per 100 ml solution.

2. The ammonium sulphate suspension was stirred for 60 minutes at 4 °C.
3. The solution was then centrifuged at 13 000 x g_n for 30 minutes and the supernatant kept for further protein purification.

3.2.3 Reversed Phase Chromatography

Principle:

Reversed phase chromatography (also known as hydrophobic chromatography) is a form of adsorption chromatography in which the adsorption to the stationary phase occurs through hydrophobic interactions. The polarity of the protein surface will determine to which degree a protein will interact with and be retarded by the matrix. Proteins can in this way be separated according to their surface polarity (Greibrokk *et al.*, 1994).

High salt concentrations will increase the ionic strength of the mobile phase and promote the hydrophobic interactions between the solutes and the column matrix and thus induce binding of the proteins to the matrix.

Procedure:

1. A phenyl sepharose column (1.5 cm x 30 cm) was equilibrated with 1.5 M $(NH_4)_2SO_4$ in 10 mM Tris HCl buffer at pH 8.2.
2. The supernatant was applied to the column. The protein was desorbed in the upper layer of the column because of the high salt concentration.
3. The desorbed protein was separated by elution with a linear gradient from 1.5 M to 0 M $(NH_4)_2SO_4$ in a 10 mM Tris HCl buffer at pH 8.2.
4. The eluted fractions that had significant absorption at 410 nm were collected and concentrated in an Amicon® ultrafiltration cell. A regenerated cellulose membrane with 10 kDa filter cut-off were used in the ultrafiltration cell.

3.2.4 Cation-Exchange Chromatography

Principle:

Water-soluble proteins have surfaces highly covered with charged and polar amino acid residues. In cation exchange chromatography positive charges on the protein surface interact with an anionic column matrix (Greibrokk *et al.*, 1994). Proteins with different surface properties will make different interactions with the matrix and can therefore be separated. Strongly cationic proteins binds stronger to the matrix than proteins with less positive charge.

High salt concentrations in the mobile phase give counter-ions that will replace the proteins on the column matrix. The proteins will subsequently be eluted according to their surface charge.

Procedure:

1. The ionic strength of the protein solution was decreased using a NAP-10 column, which is a standardized Sephadex G-25 gel filtration column (the principle of gel filtration is explained in 3.2.5).
2. The protein solution was applied to a CM sepharose column (1.5 cm x 20 cm) equilibrated with 10 mM Tris HCl buffer at pH 8.2.
3. The proteins were separated by elution with a linear gradient from 0 to 0.15 M KCl in 10 mM Tris HCl buffer at pH 8.2.

4. The eluted fractions that had significant absorption at 410 nm were collected and concentrated in an Amicon® ultrafiltration cell with a filter with a cut-off value of 10 kDa to a final volume of about 3 ml.

3.2.5 Gel Filtration

Principle:

Gel filtration chromatography separates molecules by their size (Greibrokk *et al.*, 1994). The column matrix consists of spherical particles with pores of size similar to the size of the molecules that are to be separated. Small molecules, which can access all the pores, will experience a greater solvent volume than big molecules (e.g. proteins). Proteins larger than the pores will be completely excluded from the matrix and be eluted in the void volume in the equilibration buffer. Molecules in the size range between total exclusion from the pores and total access to the pore volume will be separated according to their molecular weight. Smaller molecules will have a larger solvent volume to go through and will use more time to pass through the column.
The separation range varies depending on the selected gel filtration matrix. Sephadex matrixes are bead-formed gels made of cross-linked dextran where the degree of cross-linking determines the separation properties. Physical properties of some Sephadex matrixes are shown in table 3-1 (Pharmacia Biotech., 1993).

Gel type	Dry bead size	Fractionation range
Sephadex G-25 M	50 – 150 µm	1 000 – 5 000 Da
Sephadex G-50 M	50 – 150 µm	1 500 – 30 000 Da
Sephadex G-75	40 – 120 µm	3 000 – 80 000 Da
Sephadex G-100	40 – 120 µm	4 000 – 150 000 Da

Table 3-1 Properties of Sephadex matrixes and their fractionation range for globular proteins.

Procedure:

1. The protein solution was applied to a Sephadex G-75 gel filtration column (1.5 cm x 30 cm) equilibrated with 10 mM Tris HCl buffer at pH 8.2.
2. The proteins were separated by elution with 10 mM Tris HCl buffer at pH 8.2.
3. The eluted fractions that had significant absorption at 410 nm were collected and concentrated in an Amicon® ultrafiltration cell with a filter with a cut-off value of 10 kDa. The concentrated protein solution was stored at –80 °C.

3.2.6 SDS Polyacrylamide Gel Electrophoresis

The purity of the purified protein was analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using the Pharmacia PhastSystem.

Principle:

SDS-PAGE separates denatured macromolecules according to their molecular weight. Anionic sodium dodecyl sulphate (SDS) molecules combine with polypeptide chains in a number proportional to the size of the polypeptide, leading to a linear size to charge relation. Mercaptoethanol will cleave disulphide bonds covalently attaching polypeptides to each other. Hence, only monomers of the proteins will be present.

Polypeptides are thus separated according to their molecular weight rather than their native charge (Laemmli, 1970;O'Farrell, 1975).

Procedure:

1. 7 µl of the protein solution was mixed with 4 µl Phast loadmix in an Eppendorf tube. The Phast loadmix contains SDS, mercaptoethanol and other denaturing agents.
2. Sealed Eppendorf tubes were boiled for 10 minutes and insoluble material spun down in a micro centrifuge.
3. The PhastSystem instrument was prepared with a PhastGel 8 – 25% gradient gel and buffer strips and set to a temperature of 16° C.
4. An application comb was used to position 1 µl of each sample above the gel in the instrument.
5. The PhastSystem instrument was started using pre-programmed separation and development programs. The automated development program is based on protein staining with Coomassie Brilliant Blue. Coomassie Brilliant Blue binds non-specifically to proteins, thus staining the protein bands in the gel.

3.3 Reduction / Oxidation of the Haem Centre

Principle:

The haem iron naturally exists predominately in two different oxidation states, Fe^{2+} (ferrous) and Fe^{3+} (ferric), which have markedly different spectroscopic properties (Meyer and Kamen, 1982). After purification the cytochromes exist in both redox states. For spectroscopic investigations it is necessary that the majority of proteins in the sample are in the same redox state.

Procedure:

Reduction:

1. An excess of solid sodium dithionite ($Na_2S_2O_4$), which is a strong reductant, was added to the protein sample, thus reducing ferric haem iron to the ferrous form.
2. The reduced protein was separated from the remaining salt using a NAP-5 column, which is a standardized Sephadex G-25 gel filtration column (principle explained in 3.2.5), equilibrated with a cupper free Tris HCl buffer at pH 8.2. The equilibration buffer was varied according to the kind of experiments planed for the sample. No auto-oxidation is observed as the six-coordinated haem iron is inaccessible to molecular oxygen.

Oxidation:

1. An excess of the oxidizer $K_3Fe(CN)_6$ was added to the protein sample, thus oxidizing ferrous haem iron to the ferric form.
2. The oxidized protein was separated from the remaining salt using a NAP-5 column, equilibrated with a cupper free Tris HCl buffer at pH 8.2.

3.4 Haem Cleavage

Principle:

The haem groups in c-type cytochromes are covalently bound to the peptide through two thioether bonds with cysteines. In order to remove the haem without modifying the protein extensively, haem cleavage with 2-nitrophenylsulfenylchloride (2-NPS), described by Fontana et al. (1973), was used with some modifications. This treatment cleaves the thioether bonds between the haem group and cysteines in the protein leaving a disulfide bond between cysteine and a 2-nitrophenylsulfenyl group (Figure 3.1). The new disulfide bond can be cleaved by treatment with mercaptoethanol.

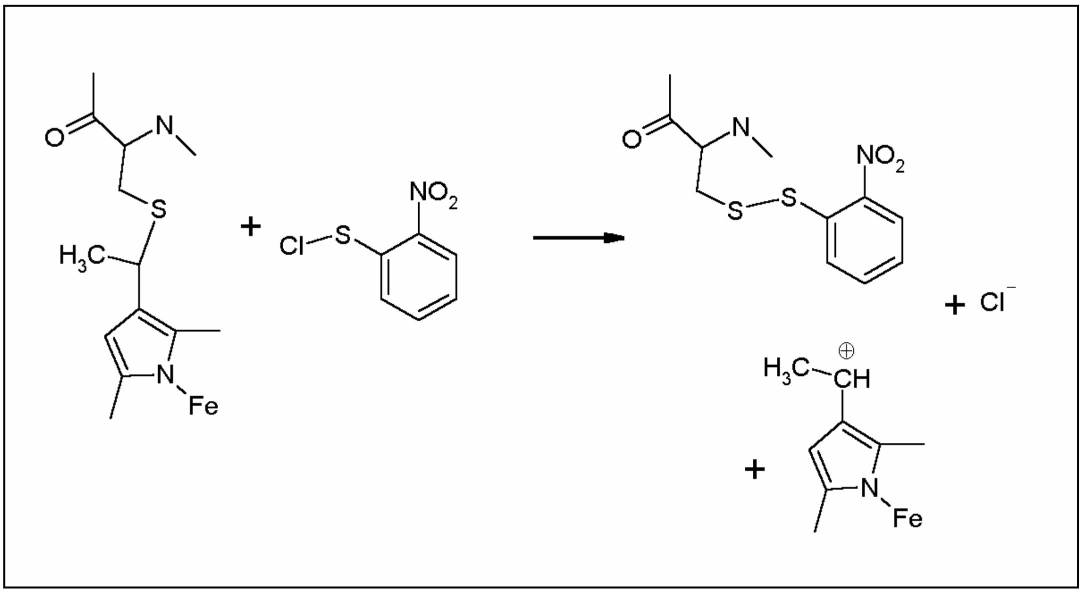


Figure 3.1 The reaction of 2-nitrophenylsulfenylchloride with the thioether bridges that link the haem group to the polypeptide chain.

Treatment with 2-nitrophenylsulfenylchloride is fairly mild, but modifies tryptophan as shown in figure 3.2. The modified tryptophans are chromophores with absorption maxima at 365 nm.

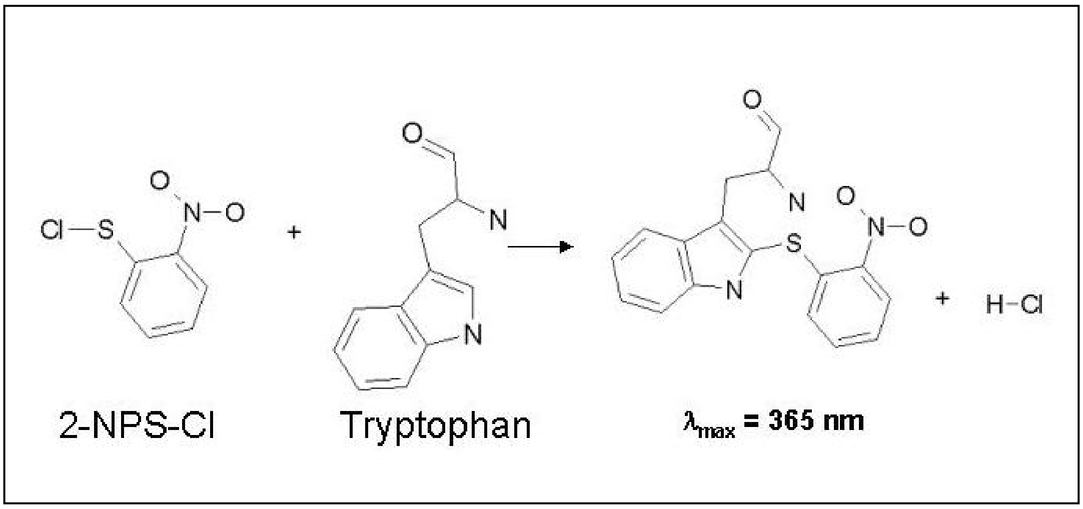


Figure 3.2 The reaction of tryptophan with 2-nitrophenylsulfenylchloride

Procedure:

1. 40 µl protein solution was mixed with 40 µl acetic acid and 40 µl 4mM 2-NPS-Cl dissolved in glacial acetic acid. The mixture was left to incubate for 10 minutes.
2. 80 µl 0.4% mercaptoethanol in a 10 M urea solution was added in order to remove the 2-NPS groups from the cysteine. The solution was incubated for another 10 minutes.
3. The modified protein was separated from the haem groups and the other reagents on a NAP-5 column, equilibrated with 10% acetic acid.
4. The modified protein was collected and kept at –20 °C. The modified protein contains modified tryptophan residues, which give rise to a characteristic absorption maximum at 365 nm that can be used to identify and detect the modified protein.

3.5 Protein Sequencing

Principle:

The primary structure of a protein (i.e. the amino acid sequence) can be found using the Edman degradation method (Edman, 1950). Amino acid sequence analysis using Edman degradation has historically been one of the most important techniques for the investigation of proteins at the molecular level. Amino acid derivatives are sequentially cleaved one at a time from the N-terminal end of the protein and analysed, leaving the remaining polypeptide chain unchanged. The amino acid derivatives can be identified e.g. with chromatographic methods or melting point. A prerequisite for the Edman degradation is that the protein has a chemically accessible alpha-amino group (Voet and Voet, 1995).
The mechanism of the Edman degradation consists of three separate steps, each requiring quite different conditions:

1. The reaction of the polypeptide with phenylisothiocyanate (PITC), known as Edman's reagent, under mildly alkaline conditions to form their phenylthiocarbamyl adduct. PITC reacts specifically with primary and secondary amines.
2. Treatment with anhydrous trifluoroacetic acid cleaves off the N-terminal residue as its thiazolinone derivative without hydrolysing any other peptide bonds.
3. The unstable thiazolinone derivative is converted to the more stable phenylthiohydantion (PTH) derivative. This PTH-amino acid can be analysed and identified.

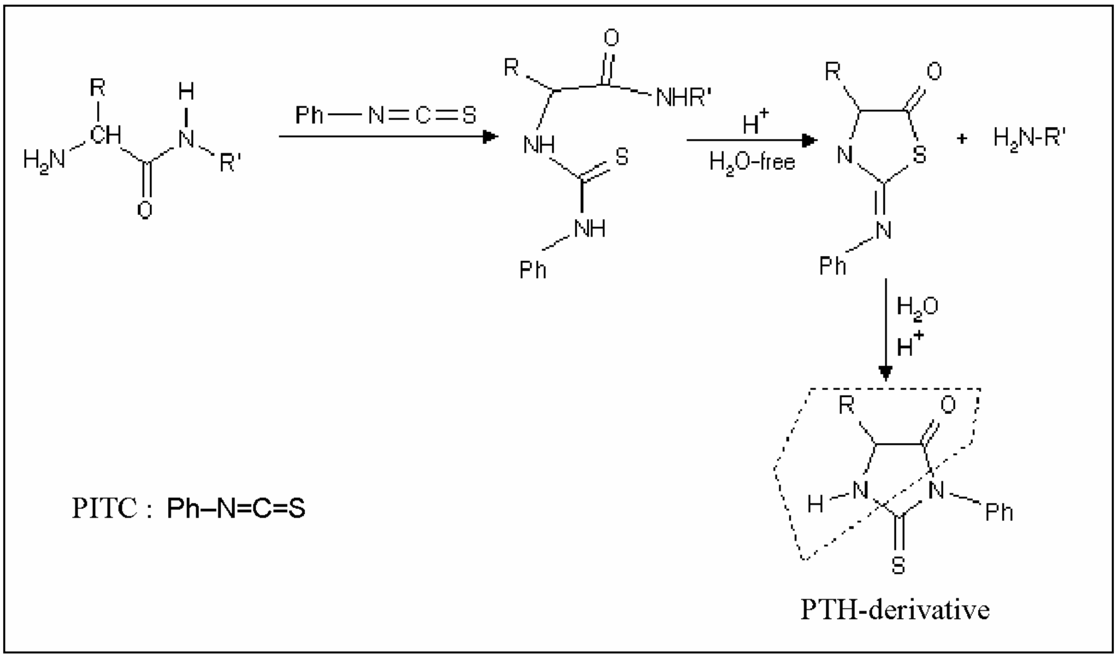


Figure 3.3 Edman degradation of the N-terminal amino acid of a peptide. The marked part of the final product corresponds to the amino acid that is cleaved off.

This procedure can be repeated on the remaining polypeptide and further amino acids investigated. Normally the first 30 – 80 residues from the N-terminal end can be identified using this method.

Procedure:

1. A sample containing 1 nmol of protein was dried and redissolved in 70% formic acid (COOH) to a protein concentration of 1 mg/ml.
2. This sample was inserted to an automated amino acid sequencer and analysed. The sequencer was an Applied Biosystems model 477A with a 120A PTH-analysator (detection wavelength 270 nm).
3. For each step of the sequencing, the detected levels of each of the twenty standard amino acids were reported. The amino acid sequence was assigned from the collected data by manual analysis.

3.6 Protein Cleavage

Principle:

Protein sequencing of long polypeptides cannot be done to completion. The polypeptide needs to be cleaved into shorter peptide chains, which can be analysed separately by Edman degradation. Methods for specific protein cleavage include enzymatic digestion and treatment with cyanogen bromide. Cyanogen bromide (CNBr) specifically cleaves peptide bonds on the C-terminal side of internal methionine residues leaving the methionine as a C-terminal homoserine lactone (Inglis and Edman, 1970).

Procedure:

WARNING! Cyanogen bromide is highly toxic (R 26/27/28-34) and has a high vapour pressure at room temperature.

1. A sample containing 50 nmol of protein was dried and dissolved in 70% formic acid.
2. A 50-fold excess of CNBr (compared with the expected molar quantity of methionine residues) was added to the sample.
3. The mixture was incubated in the dark at 20° for 24 hours.
4. CNBr and formic acid was removed by evaporation before sequencing.

3.7 Sequence Alignment using pBLAST

Principle:

pBLAST (protein Basic Local Alignment Search Tool) is a net-based set of similarity search programs that allows you to compare the primary structure of a proteins with all structures in a protein or nucleotide database (Altschul *et al.*, 1997;Wheeler *et al.*, 2001). As the sequences of similar proteins tend to differ greatly a local alignment algorithm, which finds statistically significant similarities in small regions regardless of insertions or deletions, helps finding relevant sequences (Altschul and Gish, 1996). pBLAST allows sequence comparisons with both sequence and structural databases. As sequence homologies often reflect structural homology, comparing the primary sequence of a protein with a structural database can give insight into the 3D structure of the protein (Chothia and Lesk, 1986).

Procedure:

1. An Internet browser was used to open a BLAST site (e.g. http://www.ncbi.nlm.nih.gov/BLAST/).
2. The protein sequence was inserted into the search window and the type of sequences defined (e.g. protein-protein or protein-nucleotides). After selection of a database (e.g. SwissProt or PDB) or a genome, the comparison was started.
3. The search tool returned the sequence of all similar proteins found together with statistics on the quality of each result.

3.8 MALDI-TOF Mass Spectroscopy

Principle:

Mass spectroscopy (MS) is a useful and accurate technique for determination of molecular masses. The basic principle in mass spectroscopy is to ionise the sample and accelerate it with a magnetic field through a vacuum tube. The methods of ionisation were for a long time very destructive leading to fragmentation of the sample molecules. This gave much information about the molecules, but made mass spectroscopy useless for the study of larger molecules like proteins. The advent of matrix-assisted laser desorption-ionisation (MALDI) and electrospray ionisation (ESI) made mass spectroscopy a useful tool for the study of larger molecules due to their non-destructive nature[1]. Molecular weights of proteins with masses over 150.000 Da have now been determined accurately (Van Holde *et al.*, 1998).

In MALDI instruments the sample is mixed with an acidic matrix that is easily vaporized by a laser inside the instrument. This vaporization will throw the sample molecules into the vacuum tube without being destroyed by the laser. The sample molecules will to a large extent exist as protonated cations. A magnetic field accelerates the cations. The speed that the molecules obtain will be proportional to the ionic charge and inversely proportional to the molecular weight. The most prominent feature in the spectrum will correspond to a single charged protein molecule. A divalent protein cation will give rise to a peak at half the molecular weight to charge ratio. The presence of divalent protein cations and singly charged dimers of the protein, can confirm the measured molecular weights.

Instruments with a TOF (Time of Flight) detector accelerates the ions and let them travel a certain path to the detector. Measuring the flight time of an ion through this path allows the determination of its mass.

The isotope distribution of the atoms in the molecules will cause individual molecules to have slightly varying molecular weights. For molecules containing a large number of atoms this will lead to a broadening of the peaks in the mass spectrum.

Procedure:

1. The matrix was a 1:1 mixture of water and acetonitrile containing 0.1% TFA (trifluoro-acetic acid) and saturated with sinapinic acid (3,5-dimethoxy-4-hydroxy cinnamic acid).
2. Equal volumes of the matrix solution and the protein sample were mixed. 5 µl of a 10 nM solution will be sufficient to obtain data. Different dilutions of the protein sample were measured in order to obtain the best possible conditions.
3. 1 µl of the mixture was transferred to the sample plate and left to dry.
4. Dried samples were inserted into the mass spectrometer for analysis.

---

[1] For the development of these two methods John B. Fenn (USA) and Koichi Tanaka (Japan) was awarded the Nobel Prize in Chemistry for 2002 *"for their development of soft desorption ionisation methods for mass spectrometric analyses of biological macromolecules*"

3.9 Optical Spectroscopy

Principle:

Molecules have numerous electronic quantum states differing in energies. When a molecule absorbs light it usually occurs through an excitation of an electron from its ground state to a state of higher energy. For absorption to occur the energy difference between the two electronic states has to match the energy of the incoming photon (equation 3.1). As the energy of a photon is directly dependent on the wavelength of the photon, the wavelength of absorption gives information about the electronic structure of the molecule studied.

$$\Delta E = h\nu = \frac{hc}{\lambda}$$

Equation 3.1

ΔE is the energy difference between two orbitals, h is the Plank constant, ν is the frequency of a photon, λ is the wavelength of a photon and c is the speed of light in vacuum.

The Beer-Lambert law describes how the intensity of absorption varies with the molar concentration C of the sample and the light path l at a given wavelength (Atkins, 1998):

$$A = \log\left(\frac{Io}{I}\right) = \varepsilon \times l \times C$$

Equation 3.2

A is the absorbance, I_0 is the intensity of the incident light, I is the intensity of the transmitted light, ε is the molar extinction coefficient of the sample molecules, l is the length of the light path through the sample and C is the molar concentration of the sample.

Normal single covalent C-C bonds absorb light around 200 nm. Double bonds require less energy to excite an electron; hence they absorb light at higher wavelengths. In general highly conjugated double bond systems (aromatic) or transition metal complexes are needed to get absorption in the visible part of the electromagnetic spectrum (400 nm to 700 nm).

Haem, which is an aromatic system coupled to a transition metal, has a strong and characteristic absorption spectrum consisting of several absorption bands. These absorption bands are all sensitive to the oxidation state and the ligation of the haem iron, which makes haem proteins ideally suited for spectroscopic studies.

Low-spin ferrous haem, i.e. haem with the oxidation state Fe^{2+}, has two sharp peaks in the visible region at around 550 nm (the α-band) and around 520 nm (the β-band) which are responsible for the strong colour. The Soret peak (the γ-band) has its maximum around 415 – 420 nm and a high extinction coefficient.

The low-spin form of oxidised ferric haem, with the oxidation state Fe^{3+}, has only a single broad maximum in the visible region stretching from around 520 nm to 550 nm. The Soret band in ferric haem is blue-shifted compared to the ferrous form to around 410 nm, and has a slightly smaller extinction coefficient. Ferric haem can also exhibit an absorption band in the UV region around 360 nm (the δ-band).

In addition to these *d-d** transitions, ligand-to-metal or metal-to-ligand charge-transfer bands might be observed depending on the ligands coordinating with the iron. If these charge transfer bands occur they are extremely weak compared to the *d-d** transitions.

High-spin haems have different optical spectra with distinct charge-transfer bands. Binding of small ligands (e.g. O_2, CO or CN^-) does also change the optical absorption spectrum.

3.9.1 Pyridine Haemochrome

The extinction coefficients of haem containing proteins vary from protein to protein. The pyridine haemochrome method allows spectral characterisation of the haem group without interference from the protein.

Principle:

By coordinating the haem iron to two molecules of a base, characteristic spectra can be obtained that are independent of the corresponding peptide chain. The extinction coefficients for pyridine haemochromes of many different types of haem are available and is often used for identification and determination (Falk, 1964). Pyridine haemochrome spectra are measured in aqueous alkaline solutions after reduction with sodium dithionite. A final concentration of 0.075 M NaOH and 2.1 M pyridine has been recommended (Falk, 1964).

The characteristic extinction coefficients for the absolute spectrum of haem c pyridine haemochrome are given in Table 3.2.

Haem c pyridine haemochrome	β maximum	α maximum
Wavelength, λ (/nm)	522	551
Extinction coefficient, ε (/$mM^{-1}cm^{-1}$)	18.6	29.1

Table 3-2 Pyridine haemochrome characteristics for haem c.

Procedure:

1. A stock solution of 0.15 M NaOH and 4.2 M pyridine was made by mixing 1.5 ml 1M NaOH, 1.7 ml pyridine and distilled water to a total volume of 10 ml.
2. The protein sample was reduced with a few grains of sodium dithionite before it was mixed with an equal volume of the stock solution.
3. The pyridine haemochrome obtained was studied by optical spectroscopy.

3.10 Circular Dichroism (CD)

Circular Dichroism (CD) is the phenomenon in which light absorption differs with the handedness of circularly polarized light (Van Holde *et al.*, 1998). The electric vector of circularly polarized light describes a helix (see Figure 3.4), which can either be left-handed or right-handed. Inherently asymmetric chromophores or chromophores in asymmetric environments will interact differently with right- and left-circularly polarized light leading to features in the CD spectrum.

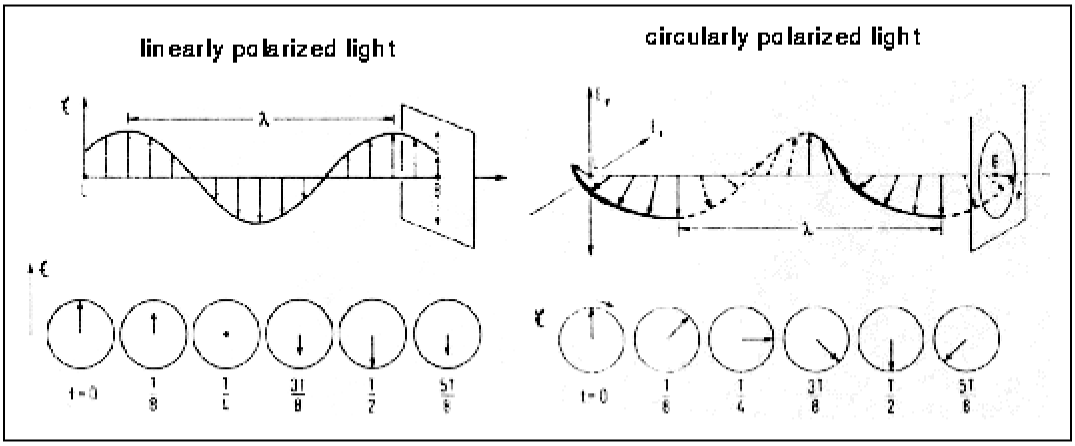


Figure 3.4 The propagation of the electric vector of linearly and circularly polarized light. (Figure taken from http://broccoli.mfn.ki.se/pps_course_96/ss_960723_21.html which is written by Kurt D. Berndt)

Circular Dichroism is expressed as either the absorption difference between the left and the right circularly polarized light $\Delta A = A_{left} - A_{right}$ or more commonly for historical reasons as the ellipticity (θ). Ellipticity in degrees is related to the difference in absorbance by $\Delta A = \theta / 32.98$ (Van Holde *et al.*, 1998). By taking concentration and path length into account ΔA can be converted to the more consistent $\Delta\varepsilon$ (difference in extinction coefficients between the left and the right circularly polarized light).

CD spectroscopy is complementary to optical spectroscopy as a weak band in the optical spectrum can have large amplitude in the CD spectrum and vice versa. CD can thus resolve peaks that are overlapping or not clearly defined in the visible spectrum. CD bands can be both positive and negative.

When polypeptides have secondary structure their peptide bonds are repeatedly in identical asymmetric environments. For example will all peptide bonds in an α-helix (or a β-sheet) experience the same asymmetric environment leading to a distinct feature in the CD spectrum. The amount of secondary structure of a protein can thus be estimated by using CD spectroscopy.

Metal active sites are almost exclusively asymmetric and the absorption bands arising from *d-d** transitions, and metal-to-ligand charge-transfer can exhibit strong CD signals. Most oxidised c-type cytochromes with Met-His ligation exhibit a weak ligand-to-metal charge-transfer band at 695 nm. This corresponds to an electron moving from the methionine ligand to the iron atom.

3.11 Electron Paramagnetic Resonance (EPR)

Electron Paramagnetic Resonance specifically probes the magnetic spin of unpaired electrons in an external magnetic field. In chemistry and biochemistry, species with unpaired electrons include inorganic and organic free radicals, triplet states (like the naturally occurring form of dioxygen), and systems that contain transition metals. EPR thus provides a very specific and sensitive tool for studying the "heart" of metalloproteins and free radical intermediates in biochemical reactions (Palmer, 2000). As metal ions and radicals often play a key role in enzyme- and electron transfer reactions, EPR can give information about the geometry of ligand binding sites, chemical and structural changes involved in biochemical function, and redox changes in electron transfer components.

EPR has been crucial in the study of many important systems like cytochrome c oxidase of the mitochondrial respiratory chain, photosystem I and II, and ribonucleotide reductase.

3.11.1 Introduction to EPR Theory

An electron is an elementary particle that can be described by its mass, charge and four quantum numbers (Atkins and Friedman, 1997). The state of an electron orbiting around an atomic nucleus is defined by the values of the four quantum numbers n, l, m_l, and m_s. The first three defines the atomic orbital for the electron, and the latter, the spin magnetic quantum number (m_s), defines the magnetic spin state of the electron. According to Pauli exclusion principle no more than two electrons may occupy any given orbital, and if two electrons occupy the same orbital, then their spin must be antiparallel (Pauli, 1925).

What we observe in EPR spectroscopy is net absorption of microwave radiation due to excitation of electrons from one electronic spin state to another. Absorption of microwaves is only observable if there is a population difference between the two possible energy states of the electron. A large population difference secures that excitation is the predominant phenomenon. If the two energy states are equally populated the rate of excitation and the rate of stimulated emission will be equal and no net absorption will be observed. According to the Boltzmann distribution the population difference can be described as a function of the energy difference and the temperature. The Boltzmann distribution and some of its implications are discussed in 3.11.1.3.

3.11.1.1 The Zeeman Effect

In an external magnetic field, unpaired electrons will align their magnetic dipole moment along or against the magnetic field lines. As the direction of the magnetic dipole in the external magnetic field determines the energy level of the electron, these two possible orientation of the magnetic dipole will have different energy. The energy difference between these two states is determined by the magnetic field the electron is experiencing (Figure 3.5). This is in turn determined by the applied magnetic field and magnetic dipoles in the immediate surroundings of the electron. The splitting of the two energy levels can be described as a function of the applied magnetic field:

$$\Delta E = g * \beta * B_0 \qquad \text{Equation 3.3}$$

where β is the Bohr magneton, B_0 is the external magnetic field and g is the electron g factor or spectroscopic splitting factor.

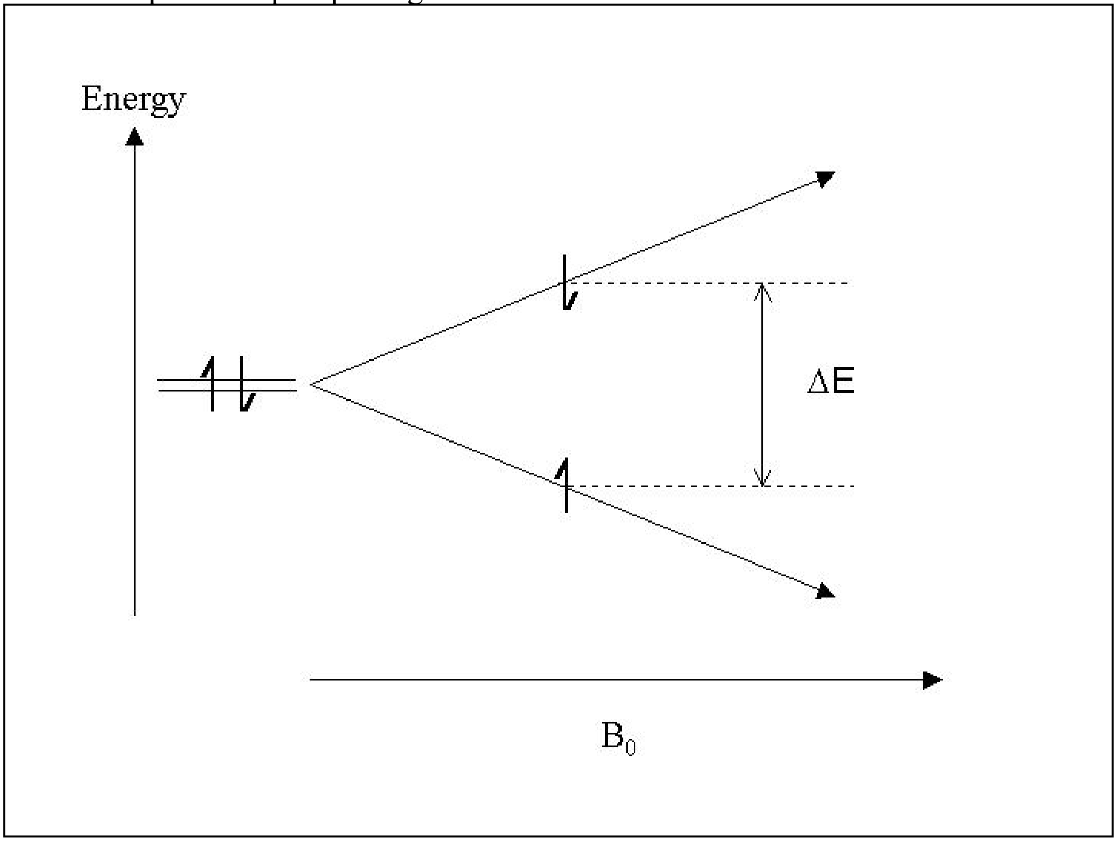


Figure 3.5 The energy levels of an unpaired electron are split by a magnetic field.

When the sample is irradiated with photons with an energy corresponding to the energy difference between the two states, resonance is achieved. This is shown in equation 3.1, which is known as the resonance condition. When there is a population difference between the two states photons with the resonance energy will be absorbed leading to a peak in the absorption spectrum.

Due to technical constraints most EPR spectrometers use a fixed microwave frequency and vary the magnetic field. Continuous wave EPR (cwEPR) spectra are predominantly recorded as 1st derivative of the absorption.

3.11.1.2 The g-Value

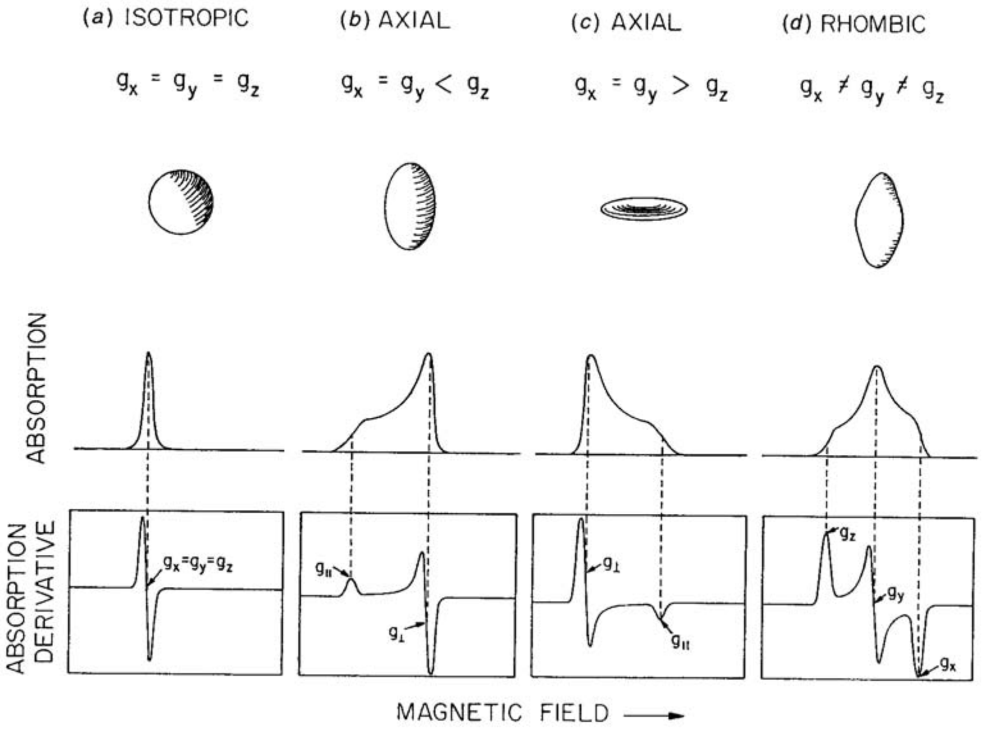
The absolute magnetic field positions of the lines in an EPR spectrum are characterized by the g-values. For a free electron $g_e = 2.00232$ (Øgrim and Lian, 1994), but the presence of nearby magnetic dipoles, which arises from electrons and nuclei, shift this value. From this follows that the electron g factor is a value determined by the immediate surroundings of the unpaired electron. This is similar to the chemical shift observed in NMR spectroscopy.

Whereas free radicals have resonance in the g = 2 region, not deviating strongly from the free electron g-value of 2.00232, metalloproteins can exhibit a wide variety of g-values ranging from less than 0.5 to more than 10. The deviation from the free electron g-value is determined by spin-orbit coupling, which is the interaction between the electrons spin and its orbital angular momentum. The effects of zero-field splitting (ZFS) can contribute to shifts of the observed g-values.

In anisotropic cases, which are what we normally observe, the g-value depends on the direction of the molecular axis relative to the magnetic field. All molecular orientations are present in molecules that are not in a crystal lattice, and the observed EPR spectra are the sum of all linear combinations of the anisotropic g-values. For an S = ½ (as in low-spin Fe(III)) calculations based on randomly oriented molecules lead to these four limiting cases as illustrated in Figure 3.6 (Palmer, 2000):

|

Isotropic	-	Three equal g-values, which means that the magnetic moment is independent of orientation. The system is completely symmetric.
Axial	-	Two of the g-values are equal but different from the third. If the third g-value is greater than the others the paramagnet can be represented by a prolate ellipsoid (a drawn out sphere). If the last g-value is smaller the paramagnet can be represented by an oblate ellipsoid (a squeezed sphere).
Rhombic	-	When all three g-values are different they all exhibit unique features in the spectrum. The magnetic moment has no symmetries.



(a) ISOTROPIC
$g_x = g_y = g_z$
(b) AXIAL
$g_x = g_y < g_z$
(c) AXIAL
$g_x = g_y > g_z$
(d) RHOMBIC
$g_x \neq g_y \neq g_z$
ABSORPTION
ABSORPTION DERIVATIVE
$g_x = g_y = g_z$
$g_{\parallel}$
$g_{\perp}$
$g_{\perp}$
$g_{\parallel}$
g_z
g_y
g_x
MAGNETIC FIELD ⟶

Figure 3.6 The four basic spectral envelopes found in S = ½ paramagnets (Palmer, 2000).

As the g-values give information of the paramagnet in the system, it follows that they give information about the factors affecting the paramagnet, which is the immediate surroundings of the unpaired electron. The energy levels of the orbitals are influenced by the same factors (Figure 3.7), and the energy difference between the orbitals can indeed be calculated from the g-values.

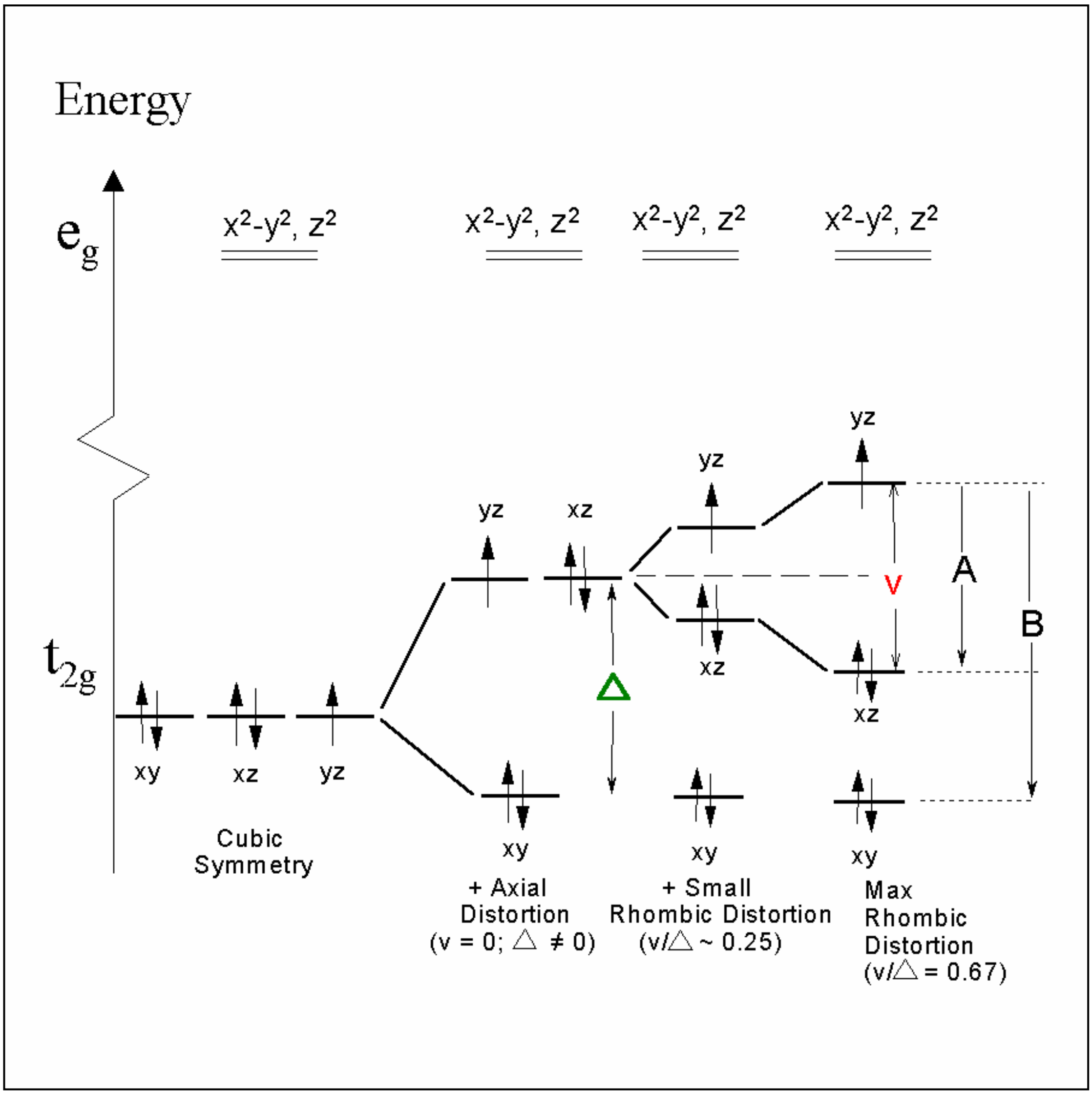

Energy
e_g
t_{2g}
x2-y2, z2
x2-y2, z2
x2-y2, z2
x2-y2, z2
yz
xz
yz
yz
xz
xy
xz
yz
xz
xy
xy
xy
Cubic Symmetry
+ Axial Distortion (v = 0; △ ≠ 0)
+ Small Rhombic Distortion (v/△ ~ 0.25)
Max Rhombic Distortion (v/△ = 0.67)
△
V
A
B

Figure 3.7 Splitting of the 3d orbitals of low-spin Fe(III) (Palmer, 2000). In low-spin haems the two e_g orbitals are too high in energy to influence the three t_{2g} orbitals.

For paramagnets with S > ½, i.e. systems with more than one unpaired electron, interactions between the unpaired electrons result in splitting of the energy levels without any external magnetic field, so-called zero-field splitting (ZFS) (Palmer, 2000). The zero-field splitting gives raise to strong anisotropic splitting and the g-values will be spread out in a large part of the magnetic field.

Whereas organic radicals have g-values only around 2, paramagnetic metal centres in proteins exhibits a great variety of g-values. EPR signatures at microwave frequency of 9.5 GHz (so-called X-band EPR, see Chapter 3.11.4) for some biologically interesting transition-metal centres is shown in Table 3.3.

Metal	Observed g-values
Mn(II)	2 – 6
Cu(II), type 1	2.036, 2.058, 2.277
Cu(II), type 2	2.053, 2.242
Cu(II), type 3	EPR silent
Mo(V)	1.95 – 2.0
Co(II)	1.8 – 6
Fe(III), high-spin, non-haem	0.25 - 10
Fe(III), high-spin, haem	1.8 - 8
Fe(III), low-spin, haem	0.25 – 3.9

Table 3-3 EPR properties of some metal ions found in biological systems (Lippard and Berg, 1994;Palmer, 2000)

3.11.1.3 The Boltzmann Distribution, Relaxation and Saturation

The energy difference between the two states where the EPR transitions occur is generally quite small, so the population difference, which is needed for detection, is also quite small. The population difference is described by the Boltzmann distribution and is a function of the energy difference between the two states (ΔE) and the temperature (T) as shown in equation 3.4.

$$\frac{N_{upper}}{N_{lower}} = e^{\frac{-\Delta E}{k_b T}} \qquad \text{Equation 3.4}$$

N_{upper} and N_{lower} is the population on the upper and lower energy states and k_b is the Boltzmann constant. If the populations in the two energy states are equal, no absorption can be observed, and the system is said to be saturated.
At room temperature the population difference may be too low for detection; hence EPR often has to be done under cryogenic conditions.
The relaxation rate is usually described by the spin-lattice (T_1) and spin-spin (T_2) relaxation times. Spin-lattice relaxation is caused by vibrational interactions between the paramagnetic centre and the molecular framework, while spin-spin relaxation involves interactions between two or more paramagnetic centres (Pilbrow, 1990).

In an unsaturated system, the double integral of the EPR first derivative spectrum is proportional to the square root of the microwave power.
At cryogenic temperatures the relaxation of the electrons from the excited state will be slow, making the signal easily saturated. An EPR signal is said to be saturated when the applied microwave radiation induces a non-equilibrium distribution between the energy states. This leads to a non-linear relation between the square root of the applied microwave power and the doubly integrated EPR signal. The saturation behaviour can give information about paramagnetic species in proteins. The saturation is temperature dependent and can be expressed by the characteristic parameter $P_{½}$,

also known as the half-saturation value. $P_{½}$ is defined as the microwave power at which half of the EPR signal is saturated.

3.11.2 EPR states observed in haems

Fe^{2+} (ferrous iron) and Fe^{3+} (ferric iron) are the two most commonly encountered oxidation states of haem iron. The ground state electron configuration is $[Ar]4s^2 3d^6$ for Fe (the ground state of iron), $[Ar]3d^6$ for Fe^{2+}, and $[Ar]3d^5$ for Fe^{3+} ([Ar] being the electron configuration of Argon). With six electrons in the d-orbitals Fe^{2+} can only achieve integer spin states or zero spin, whereas the five d-electrons in Fe^{3+} can adopt ground spin state of 1/2, 3/2, or 5/2. The integer spin states in ferrous iron are only observable under special conditions. Ferric iron is on the other hand almost exclusively found in the two EPR active states 1/2 (low-spin) and 5/2 (high spin) which are easily distinguished from each other.

3.11.3 EPR sample preparation

The quartz tubes used in the X-band EPR spectrometer (Wilmad Glass No. 707) have the dimensions 25 cm x 3.8 mm (inner diameter) and the sample volume were in the range between 180 and 200 µL. For good signals the concentration of the paramagnetic species studied should be 100 µM or higher.

Samples were prepared in eppendorf tubes and transferred to the EPR tubes with a Hamilton syringe. The preparation varied according to the experiments. Special care was taken to avoid traces of metal ions as they can exhibit EPR signals. All solutions were prepared from distilled water treated with Chelex 100. Prepared samples were stored in liquid N_2 at 77 K.

3.11.4 EPR instrumentation

Standard commercial EPR spectrometers operate in the frequency range of 8.5 to 10 GHz. Spectrometers with this microwave frequency is called X-band spectrometers. Other microwave frequencies that can be used are L-band (1-2 GHz), S-band (2-4 GHz), P-band (~15 GHz), Q-band (~35 GHz), and so-called High Field EPR which includes W-band (~95 GHz) and higher frequencies. Lower frequencies are useful in order to resolve hyperfine interactions (not described here) due to higher precision in the applied magnetic fields, and higher frequencies are useful for finding g-value anisotropies. X-band instruments are most commonly used as they provide a compromise between sensitivity and ease of sample and instrument handling (Palmer, 2000).

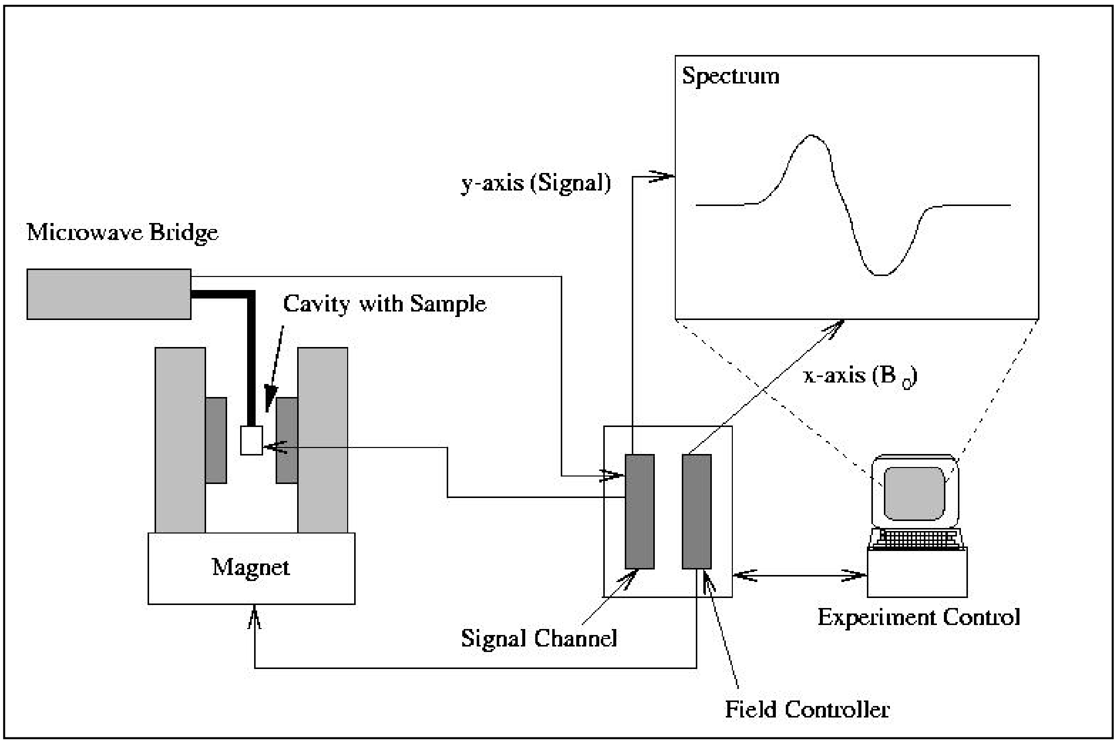

Spectrum
y-axis (Signal)
Microwave Bridge
Cavity with Sample
x-axis (B_0)
Magnet
Signal Channel
Experiment Control
Field Controller

Figure 3.8 Schematic presentation of an EPR instrument. (Figure taken from http://www.phys.chem.ethz.ch/students/esr/de/EPR/Instrumentation.htm)

The spectrometer used in these studies is a BRUKER ESP 300E 10/12 X-band spectrometer with a BRUKER ER4116DM dual mode resonator cavity. An ESR900 Liquid Helium control system from Oxford Instruments allowed temperature regulation of the sample in the range 3.6 K to 100 K.

4 Results

4.1 Cytochrome c-554 from *Methylosinus trichosporium* OB3b

Cytochrome c-554 from *Methylosinus trichosporium* OB3b, named after it optical spectroscopic properties, was probably first observed by Tonge *et al.* in 1977. It was assumed to be a part of the soluble methane monooxygenase enzyme system (Tonge *et al.*, 1977). It was rediscovered by my predecessor Siw Fauchald in 1998 and found to exhibit an unusual ferric HALS (Highly Axial Low-Spin) EPR signal when subjected to low temperature EPR spectroscopy (Fauchald, 2000). This kind of EPR signal has earlier been reported in a similar cytochrome c-552 from the bacterium *Nitrosomonas europaea* (Arciero *et al.*, 1994). The description of more c-type cytochromes exhibiting a HALS-type EPR signal will eventually be helpful in explaining the physical basis for this unusual signal.

In this project, spectroscopic features from the cytochrome c-554 from *Methylosinus trichosporium* OB3b was compared with spectroscopic features exhibited by cytochrome c-552 from *Nitrosomonas europaea* and cytochrome c-553 from *Bacillus pasteurii*. The latter of these two caught our interest as it exhibits NMR-shifts similar to that of cytochrome c-552 from *Nitrosomonas europaea*. They are both good models as their 3D structure have been determined (Timkovich *et al.*, 1998;Benini *et al.*, 2000). A fourth c-type cytochrome from the methane utilizing bacterium *Methylococcus capsulatus* Bath, also purified in our laboratory, was studied by the same means.

In addition to spectroscopic studies, protein sequencing and chemical modifications have been used to learn more about cytochrome c-554 from *Methylosinus trichosporium* OB3b.

The information available about this cytochrome c-554 from *Methylosinus trichosporium* OB3b is limited. In addition to the purification procedure and the HALS EPR spectra, SDS polyacrylamide gel electrophoresis and potentiometric studies had been performed (Fauchald, 2000). From the SDS-PAGE it was known that the molecular weight of the cytochrome was around 13 kDa, and that the purification procedure was seemingly successful. Spectropotentiometric titration performed at pH 7.0 showed a mid-point redox potential higher than 200 mV and indicated haem heterogeneity (Fauchald, 2000).

The HALS EPR spectra had been recorded at pH 8.2 and pH 7.0. At pH 7.0 the EPR signal was more heterogeneous than at pH 8.2 (Fauchald, 2000). This indicates that the sample contains at least two different spin species, which can arise from different haem groups.

The purification procedure includes gel filtration, so any impurities are expected to have similar molecular weight as the cytochrome. Hence, SDS-PAGE is not ideally suited for examination of the purification procedure.

4.2 Protein Purification

Cytochrome c-554 was purified from *Methylosinus trichosporium* OB3b which was cultivated as described in Chapter 3.1. The purification procedure for cytochrome c-554, which is described in Chapter 3.2, has routinely been employed in our laboratory since it was established by my predecessor Siv Fauchald in 1998.

The purity was monitored throughout the purification by watching the ratio $A_{(410\ \mathrm{nm})}$ / $A_{(280\ \mathrm{nm})}$, which is a measure of the haem to protein ratio. At the end of the purification procedure the $A_{(410\ \mathrm{nm})}$ / $A_{(280\ \mathrm{nm})}$ ratio approached 8, and the yield was roughly 0.5 mg from 30 g of cells. The purity and yield was essentially the same for all purifications. Figure 4.1 shows an acrylamide gel of purified protein from different purifications.

The purity of the samples was found satisfactory in all experiments performed.

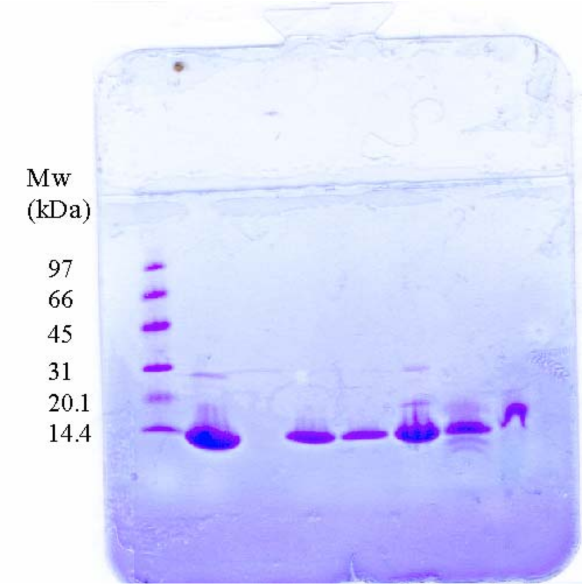


Figure 4.1 SDS-PAGE gel showing the purity of protein samples from six different purifications. The left lane is a low-molecular weight standard from Amersham with the bottom band having a molecular weight of 14.4 kDa.

4.3 Optical spectroscopy

Excitations of the aromatic π electrons in the haem group, which interact with the *d*-electrons on the iron, are responsible for the distinct colours observed in cytochromes. The interaction of the π electrons with the valence electrons of the chelate metal tunes the absorption bands of the haem group making it different from other metal-tetrapyrrole complexes (e.g. chlorophyll (Mg^{2+}) and cobalamine ($Co^{2+/3+}$)). The oxidation state of the iron influences the spectra significantly. This can be used for determination of the redox state of the haem iron.

Individual low-spin c-type cytochromes are traditionally named according to the position of the α-peak maximum of the reduced state. For the c-type cytochrome from *Methylosinus trichosporium* OB3b studied here, this maximum is found at 554 nm (Figure 4.2). Consequently this protein has been named cytochrome c-554, which is the name used throughout this thesis. Similarly, the name of the other cytochromes in this study reflects the absorption wavelength of their α-peak maximum in their reduced state.

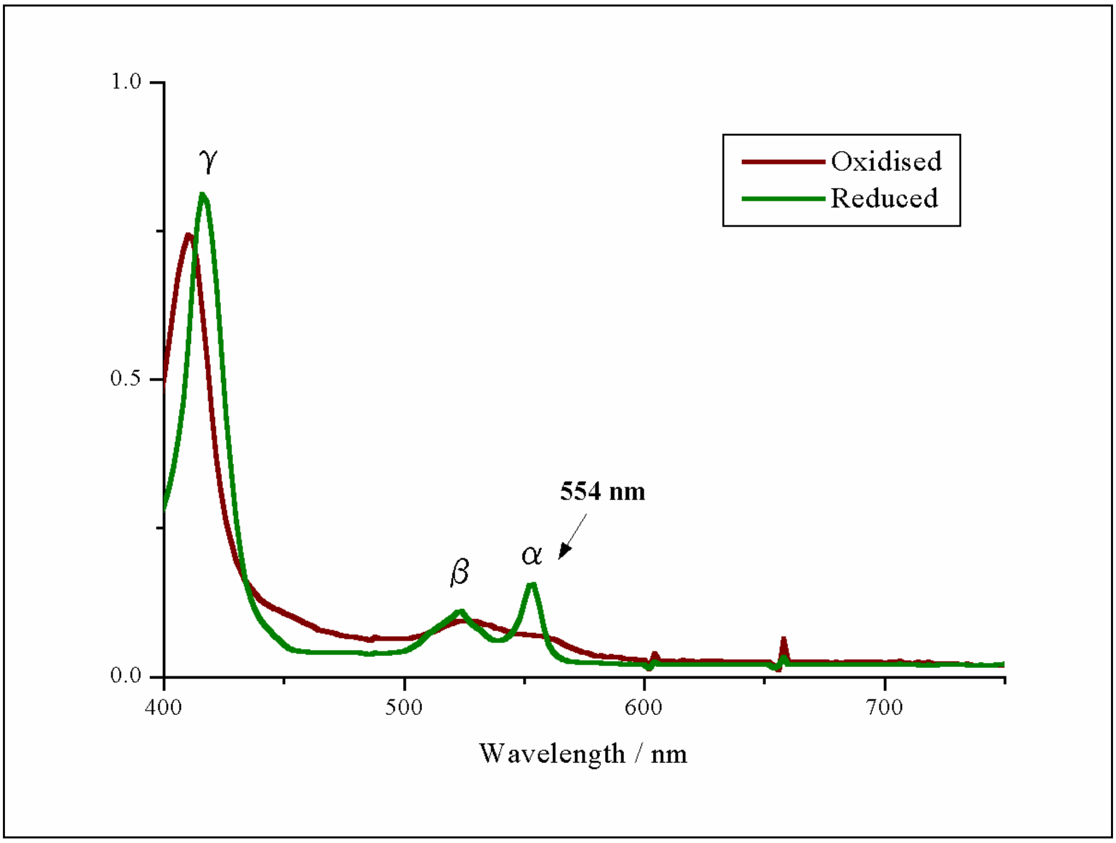


Figure 4.2 Visible absorption spectra of 7 μM solutions of the reduced and oxidised form of cytochrome c-554 from *Methylosinus trichosporium* OB3b. The distinct α-band at 554 nm observed in the reduced spectrum, after which the cytochrome is named, is shown. A shift in the absorbance wavelength of the Soret band (γ-band) is also observed.

Initial concentration measurements were done by turning the cytochrome c-554 into its corresponding pyridine haemochrome. Pyridine haemochromes have distinct optical features that only varies with the nature of the prosthetic haem group (Falk, 1964). The extinction coefficients for haem c pyridine haemochromes are shown in Table 3.2.

The optical spectrum of native cytochrome c-554 from *Methylosinus trichosporium* OB3b and its pyridine haemochrome is shown in Figure 4.3.

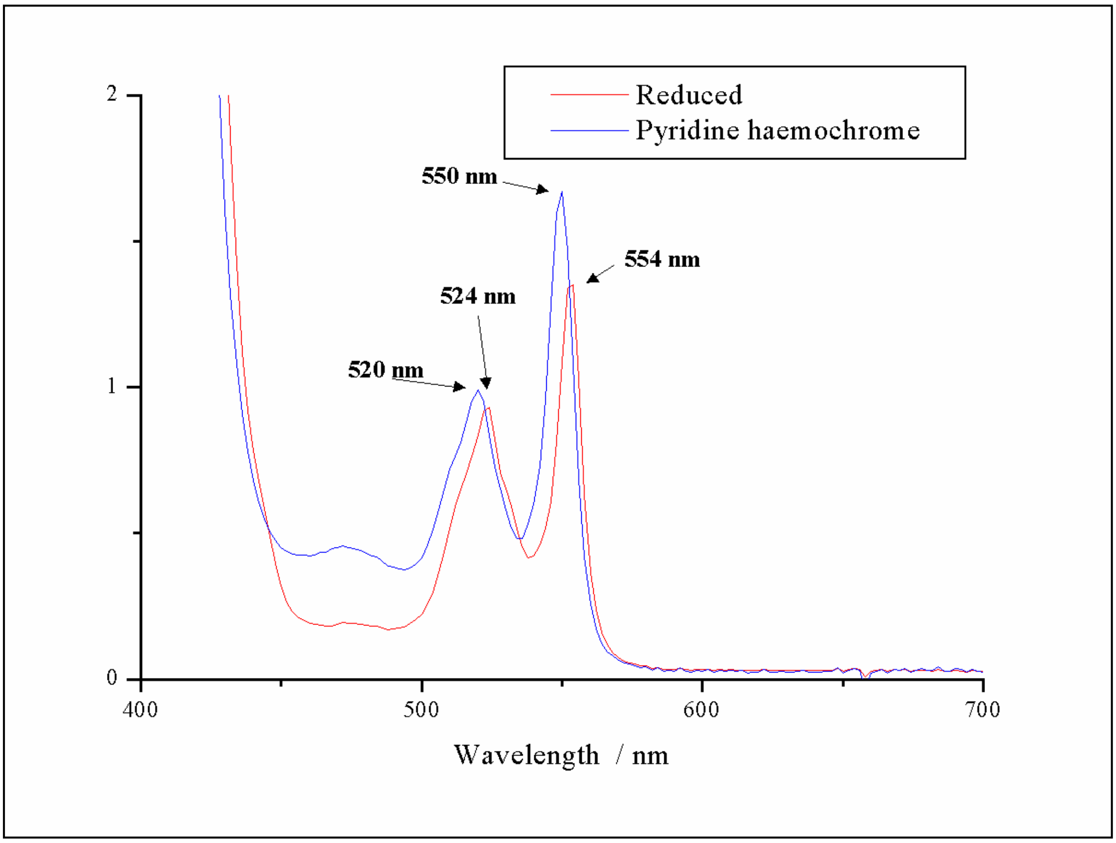


Figure 4.3 Optical absorption spectra of the reduced form of cytochrome c-554 from *Methylosinus trichosporium* OB3b and its corresponding pyridine haemochrome. The pyridine haemochrome spectrum was multiplied by a factor two to adjust for the dilution by the pyridine solution. The concentration was calculated to be 56 μM from the pyridine haemochrome absorption maxima.

The concentration measured for the pyridine haemochrome solution was used to calculate the extinction coefficients for the native protein (Table 4.1 and 4.2). As the spectral features are less distinct in the oxidised form, no wavelength was assigned to the α-band maximum.

Reduced cytochrome c-554	γ maximum	β maximum	α maximum
Wavelength, λ (/nm)	416	524	554
Extinction coefficient, ε (/$mM^{-1}cm^{-1}$)	126.4	16.6	24.1

Table 4-1 The absorption wavelengths and their corresponding extinction coefficients for the reduced form of cytochrome c-554 from *Methylosinus trichosporium* OB3b.

Oxidised cytochrome c-554	γ maximum	β maximum	α maximum
Wavelength, λ (/nm)	410	526	Not assigned
Extinction coefficient, ε (/$mM^{-1}cm^{-1}$)	115.7	14.2	-

Table 4-2 The absorption wavelengths and their corresponding extinction coefficients for the oxidised form of cytochrome c-554 from *Methylosinus trichosporium* OB3b.

The extinction coefficients for the native protein were used for determination of the cytochrome c-554 concentration.

4.3.1 The Alkaline Transition

Studies of the pH-effects on ferric horse heart cytochrome c show that it can exist in five different spectral states linked by protonation events corresponding to four distinct pK values (Theorell and Åkesson, 1941). One of these protonation events, with a pK value in the range of 8.9 to 9.5 depending on solvent conditions, has been termed the alkaline transition. The alkaline transition is associated with the bleaching of a weak absorption band centred at 695 nm. This absorption band has been found universally in both prokaryote and eukaryote ferric cytochromes c that have histidine and methionine as axial ligands.

The loss of this absorption band coincides with the displacement of the methionine ligand. This fact has led to the view that this absorption band is diagnostic of methionine ligation (Wilson and Greenwood, 1996).

The extinction coefficient of the 695 nm absorption band is in the order of 800 M^{-1} cm^{-1}, and it is hence easily overshadowed by the much stronger α-, β-, and γ-bands. Spectroscopic studies on a highly concentrated sample revealed that this absorption band is present in cytochrome c-554 from *Methylosinus trichosporium* OB3b (Figure 4.4). The peaks observed at 600 nm and 655 nm are artefacts produced by the spectrophotometer.

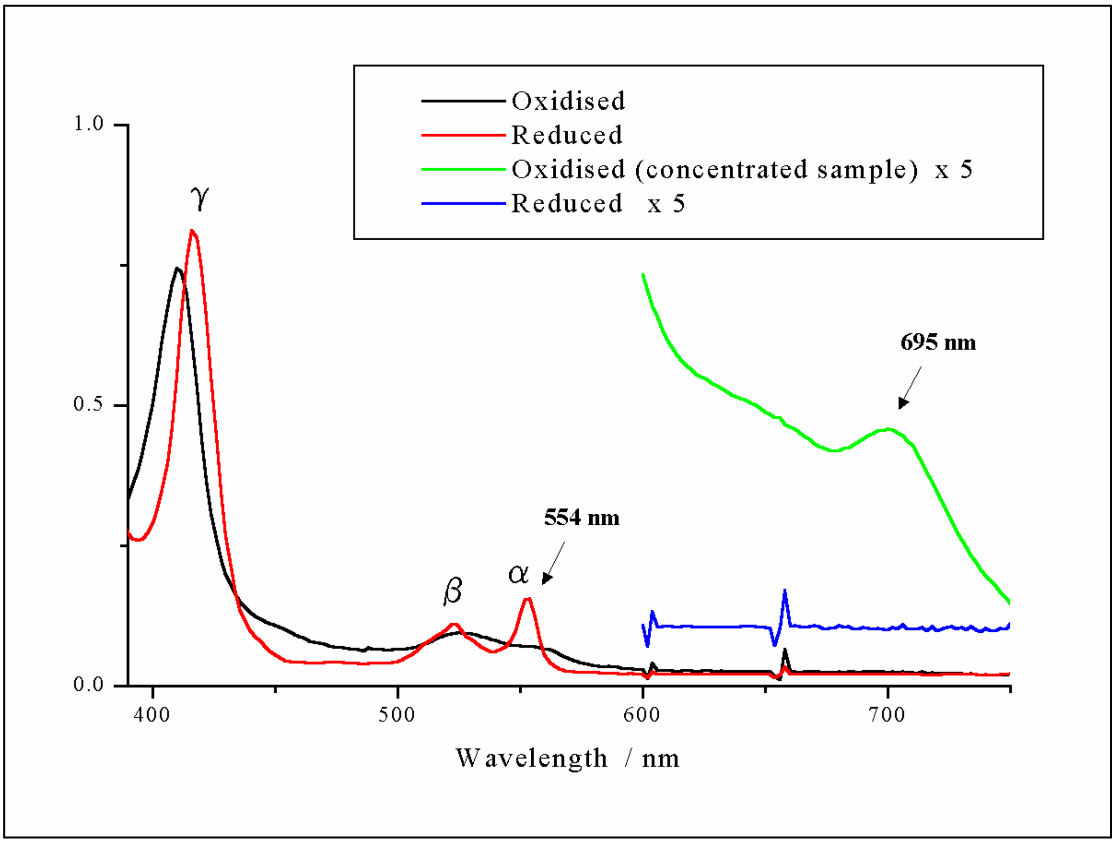


Figure 4.4 Visible absorption spectra of 7 μM solutions of the reduced and oxidised form of cytochrome c-554 from *Methylosinus trichosporium* OB3b at pH 8.2. A magnified section of a spectrum of a 83 μM sample of the ferric (oxidized) form of the protein shows the presence of the 695 nm band. This absorption band is indicative of methionine as the sixth ligand to the haem iron.

4.4 Circular Dichroism (CD)

In eukaryote cytochrome c the optical absorption band at 695 nm has a well-defined circular dichroic (CD) spectrum. This signal is abolished at high pH values similarly to the behaviour of the optical absorption band.
Horse heart cytochrome c has a histidine and a methionine ligated to the haem iron at physiological conditions. The alkaline transition, where the methionine is displaced by a different ligand, is observable in both the optical absorption band at 695 nm and the corresponding peak in the CD spectrum (Greenwood and Wilson, 1971). Figure 4.4 shows how this transition is visible in both the optical absorption spectrum and the CD spectrum of cytochrome c from horse heart. The CD spectrum contains a negative peak that is found at the same wavelength as the absorption band in the optical spectrum. This well defined dichroism has been assigned to the asymmetrical nature of the ligand field about the haem iron (Greenwood and Wilson, 1971).

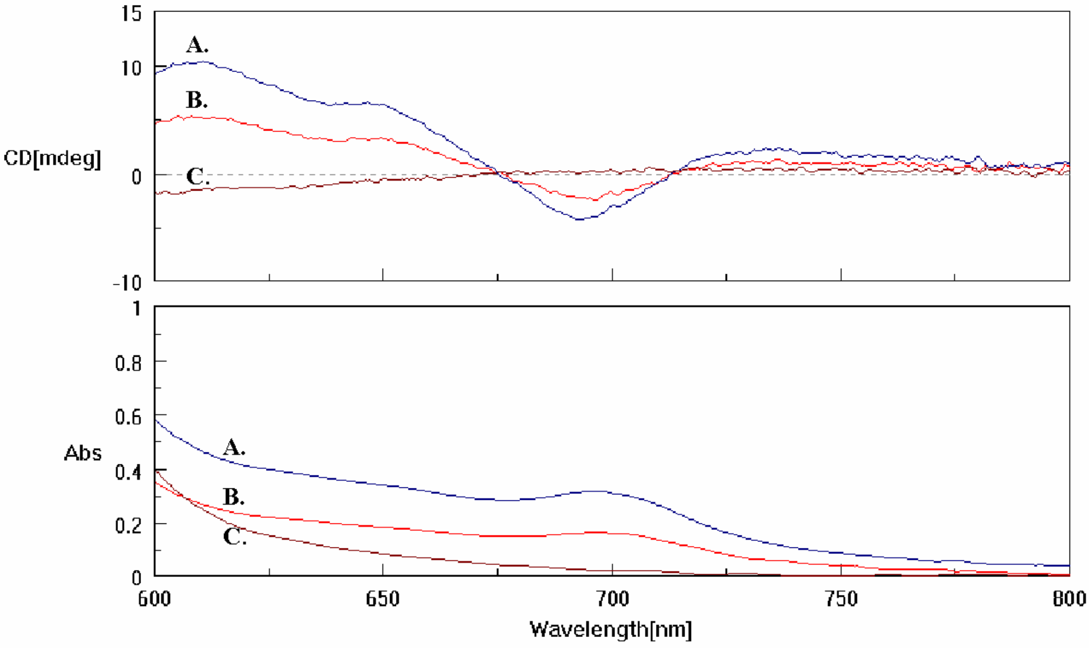


Figure 4.5 Circular Dichroic spectra (top) and optical absorption spectra (bottom) of 0.5 mM solutions of horse heart cytochrome c at pH 7.0 (A.), pH 8.2 (B.), and pH 11.0 (C.)

Cytochrome c-554 from *Methylosinus trichosporium* OB3b and cytochrome c-552 from *Nitrosomonas europaea* are found to both exhibit the optical absorption band at 695nm. The negative CD band corresponding to the 695 nm absorption band is however red-shifted. In Figure 4.6 the maximum absolute amplitude in the CD spectrum of cytochrome c-552 from *Nitrosomonas europaea* is found at 715 nm. The amount of cytochrome c-554 purified from *Methylosinus trichosporium* OB3b was not sufficient to make a highly concentrated CD sample due to the large sample volume needed. Thus the CD spectrum of this protein does not exhibit distinct spectral features. In the region of the 695 nm band it does have a minimum found at 715 nm. The occurrence of this band at a position similar to that of cytochrome c-552 from *Nitrosomonas europaea* suggests that the chromophores exist in comparable asymmetric environments.

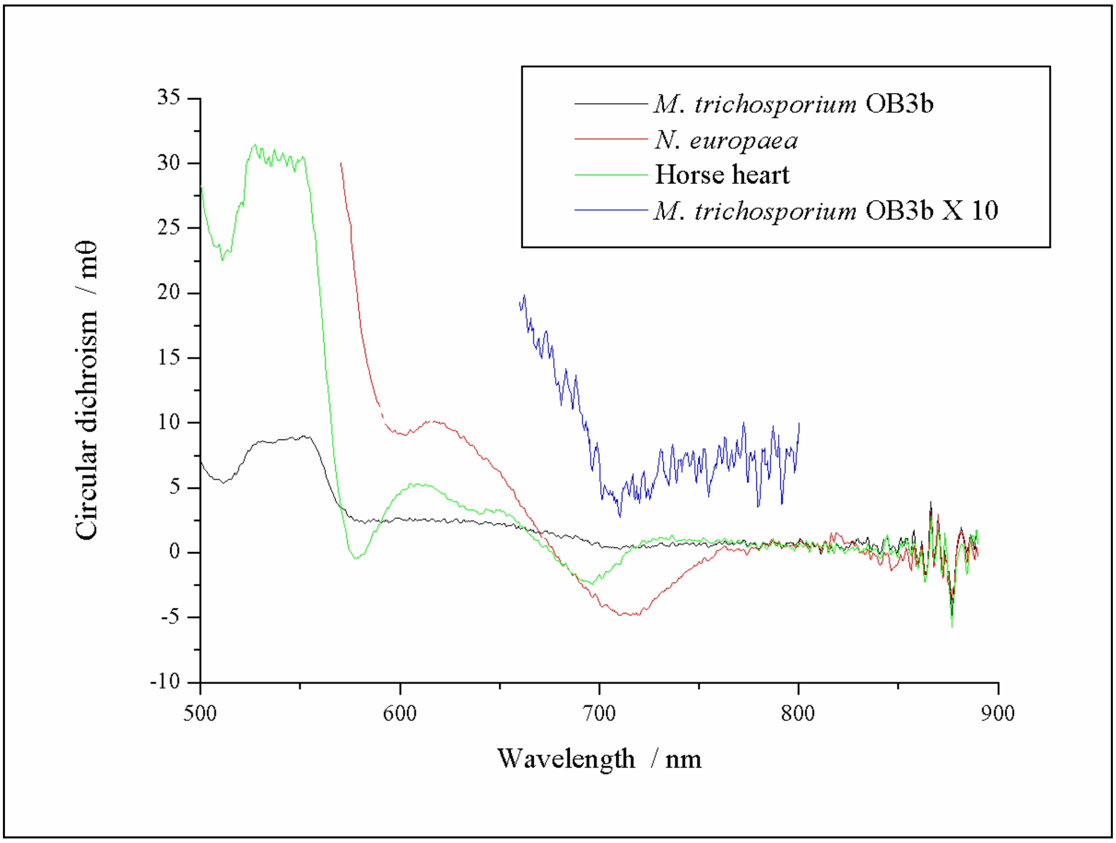


Figure 4.6 Circular Dichroic spectra of cytochrome c from *Methylosinus trichosporium* OB3b (60 μM), *Nitrosomonas europaea* (790 μM), and horse heart (200 μM), at pH 8.2. A part of the *Methylosinus trichosporium* OB3b spectrum is magnified in order to reveal the spectral features in the region of the 695 nm absorption band.

Circular dichroism is for historical reasons commonly expressed as ellipticity, θ. It relates to the more formally correct molar circular dichroism, Δε, as

$$\Delta\varepsilon = 3.298 \cdot l \cdot c \cdot \theta \qquad \textbf{Equation 4.1}$$

where l is the light path length and c the molar concentration.
Ellipticity has been used here as drift in the instrument gave the different spectra different baselines. Correct adjustment of the baseline for the weak signal from cytochrome c-554 was unachievable due to the high signal to noise ratio. The use of ellipticity does not cause any problem, as no quantitative measurements are needed.

Direct calculation of the molar circular dichroism did not result in overlaying spectra, but gave a large offset due to the baseline not being calibrated.

Figure 4.7 shows the molar circular dichroism for *Nitrosomonas europaea* and *Methylosinus trichosporium* OB3b. This confirms that the amplitude of the signal is similar for the two proteins.

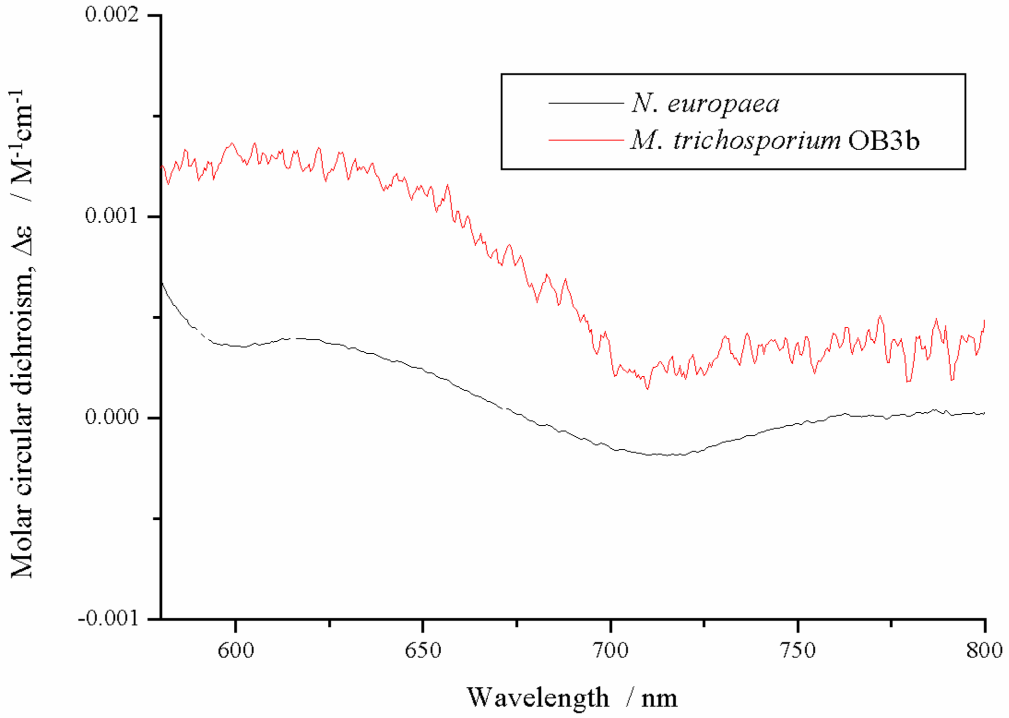


Figure 4.7 Molar circular dichroism spectra for cytochrome c from *Nitrosomonas europaea* and *Methylosinus trichosporium* OB3b. The minimum at 715 nm indicates that chromophores responsible for the 695 nm absorption band exist in comparable asymmetric environments in these two proteins.

4.5 Haem Cleavage

Cytochrome c-554 from *Methylosinus trichosporium* OB3b has earlier been studied by spectropotentiometric analysis and found to have a mid-point potential higher than 200 mV (Fauchald, 2000). This plot did however indicate haem heterogeneity suggesting that the protein contained more than one haem group.
Haem quantitation was done using mass spectrometry of the protein with and without haem. To obtain the haem-free peptide, the thioether bonds linking the haem group to the peptide must be cleaved.
Treatment with 2-nitrophenylsulfenylchloride (2-NPS) cleaves the thioether bonds that attach the haem group to the protein (Figure 3.1) and leaves a 2-NPS molecule attached via a S-S bridge to each cysteine residue in the protein. The only side effect, which is well known and easily accounted for, is modification of tryptophan residues (Figure 3.2). The modification of tryptophan residues by 2-NPS results in a chromophore with an absorption maximum at 365 nm. The extinction coefficient of this chromophore (4.1 $mM^{-1}cm^{-1}$) is significantly lower than those of the haem absorption bands. This chromophore makes the modified polypeptide easily detectable and provides a way of determining the amount of tryptophans in the polypeptide. After the treatment with 2-NPS and desalting (as described in chapter 3.4), three fractions of roughly 500 µl were collected. These fractions were all sufficiently concentrated for analysis by Mass Spectroscopy.

The UV/visible spectrum of the modified protein samples (Figure 4.8) showed an absence of the α-, β- and γ-bands characteristic for haem groups and the expected peak in the UV region around 365 nm from the tryptophan/2-NPS adduct. The extinction coefficient for this chromophore is significantly lower than those observed in haem. Due to the absence of other haem characteristics, this peak cannot be assigned to the δ-band of haem groups. All haem groups where successfully removed from the protein.

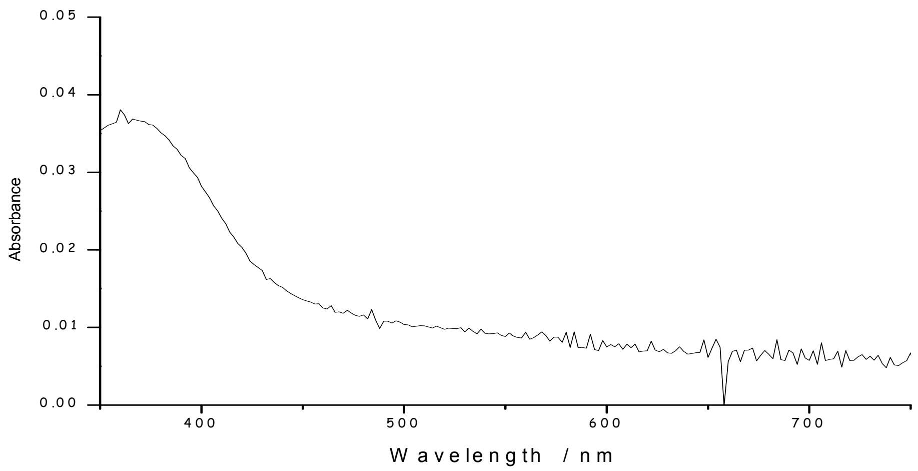


Figure 4.8 Absorption spectrum of the cytochrome c-554 sample after removal of haem using 2-NPS. The concentration of the sample is 4.0 µM (calculated from the absorbance at 365 nm and the assumption that the protein contains two tryptophan residues).

4.6 Mass Spectroscopy

Mass spectra of both the native form of the cytochrome and the modified form lacking haem were recorded as shown in Figure 4.7.

The mass spectrum of the native haem containing form of the protein (Figure 4.9 A) determined the molecular weight to be around 12 230 ± 15 Da. The uncertainty associated with these measurements is claimed to be within 0.05% when calibrated with an external standard. This would lead to an uncertainty of ± 6 Da. The observed variation was slightly higher than this, which can be due to the calibration not being optimised, or to variations in the isotopic distribution between the individual protein molecules.

2-NPS (2-nitrophenylsulfenylchloride) treatment of the sample is known to cleave off haem groups and modify specific amino acid residues (see chapter 3.4). In order to correctly characterize the spectra, other proteins with known primary structure was subjected to the same treatment with 2-NPS and mass spectra of both native protein and modified protein were recorded. These proteins were myoglobin from horse skeletal muscle and cytochrome c from horse heart. The general information obtained was that both cysteine and tryptophan residues were modified but none of them to completion. The samples were treated with mercaptoethanol which reduced the level of cysteine/2-NPS adducts.

The spectrum of the 2-NPS treated protein (Figure 4.9 B) contains several peaks corresponding to protein molecules with different number of 2-NPS molecules covalently attached to tryptophan and cysteine residues. The relative magnitude of the peaks varied from sample to sample, and it varied greatly with the supporting matrix used in MALDI-TOF MS (see chapter 3.8). Thus, no definite quantitative conclusions can be made by comparing the height of the peaks.

The molecular weight of a haem group is 615 Da and the molecular weight gain of each modification with a single 2-NPS molecule is 154 Da.

The mass spectrum of 2-NPS treated protein (Figure 4.9 B.) does not contain a peak corresponding to the native haem containing protein which is in good accordance with the expectations from the visible spectrum (Figure 4.8). The lowest obtained molecular weight in this mass spectrum is 11 630 Da, which corresponds to an intact protein only missing one haem group. The other observed peaks correspond to peptide chains with different number of 2-NPS modified amino acid residues.

In all samples a low peak was observed with higher m/z ratio than the other peaks. This artefact, which is observed in all samples, can be explained as a singly charged sample molecule in complex with a molecule from the supporting matrix.

The absence of any molecules with mass less than 600 Da below the molecular weight of the native enzyme gives a strong indication that this protein only contains one haem group. Furthermore, the five peaks indicate that the protein contains four amino acid residues that are possible targets for modification with the reagent 2-NPS. These residues are believed to be two tryptophans and two cysteines. The protein is known to contain at least two cysteine residues, which are necessary for covalent attachment of the haem group to the protein. The chromophore detected after haem removal (Figure 4.6) confirms the presence of tryptophan. The suggestion that two of the modifiable residues are tryptophans is based on comparisons with mass spectra of proteins with known primary structure that was treated identically with 2-NPS.

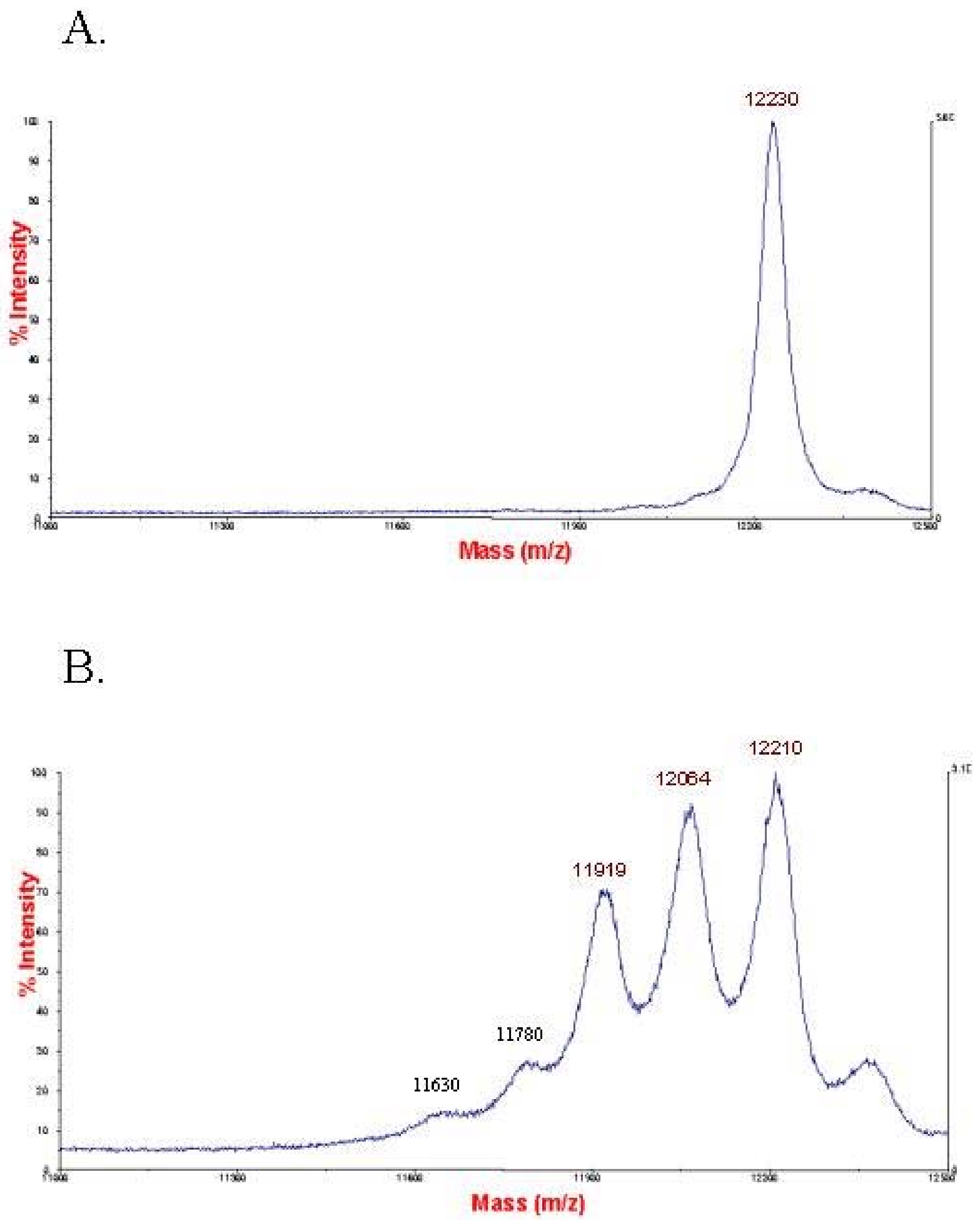


Figure 4.9 A. Mass spectrum of the native cytochrome c-554 from *Methylosinus trichosporium* OB3b. B. Mass spectrum of cytochrome c-554 after treatment with 2-NPS. The number of peaks are due to proteins with different number of added 2-NPS molecules.

4.7 EPR studies

4.7.1 The Highly Axial Low-Spin (HALS) EPR Signal

Cytochrome c-554 from *Methylosinus trichosporium* OB3b exhibits a ferric HALS (Highly Axial Low-Spin or Highly Anisotropic Low-Spin) EPR signal when subjected to low temperature EPR spectroscopy (Figure 4.10 C). Low-spin ferric c-type cytochromes normally exhibit rhombic EPR signals (see Figure 3.5 and 4.10 D). HALS EPR signals are unusual and are only observed in a small number of very diverse cytochromes (Arciero *et al.*, 1994). HALS EPR signals are also known as "large g-max" signals. The main characteristic of these signals is a g-value higher than 3.3 (Walker, 1999). The physical basis for this signal is not completely understood.

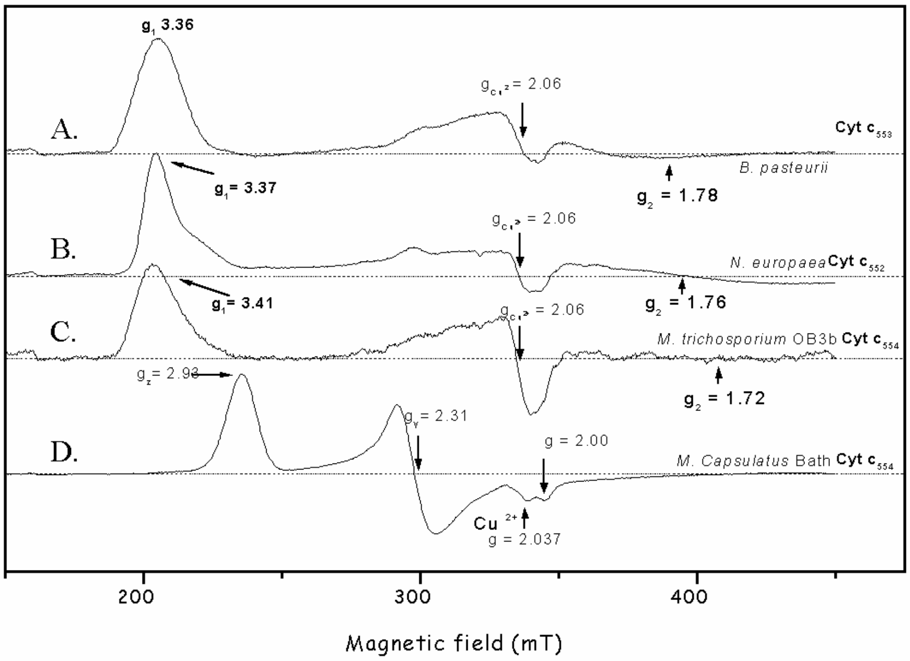


Figure 4.10 EPR spectra of four c-type cytochromes recorded at 10 K: A. Cytochrome c-553 from *Bacillus pasteurii* exhibiting HALS lineshape, B. cytochrome c-552 from *Nitrosomonas europaea* exhibiting HALS lineshape, C. Cytochrome c-554 from *Methylosinus trichosporium* OB3b exhibiting HALS lineshape, D. Cytochrome c-554 from *Methylococcus capsulatus* Bath exhibiting a rhombic lineshape. The spectral features at g=2.06 arise from contaminating Cu^{2+}.

Direct assignment of g-values other than large g-max (g_Z) is not possible for cytochrome c-554 from *Methylosinus trichosporium* OB3b. The integrated form of the HALS EPR spectrum (i.e. the microwave absorption spectrum) is weak and broad. The absorption amplitude decays slowly and extends to high magnetic fields (not shown). This results in a weak slope, and hence small amplitudes in the 1st derivative spectrum. The g-values other than the g-max are therefore hard to assign to spectral features (Salerno and Leigh, 1984). The g-values of low-spin haems, which exhibit large g-max values, have been proposed to follow the general rule (Walker, 1999)

$$g_x^2 + g_y^2 + g_z^2 = 16 \qquad \text{Equation 4.1}$$

This empirical rule has been widely used and has recently been shown to be applicable to systems having a fairly pure $(d_{xy})^2(d_{xz},d_{yz})^3$ electron configuration. This is indeed the predominant electron configuration found in model compounds that exhibit HALS (large *g*-max) EPR spectra (Walker, 1999).

The intermediate g-value of cytochrome c-554 from *Methylosinus trichosporium* OB3b can be estimated to be 1.72 based on the g-max value and comparisons with other HALS EPR spectra (Harbitz *et al.*, 2001). The linewidth of this spectral feature is very broad; hence, the g-value determined is associated with a high uncertainty. Using this estimation it is possible to get an approximation of the third g-value by using Equation 4.1. The g-values in the HALS EPR cytochrome c-554 from *Methylosinus trichosporium* OB3b are:

$g_z = 3.41 \qquad g_x = 1.72 \qquad g_y = 1.19$

For the two other cytochromes exhibiting HALS EPR signals, assignment of both the large g-max and the intermediate field feature was possible.

	g_z	g_x	g_y	v/λ	Δ/λ	v/Δ
Cytochrome c-552 *B. pasteurii*	**3.36***	**1.94***	**0.97**	1.00	3.16	0.32
Cytochrome c-553 *N. europaea*	**3.34***	**1.77***	**1.31**	1.13	6.94	0.16
Cytochrome c-554 *M. trichospurium* OB3b	**3.41***	**1.72**	**1.19**	1.01	6.19	0.16
Cytochrome c-554 *M. capsulatus* Bath	**2.93***	**2.31***	**1.48***	1.87	3.09	0.61

Table 4-3 The g-values and the correlation between axial and rhombic ligand field parameters for the low-spin haem centres derived from EPR measurements performed at 10 K (Harbitz *et al.*, 2001). * These g-values were taken directly from the respective EPR spectra.

Table 4.3 shows the observed and estimated g-values for all four cytochromes, and the correlation between axial and rhombic ligand field parameters. The physical meaning of the ligand field parameters can be understood by comparison with Figure 3.7. The parameter Δ/λ measures the strength of the axial field in units of the spin-orbit coupling constant. The parameter v/Δ measures the ratio of the rhombic to axial components. These two ratios can be calculated directly from the g-values (Walker, 1999).

The v/Δvalue for cytochrome c-554 from *Methylosinus trichosporium* OB3b is relatively small and similar to that of cytochrome c-552 from *Nitrosomonas europaea*. This low value indicates that these systems are nearly axial. Cytochrome c-553 from *Bacillus pasteurii* exhibits a higher v/Δ value indicating that it is subjected to a stronger rhombic distortion compared with the other two. Cytochrome c-554 from *Methylococcus capsulatus* Bath exhibits a normal rhombic EPR spectrum, which is reflected in these parameters.

4.7.2 pH Dependence of the HALS EPR Signal

The purification of cytochrome c-554 from *Methylosinus trichosporium* OB3b is performed at pH 8.2, and all the EPR samples was prepared at this pH. Ferric cytochrome c is known not to exhibit any optical spectroscopic changes in this pH interval (Wilson and Greenwood, 1996), but an earlier study suggested that this HALS EPR signal was more heterogeneous at pH 7.0 than at pH 8.2 (Fauchald, 2000).

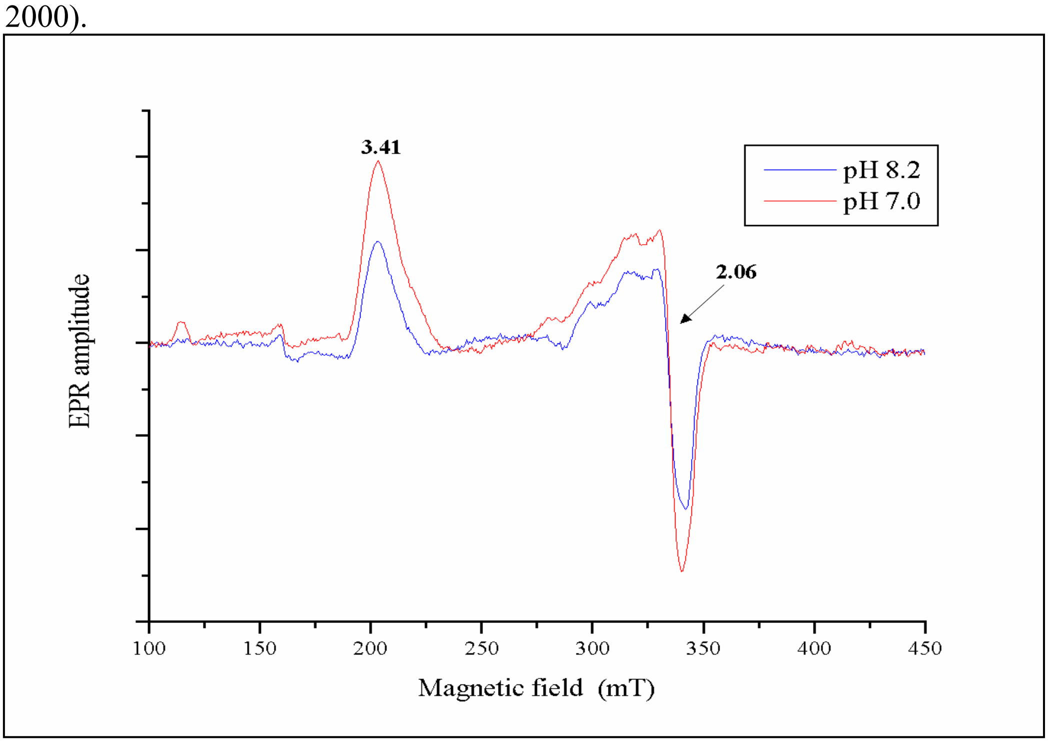


Figure 4.11 EPR spectra of cytochrome c-554 from *Methylosinus trichosporium* OB3b at pH 7.0 and pH 8.2.

EPR spectra of cytochrome c-554 from *Methylosinus trichosporium* OB3b at pH 7.0 and pH 8.2 is shown in Figure 4.11. Both are recorded at 10 K and a microwave power of 8 mW.

The spectral characteristics, i.e. the g-values, are identical in these two spectra. The only observable difference is the width of the "HALS peak". At pH 7.0 the peak has a weak shoulder that is not present at pH 8.2 is visible on the high field slope (Figure 4.12).

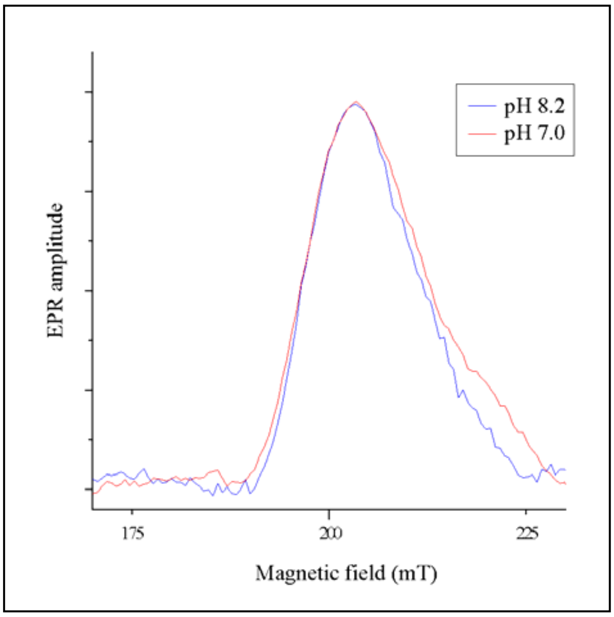


Figure 4.12 EPR spectra of cytochrome c-554 from at pH 7.0 and pH 8.2. The amplitude of the pH 8.2 spectrum is magnified to fit with the pH 7.0 spectrum.

4.7.3

Horse Heart Cytochrome c and the Alkaline Transition

Mammalian cytochrome c is known to exhibit rhombic EPR spectra at physiological pH (Lambeth *et al.*, 1973). The displacement of the methionine ligand that occurs at the alkaline transition (see Chapter 4.3.1), changes the EPR spectrum. At pH 7.0 horse heart cytochrome c (purchased from Sigma Aldrich) has a typical rhombic lineshape with a g-max value of 3.04. The EPR spectrum at pH 11.0 is made up of at least two different EPR states (Figure 4.13). The large g-max value at 3.37 in the EPR spectrum of cytochrome c at high pH is part of a HALS signal.

The identity of the replacement ligand in the alkaline form is not known, but lysine and arginine residues have been pointed out as likely candidates (Mauk and Scott, 1996).

At pH 11.0 the spectrum is more complex and consist of several different species. The g-value of 3.15 is a low-spin signal that is not a HALS signal. One HALS signal is found at g = 3.37, and another HALS signal is observed as a shoulder at a higher g-value (lower magnetic field). The feature at g = 4.3 corresponds to free Fe^{3+} in solution (Fee, 1978).

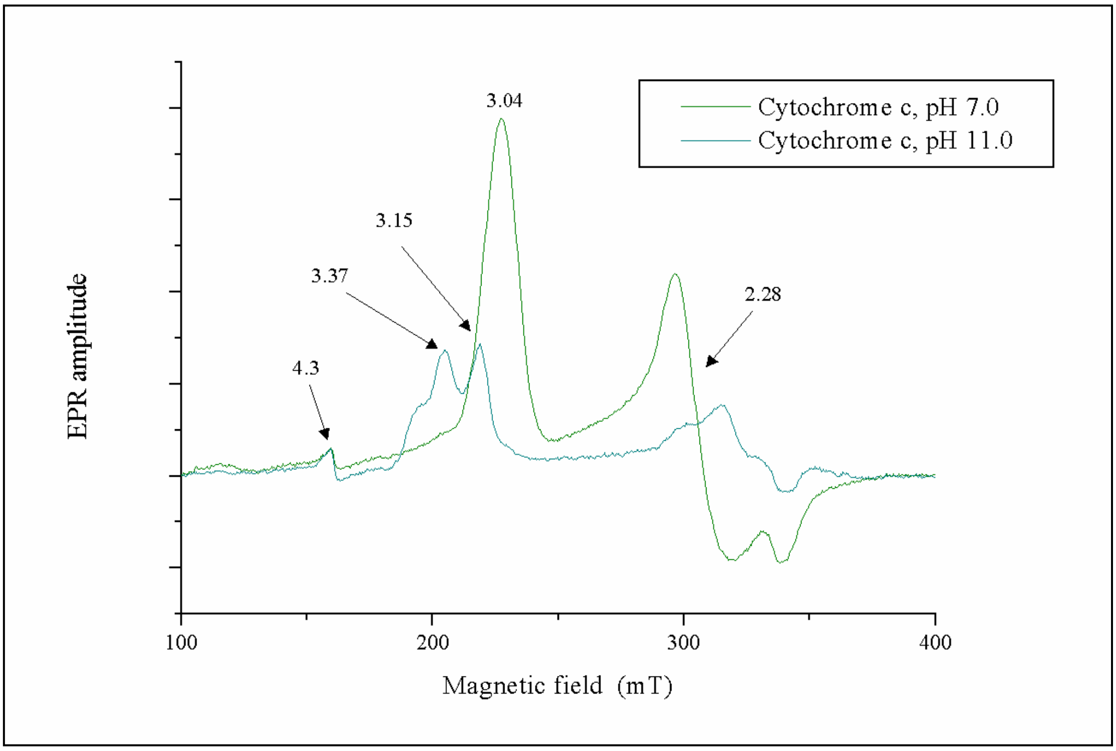


Figure 4.13 EPR spectra of horse heart cytochrome c at pH 7.0 and pH 11.0. At pH 7.0 the lineshape is rhombic with a g-max value of 3.04. At pH 11.0 the spectrum consists of at least two different low-spin states, one of which exhibits HALS lineshape (see text).

4.7.4 Microwave Power Saturation

An EPR signal is saturated when no microwave absorption can be observed. This can occur if the resonant energy states are evenly populated or if the relaxation rates of the unpaired electron is slow compared with the excitation, enough to affect the overall population difference. The latter phenomenon results in an imbalance between the excitation and relaxation, i.e. the excitation occurs faster than the relaxation is able to compensate, and the equilibrium population difference will be altered. The manifestation of this is that the EPR signal no longer behaves as an exponential function of the applied microwave power.

Microwave power saturation at a fixed temperature is used to determine the saturation midpoint ($P_{½}$) at that temperature of the paramagnetic system studied. Saturation midpoint values of several different species obtained at the same temperature can be directly compared revealing differences in relaxation properties (Sahlin *et al.*, 1986). The saturation midpoint values ($P_{½}$) is determined using this empirical equation:

$$\frac{\iint \text{EPR 1st derivative spectrum}}{\sqrt{\text{P}}} = \frac{\text{K}}{\left(1+\left(\frac{\text{P}}{\text{P}_{½}}\right)\right)^{\frac{b}{2}}} \qquad \text{Equation 4.2}$$

where K is a normalisation factor, P the variable microwave power, and b an inhomogeneity parameter. The b factor is fixed to 1 when the dominating lineshape is Gaussian, or 3 if the dominating lineshape is Lorentzian. Intermediate values of b ($3 \geq b \geq 1$) are allowed when a mixture of lineshapes is observed. A Lorentzian lineshape indicates that the relaxation is purely spin-lattice relaxation, whereas a Gaussian lineshape indicates an equal amount of spin-lattice and spin-spin relaxation.

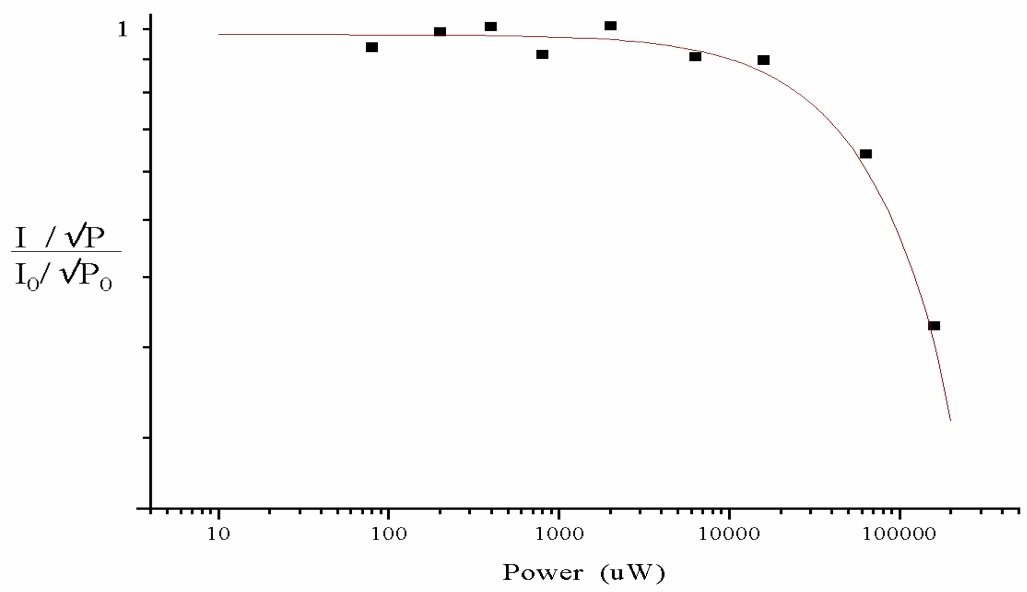


Figure 4.14 Power saturation behaviour of the HALS EPR signal from Cytochrome c-554 from *Methylosinus trichosporium* OB3b recorded at 10 K. The line is the curve fit using equation 4.2, from which the $P_{½}$ value was determined.

When determining the $P_{½}$ values b was fixed to 1 (Gaussian lineshape). The reason for locking the b-value to a distinct lineshape is to avoid mathematical variations caused by experimental error or the nature of the semi-empirical equation used to determine the $P_{½}$ values.

The $P_{½}$ value determined for the HALS EPR signal from cytochrome c-554 from *Methylosinus trichosporium* OB3b (using Equation 4.2) was 67 ± 7 mW at 10 K. This large $P_{½}$-value implies that this HALS EPR signal is difficult to saturate.

All the studied cytochromes that exhibit HALS EPR spectra show $P_{½}$ values above 40 mW at 10 K, indicating that high saturation midpoint may be a general feature of the HALS EPR signal. The c-type cytochrome from *Methylococcus capsulatus* Bath exhibits a normal rhombic EPR signal, which is more easily saturated.

Methylosinus trichosporium OB3b cytochrome c-554	67 ± 7 mW
Bacillus pasteurii cytochrome c-553	55 ± 5 mW
Nitrosomonas europaea cytochrome c-552	42 ± 5 mW
Methylococcus capsulatus Bath cytochrome c-554	37 ± 2 mW

Table 4-4 Saturation midpoint for cytochromes with HALS EPR signal. The power saturation studies were performed at 10 K.

4.7.5 Temperature dependence

As stated by the Boltzmann distribution (Equation 3.4), the population of energy levels is a direct function of the energy difference between the two energy levels and the temperature. At low temperatures the population of electrons will predominantly be in the most favourable energetic state. Increasing temperature will lead to a situation where the energy levels are more equally populated. Such a decrease in the difference between the populations will lead to weaker EPR signals. HALS EPR signals are known to be observable only below approximately 30 K (Walker, 1999). EPR spectra of the “g-max” signal were recorded at seven different temperatures ranging from 7 K to 36 K with all other instrumental parameters kept constant. At 36 K the EPR signal was barely visible.

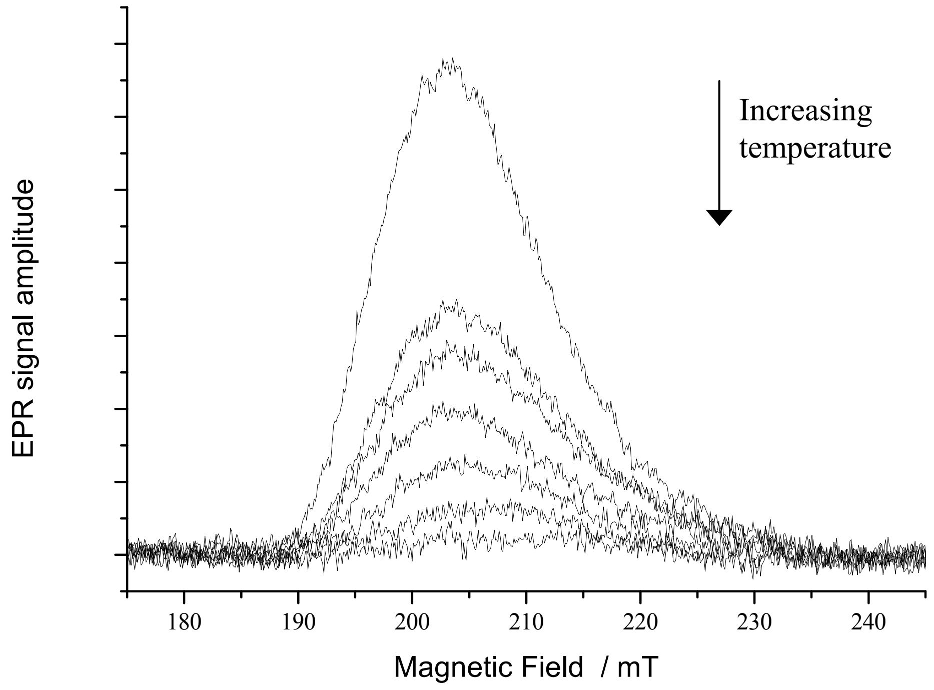


Figure 4.15 Temperature dependence of the HALS EPR spectrum. The spectra were recorded at 800 μW and the temperatures were 7, 13, 16, 21, 26, 31 and 36 respectively. Higher temperatures give weaker EPR signals.

Whereas the rhombic EPR spectrum normally seen in c-type cytochromes can be observed at 77 K (liquid N_2), the HALS EPR signal observed in cytochrome c-554 from *Methylosinus trichosporium* OB3b disappears at much lower temperatures. This is similar to the temperature behaviour of the other HALS EPR signals studied and consistent with the former notion that these signals are not readily observed above 20 K (Walker, 1999).

The magnetisation of a paramagnetic system is temperature dependent and follows the Curie-Weiss law. This law states that the magnetisation of a paramagnet is proportional to the inverse of the absolute temperature (Girerd and Journaux, 2000). Figure 4.16 shows how the double integral of the EPR signal varies with the inverse absolute temperature. The linear behaviour shows that the magnetisation follows the Curie-Weiss law in this temperature interval. This paramagnetic system behaves like an ideal paramagnet in the temperature range 7 K to 36 K.

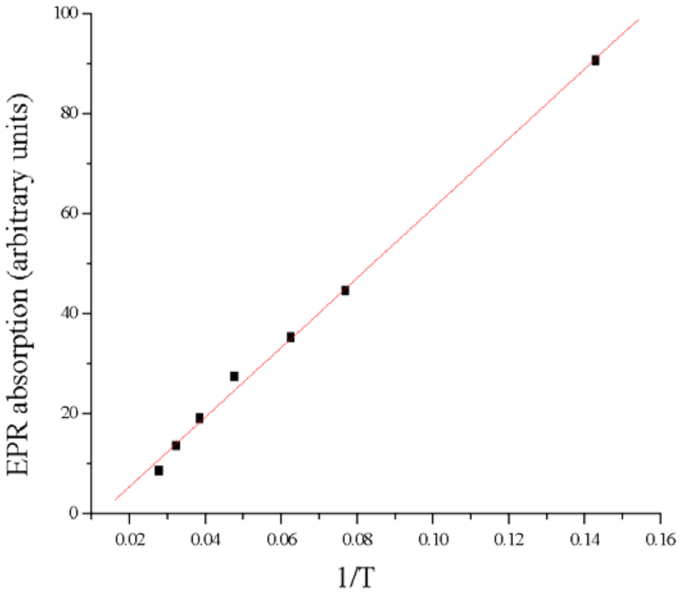


Figure 4.16 Temperature dependence of the EPR signal from cytochrome c-554 plotted as the double integral of the spectrum versus the inverse absolute temperature. The linear behaviour confirms that this system obeys the Curie-Weiss law in this temperature interval (7 – 36 K).

4.8 Amino acid sequencing

Cytochrome c-554 from *Methylosinus trichosporium* OB3b is expected to consist of 100-110 amino acids. This is based on the determined mass of the peptide, which is 11 630 Da, and an average molecular weight of the amino acids of around 119 Da (Creighton, 1993). This is not a direct average but weighted by the amino acids frequency of occurrence in globular proteins.
Sequencing of cytochrome c-554 was performed in two stages. Firstly, the native protein was sequenced from the N-terminal end using Edman degradation (Chapter 3.5). Secondly, a mixture of fragments from cyanogen bromide (CNBr) cleavage was sequenced. Cyanogen bromide cleaves specifically on the C-terminal side of methionine residues. This results in fragments where the N-terminal residue either is the N-terminal residue from the native protein, or was connected through a peptide bond to the carboxyl group of a methionine. This mixture of fragments was sequenced using Edman degradation primarily to determine the number of fragments produced.

The N-terminal sequencing of the native protein did successfully determine the first 39 amino acids. At this point, the automatic sequencer was turned off leaving a gap of five amino acids, before another eight residues could be determined. The following sequence was obtained:

```
1  Ala - Gly - Asp - Pro - Ala - Ala - Gly - Glu - Lys -
10 Val - Phe - Asn - Lys -  X  - Lys - Ala -  X  - His -
19 Gln - Val - Gly - Glu - Thr - Ala - Lys - Asn - Ala -
28 Val - Ala - Pro - Glu - Leu - Asn - Gly - Ile - Asp -
37 Gly - Arg - Lys -  X  -  X  -  X  -  X  -  X  - (A/K)-
46 Met - Ile - Phe - Pro - Gly - Leu - Ser -
```

Figure 4.17 The N-terminal amino acid sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b.

The sequence contains a characteristic haem-binding motif (underlined). This is a well-known sequence found in all c-type cytochromes (Ambler, 1991). The haem-binding motif is known to normally occur near the N-terminal end of the primary sequence in small globular Type I cytochrome c's. The cysteines in this haem-binding motif are covalently bound to the prosthetic haem group and will not yield free cysteines residues when cleaved by Edman degradation. No amino acids were detected in position 14 and 17 but they were assumed to be the conserved haem-binding cysteines as it is likely that these cysteines are not detected. The first part of the sequence is associated with little uncertainty and all other positions in this part of the sequence had well-defined amino acids.

The sequencing of the mixture of peptides obtained after the cyanogen bromide cleavage resulted in high levels of two amino acids for each step of the sequencing procedure. This indicates that the mixture consists of two polypeptide chains and suggests that the protein contains only one methionine residue at which the peptide was cleaved.

Subtracting the native sequence from this doublet of sequences, lead to a unique, well defined sequence for the first 27 amino acids of the second polypeptide chain. The following sequence was found:

```
1  Ile - Phe - Pro - Gly - Leu - Ser - Ser - Glu - Asn -
10 Asp - Gln - Ala - Asn - Val - Trp - Ala - Tyr - Leu -
19 Ser - Gln - Phe - Gly - Ala - Asp - Gly - Lys - Lys -
```

Figure 4.18 The amino acid sequence of the second fragment obtained after cyanogen bromide cleavage of cytochrome c-554 from *Methylosinus trichosporium* OB3b. Shaded residues were poorly defined and are associated with some uncertainty.

The first part of this sequence is identical to the latter part of the first sequence proving that the protein has been cleaved after the methionine in position 46.

```
1  Ala - Gly - Asp - Pro - Ala - Ala - Gly - Glu - Lys -
10 Val - Phe - Asn - Lys - Cys - Lys - Ala - Cys - His -
19 Gln - Val - Gly - Glu - Thr - Ala - Lys - Asn - Ala -
28 Val - Ala - Pro - Glu - Leu - Asn - Gly - Ile - Asp -
37 Gly - Arg - Lys -  X  -  X  -  X  -  X  -  X  - (A/K)-
46 Met - Ile - Phe - Pro - Gly - Leu - Ser - Ser - Glu -
55 Asn - Asp - Gln - Ala - Asn - Val - Trp - Ala - Tyr -
64 Leu - Ser - Gln - Phe - Gly - Ala - Asp - Gly - Lys -
73 Lys -
```

Figure 4.19 N-terminal sequence of cytochrome c from *Methylosinus trichosporium* OB3b obtained by combining the results from Edman degradation of native protein with the results from Edman degradation of the cyanogen bromide treated protein. The underlined region marks the haem-binding motif. Shaded residues were poorly defined and are associated with some uncertainty.

4.8.1 Sequencing of Cytochrome c from *Methylococcus capsulatus* Bath

Cytochrome c from *Methylococcus capsulatus* Bath has lower molecular weight (~ 11 000 Da) than cytochrome c-554 from *Methylosinus trichosporium* OB3b, and was expected to consist of 95-105 amino acids.
The N-terminal sequence of this protein is shown in Figure 4.20.

```
1  Ala - Pro - Val - Asp - Gln - Ala - Thr - Tyr - Asn -
10 Gly - Phe - Lys - Ile -
```

Figure 4.20 N-terminal sequence of cytochrome c from *Methylococcus capsulatus* Bath.

After 13 amino acids the sequencing was stopped because it was recognised as an earlier reported sequence (Ambler *et al.*, 1986). This successfully determined the protein to be cytochrome c-555, which consists 96 amino acid residues.

4.9 Sequence Alignment

Sequence homologies do indicate evolutionary relationship as well as structural similarity (Chothia and Lesk, 1986). For long sequences an identity greater than 20% probably indicates homology (Creighton, 1993). In homologue proteins (i.e. evolutionary related proteins) structure is known to be more conserved than amino acid sequence (Salemme, 1977;Creighton, 1993).
Homologue sequences to the partial sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b was found using the program BLAST. Searches for similar sequences were performed against a protein sequence database, a structural database and the genomes of *Methylococcus capsulatus* Bath and *Nitrosomonas europaea*.

4.9.1 Protein Sequence Database

The protein sequence database used was a combination of all non-redundant GenBank CDS (Coding Domain Sequence) translations and the four protein databases PDB, SwissProt, PIR and PRF. When the search was performed, this database consisted of 1,229,457 sequences, consisting of 391,140,982 amino acids.
When executing the search, position 45 was left open (i.e. as an X in the sequence), as it was not uniquely defined. The sequence alignment tool retrieved 101 sequences, which produced significant alignments from the database. All matching sequences corresponded to c-type cytochromes or hypothetical proteins. Best alignment was found with a cytochrome *c_2* from *Rhodopseudomonas acidophila*, in which 44 out of 107 amino acids (41%) was identical and 7 more showed positive correlations. Such an extent of identical amino acids indicates that these proteins are homologues.

Cytochrome c-554 from *Methylosinus trichosporium* OB3b:

```
1  AGDPAAGEKVFNKCKACHQVGETAKNAVAPELNGIDGR------------- 38
   AGDP AG+KVF KC ACH++G  AKN V P LNG+  R
1  AGDPDAGQKVFLKCAACHKIGPGAKNGVGPSLNGVANRKAGQAEGFAYSDA 51
```

Cytochrome c_2 from *Rhodopseudomonas acidophila*

Cytochrome c-554 from *Methylosinus trichosporium* OB3b:

```
39 ---------------------KXXXXXXMIFPGLSSENDQANVWAYLSQF  67
                        K      M FPGL +E D+ N+WAYLSQF
52 NKNSGLTWDEATFKEYITAPQKKVPGTKMTFPGLPNEADRDNIWAYLSQF 101
```

Cytochrome c_2 from *Rhodopseudomonas acidophila*

Cytochrome c-554 from *Methylosinus trichosporium* OB3b:

```
68  GADGKK 73
     ADG K
102 KADGSK 107
```

Cytochrome c_2 from *Rhodopseudomonas acidophila*

Figure 4.21 Sequence alignment between cytochrome c-554 from *Methylosinus trichosporium* OB3b and cytochrome c_2 *Rhodopseudomonas acidophila* based on sequence databases.

As the query sequence only consisted of 67 amino acids and the rest of the sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b is unknown, the identity might be much higher than 41% . This might not be case as the unknown part of the sequence from cytochrome c-554 from *Methylosinus trichosporium* OB3b is at the C-terminal end of that protein, whereas the known sequence aligns well with the C-terminal part of cytochrome c_2 from *Rhodopseudomonas acidophila*.

The gap put into the cytochrome c-554 sequence to get best possible alignment with the cytochrome c_2 from *Rhodopseudomonas acidophila* is close to the six amino acids which where lost during the first part of the sequencing procedure. The uncertainty in the number of lost amino acids is said to be small, and cannot account for the large gap observed.

The complete sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b needs to be determined in order to establish the total degree of identity between these cytochromes.

4.9.2 Structural Database

3D-structure is known to be more conserved than the primary sequence in homologue proteins (Salemme, 1977;Creighton, 1993). Thus proteins with large sequence similarities are expected to have similar structures. Structural databases are more limited than sequence databases, and the three dimensional structure cytochrome c_2 from *Rhodopseudomonas acidophila*, which gave the best alignment from the sequence database, is not solved.

The structural database used was the Protein Data Bank (PDB), which consists of the primary sequence of all protein structures solved (Berman *et al.*, 2000). When the search was performed, this database contained 36,638 sequences, consisting of 8,548,707 amino acids.

When executing the search, position 45 was left open (i.e. as an X in the sequence). pBLAST retrieved 51 matches, which produced significant alignments. These matches all corresponded to c-type cytochromes.

The c-type cytochromes that showed significant alignments came from a diverse set of organisms, including plants (rice), mammals (horse), fish (tuna), yeast and bacteria. These findings should not be surprising as c-type cytochromes are known to be well conserved between different species. The endosymbiont hypothesis describes how aerobic prokaryotes might have been engulfed by early eukaryotic cells to form their mitochondria (Alberts, 1994). The mitochondria are indeed the eukaryotic organelle where cytochrome c is most prevalent. Cytochrome c has also been suggested to change slowly over evolutionary time (Alberts, 1994). The time required for one “acceptable” amino acid change to appear in a sequence of 100 amino acids is 21 million years. From this it can be estimated that about 97% of the amino acid changes in cytochrome c are harmful mutations, which has been eliminated by natural selection (Alberts, 1994). The evolutionary stability is a sign of the importance of the conserved structure and the vital role of this protein.

The best alignment found in the structural database corresponds to cytochrome c_2 from *Rhodopseudomonas viridis*. The alignment was found with a 105 amino acid segment of the 107 amino acid residues in cytochrome c_2 from *Rhodopseudomonas viridis*. Of these 105 amino acids, 33 was identical (31 %), and further 10 amino acids showed positive correlation. The subject query only consisted of 67 defined amino acids; hence, the similarity could be higher when the whole sequence for cytochrome c-554 from *Methylosinus trichosporium* OB3b is known.

The sequence of cytochrome c_2 from *Rhodopseudomonas viridis* consists of 107 amino acids, which is similar to the expected size of cytochrome c-554 from *Methylosinus trichosporium* OB3b. The N-terminal part of the two sequences align well, and the last part of the known sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b aligns well with the C-terminal part of cytochrome c_2 from *Rhodopseudomonas viridis* as for *Rhodopseudomonas acidophila*. The additional 33 amino acid residues in the cytochrome c_2 sequence are located between these two parts and not at the end as expected. This part is located at the exact same place as the 6 amino acid gap in the primary structure of cytochrome c-554.

Cytochrome c-554 from *Methylosinus trichosporium* OB3b:

```
3   DPAAGEKVFNKCKACHQVGETAKNAVAPELNGIDGRKXXX------------- 42
    D A+GE+VF +C  CH +G  AKN V P LNG+ GR
2   DAASGEQVFKQCLVCHSIGPGAKNKVGPVLNGLFGRHSGTIEGFAYSDANKNS 54
```

Cytochrome c_2 from *Rhodopseudomonas viridis*

Cytochrome c-554 from *Methylosinus trichosporium* OB3b:

```
43  ---------------------XXXMIFPGLSSENDQANVWAYLSQFGADGKK  73
                            MIF G+  E   +++ AY+ QF ADG K
55  GITWTEEVFREYIRDPKAKIPGTKMIFAGVKDEQKVSDLIAYIKQFNADGSK 106
```

Cytochrome c_2 from *Rhodopseudomonas viridis*

Figure 4.22 Sequence alignment found between cytochrome c-554 from *Methylosinus trichosporium* OB3b and cytochrome c_2 from *Rhodopseudomonas viridis* based on a structural database

4.10 Mass determination of a cyanogen bromide fragment of cytochrome c-554

The sequence alignment indicated that the last part of the determined sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b represents the C-terminal end of the protein. This suggests that the gap in the middle of the sequence is much larger than expected from the sequencing.

The methionine we observe in position 46 aligns with Met79 in the cytochromes of *Rhodopseudomonas viridis* and *Rhodopseudomonas acidophila*. This indicates that an unexpected, specific cleavage have occurred, and that the missing sequence is in the middle of the peptide chain rather than at the C-terminal end.

Cyanogen bromide cleavage of the protein would give different sized fragments depending on where the methionine residue is located. Assuming there is only one methionine residue in the protein, cleavage with cyanogen bromide will yield two roughly equally sized fragments for Met46, whereas Met79 will give one fragment of ~9000 Da and one of 2943 Da.

Mass spectrometric studies on the peptide fragments obtained after cyanogen bromide cleavage revealed only one peptide fragment, which had a molecular weight of 3125 Da. A fragment of this size can only be explained as the C-terminal fragment after a cleavage at Met79. Two methionine residues would yield three fragments, but only the C-terminal fragment would be of this size.

This is consistent with the finding of only two amino acid sequences after cyanogen bromide cleavage, and in accordance with the sequences of the homologous cytochromes from *Rhodopseudomonas viridis* and *Rhodopseudomonas acidophila*.

4.11 Protein Modelling

The degree of identity between the amino acid sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b and cytochrome *c*$_2$ from *Rhodopseudomonas viridis* generated interest for modelling the structure of cytochrome c-554.
The online service Swiss-Model was used to find a structure model for cytochrome c-554. Swiss-Model is a fully automated protein structure homology-modelling server, accessible via the ExPASy (Expert Protein Analysis System) web server (http://www.expasy.org/) (Peitsch, 1995;Peitsch, 1996;Guex and Peitsch, 1997). The Swiss-Model program takes an amino acid sequence as input, and makes a model based on similar proteins found in the Swiss-Model repository. The SWISS-MODEL repository is a database of protein models generated from the Protein Data Bank (PDB), and it is continuously updated when new sequences or structures get available.
The gap in the sequence of cytochrome c-554 was filled with the corresponding sequence found in cytochrome c$_2$ from *Rhodopseudomonas acidophila*. The sequence used for structure modelling thus consisted of 107 amino acids, of which 67 amino acids comes from cytochrome c-554 from *Methylosinus trichosporium* OB3b and 40 amino acids comes from the homologue cytochrome c$_2$ from *Rhodopseudomonas acidophila* (Figure 4.22).

```
1   AGDPAAGEKVFNKCKACHQVGETAKNAVAPELNGIDGRKA   40
41  GQAEGFAYSDANKNSGLTWDEATFKEYITAPQKKVPGTKM   80
81  IFPGLSSENDQANVWAYLSQFGADGKK               107
```

Figure 4.23 Amino acid sequence used for structure modelling of cytochrome c-554. The gap in the known sequence of cytochrome c-554 was filled with the corresponding sequence from cytochrome c$_2$ from *Rhodopseudomonas acidophila*. (underlined).

Using this sequence, the modelling program found five homologue sequences with an identity of ~58%, from which it created a structure. Of these five structures, three were of cytochrome *c*$_2$ from *Rhodopseudomonas viridis* and two structures were of cytochrome *c*$_H$ from *Methylobacterium extorquens*. From these structures an average structure was created, and this was in turn optimised using the energy minimisation program GROMOS96 (Scott *et al.*, 1999). A model structure of cytochrome c-554 from *Methylosinus trichosporium* OB3b without hetero atoms was obtained. Figure 4.24 shows a cartoon presentation of the putative cytochrome c-554 structure model.

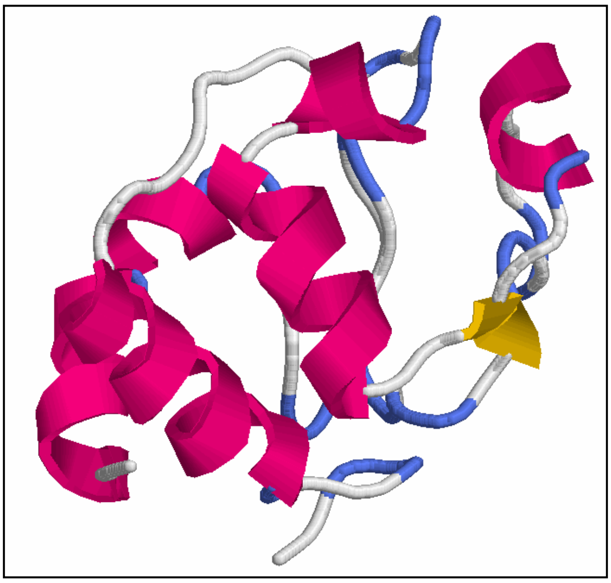


Figure 4.24 Putative 3D structure model of cytochrome c-554 from *Methylosinus*

5 Discussion

The work presented in this thesis was initiated by the finding of a c-type cytochrome in *Methylosinus trichosporium* OB3b that exhibited an unusual highly axial low-spin (HALS) EPR signal (Fauchald, 2000). Furthermore, spectropotentiometric analysis had been performed in order to establish the redox potential of the cytochrome. The absorbance / potential plot obtained at pH 7.0 could not be fitted to a one-electron Nernst plot indicating haem heterogeneity (Figure 5.1).

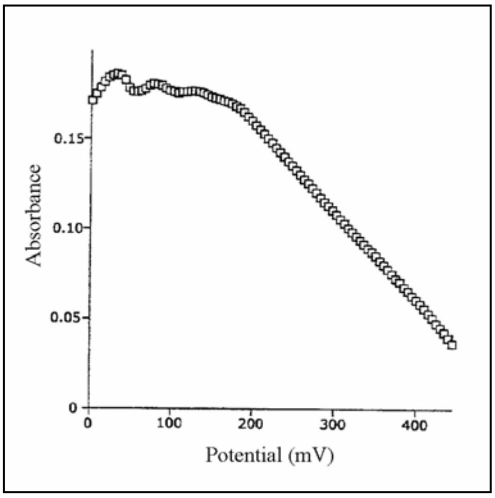


Figure 5.1 Absorbance / potential plot for the c-type cytochrome from *Methylosinus trichosporium* OB3b

The main focus in this thesis has been to characterise this protein and investigate the unusual EPR signal. Two well-characterised c-type cytochromes from *Nitrosomonas europaea* and *Bacillus pasteurii* that exhibit the same unusual EPR signal were used for comparison. A similarly sized cytochrome c from the bacterium *Methylococcus capsulatus* Bath was also investigated. Similarities are found with cytochrome c_1 from the mitochondrial respiratory chain, which also exhibits a HALS EPR signal. Cytochrome c_1 has histidine and methionine ligation of the haem iron similar to what is found in the c-type cytochromes from *Nitrosomonas europaea* and *Bacillus pasteurii*. The redox properties of the four c-type cytochromes that exhibit HALS EPR signals are found in Table 5.1. Many cytochromes that exhibit HALS EPR signals have high redox potentials. The spread of midpoint potentials in Table 5.1 does however make it clear that there is no direct correlation between this and the HALS EPR signal.

	$E^{0'}_{\text{midpoint}}$
Cytochrome c-554 from *Methylosinus trichosporium* OB3b	> 200 mV
Cytochrome c-552 from *Nitrosomonas europaea*	≈ 250 mV
Cytochrome c-553 from *Bacillus pasteurii*	47 mV
Cytochrome c peroxidase from *Nitrosomonas europaea*	450 mV
Cytochrome c_1 (bovine mitochondrial respiratory chain)	220 mV

Table 5-1 Midpoint redox potentials for c-type cytochromes that exhibit HALS EPR signals.

5.1 Optical Properties of Cytochrome c-554 from *Methylosinus trichosporium* OB3b

The extinction coefficients of the α-, β- and γ-bands of the reduced form and the β- and γ-band of the oxidised form of cytochrome c-554 from *Methylosinus trichosporium* OB3b was determined. Although these are associated with some uncertainty, they still provide the most accurate method for concentration measurements.

The absorption wavelength of the α-peak maximum in the reduced state has traditionally been used to name low-spin c-type cytochromes. Consequently, this protein has been named cytochrome c-554.

An additional absorption band with a low extinction coefficient is observed with a peak maximum at 695 nm. This absorption band is considered diagnostic of methionine ligation (Wilson and Greenwood, 1996); hence, one of the axial haem ligands is proven methionine. This is confirmed using circular dichroism. The circular dichroism of cytochrome c-554 in this part of the spectrum is almost identical to that of cytochrome c-552 from *Nitrosomonas europaea*. This suggests that the chromophores responsible for the 695 nm absorption band in these two proteins exist in similar asymmetric environments, i.e. the geometries around the haem irons are similar.

The 695 nm absorption band is also found for cytochrome c-553 from *Bacillus pasteurii*, but the circular dichroic spectrum differs from that of cytochrome c-554 from *Methylosinus trichosporium* OB3b and cytochrome c-552 from *Nitrosomonas europaea*.

5.2 Molecular Mass and Haem Content

The molecular weight of cytochrome c-554 from *Methylosinus trichosporium* OB3b was determined to be 12 230 ± 15 Da using mass spectrometry. The observed uncertainty is higher than what is claimed for the instrument. This can be a result of different isotopic distributions in the individual sample molecules, or the use of an external standard. An internal standard gives higher accuracy.

Mass spectrometry of haem-free peptides established that the protein only contained one haem group. This rules out several haem groups with different redox potentials as cause for the haem heterogeneity observed by my predecessor (Fauchald, 2000).

The method used for removal of haem groups irreversibly modifies tryptophan residues and attaches 2-NPS to cysteine residues through disulphide bridges. Several haem containing proteins with known primary structures were treated using the same procedure. It was found that the mercaptoethanol used in the procedure did not cleave the disulphide bridges between cysteines and 2-NPS to completion, leaving some modified cysteines in the protein.

The mass spectrum of cytochrome c-554 after removal of haem contained five peaks corresponding to the peptide chain with different number of modifications. This indicates that there are four amino acid residues in the protein that is prone to modification by the reagent 2-NPS. Of these four residues, two must be the cysteines that bind the haem group to the protein. The two others are suggested to be tryptophans based on comparison with known proteins treated in the same manner. The presence of tryptophans is confirmed by the appearance of a new spectral band at 365 nm after the reaction with 2-NPS (Figure 4.8).

5.3 The Highly Axial Low-Spin EPR Signal

HALS EPR signals are reported in several haem systems that have histidine and methionine ligation. Figure 5.2 shows c-type haems from four different sources that all exhibits HALS EPR signals. From the structures no common pattern can be observed that distinguish them from other similarly ligated c-type haems.

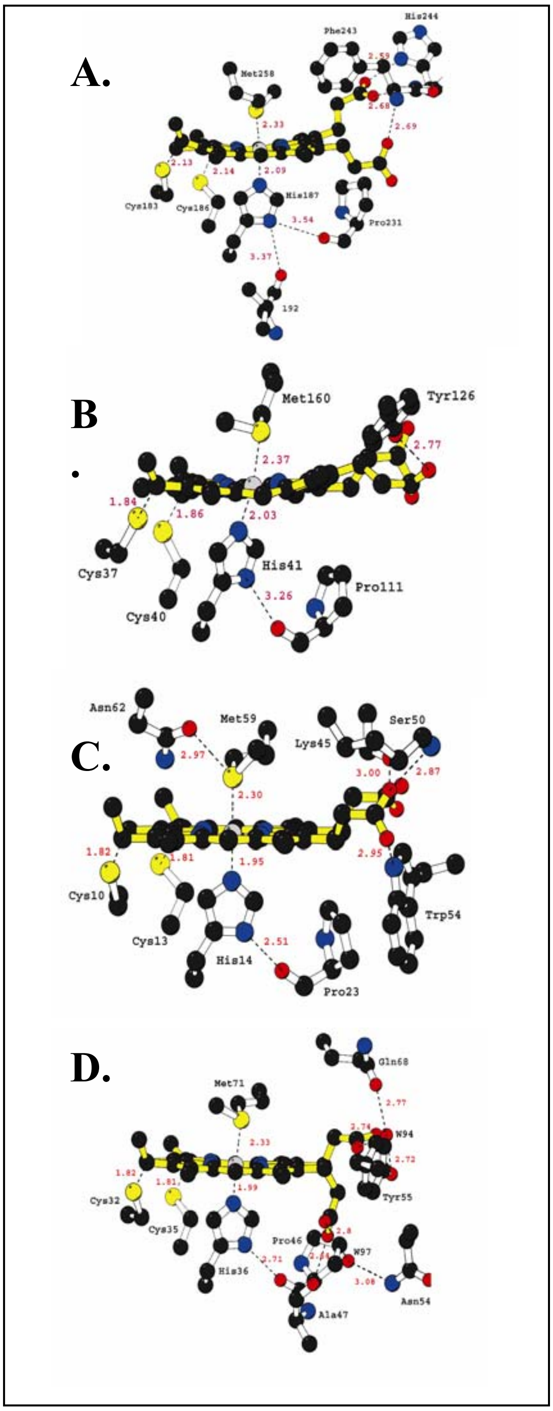

A.
Phe243
His244
Met258
2.59
2.68
2.33
2.69
2.13
2.14
2.09
His187
Cys183
Cys186
3.54
Pro231
3.37
192
B
.
Met160
Tyr126
2.77
2.37
1.84
1.86
2.03
Cys37
His41
3.26
Pro111
Cys40
C.
Asn62
Met59
Lys45
Ser50
2.97
3.00
2.87
2.30
1.82
1.81
1.95
2.95
Cys10
Trp54
Cys13
2.51
His14
Pro23
D.
Gln68
Met71
2.77
2.74
W94
2.33
2.72
1.82
1.81
1.99
Cys32
Tyr55
Cys35
2.8
Pro46
W97
2.71
3.08
His36
Asn54
Ala47

Figure 5.2 c-type haems that exhibit HALS EPR signal. A. Cytochrome c peroxidase from *Nitrosomonas europaea*. B. Haem c_1 from complex III in the bovine mitochondrial respiratory chain. C. Cytochrome c-552 from *Nitrosomonas europaea*. D. Cytochrome c-553 from *Bacillus pasteurii*.

Horse heart cytochrome c has histidine and methionine ligation of the haem iron but exhibits a rhombic EPR signal. When horse heart cytochrome c exhibits HALS EPR signals at high pH the methionine ligand has been replaced by a lysine or arginine (Wilson and Greenwood, 1996), suggesting similarities with the electronic structure in cytochrome f from the cytochrome *bf* complex (Schunemann *et al.*, 1999).

The crystal structure of cytochrome c_2 from *Rhodopseudomonas viridis* is shown in Figure 5.3. This is the structure with the highest degree of identity with cytochrome c-554 from *Methylosinus trichosporium* OB3b.

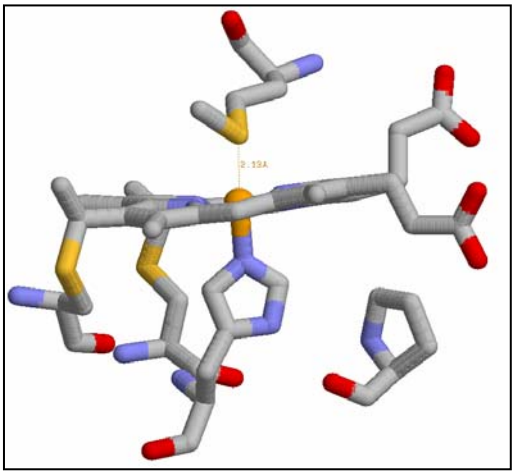


Figure 5.3 The haem environment in ferric cytochrome c_2 from *Rhodopseudomonas viridis*.

The haem ligation closely resembles those of haem c_1 from complex III and cytochrome c-553 from *Bacillus pasteurii*. The haem environment in cytochrome c-554 from *Methylosinus trichosporium* OB3b might be similar to that of cytochrome c_2 from *Rhodopseudomonas viridis* but no certain conclusions can be made.

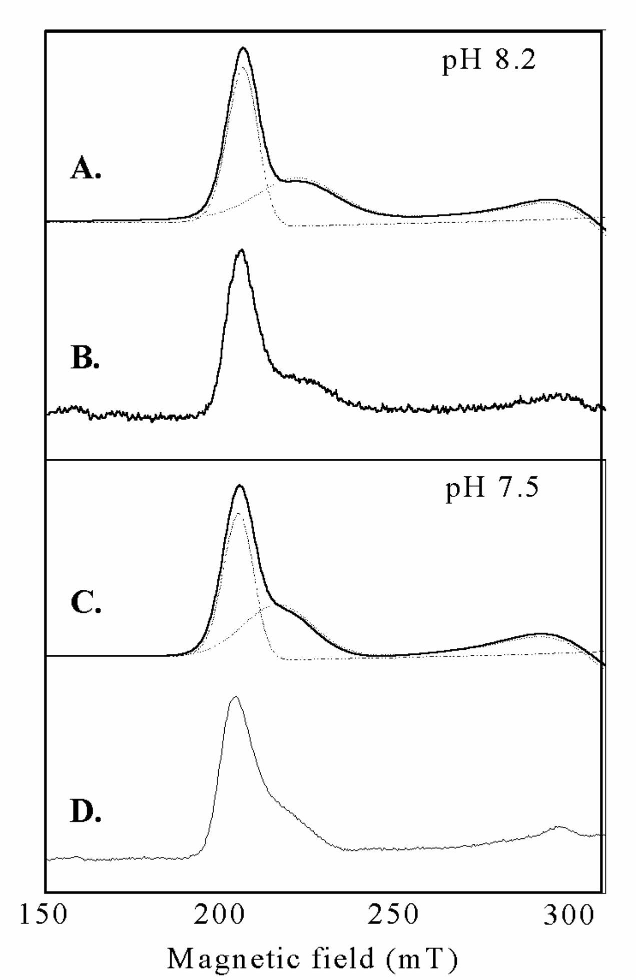


Figure 5.4 Simulation of the EPR spectra of cytochrome c-552 from *Nitrosomonas europaea* performed at 10 K. A. Simulation of spectrum recorded at pH 8.2. B. EPR spectrum recorded at pH 8.2. C. Simulation of spectrum recorded at pH 7.5. D. EPR spectrum recorded at pH 7.5

Simulations of the low-field component of the EPR spectrum of cytochrome c-552 from *Nitrosomonas europaea* do indicate that this is made up of two distinct low-spin species. Figure 5.4 shows how the spectra can be mimicked as a combination of two distinct spin states. The EPR spectra are best fitted as a combination where 3/4 of the spin population exist as a HALS species and 1/4 as a normal rhombic low-spin species. The similarities in the EPR spectrum of cytochrome c-552 from *Nitrosomonas europaea* and the EPR spectrum of cytochrome c-554 from *Methylosinus trichosporium* OB3b suggest that this signal also can be made up of different spin species. Two distinct low-spin species can explain the haem heterogeneity observed in the spectropotentiometric analysis of the cytochrome c-554 from *Methylosinus trichosporium* OB3b (Figure 5.1). Furthermore, the observed temperature variations in Figure 4.12 can be explained as a difference in the distribution between the two low-spin species. The small shoulder in the spectrum recorded at pH 7.0 can indicate that there is a higher degree of the normal rhombic species present at this pH.

The high saturation midpoint ($P_{½}$-value) and the temperature dependence found for the HALS EPR signal from cytochrome c-554 from *Methylosinus trichosporium* OB3b are in accordance with what is observed for the model compounds that exhibit HALS EPR signals.
The HALS EPR signal follows the Curie-Weiss law in the temperature range 7 K to 36 K. This indicates that the ferric haem iron behaves like an ideal paramagnet in this temperature interval. This is similar to what is observed for cytochrome c-552 from *Nitrosomonas europaea* and cytochrome c-553 from *Bacillus pasteurii.*

5.4 Amino Acid Sequence Studies

The amino acid sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b was partly determined using Edman degradation. Thirty-nine amino acid residues from the N-terminal end of the protein were determined using the native protein. This revealed the characteristic haem-binding motif (-C-X-X-C-H-) found in all c-type cytochromes that have two thioether bridges to the protein. The cysteines in this sequence motif form thioether bonds with the vinyl groups on the haem, and are responsible for the covalent attachment of the haem group to the protein. The histidine in the motif is one of the axial ligands coordinating to the haem iron.

After the determination of 39 N-terminal amino acids, the automated sequencer was interrupted. When the sequencing was resumed, a gap in the amino acid sequence followed by an additional seven amino acids was revealed. The gap was assumed to represent six amino acids based on the nature of the sequencing procedure.

Sequencing of a cyanogen bromide treated protein sample revealed a 27 amino acid fragment overlapping six of the seven amino acids in the short sequence revealed after the gap. The sequence obtained was therefore supposed to represent 73 of the N-terminal amino acids in the protein, except for residues 40 to 45 that was lacking due to the interruption of the sequencer.

Sequence alignments and structure comparisons with other cytochromes did however indicate that one of the cyanogen bromide fragments located at the C-terminal end of the protein. This indicates that a site-specific cleavage of the protein had occurred during the interruption of the first sequencing, and the gap might therefore be larger than assumed.

Only one methionine residue is found in the homologous proteins studied, and sequencing of the mixture of cyanogen bromide fragments gave a strong indication that cytochrome c-554 only contained one methionine. Mass spectrometry was used to determine the size of the fragments obtained by the cyanogen bromide treatment. Due to poor solubility of the peptides, only one significant peak was observed in the mass spectrum. This corresponded to a molecular weight of 3125 Da. This is slightly higher (~180) than the theoretical molecular weight of the C-terminal fragment created by a chain break after a methionine in position 79, using the same numbering of amino acids as indicated in Figure 4.21.

Such a C-terminal fragment gives the best possible fit to the observed molecular mass, and strengthens the hypothesis of a methionine in this position.

Type I cytochrome c's generally have a single haem bound near the N-terminal end of the protein, and the sixth ligand methionine near the C-terminal end. A methionine in position 79 is in good accordance with the haem ligand methionine found in other Type I cytochrome c's (Mauk and Scott, 1996).

5.5 Structural Modelling

In homologue proteins, 3D structure is known to be more conserved than primary structure (Salemme, 1977;Creighton, 1993). The degree of identity that cytochrome c-554 from *Methylosinus trichosporium* OB3b exhibits when aligned with cytochrome c_2 from *Rhodopseudomonas viridis* indicates that these two cytochromes have highly similar 3D structures. The structure of cytochrome c_2 from *Rhodopseudomonas viridis* is known, and can be used to create an approximate model of cytochrome c-554 from *Methylosinus trichosporium* OB3b.

A complete approximate sequence for cytochrome c-554 was made by filling the gap between the N-terminal and the putative C-terminal fragment with the corresponding sequence from *Rhodopseudomonas acidophila* cytochrome c_2. This constructed sequence was used as template for the structure modelling.

The SWISS-MODEL Protein Modelling Server identified more than 20 different homologues, two of which were used to model the structure shown in Figure 5.5.

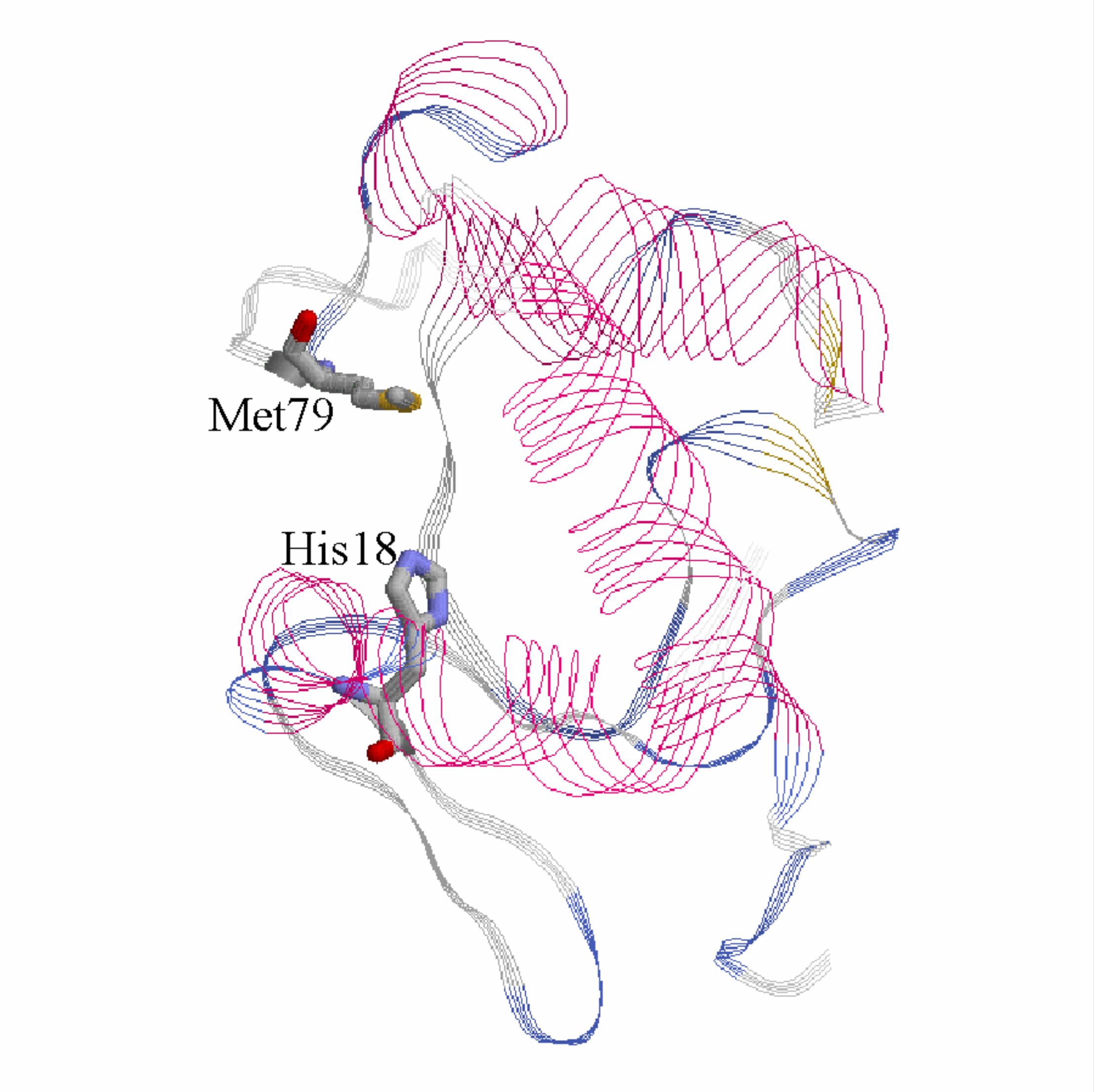


Figure 5.5 3D structure model of cytochrome c-554 from *Methylosinus trichosporium* OB3b. The crevice on the left side is the pocket in which the prosthetic haem group is normally buried. The axial ligands of the haem iron are indicated as stick models.

The SWISS-MODEL Protein Modelling Server creates a structure based solely on the primary structure of the protein; hence, no cofactors, prosthetic groups or other hetero-atoms are included in the model. In the case of cytochrome c-554 from *Methylosinus trichosporium* OB3b, this causes the haem group to be missing. The haem group normally sits in a pocket on one side of the protein, with one edge exposed to the solvent. The crevice that originates from the missing haem group is shown in Figure 5.6.

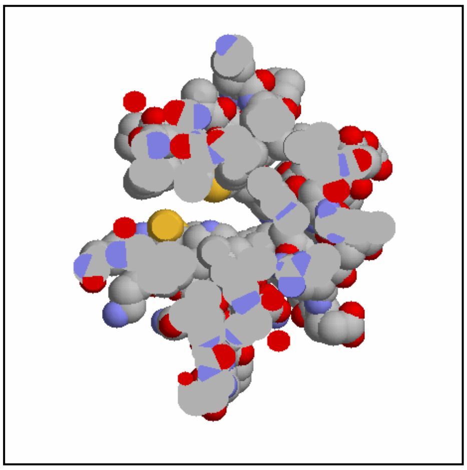


Figure 5.6 Cross-section of the spacefill model of cytochrome c-554 from *Methylosinus trichosporium* OB3b. The crevice on the left side is the haem pocket.

The haem group sets constraints to the movement and orientation of the amino acids in its vicinity. This is of particular importance for the axial iron ligands. The sulphur atom in methionine forms a dative bond to the haem iron (i.e. the shared pair of electrons which forms the bond both originates from the sulphur atom). The energy minimisation procedure used to optimise the model could not take this bond into account, as the iron was not defined in the structure. In the absence of this attraction, the energy minimisation rotates the bonds in the methionine ligand, moving the sulphur atom away from the iron. Consequently the sulphur atom in the methionine axial ligand is moved nearly one Å away from its position in the template structures upon which the model was created.

Movement of the histidine axial ligand is more restricted due to greater constraints from the protein due to the large planar sidechain. The use of the structural model is limited as the first coordination sphere around the haem group is subject to high structural uncertainty. The second coordination sphere, i.e. the hydrogen-bonded network around the haem ligands, is however more certain. This might prove useful in estimating possible electron transfer pathways in the protein.

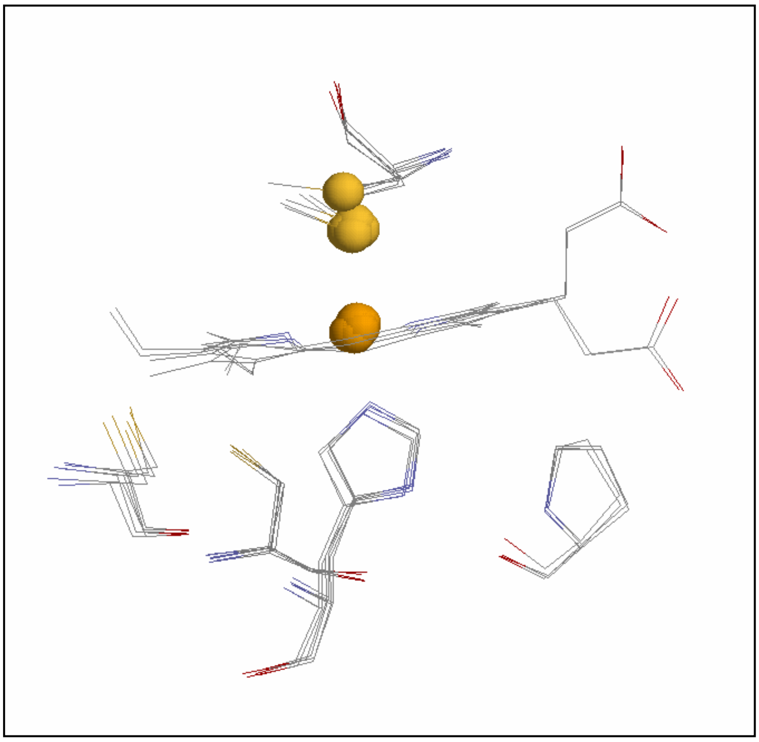


Figure 5.7 The haem environment in the structural templates and the model structures. The methionine residue in the model structure is twisted, thus moving the sulphur atom nearly 1 Å away from its expected position.

5.6 Conclusion

Cytochrome c-554 from *Methylosinus trichosporium* OB3b is a Type I cytochrome c. Optical properties and the amino acid sequence determined histidine and methionine to be the axial ligands to the haem iron. Mass spectrometry revealed that the molecular weight of cytochrome c-554 is 12 230 Da. The amino acid sequence reveals that it is homologue to many cytochrome c's from a diverse range of organisms. Best homology is found with two *Rhodopseudomonas* species where the degree of identity is more than 50 % compared with the amino acid sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b.

Cytochrome c-554 exhibits a HALS EPR signal with a g_{max}-value of 3.41. In bishistidine ligated low-spin haems this kind of EPR signal is known to be diagnostic of perpendicular alignment of the histidine planes. No similar correlation has been found for haems with histidine and methionine ligation.

5.7 Further Studies

- The complete amino acid sequence of cytochrome c-554 from *Methylosinus trichosporium* OB3b will make it possible to create better 3D models of the protein. The complete amino acid sequence is also required for the determination of the 3D structure from x-ray diffraction data.

- A complete amino acid sequence can be obtained using a different method for specific cleavage of the peptide chain followed by an Edman degradation on the purified fragments. A lysine residue is located N-terminal to the missing sequence. This suggests that trypsin cleavage might be a suitable fragmentation method. The homologous proteins contain several lysine residues in the sequence corresponding to the gap in our sequence, and it is likely that other types of fragmentation will be needed too.

- Due to the high degree of identity it is probable that cytochrome c-554 from *Methylosinus trichosporium* OB3b will crystallize under conditions similar to those found for cytochrome c_2 from *Rhodopseudomonas viridis*. An x-ray structure will prove useful when comparing this protein with other cytochromes that exhibit HALS EPR signals.

- Comparisons of other cytochromes with histidine and methionine ligation that exhibit HALS EPR signals, and investigation of the EPR signals exhibited by the homologue cytochromes is needed in order to determine the physical origin of the HALS EPR signal.

- Studies on model-compounds with histidine and methionine ligated haems will also prove useful, but this is beyond the scope of this project.

- The biological role of cytochrome c-554 from *Methylosinus trichosporium* OB3b is not established. Determination of the proteins that serve as electron donor and acceptor for cytochrome c-554 will give valuable information about the cytochrome.

Appendix

Growth Medium for Methylosinus trichosporium OB3b:

$NaNO_3$	10 mM
K_2SO_4	1 mM
$MgSO_4$	0.15 mM
$CaCl_2$	0.05 mM
KH_2PO_4	4 mM
Na_2HPO_4	6 mM
$ZnSO_4$	2 µM
$MnSO_4$	1.6 µM
H_3BO_3	2 µM
$NaMoO_4$	0.4 µM
$CoCl_2$	0.4 µM
KI	1 µM
$CuSO_4$	10 µM
H_2SO_4	36 µM
FeEDTA (or $FeSO_4$)	5 µM

The pH was adjusted to 7.0 by adding sulphuric acid.

Stock Solutions for the Growth Medium:

100 x Salt Solution:

$NaNO_3$	85 g/l	1 M
K_2SO_4	17 g/l	0.1 M
$MgSO_4$	3.7 g/l	15 mM
$CaCl_2$	0.7 g/l	5 mM

100 x Phosphate Buffer Solution:

KH_2PO_4	53 g/l	0.4 M
Na_2HPO_4	86 g/l	0.6 M

500 x Trace Metal solution:

$ZnSO_4$	288 mg/l	1 mM
$MnSO_4$	223 mg/l	0.8 mM
H_3BO_3	62 mg/l	1 mM
$NaMoO_4$	48 mg/l	0.2 mM
$CoCl_2$	48 mg/l	0.2 mM
KI	83 mg/l	0.5 mM
$CuSO_4$	1.25 g/l	5 mM
H_2SO_4 (95%)	1 ml	18 mM

1000 x Iron Solution:

FeEDTA	1.78 g/l	5 mM

Materials

Chemicals	*Source*
2-Mercaptoethanol	Sigma-Aldrich
2-Nitrohenylsulfenyl cloride	Lancaster
α-hydroxy cinamic acid	Sigma-Aldrich
Acetic Acid (100 %, p.a.)	Merck
Acetonitrile (>99.9%)	Riedel
Ammonia (25 %, p.a.)	Merck
Ammonium sulphate (p.a.)	Merck
Boric acid (p.a.)	Merck
Bovine cytochrome c	Sigme-Aldrich
Bovine DNAse I (2000 Kunitz units/ml)	Sigma-Aldrich
Bovine myoglobin	Sigma-Aldrich
Calsium cloride (p.a.)	Merck
Cobalt cloride (p.a.)	Merck
Copper sulphate (p.a.)	Merck
Cyanogen bromide	Sigma-Aldrich
Di-Potassium Hydrogen Phosphate	Merck
EDTA (> 99 %)	Sigma-Aldrich
Ferrous EDTA	Sigma-Aldrich
Formic acid (>98%)	Riedel
Helium (lq)	AGA
HEPES (> 99.5 %, Free Acid)	Sigma-Aldrich
Hydrochloric Acid (36 %, p.a.)	Prolabo
Low Molecular Weight Standard	Pharmacia
Magnesium sulphate (p.a.)	Merck
Manganese sulphate (p.a.)	Merck
Methane (g)	AGA
Methanol (>99.5 %, p.a.)	Merck
Nitric Acid (69 %, p.a.)	AppliChem
Nitrogen (lq)	AGA
PhastGel Blue R (tablets)	Pharmacia
Phast Loadmix	Pharmacia
Potassium Chloride (p.a.)	Merck
Potassium Di-Hydrogen Phosphate (p.a.)	Merck
Potassium Ferricyanide(III) (>99 %)	Sigma-Aldrich
Potassium iodide (p.a.)	Merck
Potassium Hydroxide (p.a.)	Merck
Potassium sulphate (p.a.)	Merck
Pyridine (>99.7%)	Merck
Sinapinic acid	Sigma-Aldrich
Sodium Chloride (p.a.)	Merck
Sodium carbonate (p.a.)	Merck
Sodium Dithionite (~80 %)	Sigma-Aldrich
Sodium Dodecyl Sulfate (> 99 %)	Sigma-Aldrich
Sodium Hydroxide (p.a.)	Merck
Sodium molybdate (p.a.)	Merck
Sodium nitrate (p.a.)	Merck

Sulphuric acid (95-97%)	Merck
Tris(hydroxymethyl)aminomethane (> 99.9 %)	Sigma-Aldrich
Trifluoro acetic acid	Sigma-Aldrich
Zinc sulphate (p.a.)	Merck

Instruments	*Manufacturer*
ER 4113 HV Liquid Helium Control System	Oxford Instruments
HP8452A Spectrophotometer	Hewlett Packard
pH meter (420A)	Orion
Spectropolarimeter (J-810)	Jasco
PhastSystem	Pharmacia
ESP 300E 10/12 X-band spectrometer	Bruker
Voyager DE-RP MALDI-TOF MS	Applied Biosystems

Terms and Abbreviations

2-NPS	2-nitrophenylsulfenylchloride
A	Absorbance
β	The Bohr magneton
B	Magnetic field
BLAST	Basic local alignment search tool
CD	Circular Dichroism
CDS	Coding domain sequence
ΔE	Energy difference between two orbitals
DNase	Deoxyribonuclease
ε	Molar extinction coefficient
EPR	Electron paramagnetic resonance
ESI	Electrospray ionisation
ExPASy	Expert Protein Analysis System
Ferric	Fe(III) oxidation state of iron
Ferrous	Fe(II) oxidation state of iron
g	Electron g factor / Spectroscopic splitting factor
g_e	Free electron g-value
g_n	Standard gravitational acceleration, ($\approx$ 9,807 m/s^2)
HALS	Highly axial low-spin
λ	Wavelength of electromagnetic radiation
MALDI	Matrix-assisted laser desorption ionisation
MMO	Methane monooxygenase
MS	Mass spectroscopy
OD	Optical density
OD_{540}	Optical density at wavelength $\lambda = 540$ nm
pBLAST	Protein basic local alignment search tool
PDB	Protein Data Bank
PIR	Protein Information Resource
PITC	Phenylisothiocyanate (also known as Edman's reagent)
pMMO	Particulate methane monooxygenase
PRF	Protein Research Foundation
PTH	Phenylthiohydantion
rpm	Revolutions per minute
S	Spin angular momentum quantum number
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sMMO	Soluble methane monooxygenase
T	Absolute temperature (Kelvin)
TFA	Trifluoro-acetic acid
TOF	Time of flight
tBLASTn	Translated basic local alignment search tool nucleotides
UV	Ultraviolet
ZFS	Zero-field splitting

Standard Amino Acids

Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic Acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamine	Gln	Q	Serine	Ser	S
Glutamic Acid	Glu	E	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

Reference List

Alberts,B. (1994). Molecular Biology of the Cell. New York: Garland Pub.

Altschul,S.F., Gish,W. (1996). Local alignment statistics. Computer Methods for Macromolecular Sequence Analysis *266*, 460-480.

Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J.H., Zhang,Z., Miller,W., Lipman,D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research *25*, 3389-3402.

Ambler,R.P. (1991). Sequence Variability in Bacterial Cytochromes-C. Biochimica et Biophysica Acta *1058*, 42-47.

Ambler,R.P., Dalton,H., Meyer,T.E., Bartsch,R.G., Kamen,M.D. (1986). The Amino-Acid-Sequence of Cytochrome Cis-555 from the Methane-Oxidizing Bacterium Methylococcus-Capsulatus. Biochemical Journal *233*, 333-337.

Arciero,D.M., Hooper,A.B. (1994). A Di-Heme Cytochrome-C Peroxidase from *Nitrosomonas-europaea* Catalytically Active in Both the Oxidized and Half-Reduced States. Journal of Biological Chemistry *269*, 11878-11886.

Arciero,D.M., Peng,Q.Y., Peterson,J., Hooper,A.B. (1994). Identification of Axial Ligands of Cytochrome C(552) from *Nitrosomonas-europaea*. Febs Letters *342*, 217-220.

Atkins,P.W. (1998). Physical Chemistry. Oxford: Oxford University Press.

Atkins,P.W., Friedman,R.S. (1997). Molecular Quantum Mechanics. Oxford: Oxford University Press.

Benini,S., Gonzalez,A., Rypniewski,W.R., Wilson,K.S., Van Beeumen,J.J., Ciurli,S. (2000). Crystal structure of oxidized *Bacillus pasteurii* cytochrome c(553) at 0.97-angstrom resolution. Biochemistry *39*, 13115-13126.

Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N., Bourne,P.E. (2000). The Protein Data Bank. Nucleic Acids Research *28*, 235-242.

Chothia,C., Lesk,A.M. (1986). The Relation Between the Divergence of Sequence and Structure in Proteins. Embo Journal *5*, 823-826.

Coutinho,I.B., Xavier,A.V. (1994). Tetraheme Cytochromes. Inorganic Microbial Sulfur Metabolism *243*, 119-140.

Cowan,J.A. (1993). Inorganic Biochemistry
an Introduction. New York: VCH.

Creighton,T.E. (1993). Proteins
Structures and Molecular Properties. New York: W. H. Freeman.

Degtyarenko,K.N., North,A.C.T., Findlay,J.B.C. (1999). PROMISE: a database of bioinorganic motifs. Nucleic Acids Research *27*, 233-236.

Dou,Y., Admiraal,S.J., Ikedasaito,M., Krzywda,S., Wilkinson,A.J., Li,T.S., Olson,J.S., Prince,R.C., Pickering,I.J., George,G.N. (1995). Alteration of Axial Coordination by Protein Engineering in Myoglobin - Bisimidazole Ligation in the His(64)-]Val/Val(68)-]His Double Mutant. Journal of Biological Chemistry *270*, 15993-16001.

Edman,P. (1950). Method for determinatin of the amino acid sequence in peptides. Acta Chem.Scandanavica *4*, 283-290.

Emsley,J. (2001). Nature's Building Blocks - an A-Z Guide to the Elements. Oxford: Oxford University Press.

Falk,J.E. (1964). Haems. I. Determination As Pyridine Haemochromes. In: *Porphyrins and Metalloporphyrins*Amsterdam: Elsevier Publishing, 181-188.

Fauchald,S. (2000). Preparative Isolation and Spectroscopic Characterization of a CO-Binding C-Type Cytochrome and a Novel C-Type Cytochrome From the Methanotroph Methylosinus Trichosporium OB3b. Cand. Scient Thesis, University in Oslo.

Fee,J.A. (1978). Transition Metal Electron Paramagnetic Resonance Related to Proteins. In: Methods in Enzymology 512-528.

Friden,H., Cheesman,M.R., Hederstedt,L., Andersson,K.K., Thomson,A.J. (1990). Low-Temperature Epr and Mcd Studies on Cytochrome-B-558 of the *Bacillus-Subtilis* Succinate - Quinone Oxidoreductase Indicate Bis-Histidine Coordination of the Heme Iron. Biochimica et Biophysica Acta *1041*, 207-215.

Girerd,J.J., Journaux,Y. (2000). Molecular Magnetism in Bioinorganic Chemistry. In: Physical Methods in Bioinorganic Chemistry, ed. L.Que Jr. University Science Books, 321-374.

Green,A.A. (1932). Studies in the physical chemistry of the proteins - The solubility of hemoglobin in soluitons of chlorides and sulfates of varying concentration. Journal of Biological Chemistry *95*, 47-66.

Greenwood,C., Wilson,M.T. (1971). Ferricytochrome C .1. Effect of Ph, Ionic Strength and Protein Denaturants on Spectra of Ferricytochrome C. European Journal of Biochemistry *22*, 5-&.

Greibrokk,T., Lundanes,E., Rasmussen,K.E., Karlsen,J. (1994). Kromatografi,Separasjon Og Deteksjon. Oslo: Universitetsforlaget.

Guex,N., Peitsch,M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis *18*, 2714-2723.

Hanson,R.S., Netrusov,A.I., Tsuji,K. (1992). The Obligate Methanotrophic Bacteria *Methylococcus*, *Methylomonas* and *Methylosinus*. In: The Prokaryotes

a handbook on the biology of bacteria : ecophysiology, isolation, identification, applications, ed. A.BalowsNew York: Springer-Verlag, 2350-2362.

Harbitz,E., Zoppellaro,G., Teschner,T., Fauchald,S., Katterle,B., Schunemann,V., Trautwein,A.X., Arciero,D., Hooper,A., Ciurli,S., Andersson,K.K. (2001). Spectroscopical studies of cytochrome cs, some exhibiting HALS EPR signals. Journal of Inorganic Biochemistry *86*, 249.

Hederstedt,L., Andersson,K.K. (1986). Electron-Paramagnetic-Resonance Spectroscopy of *Bacillus- Subtilis* Cytochrome-B558 in *Escherichia-Coli* Membranes and in Succinate-Dehydrogenase Complex from *Bacillus-Subtilis* Membranes. Journal of Bacteriology *167*, 735-739.

Inglis,A.S., Edman,P. (1970). Mechanism of Cyanogen Bromide Reaction with Methionine in Peptides and Proteins. 1. Formation of Imidate and Methyl Thiocyanate. Analytical Biochemistry *37*, 73-&.

Keilin,D. (1925). On Cytochrome, a Respiratory Pigment, Common to Animals, Yeast, and Higher Plants. Proceedings of Royal Society (London) B.Biological Sciences *98*, 312-339.

Kunitz,M. (1950). Crystalline Desoxyribonuclease. I. Isolation and General Properties, Spectrophotometric Method for the Measurement of Deoxyribonuclease Activity. II. Digestion of Thymus Nucleic Acid (Desoxyribonucleic Acid). The Kinetics of the Reaction. Journal of General Physiology *33*, 349 and 363.

Laemmli,U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lambeth,D.O., Campbell,K.L., Zand,R., Palmer,G. (1973). Appearance of Transient Species of Cytochrome C Upon Rapid Oxidation Or Reduction at Alkaline Ph. Journal of Biological Chemistry *248*, 8130-8136.

Lippard,S.J., Berg,J.M. (1994). Principles of Bioinorganic Chemistry. Mill Valley, Calif: University Science Books.

MacMunn,C.A. (1886). Researches Myohæmatin and the Histohæmatins. Philosophical Transactions *177*, 267.

Magnusson,K.E., Edebo,L. (1976). Influence of Cell Concentration, Temperature, and Press Performance on Flow Characteristics and Disintegration in Freeze-Pressing of Saccharomyces-Cerevisiae with X-Press. Biotechnology and Bioengineering *18*, 865-883.

Mauk,A.G., Scott,R.A. (1996). Cytochrome C - a Multidisciplinary Approach. Sausalito, Calif: University Science Books.

Meyer,T.E., Kamen,M.D. (1982). New Perspectives on C-Type Cytochromes. Advances in Protein Chemistry *35*, 105-212.

Meyer-Arendt,J.R. (1972). Introduction to Classical and Modern Optics. Englewood Cliffs, N.J: Prentice-Hall.

Moore,G.R. (1991). Bacterial 4-Alpha-Helical Bundle Cytochromes. Biochimica et Biophysica Acta *1058*, 38-41.

Moore,G.R., Pettigrew,G.W. (1990). Cytochromes C. Evolutionary, Structural and Physicochemical Aspects. Berlin: Springer-Verlag.

Morris,C. (1992). Academic Press Dictionary of Science and Technology. San Diego: Academic Press.

O'Farrell,P.H. (1975). High resolution two-dimensional electrophoresis of proteins. Journal of Biological Chemistry *250*, 4007-4021.

Øgrim,O., Lian,B.E. (1994). Størrelser Og Enheter i Fysikk Og Teknikk. Oslo: Norges Standardiseringsforbund : Universitetsforlaget.

Ogura,H., Yatsunyk,L., Medforth,C.J., Smith,K.M., Barkigia,R.M., Renner,M.W., Melamed,D., Walker,F.A. (2001). Molecular structures and magnetic resonance spectroscopic investigations of highly distorted six-coordinate low-spin iron(III) porphyrinate complexes. Journal of the American Chemical Society *123*, 6564-6578.

Palmer,G., Reedijk,J. (1991). Nomenclature of Electron-Transfer Proteins - Recommendations 1989. European Journal of Biochemistry *200*, 599-611.

Palmer,G. (2000). Electron Paramagnetic Resonance of Metalloproteins. In: Physical methods in bioinorganic chemistry - spectroscopy and magnetism, ed. L.Que Jr. University Science Books, 121-185.

Pauli,W. (1925). On the Connexion between the Completion of Electron Groups in an Atom with the Complex Structure of Spectra. Z.Physik *31*, 765.

Peitsch,M.C. (1995). Protein Modeling by E-Mail. Bio-Technology *13*, 658-660.

Peitsch,M.C. (1996). ProMod and Swiss-model: Internet-based tools for automated comparative protein modelling. Biochemical Society Transactions *24*, 274-279.

Pettigrew,G.W., Moore,G.R. (1987). Cytochromes C. Biological Aspects. Berlin: Springer-Verlag.

Pharmacia Biotech. Gel filtration. Principles and Methods. Product information . 1993.

Pilbrow,J.R. (1990). Transition Ion Electron Paramagnetic Resonance. Oxford: Clarendon Press.

Prince,R.C., George,G.N. (1997). The remarkable complexity of hydroxylamine oxidoreductase. Nature Structural Biology *4*, 247-250.

Sahlin,M., Graslund,A., Ehrenberg,A. (1986). Determination of relaxation times for a free radical from microwave saturation studies. Journal of magnetic resonance *67*, 135-137.

Salemme,F.R. (1977). Structure and Function of Cytochromes-C. Annual Review of Biochemistry *46*, 299-329.

Salerno,J.C., Leigh,J.S. (1984). Crystal-Field of Atypical Low-Spin Ferriheme Complexes. Journal of the American Chemical Society *106*, 2156-2159.

Schunemann,V., Trautwein,A.X., Illerhaus,J., Haehnel,W. (1999). Mossbauer and electron paramagnetic resonance studies of the cytochrome bf complex. Biochemistry *38*, 8981-8991.

Scott,W.R.P., Hunenberger,P.H., Tironi,I.G., Mark,A.E., Billeter,S.R., Fennen,J., Torda,A.E., Huber,T., Kruger,P., van Gunsteren,W.F. (1999). The GROMOS biomolecular simulation program package. Journal of Physical Chemistry A *103*, 3596-3607.

Smith,A.D. (1997). Oxford Dictionary of Biochemistry and Molecular Biology. Oxford: Oxford University Press.

Stanley,S.H., Prior,S.D., Leak,D.J., Dalton,H. (1983). Copper Stress Underlies the Fundamental Change in Intracellular Location of Methane Mono-Oxygenase in Methane-Oxidizing Organisms - Studies in Batch and Continuous Cultures. Biotechnology Letters *5*, 487-492.

Theorell,H., Åkesson,Å. (1941). The Optical Properties of Pure Cytchrome c and Some of Its Derivatives. Journal of the American Chemical Society *63*, 1812-1818.

Timkovich,R., Bergmann,D., Arciero,D.M., Hooper,A.B. (1998). Primary sequence and solution conformation of ferrocytochrome c-552 from *Nitrosomonas europaea*. Biophysical Journal *75*, 1964-1972.

Tonge,G.M., Harrison,D.E.F., Higgins,I.J. (1977). Purification and Properties of the Methane Mono-oxygenase Enzyme System from *Methylosinus trichosporium* OB3b. Biochem.J. *161*, 333-344.

Tsai,A.L., Palmer,G. (1982). Purification and Characterization of Highly Purified Cytochrome-B from Complex-Iii of Baker Yeast. Biochimica et Biophysica Acta *681*, 484-495.

Tsai,A.L., Palmer,G. (1983). Potentiometric Studies on Yeast Complex-Iii. Biochimica et Biophysica Acta *722*, 349-363.

Van Holde,K.E., Ho,P.S., Johnson,W.C. (1998). Principles of Physical Biochemistry. Upper Saddle River, N.J: Prentice Hall.

Voet,D., Voet,J.G. (1995). Biochemistry. New York: Wiley.

Walker,F.A. (1999). Magnetic spectroscopic (EPR, ESEEM, Mossbauer, MCD and NMR) studies of low-spin ferriheme centers and their corresponding heme proteins. Coordination Chemistry Reviews *186*, 471-534.

Wardman,P., Candeias,L.P. (1996). Fenton chemistry: an introduction. Radiat.Res. *145*, 523-531.

Weichsel,A., Andersen,J.F., Roberts,S.A., Montfort,W.R. (2000). Nitric oxide binding to nitrophorin 4 induces complete distal pocket burial. Nature Structural Biology *7*, 551-554.

Wheeler,D.L., Church,D.M., Lash,A.E., Leipe,D.D., Madden,T.L., Pontius,J.U., Schuler,G.D., Schriml,L.M., Tatusova,T.A., Wagner,L., Rapp,B.A. (2001). Database resources of the National Center for Biotechnology Information. Nucleic Acids Research *29*, 11-16.

Wilkins,P.C., Wilkins,R.G. (1997). Inorganic Chemistry in Biology. Oxford: Oxford University Press.

Wilson,M.T., Greenwood,C. (1996). The Alkaline Transition in Ferricytochrome c. In: Cytochrome c a multidisciplinary approach. University Science Books, 611-634.