Crystallographic studies of metalloproteins: cytochrome c₂ from *Rhodopseudomonas palustris* and mammalian transcobalamin
Crystallographic studies of metalloproteins: cytochrome $c_2$ from *Rhodopseudomonas palustris* and mammalian transcobalamin

Il mio più sincero ringraziamento a Lucio Randaccio e Silvano Geremia cui devo il lavoro di ricerca. Ringrazio Giorgio Nardin, Ennio Zangrando, Mario Calligaris e tutti del CEB. E' stato bello dar vita al progetto del Centro di Eccellenza di Biocristallografia a Trieste. Grazie a Daniele, Mara e Luigi, spero di avere la possibilità di lavorare ancora assieme a voi.
Preface

This thesis mainly direct to a few transporter metalloproteins, in particular those containing metallo-organic prosthetic groups such as the iron porphyrin and the cobalt corrin complexes. Porphyrin and corrin groups act as a tetradeutate macrocyclic ligand which can tightly bind metal ions.

The resulting complexes are among the most common and best-known bioinorganic compounds. For example, the heme group consists on a iron center and a substituted porphyrin ligand. Heme is an important biomolecule which is involved in a wide range of important processes, including the electron transfer. Cytochromes of c-type are the most extensively studied family of electron transfer hemoproteins. The other compound, cobalamin, contains cobalt and a partially conjugated 'corrin' ring π system, which has one ring member less than the porphyrin macrocycles. In mammalian cells, the uptake of cobalamin is due to a specific vitamin B₁₂-trasporter, namely trascobalamin.

The interaction between the apoprotein and the metal bonded macrocyclic cofactor is essential for the structure and functions of these trasporter proteins.
# Index

1. **Preface** ........................................................................................................... I

2. **Index** .............................................................................................................. II

3. **Cytochrome c₂ from Rhodopseudomonas palustris** .................. 1

1. **Introduction** .................................................................................................. 1

   1.1 **Electron transfers in biology** ................................................................. 1
       Photosynthesis and respiration ................................................................ 1
       Electron transfer proteins ...................................................................... 5

   1.2 **c-type cytochromes** .............................................................................. 6
       The protein and its prosthetic group ...................................................... 6
       Mitochondrial c-type cytochromes ......................................................... 9
       Bacterial cytochromes c₂ ...................................................................... 12
       The reduction potential of c-type cytochromes .................................. 14
       The folding/unfolding process in c-type cytochromes ...................... 16

3. **Material and methods** ................................................................................. 18

   Synthesis and purification of the protein material .................................. 18
   Mass spectrometry of the protein material .............................................. 18
   Crystallization of the oxidised form .......................................................... 18
   X-ray data collection of the oxidised form .............................................. 19
   Structure determination and refinement of the oxidised form .............. 20
   Crystallisation of the reduced form ........................................................... 21
   X-ray data collection of the reduced form .............................................. 22
   Structure determination and refinement of the reduced form .............. 23
   Crystallization of the alkaline form ........................................................... 24
   X-ray data collections of the alkaline form ............................................ 26
   Structure determination and refinements of the alkaline form ............. 27
   Analysis of structures and structural comparison .................................. 28
Index

4. Results and discussion ........................................30
  4.1 Primary sequence .........................................30
  4.2 Structural analysis of the oxidised and reduced forms ....33
    Packing ...................................................................33
    Molecular structures of the oxidised and reduced form ......35
    Redox-coupled structural changes ................................43
  4.3 Relationship between hydrogen-bonding network and reduction potential 45
  4.4 Structural analysis of the ammonia complex .............50
    Molecular structure of the alkaline form .......................50

5. Conclusions .......................................................62

Mammalian transcobalamin ........................................65

1. Introduction ......................................................66
  1.1 Vitamin B\textsubscript{12} in higher organisms metabolism 66
  1.2 Vitamin B\textsubscript{12}-transporters in mammalians .... 68

3. Material and methods ...........................................71
   Preparation of the genetic material for human transcobalamin....71
   Expression and purification of human transcobalamin ..........71
   Crystallisation of human transcobalamin ........................74
   Attempts for growing crystals of isomorphous derivatives ....75
   X-ray data collections and analysis of human transcobalamin....76
   Preparation of the genetic material for bovine transcobalamin....77
   Expression and purification of bovine transcobalamin ..........77
   Crystallization of bovine transcobalamin ........................78
   X-ray data collections of bovine transcobalamin ................80

4. Results and discussion ........................................81
  4.1 Crystallisation of human transcobalamin ..................81
  4.2 Preliminary X-ray analysis of human transcobalamin .... 82
  4.3 Cristallization of bovine transcobalamin ...................84
  4.4 Preliminary X-ray analysis of bovine transcobalamin .... 85
5. Conclusions......................................................88
References of part I...............................................89
References of part II ............................................96
PART I

Cytochrome $c_2$ from *Rhodopseudomonas palustris*
1. Introduction

1.1 Electron transfers in biology

Photosynthesis and respiration

From the point of view of global bioenergetics, aerobic photosynthesis and respiration are complementary processes (Figure 1.1).

Photosynthesis is the process by which radiant solar energy is converted into chemical energy in the form of ATP and NADPH, which are then used in a series of enzymatic reactions to convert CO₂ into organic compounds (Clayton 1980). Certain bacteria and algae are photosynthetically active as well as green plants, and possess a simple photosynthetic apparatus without the ability to oxidise
Introduction

These bacteria use photosynthesis primarily to separate charges and thus create a transmembrane proton gradient which is then used to synthesise ATP.

The absorption of two photons of light in the photosynthetic reaction centre of purple bacteria (Hall et al. 1987a,b; Deisenhorfer et al. 1991; El-Kabbani et al. 1991), exciting the bacteriochlorophyll (BClfs), leads to cytochrome c\textsubscript{2} oxidation in the periplasm and the reduction of a quinone (Q) to hydroquinone (quinol, QH\textsubscript{2}). This redox reaction is coupled to the uptake of two protons from the cytoplasm (Figure I.2a). The photosynthetic reaction centre and cytochrome c\textsubscript{2} contain four and one heme prosthetic groups, respectively. The tetra-hemes of the photosynthetic reaction centre are arranged perpendicular to the membrane in an approximately linear manner in the cytochrome subunit, which protrudes into the periplasmic space of the membrane. The quinol then leaves its binding site, diffuses in the photosynthetic membrane and is reoxidised by a second membrane protein complex, the cytochrome bc\textsubscript{1} complex, which results in proton release to the periplasm. The electrons are transferred to reduce the cytochrome c\textsubscript{2} in the periplasm. This net proton transport produces a transmembrane electrochemical gradient that can drive ATP synthesis through a third membrane-spanning complex, the ATP synthase (Lancaster and Michel 2002). The result of the photosynthetic process is to convert solar energy into biochemically amenable energy, and this conversion is made possible by transmembrane charge separation and by a precisely adjustment of the midpoint reduction potential of all cofactors involved in the process. Figure I.2b shows a schematic overview of the electron transfer pathways over components of the photosynthetic reaction centre in bacteria and their reduction potentials.
Introduction

Figure 1.2. a) Main components of photosystem involved in the photosynthesis process of bacteria. b) Schematic overview over the electron-transferring components of the photosynthetic reaction centre in bacteria and their reduction potentials. BC1f: bacteriochlorophyll; BPh: bacteriopheophytin; QA, QB, Q, QH2: quinones.
Respiration is the process by which the oxidation of nutrient is converted to the energy of the anhydride bond of ATP. The respiratory chain features a multitude of components found especially in the mitochondrial membrane. The respiratory components function as electron transfer agents according to their reduction potentials (Figure I.3). Complex I (NADH-Q reductase) contains five iron-sulphur clusters and FMN. Complex II (succinate-Q reductase) contains several iron-sulphur clusters, FAD and cytochrome b_{568}. Complex III (ubiquinol-cytochrome c reductase) contains a [2Fe-2S] iron-sulphur centre and cytochromes b_{562}, b_{566} and c'. Complex IV (cytochrome c oxidase) contains at least two copper ions and cytochrome a and a_{3}. Mitochondrial cytochrome c transfers electrons from the ubiquinol-cytochrome c oxidoreductase to the cytochrome c oxidase (Lange and Hunte 2002).

Electron flow through these complexes is associated with the release of relatively large amounts of energy, which is coupled to proton translocation, and therefore ATP production. The reduction potentials of the electron carriers play an essential role in determining the pathway of electron flow through the electron-transport chain (Pelletier and Kraut 1992).

Figure I.3. Main components involved in the respiratory process. Mitochondrial cytochrome c (cyt c) transfers electrons from the ubiquinol-cytochrome c oxidoreductase to the cytochrome c oxidase.
Electron transfer proteins

Electron transfer processes are of great importance in many metabolic pathways of living organisms. In particular, they are essential for photosynthesis and respiration in which energy gained from capturing light or by oxidation of nutrients is converted in ATP. Energy conversion is achieved by coupling the transfer of electrons to the translocation of protons across a lipid membrane. The transfer of electrons is made possible by a series of specific transporter proteins containing metallic redox active centres. Electron carriers can be distinguished in three classes: iron sulphur proteins, blue copper proteins, and cytochromes. The iron sulphur proteins have a FeₐSₐ unit as coenzyme which facilitates the electron transfer. In the [Fe₂-S₂] cluster of ferrodoxines (Figure I.4a) the two irons are tetrahedrally coordinated by the two inorganic sulfur atoms and by sulfurs donated by two cysteine residues (Thomson et al. 1985). The blue copper proteins have a copper coordinated by two histidine, one or two cysteine, and one methionine residues, in a distorted tetrahedral environment, as in plastocyanin (Figure I.4b). Several structures of blue copper proteins are known and they exhibit a quite similar protein fold. Cytochromes, like hemoglobin and cytochrome P-450, contain iron porphyrins as the redox-active centres. Whereas hemoglobin and cytochrome P-450 have only one protein-derived ligand to iron, most cytochromes have two, affording an octahedral iron centre (Figure I.4c).

![Coordination environment of metallic redox active centres in blue copper (a), ferrodoxin (b) and cytochrome proteins (c).](image-url)
1.2 c-type cytochromes

The protein and its prosthetic group

A cytochrome is a protein which contains one or more heme cofactors (iron-protoporphyrin IX). Cytochromes are typically classified on the basis of heme type (Figure 1.5): heme $a$ possesses a long phytol "tail" and is found in cytochrome $c$ oxidase (cytochrome aa$_3$); heme $b$ is found in b-type cytochromes and globins; c-type cytochrome family contains heme $c$ and it is the most extensively studied family of electron transfer proteins. heme-$d_1$ is found in the heme-containing nitrate reductase designed as cytochrome cd$_1$.

In addition to the heme group itself, the presence and nature of fifth and six ligands to iron is crucial in determining the reactivity of any hemoprotein. In cytochromes $c$ (mitochondrial), $c_6$ (from chloroplast, algae and cyanobacteria), $c_2$ (bacterial), $c_{551}$ (namely $c_8$), $c_{552}$, $c_{553}$, and $c_4$ (containing two heme groups), the axial ligands are a histidine and a methionine residue (Figure 1.6). Bacterial
cytochrome c' has one histidine and a vacant distal axial position. The other cytochromes such as c₁ (part of the multisubunit membrane complex bc₁), c₇ (containing 3 heme groups), c₅₅₄ (containing 4 heme groups), have two histidine as axial ligands. A histidine as well as the N-terminal amino group of an aromatic residue provide the heme axial ligands in cytochrome c₇ (part of the multisubunit membrane complex b₆f).

Cytochrome c-type electron carriers are widely spread in nature:

- Mitochondrial cytochrome c, c₁
- Chloroplast, algae and cyanobacteria cytochrome, c₆, c₇
- Bacterial cytochromes c₂, c₄, c₇, c₁, c', c₅₅₀, c₅₅₁, c₅₅₂, c₅₅₃, c₅₅₄

All cytochromes of c-type have the heme c prosthetic group covalently bound to the apoprotein via porphyrin/cysteine bonds (Figure I.6), which are formed from two vinyl groups of heme and two cysteine thiol groups provided by the sequence motif CXXCH (single-letter amino acid code where X is any amino acid). Remarkably, in all of the cytochrome c structures the heme is in the same orientation with respect to the CXXCH attachment sequence. That is, the 2-vinyl group is always attached to the N-terminal cysteine of the motif, and the 4-vinyl group is always attached to the C-terminal cysteine of the motif. As a consequence the same face of the heme is always presented towards the CXXCH motif.

Wood (Wood 1981) suggested that the covalent attachment in mitochondrial c-type cytochrome originated as a device for preventing loss of heme to the external media from proteins in the bacterial periplasm, which is indeed the location of c-type cytochromes in bacteria. Implicit in the previous hypothesis is that c-type cytochromes exist in mitochondria principally because of their recruitment from bacteria by an early symbiotic event (Backer and Ferguson 1999).

The heme group in cytochrome c is not only the redox center, but also critical for maintaining the native structure. Its removal causes disruption of the
native fold and loss of most of the secondary structure under physiological conditions (Stellwagen et al. 1972; Hamada et al. 1993).

**Figure 1.6.** Coordiantion enviroment of heme group in c-type cytochromes having a metionine axial ligand.
Mitochondrial c-type cytochromes

Mitochondrial cytochrome c mediates single electron transfer between integral membrane complexes in the respiratory chain of eukaryotes. Cytochrome c also provides electrons to intermembrane redox partners, such as cytochrome b₅ and sulphite oxidase in animals, or cytochrome c peroxidase and cytochrome b₂ in yeast. It is a protein ubiquitous to all organisms and at the cellular level is found associated with the inner mitochondrial membrane. In respiration, the mobile electron carrier cytochrome c shuttles electrons from the complex cytochrome bc₁ (ubiquinol-cytochrome c oxidoreductase) to the complex cytochrome c oxidase. (Figure 1.3).

The polypeptide chain contains 100-120 amino acids. Their alignment shows that the primary sequence is very highly conserved, with residue identity ranging between 45 and 100% across all eukaryotes. Cytochrome c contains a c-type heme prosthetic group which is linked to the polypeptide chain and is located near the N-terminus. Thioether bonds connect the side chains of two Cys to the heme atoms CAB (pyrrole ring B) and CAC (pyrrole ring C). The heme contains an iron ion which is six-coordinated and has two physiologically relevant oxidation states: Fe^{II} and Fe^{III}.

The UV/vis spectra of cytochrome c are dominated by the absorption of the heme moiety (Banci and Assfalg 2002). Two allowed porphyrin $\pi \rightarrow \pi^*$ transitions are mixed together by interelectronic repulsion to give an intense signal at about 410 nm (called $\gamma$ or Soret band) and two less intense signals between 500 and 600 nm (called $\alpha$- and $\beta$-bands). In the absorption spectra, a band at about 280 nm originates from aromatic amino acids. The reduced and oxidised protein can be easily identified by the absorption spectra. The reduced form shows a sharp $\alpha$-band at 550 nm ($\varepsilon = 29 \text{ mM}^{-1}$), a sharp $\beta$-band at 521 nm, ($\varepsilon = 15.5 \text{ mM}^{-1}$), with a ratio in the absorbances $\alpha/\beta$ of 1.87, and a $\gamma$-band at 413 nm (0.1 M phosphate, pH 7, and sodium dithionite). In the oxidised form, the Soret band moves to 408 nm, the $\alpha$- and $\beta$-bands collapse to one broad band at 530 nm, with a shoulder at 565
nm, and a typical band associated with the porphyrin→Fe charge-transfer transition appears at 695 nm. Electronic spectra are almost identical for cytochromes c from different eukaryotes and a typical absorption spectra for mitochondrial cytochrome c is shown in Figure I.7.

![Figure I.7](image)

**Figure I.7.** A typical spectra of a mitochondrial cytochrome c. Full line refers to the oxidised form; dashed line refers to the reduced form.

The first X-ray structure determination was done by Dickerson in 1971 on oxidised *horse* heart cytochrome c, at 2.8 Å resolution (Dickerson et al. 1973). Results of high resolution crystallographic refinement of *tuna* ferricytochrome c and ferrocytochrome c were reported in 1980 by Takano and Dickerson (Takano and Dickerson, 1980). The comparison of these two structures was made and it was pointed out that the differences between the conformations of the oxidised and reduced mitochondrial cytochrome are small.
Figure I.8. Structure representation of *tuna* cytochrome c.

The overall fold of mitochondrial cytochrome c is highly conserved for all the characterised organisms. It has an almost spherical shape and is characterised by a large helical content and extended loop structures (Figure I.8). The protein fold shows a N-terminal helix involving residues 1-15, approximately. The helix is followed by a loop, containing the histidine coordinated to the iron. A long stretch of residues without secondary structure elements containing a type II β-turn for residues 22-27, is then present. It is followed by a short loop, and then by the third helix, involving residues 61-69, and the fourth helix, involving residues 71-75.
Introduction

The loop after this helix contains the other ligand to the iron, Met80. Finally, the C-terminal helix spans from residue 90 to the end of the polypeptide chain. Figure I.8 shows a structure representation of tuna cytochrome c.

Bacterial cytochromes c₂

Bacterial cytochromes c₂ are biological electron shuttles involved in cyclic redox processes. Cytochrome c₂ is found not only in non-sulphur purple photosynthetic bacteria, such as *Rhodopseudomonas palustris*, *Blastocloris viridis* and *Rhodobacter capsulatus*, but also in non-photosynthetic bacteria, such as *Paracoccus denitrificans*. At the cellular level, the protein occurs in the periplasmic space of the membrane. In the photosynthetic process, cytochrome c₂ functions as a water-soluble electron carrier between two membrane proteins, the ubiquinol-cytochrome c oxidoreductase (the cytochrome bc₁ complex; Lange and Hunte 2002) and the photosynthetic reaction centre (Michel et al. 1986). The electrons serve to reduce the photo-oxidised bacteriochlorophyll dimer, which is the primary electron donor in photosynthetic reaction centre (Figure I.2). On the other hand, one of the four hemes in the cytochrome subunit of the photosynthetic reaction centre is directly reduced by cytochrome c₂, and then electrons are transferred to the bacteriochlorophyll dimer in photosynthetic reaction centre (Miki and Sogabe 2002).

The molecular mass of cytochromes c₂ ranges from 12 kDa to 13-14 kDa (Ambler et al. 1979) and the identity in amino-acid sequence between cytochromes c₂ from different bacteria ranges between 20 and 60 %, with significant differences also observed between the primary sequence of proteins isolated from different strains of the same bacterium (Samyn et al. 1998). For example the *Rhodopseudomonas palustris* strains 2.1.6 (ATCC 17001) and 2.1.37 (ATCC 17007) show only 89% homology (Ambler et al. 1979).

The His is required as one of the heme-iron ligands near the N-terminus. The sixth axial ligand is a methionine residue near the C-terminus. The iron metal
centre is characteristically low-spin. The strictly invariant primary sequences of the eukariotic cytochromes c have a preponderance of positively charged aminoacids with respect to the negatively charged ones (net charge +6) (Tiede et al. 1993), while the prokaryotic cytochroms c2 normally do not show this large disparity between positively and negatively charged residues (Salemme et al. 1973a). Anyhow the primary sequence of cytochrome c2 from *Rhodopseudomonas palustris* (Ambler et al. 1979) shows a large disparity between positively (15) and negatively (10) charged amino acids.

To date, X-ray structures have been determined for the cytochromes c2 isolated from *Rhodospirillum rubrum* (Salemme et al. 1973b), *Rhodobacter capsulatus* (Benning et al. 1991), *Paracoccus denitrificans* (Timkovich and Dickerson 1976; Benning et al. 1994), *Rhodobacter sphaeroides* (Axelrod et al. 1994), *Blastocloris viridis* (Sogabe and Miki 1995; Sogabe and Miki 2002), *Rhodopila globiformis* (Benning et al. 1996), *Methylobacterium extorquens* (Read et al. 1999). The prokariotic cytochromes c2 have been recognized to be structurally similar and evolutionarily related to the mitochondrial cytochromes c (Moore and Pettigrew 1990). Figure 1.9 shows a structure representation of *Blastocloris viridis* cytochrome c2.

The oxidation and reduction of cytochrome c2 are detected from UV/vis absorption spectra with oxidants and reductants, such as ferri- and ferro-cyanide. The oxidised form has only a single broad maximum in the visible region, while the reduced form has two sharp peaks (α and β) in the visible region. Furthermore, on reduction, the Soret peak is shifted to red and has an increased absorptivity. The wavelenght maxima and absorptivities are typical of cytochromes c2 and c except in the region of the protein absorbance at 280 nm, which is dependent on the aromatic amino acid content for each protein (Miki and Sogabe 2002).
The reduction potential of c-type cytochromes

The midpoint reduction potential of mitochondrial cytochromes c is about +260 mV, while it is usually about 100 mV higher in chloroplast cytochromes $c_6$, and it varies from +250 to +450 mV in bacterial cytochromes $c_2$. 

---

Figure 1.9. Structure representation of Blastocloris viridis cytochrome $c_2$. 

---

Diagram:

- 450(mV) corresponds to bacterial cytochrome $c_2$
- Reduction potential corresponds to chloroplast cytochrome $c_6$
- 260(mV) corresponds to mitochondrial cytochrome c
Introduction

Determination of local structural differences among the c-type cytochromes are of interest to investigate the factors controlling the reduction potentials of these soluble water proteins (Pettigrew et al. 1978). Consequently, there has been considerable crystallographic work done in detecting structural changes between reduced and oxidised c-type cytochromes.

The high-resolution determinations of several c-type cytochromes in both oxidation states have revealed that the conformational differences between the reduced and oxidised state are very small. The differences concern thermal-parameter changes, small adjustments around the heme moiety and movement of internal water molecules, rather than significant polypeptide-chain shifts, which instead are found to be minimal. A buried water molecule located at the heme binding pocket near to the iron bonded Met has been found in almost all the X-ray structures of eukaryotic and prokaryotic c-type cytochromes. The positional change of this water molecule is the most prominent structural difference observed with the change of the iron oxidation state in the eukaryotic cytochromes c (Moore and Pettigrew 1990; Lett et al. 1996).

Recently, Satoshi Sogabe and Kunio Miki (Sogabe and Miki 2001) reported a similar change for the analogous water molecule found in the bacterial Blastochloris viridis cytochrome c₂, which presents a reduction potential of +285 mV, close to that of eukaryotic cytochromes c. The authors suggested that this buried water molecule, found in various c-type cytochromes, plays an important role in adjusting the redox potential with alteration of the surrounding hydrogen bond network. However, in order to confirm the role of the internal water molecule in the modulation of reduction potential, it is highly desirable to compare the structures of c-type cytochromes having a significantly different midpoint reduction potential.
The folding/unfolding process in c-type cytochromes

Besides the structure and function of c-type cytochromes, also the folding/unfolding process of these proteins is intimately linked to the presence of the heme prosthetic group and its His and Met axial ligands (George and Schejter 1964; Stellwagen et al. 1972; Fisher et al. 1973; Damaschun et al. 1991; Hamada et al. 1993; Elöve et al. 1994; Colòn et al. 1996; Tezcan et al. 2002). The detachment of the Met ligand from iron coordination is the first step of cytochrome c unfolding observed upon increasing pH and/or temperature (Elöve et al. 1994; Banci et al. 1998). The pH-induced protein conformational transitions and changes in the ligation state of the heme iron in cytochrome c₂ from *Rhodopseudomonas palustris* have been monitored by electrochemical and spectroscopic measurements (Battistuzzi et al. 1995; Bertini et al. 1998). At pH values above 8, the appearance of an additional NMR signal pattern was interpreted as the result of the formation of new species with distinct structural properties. In these and previous works (Timkovich et al. 1984) on cytochrome c, it was hypothesised that the alkaline transition is probably due to the replacement of the methionine axially bound to the heme iron with a stronger (most probably N-donor) ligand (Ubbink et al. 1994; Battistuzzi et al. 1997). The axial Met displacement in c-type cytochromes has been widely studied also using antagonist exogenous ligands (George and Schejter 1964; Myer 1984; Shao et al. 1993; Shao et al. 1995; Banci et al. 1998; Dumortier et al. 1998). Recently, an NMR study of the conformational flexibility of the oxidized *horse* heart cytochrome c through its interaction with NH₃ has been reported (Banci et al. 1998). It was shown that NH₃ binds to iron(III) by displacing the axial Met with an affinity constant in the range 1.5 - 4 M⁻¹. The ¹H-NMR spectra of this ammonia adduct is similar to that obtained with alkaline transition. Despite the extensive studies in solution, the imidazole adduct of cytochrome c₂ from *Rhodobacter Sphaeroides* is the only crystal structure (2.2 Å resolution) so far reported among c-type cytochromes with the axial Met ligand replaced (Axelrod et al. 1994).
All these data are consistent with the conclusion that during the unfolding pathway of c-type cytochrome the methionine-iron bond is broken, but the protein partially unfolded states are not structurally determined (Figure I.10).

Figure I.10. Schematic overview of the unfolding pathway of c-type cytochrome observed in solution upon increasing pH, temperature, or using exogenous axial ligands.
3. Material and methods

Synthesis and purification of the protein material

Cytochrome c\textsubscript{2} obtained from a \textit{Rhodopseudomonas palustris} strain 42 OL of the culture collection of the Research Center on Autotrophic Microorganism (Florence, Italy) was purified as described by Bertini et al. (Bertini et al. 1998). The purity of the protein was checked from the ratio of its absorbance at 280 nm to that at 412 nm.

Mass spectrometry of the protein material

Mass spectra were recorded with a SCIEX API I single quadrupole mass spectrometer equipped with an ion-spray source. Calibration was carried out with a PPG (poly-propylene glycol) calibration solution. Cytochrome c (\textit{horse} heart type IV - Sigma) was used as protein standard. Mass spectrometric standard deviations in the determination of cytochrome c molecular weight were ± 1 Da. Proteins (about 0.1 µg/µL) were dissolved in water and diluted with CH\textsubscript{3}CN and HCOOH up to 50% and 0.1% v/v, respectively, before flow-injection into the spectrometer (flow rate 1.6 µL/min - positive mode - dwell time 0.5 ms - step size 0.1 amu - 5-10 scans accumulated). Mass data were analysed with the software supplied with the instrument.

Crystallization of the oxidised form

Crystallization experiments of the cytochrome c\textsubscript{2} from \textit{Rhodopseudomonas palustris} were performed with the hanging-drop vapor diffusion method at 291 K.
Protein solution (9.0 mg/ml buffered to pH 6.0 in 20 mM phosphate) was oxidized with an excess of ferricyanide. The oxidation state of the protein was monitored by absorbance spectrometry. 2 µl of protein solution were mixed with the same volume of the reservoir solution and equilibrated against 1 ml of the reservoir solution. Crystals grew in 3-7 days from 42-44 % saturated ammonium sulphate and 0.1 M citrate buffer pH 4.4 (Figure 1.11a).

**X-ray data collection of the oxidised form**

Data collections were carried out at the Elettra Synchrotron (Trieste, Italy) using a monochromatic radiation with wavelength of 1.000 Å and a MAR Research 345 mm imaging plate as detector. Before freezing, some crystals of ferricyanide were dissolved in the drop containing the cytochrome c2 crystals, to provide a fresh oxidation environment. After fifteen minutes the protein crystals were quickly passed through a reservoir solution containing 20% glycerol as cryoprotectant and flash-frozen in a stream of N2 at 100 K.

The reflections of oxidised crystals were indexed using a monoclinic unit-cell having dimensions of a = 50.31, b = 71.53, c = 66.66 Å and β = 93.5°, with systematic absences in agreement with the P21 space group (Garau et al. 2000). The calculated Matthews coefficient (Matthews 1968), $V_m$ is 2.45 Å³ Da⁻¹ for the oxidised form, assuming 8 molecules in the unit cell and a molecular weight of 12.2 kDa for the polypeptide chain (Ambler et al. 1979).

A complete data set from a monoclinic ferricytochrome c2 crystal was collected to 1.7 Å. The determination of unit-cell parameters, integration of reflection intensities and data scaling were performed using MOSFLM and SCALA programs from the CCP4 suite (Collaborative Computational Project, Number 4, 1994).
Structure determination and refinement of the oxidised form

The structure of the oxidised form was solved by molecular replacement using AMoRe (Navaza 1994) as implemented in the CCP4 program suite (Collaborative Computational Project, Number 4, 1994) and by using the information obtained by the Patterson map.

The starting model was the refined 2.0 Å structure of *Rhodospirillum rubrum* (Salemme et al. 1973) (PDB code: 2c2c), whose amino-acid sequence shows a 42% identity with that of ATCC 17006/2.1.37 strain of *Rhodopseudomonas palustris* cytochrome c₂ [Swiss PDB code: cy22] (Ambler et al. 1979). All residues in *Rhodospirillum rubrum* cytochrome c₂ that differ from those of *Rhodopseudomonas palustris* cytochrome c₂ were mutated, insertion and deletion region were removed and joined, respectively.

Using data in the resolution range 8 - 2 Å, only two almost equivalent rotational solutions, related by a two-fold rotor parallel to the a axis, were obtained. This suggests that the four independent molecules have two by two the same orientation (Garau et al. 2000). The presence of a very intense single peak (65% of the origin peak) in the Patterson map indicated that these pairs are related by the same translation vector (0.5, 0.34, 0.5). The rigid-body refinement of the four polypeptide chains, obtained from the AMoRe solution (Table I.1), improved the correlation coefficient and R factor to 55.9% and 51.7%, respectively. A random subset of data (5%) was omitted from all refinement calculations to provide an assessment of the progress of refinement. Restrained positional and thermal factor refinement with non-crystallographic symmetry restraints improved the R and free R factors to 39.1% and 42.0%, respectively. Rebuilding of the model and location of solvent atoms were based on the analysis of (2F₀-Fₑ) and (F₀-Fₑ) maps and were performed on Silicon Graphics workstations using the program O (Jones et al. 1991).
Material and methods

Table I.1: Roto-translation solutions obtained by molecular replacement for the four independent molecules of the oxidised form.

<table>
<thead>
<tr>
<th></th>
<th>(\alpha)</th>
<th>(\beta)</th>
<th>(\chi)</th>
<th>(T_x)</th>
<th>(T_y)</th>
<th>(T_z)</th>
<th>Corr</th>
<th>(R_{fack})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution 1</td>
<td>192.09</td>
<td>87.88</td>
<td>304.88</td>
<td>0.4423</td>
<td>0.0000</td>
<td>0.0000</td>
<td>26.8</td>
<td>60.5</td>
</tr>
<tr>
<td>Solution 2</td>
<td>192.09</td>
<td>87.88</td>
<td>304.88</td>
<td>0.9425</td>
<td>0.3439</td>
<td>0.4999</td>
<td>49.1</td>
<td>55.5</td>
</tr>
<tr>
<td>Solution 3</td>
<td>168.09</td>
<td>91.92</td>
<td>124.58</td>
<td>0.9412</td>
<td>0.0412</td>
<td>0.0000</td>
<td>45.7</td>
<td>53.7</td>
</tr>
<tr>
<td>Solution 4</td>
<td>168.09</td>
<td>91.92</td>
<td>124.58</td>
<td>0.4415</td>
<td>0.3853</td>
<td>0.4992</td>
<td>53.0</td>
<td>54.0</td>
</tr>
</tbody>
</table>

The atomic model of the oxidized native form was further refined using the primary sequence obtained from the structure of ammonia adduct (see structure determination and refinement of the alkaline form) and additional water molecules were introduced. The model was refined using the REFMAC program (Murshudov et al. 1997) with restrained positional and thermal factor refinement and non-crystallographic symmetry (NCS) restrains. These NCS restraint value of 0.05 Å was used for all the four crystallographically independent protein molecules and for the water molecules related by pseudo symmetry. The NCS restraints for specific amino acids (side chain and backbone for Thr94 and Phe95 and only side chain for Asp2, Lys4, Lys11, Arg18, Glu53, Lys76, Lys96 and Lys114) were released when large differences in the conformation among the four independent units were detected in the Fourier maps.

Table I.2 reports a summary of data-collections and crystallographic statistics of the oxidised form of *Rhodopseudomonas palustris* cytochrome c₂, together with those of the reduced form and the ammonia adduct.

Crystallisation of the reduced form

Crystallisation experiments of the cytochrome c₂ from *Rhodopseudomonas palustris* were performed with the hanging-drop vapor diffusion method at 291 K. Protein solution (9.0 mg/ml buffered to pH 6.0 in 20 mM phosphate) was reduced
with an excess of sodium dithionite. The oxidation state of the protein was monitored by absorbance spectrometry. 2 µl of protein solution were mixed with the same volume of the reservoir solution and equilibrated against 1 ml of the reservoir solution. At acid pH, crystals grew in 3-7 days from 42-44 % saturated ammonium sulphate and 0.1 M citrate buffer pH 4.4 (Figure I.11a).

X-ray data collection of the reduced form

Data collections were carried out at the Elettra Synchrotron (Trieste, Italy) using a monochromatic radiation with wavelength of 1.000 Å and a MAR Research 345 mm imaging plate as detector. Before freezing, some crystals of sodium dithionite were dissolved in the drop containing the cytochrome c2 crystals, to provide a fresh reduction environment. After fifteen minutes the protein crystals were quickly passed through a reservoir solution containing 20% glycerol as cryoprotectant and flash-frozen in a stream of N₂ at 100 K.

The reflections of the reduced form crystals were indexed using a monoclinic unit-cell having dimensions of a = 49.90, b = 71.28, c = 67.07 Å and β = 93.7°, with systematic absences in agreement with the P2₁ space group (Geremia et al. 2002). The calculated Matthews coefficient (Matthews 1968), Vm, is similar to that of the oxidised form. A complete data set from a monoclinic ferrocytochrome c2 crystal was collected to 1.9 Å resolution.

The determination of unit-cell parameters, integration of reflection intensities and data scaling were performed using MOSFLM and SCALA programs from the CCP4 suite (Collaborative Computational Project, Number 4, 1994).
Structure determination and refinement of the reduced form

The plate-like form of the reduced crystals resulted in an anisotropical diffraction. In particular, the crystal of the reduced form diffracted at least to 1.95 Å resolution when the X-ray primary beam was orthogonal to the plate-like crystal, but only to 2.2 Å resolution when the plate was parallel to the X-ray primary beam. For this reason, the R\text{merge} of the reduced form in the last shell (Table 1.2) is extremely high. However, the <I/σ(I)> for this last shell is 2.0 and we have chosen to include all statistically acceptable data for the refinement of this structure.

The structure of the reduced form was solved by molecular replacement using AMoRe (Navaza 1994) as implemented in the CCP4 program suite (Collaborative Computational Project, Number 4, 1994). The refinement of the reduced cytochrome c₂ structure was performed starting with the coordinates of the oxidized native form. The initial rigid body refinement using REFMAC gave an R factor of 31.1% for 34721 unique reflections. After rebuilding of the model and fifty cycles of positional refinement using REFMAC the R and R\text{free} factors dropped to 19.3% and 24.6%, respectively. In order to remove the bias of the starting model, further 20 cycles were performed using the SHELXL program with default restraints (Sheldrick 1997). Using these weaker restraints, the R decreased to 18.8% and the R\text{free} increased to 25.6%. Last twenty cycles of refinement, using REFMAC with NCS restraints (the same used for the oxidized protein with the further release of side chain and backbone of Ala87, Val88 and Gly89) gave the final R and R\text{free} factors reported in Table 1.2.
Material and methods

Crystallization of the alkaline form

Crystallization experiments of the cytochrome c$_2$ from *Rhodopseudomonas palustris* were performed with the hanging-drop vapor diffusion method at 291 K. Protein solution (9.0 mg/ml buffered to pH 6.0 in 20 mM phosphate) was either oxidized or reduced with an excess of ferricyanide or sodium dithionite, respectively. The oxidation state of the protein was monitored by absorbance spectrometry. 2 µl of protein solution were mixed with the same volume of the reservoir solution and equilibrated against 1 ml of the reservoir solution.

At basic pH, the crystals used for data collection grew in about three weeks but only from the reduced protein solution, in the presence of 61 % saturated ammonium sulphate and 0.1 M Tris-HCl buffer pH 8.5-9.0 (Figure I.11b).
Table 2. Data processing and refinement statistics for the native structures of *Rhodopseudomonas palustris* cytochrome c2 (oxidized and reduced state) and for the ammonia adduct structure of this cytochrome. Values in parentheses refer to the outer resolution shell (1.79 - 1.70 Å oxidized state; 2.06 - 1.95 Å reduced state; 1.21 - 1.15 Å ammonia adduct). The coordinates and structural factors have been deposited to the Protein Data Bank under the accession codes 1fjo and 1i8p, and 1i8o.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Oxidized c2</th>
<th>Reduced c2</th>
<th>Ammonia c2 adduct</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection and processing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P21</td>
<td>P21</td>
<td>P3221</td>
</tr>
<tr>
<td>Unit-cell parameters (Å, °)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a = 50.31</td>
<td>a = 50.12</td>
<td>a = 64.83</td>
<td></td>
</tr>
<tr>
<td>b = 71.53</td>
<td>b = 71.71</td>
<td>b = 64.83</td>
<td></td>
</tr>
<tr>
<td>c = 66.66</td>
<td>c = 67.36</td>
<td>c = 68.55</td>
<td></td>
</tr>
<tr>
<td>β = 93.5</td>
<td>β = 93.7</td>
<td>γ = 120.0</td>
<td></td>
</tr>
<tr>
<td>No of measurements</td>
<td>194662</td>
<td>106868</td>
<td>364867</td>
</tr>
<tr>
<td>No of unique reflections</td>
<td>51959</td>
<td>34702</td>
<td>57362</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt; (%)</td>
<td>8.7 (35.8)</td>
<td>17.5 (59.9)</td>
<td>5.4 (41.2)</td>
</tr>
<tr>
<td>&lt;L/σ(L)&gt;</td>
<td>11.0 (3.6)</td>
<td>6.4 (2.0)</td>
<td>19.0 (4.1)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.7</td>
<td>3.1</td>
<td>6.3</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9 (99.7)</td>
<td>99.7 (100)</td>
<td>99.3 (96.7)</td>
</tr>
<tr>
<td><strong>Refinement statistics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20-1.70</td>
<td>20-1.95</td>
<td>20-1.15</td>
</tr>
<tr>
<td>No of reflection used</td>
<td>49259</td>
<td>32939</td>
<td>55707</td>
</tr>
<tr>
<td>R-factor&lt;sup&gt;a&lt;/sup&gt; (%)</td>
<td>17.8</td>
<td>19.2</td>
<td>13.9</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt;-factor&lt;sup&gt;b&lt;/sup&gt; (%)</td>
<td>21.8</td>
<td>23.7</td>
<td>16.7</td>
</tr>
<tr>
<td>No of protein atoms</td>
<td>3420</td>
<td>3420</td>
<td>855</td>
</tr>
<tr>
<td>No of heme atoms</td>
<td>172</td>
<td>172</td>
<td>43</td>
</tr>
<tr>
<td>No of water molecules</td>
<td>653</td>
<td>613</td>
<td>238.5</td>
</tr>
<tr>
<td>Total number of atoms</td>
<td>4264</td>
<td>4205</td>
<td>1147.5</td>
</tr>
<tr>
<td><strong>B-factor statistics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average B main chain (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>11.9</td>
<td>13.1</td>
<td>9.8</td>
</tr>
<tr>
<td>Average B side chain (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>16.3</td>
<td>16.5</td>
<td>13.4</td>
</tr>
<tr>
<td>Average B heme atoms (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>8.7</td>
<td>9.4</td>
<td>7.0</td>
</tr>
<tr>
<td>Average B waters (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>30.7</td>
<td>32.0</td>
<td>28.3</td>
</tr>
<tr>
<td>Overall B (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>16.4</td>
<td>17.4</td>
<td>11.6</td>
</tr>
<tr>
<td><strong>RMSD from ideal geometry</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.010</td>
<td>0.028</td>
<td>0.010</td>
</tr>
<tr>
<td>Bond angle distances (°)</td>
<td>2.0</td>
<td>2.2</td>
<td>1.6</td>
</tr>
</tbody>
</table>

<sup>a</sup>R<sub>fact</sub> = ∑ |F<sub>obs</sub> - F<sub>calc</sub>| / ∑ F<sub>obs</sub>; <sup>b</sup>R<sub>free</sub> calculated for 5% of the observed reflections.
X-ray data collections of the alkaline form

Data collections were carried out at the Elettra Synchrotron (Trieste, Italy) using a monochromatic radiation with wavelength of 1.000 Å and a MAR Research 345 mm imaging plate as detector. After fifteen minutes a protein crystal grown at pH 9.0 were quickly passed through a reservoir solution containing 20% glycerol as cryoprotectant and flash-frozen in a stream of N₂ at 100 K. Analysis of the diffraction pattern from the alkaline form crystals and structure solution reveal that they belong to the trigonal space group P3₂1 with unit-cell dimensions \( a = b = 64.67 \) and \( c = 68.25 \) Å (Garau et al. 2000). A complete data set from this trigonal crystal was collected to 1.4 Å resolution and data-collection statistic is reported in Table I.3.

<table>
<thead>
<tr>
<th>Table I.3: Data collection statistics for the alkaline form obtained at pH = 9.0.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Temperature (K)</td>
</tr>
<tr>
<td>Resolution-range</td>
</tr>
<tr>
<td>Unit-cell parameters (Å)</td>
</tr>
<tr>
<td>( V (Å³) )</td>
</tr>
<tr>
<td>( Z )</td>
</tr>
<tr>
<td>Number of observations</td>
</tr>
<tr>
<td>Unique reflections</td>
</tr>
<tr>
<td>( R_{merge} )(%)*</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>Outer shell</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>Outer shell</td>
</tr>
<tr>
<td>( I/\sigma (I) )</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>Outer shell</td>
</tr>
</tbody>
</table>

\( R_{merge} = \frac{\sum |I - \langle I \rangle|}{\sum |I|} \)

*Outer-shell data is the resolution range 1.47-1.40 Å.
Afterwards another complete data set from a crystal grown at pH 8.5 was collected to 1.15 Å resolution. This time, before freezing, some crystals of ferricyanide were dissolved in the drop containing the crystal, to provide a fresh oxidation environment. Analysis of the diffraction pattern from the alkaline form crystal and structure solution revealed that it belong to the space group P3_21 with unit-cell dimensions \( a = 64.83 \) and \( c = 68.55 \) Å. Table 1.2 reports a summary of data-collection and crystallographic statistics relative to the latter atomic resolution data collection.

**Structure determination and refinements of the alkaline form**

The structure solution of the alkaline form was searched in both space groups P3_{12}1 and P3_{21} by molecular replacement, using the atomic coordinates of the oxidised form as starting model, and the data set from the trigonal crystal which was collected to 1.4 Å resolution (Garau et al. 2000). The initial R value for the molecular replacement solution obtained in the P3_{21} space group was 45.1 in the 8-2 Å resolution range. The restrained positional and thermal factor refinement improved the R and free R factors to 32.3% and 36.9%, respectively.

When the second data set at 1.15 Å resolution was collected (Geremia et al. 2002), the crystallographic refinement of the alkaline form was continued starting from the partially refined structural model at 1.4 Å resolution with an R factor of 29.1%. The anisotropic refinement protocol scheme used for the alkaline form structure at atomic resolution was:

1. Full resolution
2. Isotropic refinement
3. Evident water molecules and model correction
4. Anisotropic (Bij) refinement
5. Disorder and partial water molecules
6. Riding hydrogen atoms
7. Relax / remove restraints
8. All reflections
9. Full matrix last squares
Material and methods

Rebuilding of the model, amino acid substitutions and location of solvent atoms were performed on Silicon Graphics workstations using the O program (Jones et al. 1991) and the enhanced 2Fo-Fc and Fo-Fc maps as a guide. Refinement of the structure was performed using SHELXL program (Sheldrick 1997). Polypeptide atoms were restrained to the 1,2 and 1,3-target distances based on the study of Engh and Huber (Engh and Huber 1991) while the positional refinement of heme atoms and solvent molecules were not restrained. The positional refinement improved the R and R_free factors (Brünger 1992) to 18.1% and 19.9%, respectively. Making all atoms anisotropic resulted in a drop of the R and R_free factors to 15.0% and 17.3%, respectively. Anisotropic thermal parameters of the water molecules were restrained with an effective standard deviation of 0.1, so that their U_{ij} components approximated the isotropic behaviour while protein atoms were restrained in order to make the U_{ij}-values of neighbouring atoms similar (Sheldrick 1997). Finally, the hydrogen atoms were added at calculated positions and further minor adjustments were performed.

In order to confirm the presence of the ammonia molecule as axial ligand, the nitrogen atom was replaced by an oxygen atom during parallel last refinement cycles. While no change in the R_factor values occurred, the temperature factor of the oxygen axial ligand increased from 6.0 to 8.5 Å².

The determination of unit-cell parameters, integration of reflection intensities and data scaling were performed using MOSFLM and SCALA programs from the CCP4 suite (Collaborative Computational Project, Number 4, 1994).

Analysis of structures and structural comparison

The model quality was checked with PROCHECK (Laskowki et al. 1993). The secondary structure of the protein was analysed using the DSSP program (Kabsch and Sander 1983). Structural comparison between the native protein and its ammonia adduct was carried out by r.m.s. superposition of the protein Cα
atoms of the subunit A of the oxidized native form using the program COMPAR (Collaborative Computational Project, Number 4, 1994). The superimposition of the peptide segments involved in the conformational transition between the native form and N-ligand cytochrome c₂ complex of *Rhodopseudomonas palustris* (Figure I.23a) and of *Rhodobacter sphaeroides* (Figure I.23b) was performed by r.m.s. superposition of heme moiety of the native and N-adduct form using the lsq commands of the *O* program (Jones et al. 1991). Atomic coordinates for the imidazole adduct and native form of *Rhodobacter sphaeroides* cytochrome c₂ were obtained from the Brookhaven Protein Data Bank (Bernstein et al. 1977) using the entries 1cxa and 1cxc, respectively. Statistical analysis of Fe-NH₃ and Fe-OH bond lengths in small molecule crystal structures were performed by using data from the Cambridge Structural Database (Allen et al. 1979; CSDS 5.20 Version, Oct 2000 with 224400 entries).
4. Results and discussion

4.1 Primary sequence

The primary structure of *Rhodopseudomonas palustris* cytochrome c\(_2\) has been previously determined by Ambler and co-workers for two different strains, 2.1.6 (ATCC 17001) and 2.1.37 (ATCC 17007) (Ambler et al. 1979). For our structural studies, we have used a cytochrome c\(_2\) protein expressed from a different strain (42 OL) of the same bacteria, isolated from a sugar-refinery soil, and whose amino acid sequence was unknown (Bertini et al. 1998).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42 OL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1.12. Primary sequences of *Rhodopseudomonas palustris* cytochrome c\(_2\) isolated from the 42 OL, the 2.1.6 (ATCC 17001) and the 2.1.37 (ATCC 17007) strains. The mutations of the cytochrome c\(_2\) from 42 OL strain, with respect to the others two sequences, are evidenced on the lines between the alignments. The ‘+’, ‘-‘ and ‘0’ symbols reflect positive, negative and zero BLOSUM-62 substitution matrix scores, respectively.

In the fitting process of the electron density maps, we used the sequence of the *Rhodopseudomonas palustris* cytochrome c\(_2\) from the 2.1.6 strain (Ambler et al. 1979) as starting reference of the primary sequence for the strain 42 OL.
During manual model inspection and refinement cycles, a total of 16 residues have been mutated with respect to this starting sequence. The sequence changes are (with the starting assignment in brackets) Lys4(Ala), Ala5(Lys), Ala29(Gly), Ala60(Gln), Asp61(Glu), Val64(Ile), Pro65(Ala), Ala68(Pro), Phe73(Tyr), Glu80(Asp), Lys83(Gln), Gln86(Lys), Val88(Thr), Val90(Cys), Glu100(Asp) and Val107(Ala). Thus, the resulting primary sequence of the protein obtained from the strain 42 OL has a much higher identity with that of the 2.1.37 strain protein sequence. In fact, only five amino acids are different in the two primary sequences: Ala29(Gly), Val64(Ile), Pro65(Asn), Ala68(Asn) and Glu80(Asp). The primary sequence of *Rhodopseudomonas palustris* cytochrome c2 strain 42 OL, aligned with sequences of the protein obtained from the 2.1.37 and 2.1.6 strains (Ambler et al. 1979), is presented in Figure 1.12. The ambiguities on Asx, Glx, Thr-Val residues were resolved by checking the H-bond network. Only the ambiguities for Asx61, Glx80, Asx85 residues could not be definitively resolved. They were assigned using a conservative approach with respect to the amino acid sequences of the proteins obtained from strains 2.1.6 and 2.1.37.

The post-translational conversion of the N-terminal glutamine residue to pyroglutamic acid was clearly detected on the electron density map. It is a very common type of blockage that occurs when a terminal glutaminyl or glutamyl residue cyclizes, either spontaneously or enzymatically, to yield a pyrrolidonecarboxyl residue (Doolittle 1972).

The purified *Rhodopseudomonas palustris* cytochrome c2 used for crystallisation was a mixture of several mass forms, as determined by mass-spectrometry (Figure 1.13). The main peak of the mass-spectra (12,783.5 Da), is in good agreement with the value of 12,784.5 Da, calculated from the primary sequence obtained by X-ray analysis. The second peak (12,800.2 Da), which differs +17 Da from the highest peak, can be attributed to the protein without the post-translational modification to the N-terminal glutamine residue. Similarly, the following two peaks (12,820.6 and 12,837 Da), with a mass difference of 17 Da, are probably due to the same kind of N-terminal blockage. These two peaks are related to the previous ones by a +37 difference (12,738.5 → 12,820.6 Da, ΔDa = 37.1; 12,800.2 → 12,837.2 Da, ΔDa = 37.0), but it is difficult to assert if these
multiple mass forms correspond to different isoforms or adducts (e.g. the chloride ion adduct should have a +36.5 Da mass). While the presence of several isoforms is commonly found in the structurally related cyanobacterial cytochrome $c_6$ (Kerfeld 1997), it is quite unusual however to find multiple mass forms in mass spectra of bacterial cytochromes $c_2$ and mitochondria cytochromes $c$.

Figure 1.13. Mass spectrometry analysis of *Rhodopseudomonas palustris* cytochrome $c_2$ isolated from 42 OL strain. Reconstructed mass value (Da) are reported for each forms of cytochrome $c_2$ together with mass differences ($\Delta$ Da) from the main peak corresponding to N-terminally modified (Gln to pyroglutamic acid) cytochrome $c_2$. 
4.2 Structural analysis of the oxidised and reduced forms

Packing

The crystal structures of oxidised and reduced cytochrome c\textsubscript{2} from the *Rhodopseudomonas palustris* (Geremia et al. 2002) are isomorphous and four independent polypeptide chains (called molecules A, B, C and D; see Figure I.14) are present in the asymmetric unit. In addition, two sulphate ions, one glycerol molecule and 652 water molecules are detected in the asymmetric unit of the oxidised structure, and a total of 613 water molecules in the reduced form.

Each independent polypeptide chain has a distorted octahedral environment oriented along the crystallographic axes, formed by six globular protein chains. The distorted primitive cubic arrangement of the crystal packing is stabilised by several interchain H-bonds. In particular, along the unique axis $b$, two H-bond schemes are observed in all the four crystallographically independent molecules and involve the Glu7 OE2 and the Gly82 O of one molecule with the Gln86 NE2 and the Tyr109 OH of that related by the screw axis, respectively. Along the $a$ axis, an intermolecular H-bond scheme, common to Mol A-Mol B and Mol C-Mol D and involving the Gly55 O and Thr15 OG1 atoms, is present on both sides of the molecules. In the same zone near the exposed edge of the heme prosthetic group, hydrophobic interactions between polypeptide chains occur. The exposed heme edge and the high conservative residue Phe95 of one molecule, are at Van der Waals contact with Val57 and Thr59 of another one not related by crystallographic symmetry. Although c-type cytochromes are assumed to function as monomer, the X-ray structures of several chloroplast cytochrome c\textsubscript{6} suggested that oligomerisation may have a functional significance (Kerfeld et al. 1995; Kerfeld 1997). Even if there is no biochemical evidence that bacterial cytochromes c\textsubscript{2} form oligomers, significant contacts between monomers are
Results and discussion

present about the heme crevice in the *Rhodopseudomonas palustris* cytochrome c₂ crystal structures.

![Arrangement of the four crystallographically independent polypeptide chains in the unit cell.](image)

**Figure 1.14.** Arrangement of the four crystallographically independent polypeptide chains in the unit cell.

The two crystallographically independent co-crystallized sulphate ions are located near Mol A and Mol C, and interact with the Arg18 side chains. Of the 653 water molecules located in the oxidized structure and of the 613 in the reduced structure, 328 are related by non-crystallographic symmetry in all the four subunits.
Molecular structures of the oxidised and reduced form

Each protein unit is made up by 114 residues and one prosthetic group. The electron density maps, corresponding to the polypeptide chain backbones and to the side chains of amino acids are well definite and allowed the unambiguous recognition of the post-translational conversion of the N-terminal glutamine residue to the pyroglutamic acid. Application of non-crystallographic symmetry restraints to the four independent polypeptide chains improves the resulting Fourier maps. These look as having the resolution better than were actually used, because the phases are more accurate.

The four independent peptide chains have the same folding, with the exclusion of Thr94 and Phe95 residues in both reduced and oxidised state and of Lys21, Ala87, Val88 and Gly89 residues in the reduced state, as shown in the Kleywegt NCS plots (Kleywegt and Jones 1996) of the main-chain $\phi$ and $\psi$ angles (Figure I.15). In the oxidized form, the $\alpha$-carbon atoms of Mol A superimposes on those of Mol B, Mol C and Mol D with a root mean square (r.m.s.) deviation of 0.17, 0.06 and 0.10 Å respectively. A similar behavior is found also in the reduced form. For the sake of simplicity, the following discussion on protein folding will refer only to Mol A of the oxidised state unless otherwise indicated.
Figure 1.15. Kleywegt NCS plot of the main-chain torsion angles $(\phi, \psi)$ of the native *Rhodopseudomonas palustris* cytochrome c$_2$ structure in the oxidized (a) and reduced (b) state. The NCS-related residues are joined by straight lines.
Results and discussion

Figure 1.16. Cartoon representation of the polypeptide chain backbone of the native *Rhodopseudomonas palustris* cytochrome c2 structure, with heme and iron ligands in stick. The elements of secondary structure of the protein are five α helices (A, B, C, D, F), one 3\(\frac{1}{10}\) helix (E), six β turns and three γ turns.

Table 1.3. List of secondary structural elements found in the native form of *Rhodopseudomonas palustris* cytochrome c2.

<table>
<thead>
<tr>
<th>Amino acid residues</th>
<th>Structural element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala3-Cys13</td>
<td>α-Helix (A)</td>
</tr>
<tr>
<td>Met14-His17</td>
<td>β-Turn (type I)</td>
</tr>
<tr>
<td>Arg18-Lys21</td>
<td>β-Turn (type I)</td>
</tr>
<tr>
<td>Met23-Gly25</td>
<td>γ-Turn</td>
</tr>
<tr>
<td>Leu28-Val31</td>
<td>β-Turn (type II)</td>
</tr>
<tr>
<td>Val31-Arg34</td>
<td>β-Turn (type II)</td>
</tr>
<tr>
<td>Lys35-Thr38</td>
<td>β-Turn (type II)</td>
</tr>
<tr>
<td>Ala39-Phe42</td>
<td>β-Turn (type II)</td>
</tr>
<tr>
<td>Pro46-Glu53</td>
<td>α-Helix (B)</td>
</tr>
<tr>
<td>Thr59-Asp61</td>
<td>γ-Turn</td>
</tr>
<tr>
<td>Ile63-Ala68</td>
<td>α-Helix (C)</td>
</tr>
<tr>
<td>Pro70-Glu80</td>
<td>α-Helix (D)</td>
</tr>
<tr>
<td>Lys81-Lys83</td>
<td>γ-Turn</td>
</tr>
<tr>
<td>Ala84-Ala87</td>
<td>3(\frac{1}{10})-Helix (E)</td>
</tr>
<tr>
<td>Glu100-Thr112</td>
<td>α-Helix (F)</td>
</tr>
</tbody>
</table>
A cartoon representation of the polypeptide chain backbone, with heme and iron ligands in stick, is illustrated in Figure 1.16. The polypeptide chain of this 'native form' is composed by five α-helices, one 3_{10}-helix and tight turns that wrap around the prosthetic group (Table I.3). The B and C edges of the heme group are covalently bound to Cys13 and Cys16 residues, respectively. The Cys13 O forms two H-bonds with Cys16 N and His17 N. The His17 residue is the fifth iron-coordinating ligand, while the sixth ligand, Met93, is part of a random coil stretching from Val88 to Asn99.

The Ala84-Asp85-Gln86-Ala87 residues form the unusual isolated 3_{10}-helix secondary structural element. These residues belong also to the rare six-residue insertion Lys83-Val88 that precedes the heme binding methionine. This uncommon insertion, has been recently detected also in the primary sequence of cytochromes c_{2} Rhodospirillum centenum (Samyn et al. 1998). The 3_{10}-helix presents two H-bonds, one between Lys83 O and Gln86 N, and one between Ala84 O and Ala87 N. An isolated β-bridge connects Lys35 and Val57 residues, welding two polypeptide stretches located within a region rich in turns (Leu28-Phe42) and between B and C helices. The overall structure has four proline residues. The CO group of the Pro26 forms an important H-bond with His17 ND1, determining the characteristic orientation of the imidazole plane with respect to the heme moiety (Figure 1.16). The Pro46 and Pro70 residues initiate the B and D helices, respectively, while the Pro65 residue takes part in C-helix.

All non-Gly residues fall within the allowed regions of the Ramachandran plot (Figure I.15), except Met23 (ϕ = -147; ψ = -125) and Asp69 (ϕ = -167; ψ = 91) residues. The Asp69 residue links the C and D helices. The Met23 residue takes part in a Met23-Gly25 γ-turn, which is directly in contact with the D heme edge on the histidine ligand side (Figure I.17). This γ-turn is stabilized by two H-bonds, the first involving Met23 N and Gly25 O, the second between Gly25 N and Cys16 O (Milner-White et al. 1988). Several cytochrome c_{2} structures show a residue exposed to the solvent with a similar backbone conformation, approximately in the same position below the C heme edge, on the histidine side (Table I.4). This position is frequently occupied by a Lys residue. It has been
Results and discussion

proposed that the additional hydrogen bond made by lysine Ne gains the stability of this strained folding (Sogabe and Miki 1995; Benning et al. 1996).

Figure 1.17. Stereo front view of the heme edge exposed to the solvent in the oxidised native form of *Rhodopseudomonas palustris*. The external Met23 residue involved in a strained folding near the exposed edge of the heme, important nearby residues and the co-crystallised sulphate ion are shown.

In cytochrome c₂ from *Rhodobacter capsulatus* this position is occupied by a Gly residue that shows similar main chain torsion angles ($\phi = -141; \psi = -145$; Benning et al. 1991). The Gly residue, that essentially lacks in a side chain, can confer a high degree of local flexibility to the polypeptide chain and can allow this conformation. This local strain was observed also in several mitochondrial cytochromes c, where the key residue is a Lys, in chloroplast cytochromes c₆ and in other bacterial c-type cytochromes (Table I.4). In cytochrome c₂ from *Rhodopseudomonas palustris*, a methionine residue, whose side chain has a relatively strong steric hindrance and is unable to form strong H-bonds, occupies...
this position. Furthermore, the side chain of Met23 assumes two different orientations in Mol B of the oxidised state. The presence of a residue, near the exposed edge of the heme, with the main-chain torsion angles $\phi$ and $\psi$ of about $-140^\circ$ and $-130^\circ$ seems a typical feature of c-type cytochromes, and it could be correlated to the interaction with proteins involved in the electron transfer process.
### Results and discussion

Table I.4. Comparison of the main chain torsion angles, $\Phi$ and $\Psi$ (deg), of the Met23 residue in *Rhodopseudomonas palustris* cytochrome $c_2$ with those found in several X-ray structures of bacterial cytochromes $c_2$, mitochondrial cytochromes $c$, chloroplast cytochromes $c_6$ and other bacterial $c$-type cytochromes for the analogous residue exposed to the solvent in the approximately same position below the C heme edge, on the histidine side. The PDB code of the structures and the relative resolution limit are also reported.

<table>
<thead>
<tr>
<th>Cytochrome</th>
<th>c-type</th>
<th>PDB code</th>
<th>Res. (Å)</th>
<th>Residue</th>
<th>$\Phi$</th>
<th>$\Psi$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Rps. palustris</em></td>
<td>$c_2$</td>
<td>1fj0</td>
<td>1.7</td>
<td>Met23</td>
<td>-147</td>
<td>-125</td>
</tr>
<tr>
<td></td>
<td>$c_2$</td>
<td>1i8o</td>
<td>2.0</td>
<td>Met23</td>
<td>-153</td>
<td>-134</td>
</tr>
<tr>
<td></td>
<td>$c_2$</td>
<td>1i8p</td>
<td>1.2</td>
<td>Met23</td>
<td>-141</td>
<td>-131</td>
</tr>
<tr>
<td><em>Rps. viridis</em></td>
<td>$c_2$</td>
<td>1co6</td>
<td>1.6</td>
<td>Lys26</td>
<td>-126</td>
<td>-132</td>
</tr>
<tr>
<td><em>Rb. capsulatus</em></td>
<td>$c_2$</td>
<td>1c2r</td>
<td>2.5</td>
<td>Gly30</td>
<td>-141</td>
<td>-145</td>
</tr>
<tr>
<td><em>R. globiformis</em></td>
<td>$c_2$</td>
<td>1hro</td>
<td>2.2</td>
<td>Lys31</td>
<td>-141</td>
<td>-134</td>
</tr>
<tr>
<td><em>P. denitrificants</em></td>
<td>$c_2$</td>
<td>1cot</td>
<td>1.7</td>
<td>Lys34</td>
<td>-117</td>
<td>8</td>
</tr>
<tr>
<td><em>Rb. sphaeroides</em></td>
<td>$c_2$</td>
<td>1cxa</td>
<td>2.2</td>
<td>Lys35</td>
<td>-110</td>
<td>-77</td>
</tr>
<tr>
<td></td>
<td>$c_2$</td>
<td>1cxb</td>
<td>2.0</td>
<td>Lys35</td>
<td>-112</td>
<td>-99</td>
</tr>
<tr>
<td></td>
<td>$c_2$</td>
<td>1cxc</td>
<td>1.6</td>
<td>Lys35</td>
<td>-123</td>
<td>18</td>
</tr>
<tr>
<td><em>Rs. rubrum</em></td>
<td>$c_2$</td>
<td>3c2c</td>
<td>1.7</td>
<td>Lys27</td>
<td>-66</td>
<td>-69</td>
</tr>
<tr>
<td></td>
<td>$c_2$</td>
<td>2c2c</td>
<td>2.0</td>
<td>Lys27</td>
<td>-73</td>
<td>-63</td>
</tr>
<tr>
<td><em>Yeast iso-I</em></td>
<td>c</td>
<td>1ycc</td>
<td>1.2</td>
<td>Lys27</td>
<td>-124</td>
<td>-128</td>
</tr>
<tr>
<td><em>Rice</em></td>
<td>c</td>
<td>1ccr</td>
<td>1.5</td>
<td>Lys35</td>
<td>-132</td>
<td>-143</td>
</tr>
<tr>
<td><em>Tuna</em></td>
<td>c</td>
<td>3cyt</td>
<td>1.8</td>
<td>Lys27</td>
<td>-105</td>
<td>-132</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>5cyt</td>
<td>1.5</td>
<td>Lys27</td>
<td>-71</td>
<td>-59</td>
</tr>
<tr>
<td><em>Horse</em></td>
<td>c</td>
<td>1hre</td>
<td>1.9</td>
<td>Lys27</td>
<td>-105</td>
<td>-136</td>
</tr>
<tr>
<td><em>Bacil. pasteurii</em></td>
<td>$c_6$</td>
<td>1b7c</td>
<td>1.0</td>
<td>Ala22</td>
<td>-146</td>
<td>-112</td>
</tr>
<tr>
<td><em>Mon. brunii</em></td>
<td>$c_6$</td>
<td>1ctj</td>
<td>1.1</td>
<td>Gly23</td>
<td>-91</td>
<td>-160</td>
</tr>
<tr>
<td><em>Scen. obliquis</em></td>
<td>$c_6$</td>
<td>1c6r</td>
<td>1.9</td>
<td>Asn23</td>
<td>-99</td>
<td>-154</td>
</tr>
<tr>
<td><em>Chlam. reinhardtii</em></td>
<td>$c_6$</td>
<td>1cyi</td>
<td>1.9</td>
<td>Asn23</td>
<td>-137</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>$c_6$</td>
<td>1cyj</td>
<td>1.9</td>
<td>Asn23</td>
<td>-151</td>
<td>101</td>
</tr>
<tr>
<td><em>P. denitrificants</em></td>
<td>$c_{552}$</td>
<td>1q13</td>
<td>1.4</td>
<td>Gly25</td>
<td>-145</td>
<td>-125</td>
</tr>
<tr>
<td></td>
<td>$c_{552}$</td>
<td>1q14</td>
<td>1.5</td>
<td>Gly25</td>
<td>-145</td>
<td>-121</td>
</tr>
<tr>
<td><em>P. aeruginosa</em></td>
<td>$c_{551}$</td>
<td>351c</td>
<td>1.6</td>
<td>Met22</td>
<td>-113</td>
<td>-101</td>
</tr>
<tr>
<td></td>
<td>$c_{551}$</td>
<td>451c</td>
<td>1.6</td>
<td>Met22</td>
<td>-112</td>
<td>-98</td>
</tr>
</tbody>
</table>

41
Another interesting aspect that comes out from the Kleywegt NCS plot (Figure I.15) is the different conformation of the backbone observed for the Thr94-Phe95 residues in the four crystallographically independent molecules for both oxidation states. These exposed residues are located above the heme plane, on the methionine side (Figure I.17), with Phe95 involved in hydrophobic intermolecular interaction between couples of crystallographically independent polypeptide chains as described above. Local conformational changes involving the Thr94-Phe95 backbone could be important in optimising the interaction of this cytochrome c₂ with the reaction centre.

The intermolecular hydrophobic interactions of the exposed C heme edge and Phe95 residue involve the Val57 and Thr59 residues. These residues are connected along the backbone by the Trp58 residue. The buried side chain of this Trp58 residue points towards the propionate group of the heme A edge forming a hydrogen bond with O1A. This Trp58 together with Leu56, Val57, and Ala36 makes a hydrophobic channel that connects the prosthetic groups of two close molecules.

The Arg18-Lys21 type-I β-turn, which precedes the Met23-Gly25 γ-turn, is exposed to the solvent and is located on the same side of the γ-turn with respect to the prosthetic group (Figure I.17). In Mol A and Mol C of the oxidised form, this β-turn and the Arg18 side chain interact with a co-crystallised sulphate ion by means of the H-bond donors Asp20 N, Lys21 N, Arg18 NE and Arg18 NH₂ (Fig. 4). The close position of Arg18 and Lys21 side chains makes a positively charged surface which promotes the interaction with the negatively charged surface of the sulphate ion. This observation confirms the role of specific basic amino groups, located near the heme crevice, in facilitating the binding and orientation of cytochromes c₂ to photosynthetic reaction centre, which exposes negatively charged carboxylate groups on its surface (Long et al. 1989). On the protein surface, several charged residues (Lys11, Asp20, Lys21, Glu53, Lys76, Lys96, Lys114) show a different conformation of the side chain in the four independent molecules.
The glycerol molecule of the oxidized form is located between Mol C and Mol D, near the Lys11 side chain of Mol C and the Lys35 side chain of Mol D and it forms H-bonds with Ala8D O, Gln12D NE2 and Lys35C NZ. We do not know the functional role, if any, of the post translational conversion of the N-terminal glutamine residue to the pyroglutamic acid.

328 water molecules, out of the 653 detected in the oxidised form and 613 detected in the reduced form, are related by non-crystallographic symmetry (in all the four subunits). The use of NCS restraints during refinement considerably improves the effective data to parameter ratio. As a consequence, the phases are more accurate and the resulting Fourier maps often look as though they were calculated with higher resolution data than those actually used.

Three of these water molecules are buried within the interior of each globular protein molecule. The Wat1 molecule forms three H-bonds with Asn48 ND2, Tyr66 OH and Thr91 OG1. The Wat2 molecule is located near the propionate group bonded to the heme pyrrolic ring A and forms four H-bonds with propionate O2A, Arg34 NE, Lys35 O and Thr38 N. The Wat2 and the Wat1 molecules are evolutionarily conserved among eukaryotic cytochromes c (Takano and Dickerson, 1981a,b; Loie et al. 1990; Bushnell et al. 1990). The Wat3 molecule is located close to the Met23-Gly25 β-turn and forms three H-bonds with Arg18 N, Lys21 O and Pro25 O. The latter water molecule, bridging two carbonyl groups, stabilises the conformation of this strained portion of the polypeptide chain.

Redox-coupled structural changes

There has been considerable crystallographic work done in detecting redox-coupled structural changes between reduced and oxidised c-type cytochromes (Takano and Dickerson 1980; Matsura et al. 1982; Berghuis and Brayer 1992; Berghuis et al. 1994a,b; Schnackenberg et al. 1999; Harrenga et al. 2000). The high-resolution determinations of tuna (Takano and Dickerson 1980) and yeast
(Berghuis and Brayer 1992) cytochromes c in both oxidation states have revealed that the conformational differences between the reduced and oxidized state of mitochondrial cytochrome c are very small. In particular, the differences are expressed for the most part in terms of thermal-parameter changes, small adjustment around the heme moiety, movement of internal water molecules, rather than in terms of explicit polypeptide-chain shifts, which are found to be minimal. The three-dimensional structures, in the two redox states, of bacterial c-type cytochromes, cytochrome c₅₅₁ from *Pseudomonas aeruginosa* (Matsuura et al. 1982), cytochrome c₅₅₂ from *Paracoccus denitrificans* (Harrenga et al. 2000) and cytochrome c₂ from *Rhodobacter rubrum* (Salemme et al. 1973a) revealed conformational differences even smaller. The X-ray structures of both redox states of the analogue chloroplast cytochrome c₆ from the green alga *Scenedesmus obliquus* (Schnackenberg et al. 1999) revealed only slight differences on the protein surface and no movement of internal waters.

The present structures of *Rhodopseudomonas palustris* cytochrome c₂ (Geremia et al. 2002) confirm that the conformational differences between the two redox forms, if any, are small. The r.m.s. difference of Cα positions is 0.14 Å (0.11 Å for molecule A; 0.22 Å for molecule B; 0.13 Å for molecule C; 0.11 Å for molecule D). A significant structural change occurs only in the mol B where a flip of Val88 shifts the backbone of the Asp85-Val90 region. The movement of the Cα position in this backbone region is 0.64 Å. Only minor structural changes can be observed for the other molecules and most differences are even smaller than those observed already between crystallographically independent molecules. The heme iron displacement, the geometry of the heme group and the conformation of the ligands seem scarcely affected by the iron oxidation state of the cytochrome c₂. In agreement with the previous crystallographic works on c-type cytochromes, no large variation of the iron axial coordination distances is observed. The mean value of the iron-nitrogen bond lengths to His17 is 2.02 Å in the oxidised state and 1.99 Å in the reduced state, with both relative standard error of the mean of 0.04 Å, while the mean value of the iron-sulphur distances to Met93 is 2.33 and 2.39 Å, with standard error of the mean of 0.02 and 0.03 Å in the oxidised and reduced state, respectively.
Results and discussion

The B-factors of the peptide chain atoms in the reduced form are higher than those in the oxidised one. Furthermore, the analysis of the non-crystallographic symmetry reported in Figure I.15 suggests that the structure of oxidised cytochrome c₂ is more rigid than that of the reduced form. This is in contrast with previous findings that the peptide chain of oxidised cytochromes c is more flexible in some regions with respect to the reduced forms (Berghuis and Brayer 1992; Harrenga et al. 2000).

4.3 Relationship between hydrogen-bonding network and reduction potential

Wat1 molecule is the highly evolutionarily "conserved" solvent molecule in both prokaryotic cytochromes c₂ and eukaryotic cytochromes c. The structural comparison by superimposition of the local environment around the Wat1 molecule of the oxidised and reduced forms of eukaryotic cytochrome c from tuna and of the prokaryotic cytochromes c₂ from Rhodopseudomonas palustris and Blastocloris viridis is reported in Figure I.18. These cytochromes, which have the same residues involved in the H-bond network with Wat1, have been chosen in order to allow a homogeneous comparison. In oxidised forms of cytochromes with low reduction potential the Wat1 molecule is found in a position different from that found in the reduced forms. On the contrary, in the oxidised form of cytochrome c₂ from Rhodopseudomonas palustris, this water molecule is detected in a position close to that found in the reduced forms. This "conserved" water molecule appears to play an important role in defining the value of the reduction potential.

Recently, thermodynamic studies suggest that the higher reduction potential of the bacterial species appears to be determined entirely by a greater enthalpic stabilisation of the reduction state (Battistuzzi et al. 1997) and that the entropic term is comparable in the two oxidation states (Battistuzzi et al. 1999). In order to
support the role of the highly evolutionarily “conserved” solvent molecule in the modulation of c-type cytochrome reduction potential, a raw evaluation of the energy of the H-bond network involving the Wat1 molecule, has been performed. In Table I.5 the distances between H-bond donor and acceptor atoms are reported together with the relative electrostatic contribution of the interactions between Wat1 and the three amino acids. The electrostatic contribution was obtained by molecular mechanics calculations.

The shortening of the H-bond distances in the oxidised forms of *tuna* cytochrome c and *Blastocloris viridis* cytochrome c2, with respect to those of the corresponding reduced forms, leads to a stabilisation in energy of 4.1 and 2.5 kcal/mole, respectively. This gain in energy corresponds to a decrease in the reduction potential of about 180 mV for the cytochrome c, and 110 mV for the cytochrome c2. The oxidised form of cytochrome c2 from *Rhodopseudomonas palustris* shows H-bond distances and relative electrostatic interaction energies similar to those of the reduced forms. This is confirmed also by the calculation on the reduced form of analogue *Rhodopila globiformis* cytochrome c2 (structure at 2.2 Å resolution; Benning et. al 1996). This gives a value of -16.7 Kcal/mol for the electrostatic interaction of the “conservative” water molecule with the aminoacidic triad ( Wat1-Tyr 2.80 Å; Wat1-Asn 3.11 Å; Wat1-Thr 2.79 Å). The weakness of these H-bonds with respect to the H-bonds found in the oxidized form of cytochrome c and cytochrome c2 from *Blastocloris viridis*, is consistent with the increase in the reduction potential value. It should be noted that no redox-dependent positional change of internal water molecules has been observed also in the X-ray structures of cytochrome c6 from *Scenedesmus obliquus* (Schnackenberg et al.1999). These findings suggest that the movement of the buried solvent molecule lowers by about 100 mV the reduction potential through the stabilisation of the oxidised form.

Why the buried molecule found in *Rhodopseudomonas palustris* structures does not change its position with change of the redox state is the outstanding question. To answer this question we must look at the H-bond network that links the water molecule to the heme iron (Figure I.19). The different behaviour of the water molecule in *Rhodopseudomonas palustris* may be associated to the hydrogen
bond distance between the hydroxyl group of the conservative Tyr residue and the sulphur atom of the Met axial ligand. In the cytochrome $c_2$ from *Rhodopseudomonas palustris* this distance is 3.34 Å in the oxidised form and 3.37 Å in the reduced one, significantly longer of those found in the oxidised and in the reduced forms of *Blastocloris viridis* cytochrome $c_2$ and of *tuna* cytochrome $c$ (that ranges from 3.12 to 3.18 Å) (Figure I.19). The increase of this H-bond distance could be responsible of the low mobility of the buried water molecule with the change of the protein redox state in the *Rhodopseudomonas palustris* cytochrome $c_2$. In fact, the partial interruption of the H-bond network could disconnect the "communication" between the metal centre and the water molecule. As a consequence, its position would not be influenced anymore by the iron oxidation state (Garau et al. 2002).
Results and discussion

Figure I.18. Stereo view of the superimposition of the local environment around the Wat1 molecule of the reduced and oxidised form of eukaryotic cytochrome from *tuna* and of prokaryotic cytochromes c2 from *Blastocloris viridis* and from *Rhodopseudomonas palustris*. 
Table I.5. H-bond distances (Å) of the conservative water molecule (Wat1) detected in the reduced and oxidized form of prokaryotic cytochromes c_{2} from *Rhodopseudomonas palustris* and from *Blastocloris viridis* and of eukaryotic cytochrome c from *tuna*. The resolution limit (Å) of structures, the B factors (Å^2) of the Wat1 molecule and the electrostatic contribution (kcal/mol) of the interactions between this water molecule and the three amino acids is also reported.

<table>
<thead>
<tr>
<th></th>
<th>Resolution (Å)</th>
<th>Bfactor</th>
<th>Distance from Wat1</th>
<th>Eot (Kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tyr OH</td>
<td>Asn ND2*</td>
</tr>
<tr>
<td><em>Rps. palustris</em> (ox) Mol A</td>
<td>1.7</td>
<td>10.4</td>
<td>2.84</td>
<td>2.99</td>
</tr>
<tr>
<td>Mol B</td>
<td>9.9</td>
<td>2.73</td>
<td>3.01</td>
<td>2.92</td>
</tr>
<tr>
<td>Mol C</td>
<td>10.1</td>
<td>2.82</td>
<td>2.95</td>
<td>2.87</td>
</tr>
<tr>
<td>Mol D</td>
<td>10.2</td>
<td>2.77</td>
<td>3.01</td>
<td>2.90</td>
</tr>
<tr>
<td>Mean value (esd)</td>
<td>10.2(2)</td>
<td>2.79(4)</td>
<td>2.99(2)</td>
<td>2.90(2)</td>
</tr>
<tr>
<td><em>Rps. palustris</em> (red) Mol A</td>
<td>1.95</td>
<td>7.6</td>
<td>2.78</td>
<td>3.06</td>
</tr>
<tr>
<td>Mol B</td>
<td>9.9</td>
<td>2.98</td>
<td>2.92</td>
<td>2.80</td>
</tr>
<tr>
<td>Mol C</td>
<td>7.6</td>
<td>2.88</td>
<td>3.05</td>
<td>2.85</td>
</tr>
<tr>
<td>Mol D</td>
<td>10.1</td>
<td>2.84</td>
<td>3.02</td>
<td>2.78</td>
</tr>
<tr>
<td>Mean value (esd)</td>
<td>9.1(11)</td>
<td>2.87(6)</td>
<td>3.01(5)</td>
<td>2.81(2)</td>
</tr>
<tr>
<td><em>Rps. viridis</em> (ox)</td>
<td>2.0</td>
<td>16.4</td>
<td>2.70</td>
<td>2.72</td>
</tr>
<tr>
<td><em>Rps. viridis</em> (red)</td>
<td>1.6</td>
<td>11.2</td>
<td>2.85</td>
<td>3.03</td>
</tr>
<tr>
<td><em>Tuna</em> (ox)</td>
<td>1.8</td>
<td>20.1</td>
<td>2.62</td>
<td>2.84</td>
</tr>
<tr>
<td><em>Tuna</em> (red)</td>
<td>1.5</td>
<td>12.2</td>
<td>2.87</td>
<td>2.96</td>
</tr>
</tbody>
</table>

* OD1 in oxidized form of *tuna* cytochrome c.

Figure I.19. H-bond network that links the water molecule (Wat1) to the heme iron.
4.4 Structural analysis of the ammonia complex

Molecular structure of the alkaline form

The structure of the native form of Rhodopseudomonas palustris cytochrome c2 (Figure I.16), obtained from crystals grown at pH 4.4 has been described in the previous sections, and details of the structure compared with other c-type cytochromes.

The structure of Rhodopseudomonas palustris cytochrome c2 (Geremia et al 2002), crystallised at pH 8.5 (alkaline form), is represented in Figure I.20. The overall ternary protein structure of this alkaline form is similar to that of the native form except the 310 helix and the loop involving Met93.

In brief, the globular structure of the protein is composed by five α-helices, one 310 helix, six β turns, three γ turns and extended chains that wrap around the heme group. The unusual isolated 310-helix secondary structural element, which precedes the Met93 ligand, is not present in any other resolved structure of cytochromes c-type. It is formed by the rare six-residue insertion, Gly82-Ala87 (Samyn et al. 1998). The prosthetic group is covalently bonded to the polypeptidic chain via two thioether linkages formed by addition of two cysteinate side chains to the vinyl groups of the protoporphyrin IX. The porphyrin moiety is located close to the N-terminus of the protein chain and is surrounded by a hydrophobic sheath. The co-ordination distance of the His ligand in this low-spin cytochrome c2 complex is 2.01 Å, close to that found in the native form. As usually found, the heme group is not completely planar, but is distorted into a saddle shape. This equatorial ligand distortion is very similar within the experimental error to that of the native form. This suggests that the axial methionine ligand (see below) has no role on the deformation of the porphyrine moiety.

The Cα atoms of the native and alkaline forms superimpose with a root-mean-square deviation (rmsd) of 0.7 Å. With respect to the native form, in the
alkaline form the peptide bond of Lys92 is flipped and this forces the successive residues to turn away from the core of the protein (Figure 1.20).

Figure 1.20. View of a ribbon diagram of the polypeptide chain backbone of *Rhodopseudomonas palustris* cytochrome c₆ ammonia complex. The heme group, Cys13 and Cys16, the iron axial ligands (His17, NH₃), and Lys92-Phe95 residues are in ball and stick. The elements of secondary structure of the protein are five α helices (green), one 3₁₀ helix (purple), six β turns and three γ turns.
Results and discussion

Figure 1.21. Ramachandran plot of the main-chain torsion angles \((\phi, \psi)\) of the region involved in the conformational transition between the two forms of *Rhodopseudomonas palustris* cytochrome \(c_2\). The residues of the native form (full circles) are joined by straight lines with the same residue of the ammonia complex (triangles).

The sulphur atom of Met93 does not co-ordinate the heme iron atom and is located about 10.5 Å away from it. The main chain torsion angles of the region involved in the conformational transition between the two forms (native, alkaline) of *Rhodopseudomonas palustris* cytochrome \(c_2\) are shown in the Ramachandran plot of Figure 1.21.

The modified region of the ammonia complex structure is involved in lattice packing interactions. The displaced Met93 side chain of a protein molecule is in Van der Waals contact with the Phe95 residue of a symmetry related protein molecule.
In spite of the dramatic change in the chemical environment of the prosthetic group (Figure I.22), the residues mainly involved in this conformational transition are only three: Lys92-Met93-Thr94 (Figure I.23a). In particular, the Met93 changes the β-sheet conformation adopted in the native form into the left-handed helical conformation of the ammonia complex and both Lys92 and Thr94 lie in the right-handed α-helix region in the native structure while they adopt the beta-sheet conformation in the present structure.

**Figure 1.22.** Electron density map around the prosthetic group of *Rhodopseudomonas palustris* cytochrome c₂ ammonia complex. The 2Fo-Fc map (contour levels at 3σ) shows the NH₃ exogenous ligand coordinated to the heme iron.
A similar methionine displacement was also observed in the structure of cytochrome c$_2$ from *Rhodobacter sphaeroides*, where an exogenous imidazole molecule binds the heme iron atom (Axelrod et al. 1994). This structure was determined at 2.2 Å resolution with crystals grown in ammonium sulphate and using imidazole as buffer solution at pH 7.0. The residues of the imidazole adduct and the native form of *Rhodobacter sphaeroides* involved in the conformational transition, are depicted in Figure 1.23b for comparison. It is apparent that the conformational change observed in the imidazole derivative of *Rhodobacter sphaeroides* cytochrome c$_2$ is quite different with respect to the *Rhodopseudomonas palustris* cytochrome c$_2$. The conformation change starts at the Met ligand, but the peptide chain of the imidazole complex gradually reverts to the folding of the native backbone after 5-6 residues. Furthermore the side chain of the methionine ligand exhibits a relatively small shift if compared with the ammonia adduct. The trans conformation of the Met side chain, observed in the native forms, is maintained in the imidazole adduct, while in the ammonia complex it is changed to the gauche$^+$ conformation. An opposite behaviour is observed for the side chain conformation of the neighbourhood phenylalanine. In the native forms and in the ammonia adduct it assumes a trans conformation while in the imidazole complex it has a gauche$^+$ conformation. As a consequence, in the imidazole complex the stacking interaction between this phenylalanine side chain and the heme moiety is lost, while it is maintained in the ammonia adduct. It has been proposed that the direct interaction of the phenyl ring of this highly conserved amino acid with the π-electron system of the heme group is crucial for the rate of electron transfer reactions with redox partners (Poulos and Kraut 1980; Pielak et al. 1987; Liang et al. 1988; Louie and Brayer 1989).
Figure I.23. Superimposition of the local environment around the heme group in the native form (light sketch) and ammonia complex (dark sketch) of the cytochrome c₂ from *Rhodopseudomonas palustris* (a); and that in the native form (light sketch) and imidazole complex (dark sketch) of the cytochrome c₂ from *Rhodobacter sphaeroides* (b).
The movement of the methionine ligand, whose side chain is completely turned out towards the external solvent in the ammonia complex, provokes an extended change in the H-bond network around the heme group of the *Rhodopseudomonas palustris* cytochrome c₂ (Figure 1.20). In particular, the Thr91 side chain shows a different rotamer conformation in the two forms (χ₁ is 180° in the native form and 64° in the ammonia adduct). The three hydrogen bonds formed by OH of Thr91 with the highly conserved internal water molecule Wat1, the oxygen OD₂ of the heme propionate group and the N of Lys92, observed in the native form, are replaced by three new hydrogen bonds with O Asn71, OH Thr94 and an internal water molecule Wat5.

The backbone conformation of the key residue, Lys92 is stabilised by the O Lys92–Wat5 hydrogen bond. The Lys92 side chain shows the same gauche+ rotamer conformation in both native and present forms. The backbone of the other residues involved in the conformational transition is further stabilised by two hydrogen bonds (O Met93–Wat4 and N Phe95–Wat4) formed by the water molecule, Wat4, located in the heme pocket.

In the present structure (Figure 1.20), the isolated 3₁₀ helix that precedes the Met93 is one residue shorter than that found in the native form (Figure 1.16). The Val88 residue is the amino acid that causes this shortness. In the native structure, this residue is the last residue of the 3₁₀ helix secondary structure. In the cytochrome c₂-ammonia complex it is flipped with respect to the native form. This change of backbone direction breaks the Asp85–Val88 H-bond and results in a shortening of the 3₁₀ helix one residue earlier than the native form.

There are eight buried water molecules in the present form (Wat1–Wat8) (Figure 1.20), while only three are present in the native form (Figure 1.16). Three of these (Wat1–Wat3) are in the similar positions as the "conserved" water molecules found in the native form obtained at acid pH. Wat4, Wat5 and Wat6 are located in the space cleared by the Met93 side chain, while Wat7 and Wat8 are located near the A propionate group. A summary of the intermolecular hydrogen bond distances formed by buried water molecules is given in Table 1.6. The total number of water molecules is strictly related to the resolution, but we can exclude
that the structure at 1.7 Å resolution has other buried water molecules. Probably, the structure of the ammonia adduct shows a significantly enhanced resolution because of the kinetics of the crystallisation process.

Table 1.6: Intermolecular hydrogen bonds involving the buried solvent molecules and the exogenous axial ligand in the *Rhodopseudomonas palustris* cytochrome c₆ ammonia complex.

<table>
<thead>
<tr>
<th>Atom</th>
<th>Residue and atom involved in hydrogen bonding (H-bond distance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₃</td>
<td>Tyr66 OH (3.17 Å), Wat4 (2.87 Å), Wat6 (3.22 Å)</td>
</tr>
<tr>
<td>Wat1</td>
<td>Tyr66 OH (2.65 Å), Asn48 (2.89 Å), Wat5 (3.00 Å)</td>
</tr>
<tr>
<td>Wat2</td>
<td>HEM O2A (2.65 Å), Thr38 N (3.00 Å), Lys35 (2.88 Å), Wat7 (2.74 Å)</td>
</tr>
<tr>
<td>Wat3</td>
<td>Arg18 N (2.79 Å), Lys21 O (2.73 Å), Gly25 O (2.86 Å)</td>
</tr>
<tr>
<td>Wat4</td>
<td>NH3 (2.87 Å), Met93 O (3.01 Å), Phe95 N (2.94 Å), Wat5 (2.83 Å)</td>
</tr>
<tr>
<td>Wat5</td>
<td>Thr91 OG (2.76 Å), Lys92 O (2.78 Å), Wat4 (3.00 Å), Wat4 (2.83 Å)</td>
</tr>
<tr>
<td>Wat6</td>
<td>Phe95 O (2.79 Å), Leu67 (3.27 Å), NH₃ (3.21 Å)</td>
</tr>
<tr>
<td>Wat7</td>
<td>Val31 O (3.11 Å), Arg34 O (2.80 Å), Lys35 O (3.15 Å), Wat2 (2.74 Å)</td>
</tr>
<tr>
<td>Wat8</td>
<td>HEM O2A (2.94 Å), Ala27 O (2.80 Å), Arg34 NH2 (2.78 Å)</td>
</tr>
</tbody>
</table>

Figure 1.24. Electron density maps around heme moiety of the alkaline form obtained by the first refinement of the naive form starting model. The 2Fₒ-Fₑ map at 2σ (brown) and the Fₒ-Fₑ map at 3σ (cyan) clearly show the displacement of the Met93 ligand by a single non-hydrogen atom in the cytochrome c₆ crystallised at basic pH (alkaline form), with respect to that obtained at acid pH (native form).
Results and discussion

During the refinement of the alkaline form (ammonia adduct) structure from the data set collected at 1.4 Å resolution, we have collected a new complete data set at higher resolution (1.15 Å resolution) with minor modification of the crystallization conditions. The structure of this last form, refined almost at atomic resolution, has permitted the unambiguous assignment of the primary structure of this un sequenced cytochrome obtained from the strain 420L.

From inspection of the first electron density map of the cytochrome c2 crystallised at basic pH, it appeared that the Met93 ligand was displaced by a single non-hydrogen atom (Garau et al. 2000), with respect to that obtained at acid pH (Figure 1.24). The affinity of the iron atom for nitrogen ligands suggested that the sixth axial ligand position is occupied by an ammonia molecule produced by the precipitating agent (NH₄)₂SO₄ at alkaline pH. This assignment is now supported by the thermal factor analysis, the coordination distances and the H-bond scheme of this high resolution structure.

In fact, the final refinement gave a B factor of 6.0 Å² for the axial nitrogen atom. This value is close to those of the iron atom (6.3 Å²), the trans NE2 His (6.5 Å²) and the equatorial pyrrole nitrogen atoms (NA 6.3 Å², NB 6.2 Å², NC 6.6 Å², ND 6.7 Å²). An oxygen atom refined in this site gives a significant higher thermal factor (8.5 Å²). Such thermal factor is unlikely for an exogenous axial ligand also held by three strong H-bonds (Figure 1.25). Furthermore, the axial distance of 2.11 Å found for the exogenous ligand coordination is consistent with a Fe-NH₃ bond length. In fact, a statistical analysis of available small molecule crystal structures reveals that the mean value of the Fe-NH₃ bond length calculated from nine iron-ammonia complexes is 2.10(3) Å, while the corresponding mean value for the Fe-OH bond obtained from thirteen iron-hydroxo complexes is 1.91(2) Å. However, no X-ray structural determination of an ammonia adduct of iron porphyrins has been so far reported. The closest structurally related complex so far characterised is Fe(DBPh₂)₂(NH₃)(Py), where (DBPh₂)₂ is the macrocyclic equatorial ligand bis(dimethylglyoximato-diphenylborylated) (Vernik and Stynes 1996). In this complex the Fe-NH₃ axial distance is 2.035(5) Å. Mononuclear hydroxo-iron(III) porphyrins are normally unstable (Buchler et al. 1982) and the only two examples structurally characterised are pentacoordinated complexes, the hydroxo-iron(III)-
Results and discussion

α,γ bis(tert-butyl)-octaethyl-α,γ dihydroporphyrine (Fe-OH = 1.89(1) Å) (Buchler et al. 1982), and the hydroxo-iron(III) meso-tetraphenylporphyrine (Haryono et al. 1998). Nevertheless, unambiguous assignment of the sixth axial ligand derives from the H-bond network around the heme group. In fact, the H-bond scheme (Figure 1.25) reveals that the exogenous ligand is a three proton donor in the H-bonds that it forms with Tyr66, Wat4 and Wat6.

![Figure 1.25. H-bond scheme around the NH₃ exogenous ligand of Rhodopseudomonas palustris cytochrome c₂ ammonia complex.](image)

Thus, there are three factors supporting the presence of an ammonia molecule in the heme-iron coordination environment: the thermal factor analysis, the coordination distances and the H-bond scheme. The argument regarding the thermal factor is the weaker, but the comparison of the B-factors has been done with respect to the nearby atoms (not only between different axial species, N vs. O) and also considering that the exogenous axial ligand is held by three H-bonds. Even if most of the H-atoms of the prosthetic group were observed in the Fo–Fc maps, the hydrogen atoms of the ammonia moiety cannot be clearly detected. The
H-bond scheme (Figure 1.25) reveals that the exogenous ligand is a three proton donor in the H-bonds that it forms with Tyr66, Wat4 and Wat6. This is consistent with a coordinated ammonia molecule and not with a coordinated water molecule that could act as a two proton donor and one proton acceptor in the H-bonds.

A brief discussion on how the oxidation state influences the crystallization process of the ammonia adduct needs to be reported. At basic pH, crystals grow from reduced protein solution but not from oxidized protein solution, where a protein aggregate is obtained (see crystallization of the alkaline form). A slow oxidation of the protein could be essential for the kinetics of the crystallization. Furthermore, crystals are stable adding ferricianide in the drop containing them, but they dissolve adding ditionite.

It is generally accepted that the Met displacement in ferricytochrome is easier that in ferrocytochrome. However, not only the stability of the Fe-Met bond with respect to the oxidation state of the heme iron is important for the formation of ammonia adduct but also the relative stability of the Fe(II)-NH3 and Fe(III)-NH3 bond. Probably the nature of the Fe-S bond, that has a higher back donation contribution with respect to the σ-donating ammonia ligand, could suggest that the stronger back donation to the sulphur atom, from a Fe(II) with respect to a Fe(III) atom, increases the stability of the Fe-S with respect to the Fe-N in the reduced protein. Anyhow, this may be just speculation with respect to crystal structures refined at 1.7, 1.9 Å of resolution and in absence of the structure of the ammonia-adduct in the reduced form.

Finally, it has been frequently observed that the X-ray beam could produce a change in the oxidation state of the protein. The only conclusive method to check that the exposure of the crystals to the X-ray does not alter the redox state of the protein is the collection of the absorption spectra during data collection. Unfortunately the diffraction data were collected at the Elettra synchrotron where the facilities for this kind of measurements were not available. However, all the previous works on redox states of c-type cytochromes report that the oxidation state is not changed by the X-ray exposure even at room temperature. For example, in the recent paper of Huber R. (Schnackenberg et al. 1999), that reports
the structures of the reduced and oxidised cytochrome c₆ (with redox potential close to that of our cytochrome) obtained at room temperature, is written: “The redox state of cytochrome c₆ in the crystal was obviously not influenced by the X-ray measurements”. Even if we have used a stronger X-ray source we have collected data from crystal frozen at 100 K and surrounded by a solution containing a high concentration of the reducing or oxidising agent. On the other hand, house trials by absorbance spectrometry on crystals in both redox states, before and after X-ray exposure, suggest no influence of the X-ray on the oxidation state of this c-type cytochrome.
5. Conclusions

The complete primary structure of the *Rhodopseudomonas palustris* cytochrome c2 protein, expressed from the strain 42 OL, was obtained by X-ray analysis and checked by mass spectrometry. It confirms that sequence differences between independent strains of the same bacterial are as large as those between mitochondrial cytochrome c of mammals (Ambler et al. 1979).

The X-ray structures of the cytochrome c2 from *Rhodopseudomonas palustris* are the first high-resolution crystal structures, in both oxidation states, of a cytochrome c2 with relatively high redox potential (+365 mV). The differences in the native structure between the two redox states are small. The four crystallographically independent polypeptide chains show an unusual isolated 3_10_-helix secondary structural element formed by the rare six-residue insertion (Gly82-Ala87), which precedes the heme binding Met. The structural comparison of the present cytochrome c2 with other c-type cytochrome s has revealed that the presence of an external residue (Met23), involved in a strained folding near the exposed edge of the heme, is a typical feature of this family of proteins.

The most striking feature of the present structures is the position of the "conservative" water molecule located at the heme-binding pocket. In both redox forms, this water molecule is detected in a position close to that found in the reduced forms of other c-type cytochromes. The weakness of the H-bonds in the oxidised form of *Rhodopseudomonas palustris* cytochrome c2 with respect to the oxidised forms of c-type cytochromes with low reduction potential suggests the importance of this H-bond network in determining the value of the midpoint reduction potential of c-type cytochromes. Recently, the structure determinations of cytochrome c6 from *Scenedesmus obliquus* in both redox forms show no redox-dependent movement of the internal water molecule (Schnackenberg et al. 1999). The midpoint redox potential of cytochrome c6 is usually about 100 mV higher than that of the mitochondrial protein (Kerfeld 1997). These two protein structures with high reduction potential (the present bacterial cytochrome c2 and the chloroplast cytochrome c6), and the electrostatic calculations confirm that
"conservative" buried water molecule drops of about 100 mV the reduction potential of mitochondrial cytochromes c through the stabilisation of the oxidised form.

Finally, it should be noted that the different behaviour of Wat1 in *Rhodopseudomonas palustris* also corresponds to a lengthening of the Tyr66 OH and Met93 SD hydrogen bond distance. In the cytochrome c2 from *Rhodopseudomonas palustris* this distance is 3.34 Å in the oxidised form (mean value with standard error of the mean of 0.02 Å) and 3.37 Å in the reduced one (mean value with standard error of the mean of 0.02 Å), and it is significantly longer than those, ranging from 3.12 to 3.18 Å, found in the oxidised and in the reduced forms of *Blastocloris viridis* cytochrome c2 and of *tuna* cytochrome c. The increase of this H-bond could be responsible of the low mobility of the "conservative" water molecule with the change of the protein redox state in the *Rhodopseudomonas palustris* cytochrome c2. In fact, the partial "interruption" of the H-bond network can disconnect the communication between the metal centre and this water molecule. As consequence, its position is not more influenced by the iron oxidation state.

The most striking feature of the structure obtained at alkaline pH is the modification of the coordination sphere of the iron centre. The sulphur atom of the Met93 axial ligand does not coordinate the heme Fe-atom, but is replaced by an exogenous ammonia molecule. This is the only X-ray structure so far reported of a heme coordinated by an ammonia molecule. The detachment of the methionine ligand is helped by a very localised change in the backbone conformation with the maintenance of the overall protein folding.

It is interesting to note that the crystal structure of *Blastocloris viridis* cytochrome c2, obtained from crystals grown in similar conditions, does not show the Met ligand displacement (Sogabe and Miki 1995). Most likely, the 310-helix formed by the rare six-residues insertion Gly82-Ala87, present in the protein from *Rhodopseudomonas palustris*, increases the Met loop flexibility and consequently the Met-iron bond lability.
Conclusions

There is no significant higher B-factor distribution in the region near the $3_{10}$ helix. It is true that if this region is more flexible it should exhibit high B-factors. However, also the crystal packing can influence the thermal factors. The Met loop that starts from the $3_{10}$ helix, is involved in intermolecular contacts. The comparison between the eight crystallographic independent molecules of the native form (oxidised and reduced states) shows a significant structural change in the Mol B of the reduced form, where a flip of the Val88-Gly89 peptide plane shifts the backbone in the Asp85-Val90 region. In our opinion, this is a clear evidence that this region is particularly flexible.

In addition, comparison between the structures of the native forms of the *Rhodopseudomonas palustris* and *Blastocloris viridis* cytochromes c$_2$ shows a difference in the hydrogen bond distance between the Tyr66 OH and the coordinated SD of the Met93 ligand. In the cytochrome c$_2$ from *Rhodopseudomonas palustris* this distance is 3.34 Å (mean value with standard error of the mean of 0.02 Å), significantly longer than that found in the *Blastocloris viridis* cytochrome c$_2$ (3.12 Å). The H-bond formed by the coordinated Met ligand seems to be very important for the stabilisation of the native form of c-type cytochromes. Also in the native form of *Rhodobacter sphaeroides* cytochrome c$_2$, a very weak H-bond is present (3.48 Å). Probably, the crystal formation of the *Rhodobacter sphaeroides* cytochrome c$_2$ imidazole complex is also favoured by the presence of this weak stabilisation of the Met ligand through the Tyr H-bond.
PART II

Mammalian transcobalamin
1. Introduction

1.1 Vitamin B₁₂ in higher organisms metabolism

Vitamin B₁₂, or cianocobalamin (CNCbl), is not a biologically active specie, whereas methylcobalamin (MeCbl) and 5'-deoxy-5'-adenosylcobalamin (B₁₂ coenzyme, AdoCbl) (Scheme II.1) are important as cofactors for L-methylmalonyl-CoA mutase and methionine synthase in both mammalians and prokaryotes, and for several other enzymes found exclusively in prokaryotes (Ludwig and Evans 1999).

In mammalian, proteins binding B₁₂ can be distinguished in two classes: B₁₂-enzymes and B₁₂ transporter proteins. The B₁₂-dependent enzymes methylmalonyl-CoA mutase and methionine synthase are involved in essential metabolic processes such as the β-oxidation of odd-chain fatty acids (the former), and the synthesis of the methionine amino acid (the latter). Due to its production of tetrahydrofolate, methionine synthase is indirectly involved also in the synthesis of dTMP and thus DNA.

Scheme II.1.

X = OH → Hydroxo-Cbl
X = CN → Cyano-Cbl (vitamin B₁₂)
X = CH₃ → Methyl-Cbl
X = 5'-deoxyadenosine → Ado-Cbl (coenzyme B₁₂)
All of the currently known reactions of B_{12}-dependent enzymes involve the making and breaking of the Co-C bond. AdoCbl is the cofactor in isomerase and mutase enzymes, which catalyze the intramolecular 1,2-shift of a hydrogen and an electronegative X group. An example is methylmalonyl-coenzymeA mutase, which isomerises reversibly the methylmalonyl group to succinyl group. The rearrangement proceeds though a stepwise process initiated by the key homolysis step of the Co-C bond, with formation of two radicals, cob(II)alamin (B_{12r}) and adenosyl. The former is a relatively long-lived species, while the latter rapidly abstracts an H atom from the substrate and then it rearranges to the product restoring the Co-C bond. The MeCbl-based enzymes (methyltransferases) catalyzes the transfer of methyl groups from an N or O atom to cob(I)alamin (to form MeCbl) for onward transmission to an S atom, as in methionine synthase. The overall mechanistic scheme requires a reversible heterolytic cleavage of the Co-Me bond in the methyl carbocation and the strongly nucleophytic cob(I)alamin species. MeCbl also plays a role in the pathway of carbon dioxide fixation in several anaerobic aceto-bacteria, as well as in the reversible pathway, which results in formation of methane from acetic acid (Randaccio 1999).

![Figure 1.1. Structure representation of methionine synthase and methylmalonyl-CoA mutase. The two cofactors (MeCbl and AdoCbl, respectively) are coloured in orange.](image-url)
The X-ray structures of the AdoCbl and MeCbl free cofactors were reported in 1961 (Lenhert and Hodgkin 1961) and 1985 (Rossi et al. 1985), respectively. Before 1994, due to the lack of structural information on coenzyme-enzyme binding, most hypotheses concerning such enzymatic mechanisms, particularly for hydrolysis, were based essentially on a wide-ranging and in depth study on the simple models (Randaccio 1999).

Recently the molecular structures of several B$_{12}$-enzymes have been determined (Drennan et al. 1994; Dixon et al. 1996; Mancia et al. 1996; Shibata et al. 1999) and have furnished new insight into the enzymatic mechanism of MeCbl- and AdoCbl- dependent enzymes. The structure of L-methylmalonyl-CoA mutase enzymes and methionine synthetase active site are shown in Figure II.1.

1.2 Vitamin B$_{12}$-transporters in mammalians

For human the dietary requirement of B$_{12}$ is about 1-2 µg, and vitamin B$_{12}$ is readily available from foods of animal origin. B$_{12}$ deficiency causes fatal bone marrow failure, megaloblastic anemia and demyelinating diseases of the peripheral and central nervous system (Stabler et al. 1990; Li et al. 1993). Highly proliferating tissues (Hogenkamp et al. 1999), such as those of several tumours, are particularly susceptible to cobalamin uptake. Mammalians have developed a specific transportation mechanism (Figure II.2) with three binding proteins, intrinsic factor (IF), transcobalamin (TC) and haptocorrin (HC), for the intestinal absorption, transport and cellular uptake of the vitamin (Glass et al. 1954; DiGirolamo and Huennekens, 1975; Kräutler et al., 1988; Nexø, 1998; Rothenberg et al., 1999).
Released during digestion cobalamin molecules bind to gastric intrinsic factor and the complex IF-Cbl enters the ileal mucosal cells by a receptor mediated mechanism (Allen et al. 1978). Inside the enterocytes cobalamin is transferred from intrinsic factor to TC and transported into portal plasma, where cobalamin complexed to transcobalamin is endocytosed by membrane receptors (Cooper and Paranchych 1961; Schneirder et al. 1976; Bose et al. 1995). Inside the target cells the cobalamin molecules are metabolised to the two cofactors: 5'-deoxyadenosyl-Cbl (coenzyme for methylmalonyl-CoA mutase) and methyl-Cbl (coenzyme for methionine synthetase) (Nexo, 1998). The third Cbl-transporter,
haptocorrin, withdraws occasional Cbl-analogues from plasma circulation by the liver, preventing incorporation of the analogues into tissues, where they may inhibit B₁₂-enzymes (Kolhouse and Allen 1977).

Cbl-transporting proteins consist of a protein core with approximately 400 amino acids. Transcobalamin does not contain any carbohydrates whereas intrinsic factor and haptocorrin are highly glycosylated (15% of carbohydrates in IF and 33-40% in HC) (Alpers and Russel-Jones 1999). Alignment of the amino acid sequence of human transcobalamin with other mammalian transcobalamins (bovine, rat and mouse) shows about 70% identity, whereas the alignment with other Cbl-transporter proteins (human, rat and mouse intrinsic factor, and human and porcine haptocorrin), reveals 26-32% identity, with four regions of greater than 80% homology. These regions encompass the majority of the hydrophobic areas of these proteins and are probably involved in cobalamin binding (Li et al. 1993; Nexø 1998; Fedosov et al. 1999).

Whereas the X-ray structures of several B₁₂-enzymes are available, no structural information on B₁₂-transporting proteins has so far been reported. Recently the cloning of the cDNA encoding human and bovine transcobalamin has provided sufficient quantities of this Cbl-binding protein in pure form for X-ray structure determination (Fedosov et al. 1999; Fedosov et al. 2000). Human and bovine transcobalamins have a core of 409 and 414 amino acids, respectively, and their primary sequences are shown in Figure II.3.

![Figure II.3. Amino acid sequences of human transcobalamin (up) and bovine transcobalamin (down). The first 18 amino acids refer to the signal peptide.](image-url)
3. Material and methods

Preparation of the genetic material for human transcobalamin

The genetic material for expression of human transcobalamin was produced from commercially available human kidney cDNA. The transcobalamin encoding fragment of DNA was amplified by polymerase chain reaction (Quick-Clone preparation) using TC-specific primers and, after electroelution purification on agarose electrophoresis, the major 1.2-kilobase pair DNA band was ligated to Xhol and SpeI sites of the expression plasmid pPIC-Z. The designed sequence of the fusion protein contained the secretion α-signal, the site for yeast protease Kex2, and the mature human transcobalamin. This construction obtained by Fedosov (Fedosov et al. 2000) ensured cleavage of the NH2-terminal peptide from the recombinant protein during its secretion.

Expression and purification of human transcobalamin

I have expressed and purificated human (and bovine) transcobalamin at the protein chemistry laboratory of Prof. Torben Petersen (Science Park, Århus University). Recombinant yeast Pichia pastoris cells SMD 1168 were first grown for 24h in 2x100mL of buffered glycerol complex medium (BMGY), supplemented with 0.5 μM H2O·Cbl. 1 L of BMGY contained 10g of yeast extract, 20g of pepton, 700mL H2O, 100mL phosphate buffer (NaH2PO4/Na2HPO4) 1M pH 6.0, 50mL cas amino acids 20%, 100mL of filter sterilised yeast nitrogen base 10%, 10mL glycerol 10%. Successively cells in 2x100mL BMGY were put in 4x1L of the same medium, supplemented with 0.5 μM H2O·Cbl, and grown for 24h. Then the cells were pelletted and resuspended in 4x1L of induction medium (buffered methanol complex medium, BMMY) containing the same additives without glycerol, and 1% methanol. The growth continued for 24h. Afterwards,
the suspension was centrifuged, and the cell-free supernatant was used for isolation of the recombinant human transcobalamin.

Proteins in cell-free supernatant were precipitated by ammonium sulphate, dialysed and subjected to exchange chromatography. Batch adsorption on CM-Sepharose (12 mL) was performed in 5 mM phosphate buffer (NaH₂PO₄/Na₂HPO₄) pH 6.0, and the matrix was washed sequentially with 0.02, 0.05, and 0.1 M phosphate buffer pH 6.0. Red coloured holo-transcobalamin was eluted from CM-Sepharose with 0.2M phosphate buffer pH 7.5. The obtained preparation was treated with 2,500 units of peptide N-glycosydase F at 37 °C for 2 days to remove carbohydrates from the protein. Then, transcobalamin was subjected to gel filtration on Sephacryl S-200 (250 mL), equilibrated with 0.1 M Tris pH 7.5, 1M NaCl. The fractions with red-coloured protein were pooled and repurified on a 2-mL CM-Sepharose column. The correct NH₂-terminal sequence of human transcobalamin was confirmed on a protein sequencer 477A (Perkin-Elmer). The final preparation for each expression cycle contains about 10-15 mg of the recombinant protein saturated with H₂O·Cbl. Finally it was stored at -20 °C. All salts and materials for media were purchased from Merk, Sigma, Roche Molecular Biochemicals, and Difco.

The purity and concentration of the protein preparation was estimated on a M350 double beam spectrophotometer (Camspec), reading adsorption at 280 nm (Trp and Tyr adsorption), 361 nm (Cbl·H₂O adsorption). In fact a highly pure TC solution (1 mg/mL) has A₂₈₀/A₃₆₁ ≈ 2.3, and 1 µM (about 0.045 mg/mL) of highly pure TC has A₃₆₁ = 0.03 (ε₃₆₁ = 30,000 M⁻¹ cm⁻¹) (Fedosov et al. 2000).

The Scheme II.2 shows the transcobalamin protein expression in yeast and the protein purification procedure.
Transcobalamin protein expression in yeast

Recombinant yeast cells

- Incubation: (30°C, 300rpm, 24h)
- Induction: (30°C, 300rpm, 6h)

- Resuspension in 4 x 1L BMGY

- Incubation: (30°C, 300rpm, 24h)
- Incubation: (30°C, 300rpm, 6h)
- Incubation: (30°C, 300rpm, 6h)

- 4 x 100 mL BMGY

- 4 x 1L BMGY

- 4 x 10 mL MeOH

- 3000rpm, 15min, 20°C

- 12000rpm, 5min, 20°C (removal of the cell particles)

Supernatant with expressed protein

Transcobalamin protein purification

- 12000rpm, 20min, 20°C (removal of the cell particles)
- Pellet resuspension (60mL Pi 20mM)

- (pH 6.0 for hTC; pH 7.5 for bTC)

- Ion exchange
- CM-Sepharose column
- 0.05M Pi 0.2M Pi for hTC
- 0.1M Pi 0.5M Pi for bTC
- (pH 6.0 for hTC; pH 7.5 for bTC)

- Washing and elution
- TC preparation (~ 80-90% purity)

- Removal of carbohydrates from protein surface
- N-glycosidase F (37°C, 48h)

- Electrophoresis

Scheme II.2. (a) Scheme of transcobalamin expression procedure in yeast. (b) Scheme of transcobalamin purification procedure.
Crystallisation of human transcobalamin

The hanging-drop vapour-diffusion method was used as technique for crystallisation attempts, with 4 µl drops (2 µl protein solution plus 2 µl of reservoir solution) equilibrated against 1 ml reservoir solution. The point of saturation at which crystals of transcobalamin are formed has been studied during crystallisation attempts varying protein concentration, temperature, pH, ionic strength, precipitant agents, in the presence or not of additional organic solvents.

The human transcobalamin concentration was varied from 2 to 8 mg/mL (0.04-0.20 mM) in the final preparation solution (0.1 M Tris pH 7.5, 1M NaCl) using microcon (3) tube (Amicon). Higher protein concentration was unreachable in the final preparation solution because the protein tends to aggregate and precipitate on microcon tube filter used for protein concentration. Crystallisation experiments have been performed at temperature values of 4, 16 and 25 °C. Suitable pH values and precipitant agents for crystal nucleation were initially searched using commercially available reagents (Hampton Research) based on the sparse-matrix method (Jancarik and Kim 1991) and using a protein concentration of 4 mg/mL. It was possible to see the formation of crystal spherulites and little needle-shaped crystals in a pH range 6.5-8.5 using sodium cacodylate, sodium HEPES or Tris buffer, and using PEG4000, ammonium sulphate, and sodium citrate as precipitant agents. Following crystallisation attempts, at the same protein concentration, indicated that PEG 4000 and PEG 6000 in the range 15-25% induced nucleation rapidly (about 2-3 days) with formation of clusters of microcrystals. Ammonium sulphate in the range 0.8-2.0M and sodium acetate in the range 1.2-1.5M induce slower nucleation (5-7 days) and, only for ammonium sulphate, formation of isolated elongated microcrystals. Using seeding technique it was possible to observe a very slow growth of these microcrystals also in a range of ammonium sulphate 0.6-0.8M. After several weeks they reached maximum dimensions of about 0.5x0.1x0.1mm.
Material and methods

Best pH values were tested in the range 6.0-9.0. Values in the range 7.5-8.5 with Tris 0.1M gave best results either with PEG or ammonium sulphate as precipitant agents. The best temperature was 25°C for all crystallisation conditions. At 4°C the nucleation process was slower but the number of nuclei increased. Other crystallisation experiments using several salts (NH₄SO₄, LiCl, NaCl, NaCl, CaCl₂, Mg(CH₃COO)₂, Na(CH₃COO) ) in a range of concentration 0.05-0.4 M were then performed by adding salt into reservoir solution containing PEG4000 20% and Tris buffer 0.1 M pH 8.0. These experiments, for studying the effect of the ionic strength on crystallisation process, indicated that the tested salts increased nucleation without improving crystal dimensions. In order to study the effect of the presence of additional organic solvent, methanol, ethanol, propanol, MPD (2-methyl-2,4-pentanediol), PEG400 and glycerol have been tested in a range of concentration 5-30%, with PEG4000 20% and Tris buffer 0.1 M pH 8.0. Results of these attempts indicated that PEG400 and glycerol increased only nucleation, while alcohols induced the precipitation of the protein, rapidly for methanol and slowly for MPD. After a week, ethanol at 20% was the only which caused the growth of several clusters of thin plate transcobalamin crystals. In order to decrease the initial precipitation of the protein induced by ethanol and increasing protein crystal size, boxes were prepared at 4°C and then kept at 16°C. Optimised conditions were found for a protein concentration of 6 mg ml⁻¹, and for a reservoir solution containing 20% PEG 4000, 18% ethanol, and 100 mM Tris buffer pH 8.0.

Attempts for growing crystals of isomorphous heavy-atom derivatives

Attempts to introduce heavy atoms in crystals of human transcobalamin have been made, either using co-crystallisation method or soaking technique. In the co-crystallisation method ethylmercurithiosalicylic sodium salt (trimerosal), potassium tetrachloroplatinate(II), and potassium dicyanoaurate(I) (Aldrich) were tested in a range of concentration of 5.0-0.05 mM into the drop (drop protein concentration was 2.5 mg/mL or 0.05 mM). Reservoir solution contained 20%
Material and methods

PEG 4000, ethanol in a range of 16-20%, and 100 mM Tris buffer pH 7.5. Crystals obtained in such condition appeared often smaller and thinner than those obtained without adding heavy atom solution. There was no clear relationship between crystals growth and heavy atom solution concentration.

X-ray data collections and analysis of human transcobalamin

X-ray diffraction experiments were carried out at the Elettra Synchrotron, Trieste (Italy), using a monochromatic radiation with wavelength of 1.200 Å and a MAR CCD detector. Crystals of transcobalamin were mounted directly on a mounted 0.4 mm cryo-loop (Hampton Research) and flash-frozen in a stream of N2 at 100 K (Cryosystem).

The determination of unit-cell parameters, integration of reflection intensities and data scaling were performed using MOSFLM and SCALA from CCP4 program suite (Collaborative Computational Project, Number 4, 1994). Self-rotation functions were calculated in the resolution range 20 - 3.2 Å with a 30 Å radius of integration, using the program POLARRFN (Collaborative Computational Project, Number 4, 1994).
Material and methods

Preparation of the genetic material for bovine transcobalamin

The cDNA encoding the mature bovine transcobalamin (Fedosov et al. 1999) was amplified by polymerase chain reaction (HB Biotechnology) and then ligated to the expression plasmid (pPICZα-LB). The plasmid contains the Myc sequence, a His tag, and a factor Xa cleavage site before the protein sequence. The plasmid-transcobalamin construction was linearised inside the alcohol oxidase 1 promoter by treatment with endonuclease NsiI. The transformed methanol-metabolizing Mut+ strain (SMD 1168) of Pichia pastoris, obtained by Fedosov, synthesised high level of transcobalamin (Fedosov et al. 1999).

Expression and purification of bovine transcobalamin

Initially, I grow for 24h at 37 °C the recombinant yeast cells in 2x100mL of buffered glycerol complex medium (BMGY), supplemented with 0.5 μM H₂O·Cbl. 1 L of BMGY contained 10g of yeast extract, 20g of pepton, 700mL H₂O, 100mL phosphate buffer (NaH₂PO₄/Na₂HPO₄) 1M pH 6.0, 50mL cas amino acids 20%, 100mL of filter sterilised yeast nitrogen base 10%, 10mL glycerol 10%. Successively, cells in 2x100mL BMGY were put in 4x1L of the same medium, supplemented with 0.5 μM H₂O-Cbl, and grown for 24h. Then, the cells were pelleted and resuspended in 4x1L of induction medium (buffered methanol complex medium, BMMY) containing the same additives without glycerol, and 1% methanol. The growth continued for 24h. Afterwards, the suspension was centrifuged, and the cell-free supernatant was used for isolation of the recombinant bovine transcobalamin.

Proteins in the obtained supernatant were precipitated by ammonium sulphate at 70% of saturation. The pellet was dissolved in 20 mL of 0.02 M phosphate buffer pH 7.5. Then the protein solution was dialysed against 10 L of the same buffer overnight at 5 °C. Then, TC was adsorbed on 2 mL of CM-Sepharose (Pharmacia) at room temperature and washed in the column by
Material and methods

increasing concentrations of phosphate buffer pH 7.5: 0.02 M, 0.05 M, 0.1 M. The protein was eluted with 2mL of 0.5M phosphate buffer pH 7.5, and the collected sample contained approximately 1 mg of the fusion protein mHX-TC. The sample was treated with 30 mg of factor Xa for 30 h at 37 °C to remove the service peptides from the N terminus. The preparation was applied to a 250 mL Sephacryl S-200 column (Pharmacia) equilibrated with 0.1 M Tris, 1 M NaCl, pH 8.0 at room temperature. The eluted fractions with maximal specific Cbl adsorption at 362 nm were pooled, dialysed against 0.02 phosphate buffer pH 7.5 and concentrated by adsorption on a small CM-Sepharose column (0.3 mL). The final preparation was stored frozen.

The Scheme II.2 shows the transcobalamin protein expression in yeast and the protein purification procedure. The purity index and protein concentration were estimated from the ratio of the absorption between two wavelengths (see purification of human transcobalamin).

Crystallization of bovine transcobalamin

The transcobalamin was crystallised using the hanging-drop vapour-diffusion method with 4 µl drops (2 µl protein solution plus 2 µl of reservoir solution) equilibrated against 1 ml reservoir solution. The point of saturation at which crystals of bovine transcobalamin are formed has been studied during crystallisation attempts likewise to human transcobalamin.

The bovine transcobalamin concentration was varied from 3 to 18 mg/mL (0.06-0.36 mM) in the final preparation solution (0.1 M Tris pH 7.5, 1M NaCl). Contrary to human transcobalamin, bovine transcobalamin does not tend to aggregate and precipitate in microcon (3) tube (Amicron) used for protein concentration. Crystallisation experiments have been performed at temperature values of 4 and 25 °C.

Crystallisation conditions similar to those used for growing best crystals of human transcobalamin [protein concentration 3 mg/mL; reservoir containing PEG
Material and methods

4000 in a range 20-35%, ethanol in a range 14-24%, and Tris 100 mM buffer pH 8.0] were tested also for bovine transcobalamin, but unsuccessfullly. Following crystallisation experiments were based on the sparse-matrix method (Jancarik and Kim 1991) using commercially available reagents (Hampton Research) with a protein concentration of 3 mg/mL. It was possible to see the formation of crystal spherulites with PEG4000, PEG8000, and little needle-shaped crystals with ammonium sulphate and sodium citrate, in a pH range of 7.5-8.5.

A screening with ammonium sulphate as precipitant agent for crystallisation, [protein concentration 3 mg/mL; reservoir solution containing ammonium sulphate in a range 1.3-1.8 M, and Tris 100 mM buffer in a range pH 7.0-8.0, and salts such as LiCl and Mg(CH₃COO)₂ in a range 0.1-0.3 M] indicated only little improvement in the needle-shaped crystal growth. Afterwards, using a protein concentration of 10 mg/mL, a reservoir with PEG 8000 or PEG 4000 (20-30%), ammonium sulphate 0.2 M and Tris 0.1 M pH 8.0, crystallisation attempts indicated that the increase of protein concentration and PEG having a high molecular mass is essential for the crystals growth. When PEG8000 was used at 24-26%, crystals of bovine transcobalamin could reach maximum dimensions of about 0.3x0.3x0.1mm after 1-2 days. Crystallisation attempts without ammonium sulphate in PEG8000 reservoir solution indicated that this salt is not essential for the crystal growth but increases dimensions of crystals.

In order to protect the crystal during synchrotron X-ray diffraction analysis by cryocooling technique, several attempts to grow the crystal directly in a cryoprotectant solution were performed. Glycerol, ethylenglycol, MPD (2-methyl-2,4-pentanediol), and PEG400, were tested with a fixed concentration of 20%, in the reservoir solution [containing PEG8000 in a range 20-30%, ammonium sulphate 0.2 M and Tris 0.1 M pH 8.0]. Results of these attempts indicated that MPD 20% allows a good crystal growth and contemporarily suppresses ice formation in stream of N₂ at 100 K.

Following crystallisation experiments using several salts (LiCl, NaCl, MgCl₂, or Mg(CH₃COO)₂) with a fixed concentration 0.2 M in the reservoir solution [containing PEG8000 (in a range 15-30%), Tris pH 8.0 (0.1 M) and MPD
Material and methods

20%] indicated that magnesium acetate slightly improves the crystal growth with respect to ammonium sulphate salt.

X-ray data collections of bovine transcobalamin

X-ray diffraction experiments were carried out at the Elettra Synchrotron using a MAR CCD detector. Crystals of transcobalamin were mounted directly on a mounted 0.4 mm cryo-loop (Hampton Research) and flash-frozen in a stream of \( N_2 \) at 100 K (Cryosystem). The determination of unit-cell parameters, integration of reflection intensities and data scaling were performed using MOSFLM and SCALA from CCP4 program suite (Collaborative Computational Project, Number 4, 1994).
4. Results and discussion

4.1 Crystallisation of \textit{human} transcobalamin

Even if crystals obtained in ammonium sulphate as precipitant agent (see \textit{human} transcobalamin crystallization section) were of suitable dimensions for synchrotron X-ray analysis, no diffraction was obtained from them.

The final optimised conditions for growing crystals suitable for X-ray analysis were 20\% PEG 4000, 20\% ethanol and 100 mM Tris buffer pH 8.0, and 6 mg ml\textsuperscript{-1} protein. The red coloured crystals of transcobalamin generally grew in about two weeks emerging as clusters from which it was possible to separate the isolated plate-like crystals with maximal dimension of 0.30 x 0.30 x 0.02 mm (Figure II.4).

\textbf{Figure II.4.} Crystals of \textit{human} transcobalamin.
4.2 Preliminary X-ray analysis of human transcobalamin

Crystals of human transcobalamin diffracted to better than 3.0 Å of resolution. Figure II.5 shows a diffraction image. Ethanol in the reservoir solution acted also as cryocooling protectant for crystals. A complete data set was collected from a single crystal and the data collection statistics is reported in Table II.1.

The diffraction pattern was indexed using an orthorhombic-cell $a = 49.04$, $b = 145.27$ and $c = 164.96$ Å. Axial reflections along $h00$, $0k0$ and $00l$ were systematically absent when $h$, $k$ or $l = 2n + 1$, identifying the space group as being $P2_12_12_1$.

Figure II.5. (a) X-ray diffraction pattern from a crystal of human transcobalamin. The resolution of the outer edge of the image is at 3.1 Å; (b) an enlarged image of the indicated area in (a).
Table II.1: Data collection statistics for human transcobalamin.
Values for the highest resolution shell in parenthesis.

<table>
<thead>
<tr>
<th>Wavelength (Å)</th>
<th>1.200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>100</td>
</tr>
<tr>
<td>Resolution-range</td>
<td>32.0 - 3.2</td>
</tr>
</tbody>
</table>
| Unit-cell parameters (Å) | a = 49.04  
|                    | b = 145.27  
|                    | c = 164.96  |
| Number of observations | 105390 |
| Unique reflections | 19677 |
| Completeness (%) [I > σ(I)] * | 95.5 (93.8) |
| Averaged I/σ (I) * | 5.8 (2.2) |
| R_merge †† | 0.149 (0.434) |

* R_{merge} = \sum |I - <I>| / \sum <I>.
†† Outer-shell data is the resolution range 3.37-3.20 Å.

Assuming two crystallographic independent molecules of transcobalamin in the asymmetric, the Matthews coefficient is 3.2 Å³ Da⁻¹ (Matthews, 1968) and the solvent content is approximately 61%. Alternatively, the Matthews coefficient can be 2.1 Å³ Da⁻¹ and the solvent content 41% assuming three molecules in the asymmetric unit.

Analysis of self-rotation peaks revealed the presence of a pseudo twofold symmetry in the [011] direction. The stereographic projection (k = 180° section) of the self-rotation is showed in Figure II.6. The presence of a peak at ω = 45°, φ = 90° of height 35% of the origin peak, suggests that most likely two crystallographic independent molecules of transcobalamin are present in the asymmetric unit.
Results and discussion

Figure II.6. Plot of the self-rotation function at k=180°. The section has ω=0° or 180° at the centre, ω=90° around the edge, and ϕ as marked around the periphery. The view is down the c axis; ω=90°, ϕ=0° corresponds to the direction of the a axis, and ω=0° to the direction of the c axis.

4.3 Cristallization of bovine transcobalamin

Drops containing a protein concentration of 10 mg/mL equilibrated against a reservoir solution containing PEG 8000 24-26%, ammonium sulphate 0.2 M and Tris 0.1 M pH 8.0, let hexagonal crystals of bovine transcobalamin grew in few days with maximal dimension of 0.30 x 0.30 x 0.50 mm (Figure II.7a). The addition of 2-methyl-2,4-pentandiol (MPD) and the substitution of ammonium sulfate with Mg acetate caused the growth of new elongated monoclinic crystals which appeared in about three weeks (Figure II.7b). Optimised crystallisation
conditions were found for PEG 8000 24-26%, 20% MPD, Mg acetate 0.2 M and Tris 0.1 M pH 8.0. MPD acted also as cryoprotectant holding solution agent.

Figure II.7. a) Hexagonal crystal of bovine transcobalamin. b) Elongated monoclinic crystals of bovine transcobalamin.

4.4 Preliminary X-ray analysis of bovine transcobalamin

Hexagonal crystals of bovine transcobalamin diffracted to less than 6.0 Å and the diffraction pattern was indexed using a unit cell $a = 49.04$, $b = 145.27$ and $c = 164.96$ Å.

An elongated crystal of bovine transcobalamin obtained using a reservoir solution containing PEG 8000 28%, 20% MPD, Mg acetate 0.2 M and Tris 0.1 M pH 8.0, diffracted to better than 2.0 Å resolution (Figure II.8) and the diffraction pattern was indexed using a monoclinic-cell $a = 95.32$, $b = 100.19$, $c = 98.73$ Å and $\beta = 96.9$ Å. Systematic absences were in agreement with the space group P2$_1$. Assuming four crystallographic independent molecules of bovine transcobalamin in
the asymmetric, Vm results $2.54 \text{Å}^3\text{Da}^{-1}$ (Matthews, 1968) and the solvent content is approximately 51%.

Trascobalamin contains a cobalt corrin atom in its protein structure. The consequence of the Co atom anomalous scattering is that the intensities of a reflection $h k l$ and its Bijvoet mate $-h -k -l$ are no longer equal. The difference in intensity between Bijvoet pairs ($|F_h(+)|^2$ and $|F_h(-)|^2$) can be used for phase information in the protein structure determination by the multiple wavelength method (MAD) (Hendrickson 1991).

Three complete data sets were collected from a single crystal using three energies near the absorption edge of the Co atom: 10.3311 ($\lambda = 1.2000$ Å), 7.7210 ($\lambda = 1.6058$ Å) and 7.6960 keV ($\lambda = 1.6110$ Å). Table II.2 reports a summary of data collection and crystallographic statistics.

Figure 11.8. (a) X-ray diffraction pattern from a crystal of bovine transcobalamin. The resolution of the outer edge of the image is at 1.95 Å; (b) an enlarged image of the indicated area in (a).
Results and discussion

Table II.2: Data collection statistics for bovine transcobalamin. Values for the highest resolution shell in parenthesis.

<table>
<thead>
<tr>
<th></th>
<th>Remote</th>
<th>Peak</th>
<th>Edge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>1.200</td>
<td>1.606</td>
<td>1.611</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Resolution-range</td>
<td>50.00 - 2.00</td>
<td>50.00 - 2.40</td>
<td>50.00 - 2.60</td>
</tr>
<tr>
<td>Unit-cell parameters (Å, °)</td>
<td>a = 95.32</td>
<td>a = 95.25</td>
<td>a = 95.36</td>
</tr>
<tr>
<td></td>
<td>b = 100.19</td>
<td>b = 100.16</td>
<td>b = 100.14</td>
</tr>
<tr>
<td></td>
<td>c = 98.73</td>
<td>c = 98.76</td>
<td>c = 98.83</td>
</tr>
<tr>
<td></td>
<td>β = 96.93</td>
<td>β = 96.98</td>
<td>β = 96.98</td>
</tr>
<tr>
<td>Number of observations</td>
<td>1064679</td>
<td>872472</td>
<td>900304</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>135349</td>
<td>104952</td>
<td>104295</td>
</tr>
<tr>
<td>Completeness (%) [I &gt; σ(I)] *</td>
<td>96.9 (96.9)</td>
<td>96.7 (96.7)</td>
<td>97.1 (97.1)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.9(3.8)</td>
<td>4.2(4.3)</td>
<td>4.4(4.4)</td>
</tr>
<tr>
<td>Averaged I/σ(I) *</td>
<td>9.0 (2.2)</td>
<td>8.8 (2.3)</td>
<td>8.6 (1.7)</td>
</tr>
</tbody>
</table>

* $R_{merge} = \Sigma |I_{i}-<I_{i}>,>/\Sigma <I_{i}>$, where $\lambda^*$ is the remote wavelength.

* Outer-shell data is the resolution range 2.20-2.00 Å (remote), 2.50-2.40 Å (peak), 2.70-2.60 Å (edge).
5. Conclusions

Transcobalamin is a cobalamin binding protein in mammalian plasma that facilitates the cellular uptake of vitamin B\textsubscript{12}. Its protein ternary structure is unknown and the essential first requirement for X-ray structure determination is to grow suitable crystals.

*Human* and *bovine* transcobalamin were expressed using recombinant yeast cells and purified.

*Human* transcobalamin was successfully crystallized using polyethylene glycol and ethanol as precipitants. Crystals belong to the orthorhombic space group P2\textsubscript{1}2\textsubscript{1}2\textsubscript{1}, with unit-cell parameters $a = 49.04$, $b = 145.27$, and $c = 164.96$ Å. A complete data set to 3.2 Å resolution was collected from a single crystal using synchrotron radiation. Estimation of the crystal packing ($V_m = 3.2$ Å\textsuperscript{3} Da\textsuperscript{-1}) and self-rotation function analysis suggest the presence of two molecules in the asymmetric unit related by a pseudo-twofold symmetry.

*Bovine* transcobalamin was crystallised using polyethylene glycol and 2-methyl-2,4-pentandiol. Crystals belong to the monoclinic space group P2\textsubscript{1} (a = 95.32, $b = 100.19$, $c = 98.73$ Å and $\beta = 96.9$ Å), and diffract to better than 2.0 Å resolution. Multiwavelength anomalous diffraction data sets were collected on beamline XRD1 at Elettra (Trieste) from a single crystal of the *bovine* protein, using the cobalamin Co atom as anomalous scatterer. Attempts to find the position of Co atoms in the monoclinic cell from Bijvoet difference Patterson maps are under way.
References of part I


Read, J., Gill, R., Dales, S. L., Cooper, J. B., Wood, S. P., and Anthony, C. 1999. The molecular structure of an unusual cytochrome c\textsubscript{2} determined at 2.0 Å; the cytochrome c\textsubscript{H} from Methylobacterium extorquens. *Protein Sci.* 8: 1232-1240.


References of part II


