THE NATURE AND ROLE OF OXIDATION STATE DEPENDENT CONFORMATIONAL DIFFERENCES IN CYTOCHROME *C*

By

Albert Marinus Berghuis

B. Sc. (Chemistry) Rijks Universiteit Groningen, The Netherlands, 1983

M. Sc. (Chemistry) Rijks Universiteit Groningen, The Netherlands, 1987

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

THE DEPARTMENT OF BIOCHEMISTRY

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1993

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of Biochemistry

The University of British Columbia
Vancouver, Canada

Date April 19 1993

DE-6 (2/88)

Abstract

The objective of the work described in this thesis was to study the nature and role of conformational differences between the oxidation states of cytochrome *c*. Using x-ray crystallographic techniques, the oxidized form of yeast iso-1-cytochrome *c* was solved and compared to the previously determined reduced state. The following differences between the oxidation states were identified. Three segments of polypeptide chain, located for the most part on the Met80 side of the protein, were shown to display an increase in mobility in the oxidized state. A conserved internal water molecule, Wat166, was observed to shift 1.7 Å towards the heme iron atom and reorient its dipole moment in the oxidized state. As part of this movement several hydrogen bonds were broken including the interaction between Tyr67 OH and the Met80 SD heme ligand. Finally, differences between the two oxidation states were also observed for the conformation of the pyrrole A propionate and its associated hydrogen bond network, the distortion of the porphyrin ring plane, and the orientation of the imidazole plane of the His18 ligand.

In order to assess the function of the observed conformational differences between the two oxidation states of cytochrome *c*, the three dimensional structures of five mutants (N52A, N52I, Y67F, N52I-Y67F and I75M) were determined, of which three were completed in both oxidation states. Correlation of wild-type and variant protein structures with functional studies suggested that Wat166 was a central feature in oxidation state dependent differences, and three roles for this water molecule could be identified. First, the oxidation state dependent positioning and orientation of Wat166 appears to be particularly important for modulating the interaction between Tyr67 OH and Met80 SD. This hydrogen bond was shown to influence the electron withdrawing power of the Met80 ligand and therefore is a factor in controlling the midpoint reduction potential of cytochrome *c*. Secondly, the presence of Wat166 is necessary to maintain the spatial and hydrogen bonding relationships between residues in this region of the protein.

Finally, Wat166 also appears to mediate the oxidation state dependent flexibility of selected polypeptide chain segments. The biological function of this phenomenon is still unclear, but our results suggests that it might play a role in interactions between cytochrome *c* and its redox partners.

In conclusion, the work described in this thesis gives insight into the structure-function relationships in cytochrome *c* and provides a basis for future studies aimed at understanding the mechanism of electron transfer carried out by this protein.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

B	Thermal factor
CCP	Cytochrome *c* peroxidase
CD	Circular dichroism
DTT	Dithiothreitol
E_m	Midpoint reduction potential
EXAFS	Extended x-ray absorption fine structure
F_o, F_c	Observed and calculated structure factors
I	Intensity of a reflection
K	Absolute scale factor
NMR	Nuclear magnetic resonance
SHE	Standard hydrogen electrode
UV	Ultraviolet light
a, b, c	Crystallographic unit cell axes, or axis lengths
e.u.	entropy units; 1 e.u. = 1 cal mol^{-1} K^{-1}
f_o	Atomic scattering factor
h, k, l	Miller indices
r.m.s.	Root mean squared
x, y, z	Positional parameters
α	Phase
λ	Wavelength of the radiation used in the diffraction experiment. All the experiments described in this thesis are done with Cu K_α radiation, $\lambda = 1.54184$ Å

ϕ	One of the rotational axes which define the orientation of the crystal in the diffraction experiment. This angle is varied to obtain an estimate of the absorption effect when using a diffractometer for data collection
φ	Generic symbol for a mathematical function
σ_x	Standard deviation of x or estimated standard deviation of quantity x
θ	Diffraction angle
Å	Angstrom (0.1 nm)

The conventions of the IUPAC-IUB Combined Commissions on Biochemical Nomenclature are followed for both three letter and one letter abbreviations for amino acids [*J. Biol. Chem.* **241**, 527–533 (1966); *J. Biol. Chem.* **243**, 3557–3559 (1968)]; for designating atoms and for describing the conformational torsional angles of the polypeptide chain [*J. Biol. Chem.* **245**, 6489–6497 (1970)]. Designations for atoms of the protoheme IX group are according to the Brookhaven National Laboratory Protein Data Bank (Bernstein *et al.*, 1977; see also Figure 1.2).

The amino acid numbering scheme used for the yeast iso-1-cytochrome *c* wild-type and mutant structures described in this thesis is based on an alignment to the sequences of vertebrate cytochromes *c*. The N-terminal residue is numbered -5 and the C-terminal residue is numbered 103 (see Table 1.1). The numbering of solvent molecules for these structures is identical to that of the reduced yeast iso-1-cytochrome *c* wild-type protein (Louie & Brayer, 1990; Protein Data Bank entry 1YCC). Water molecules not found in reduced yeast iso-1-cytochrome *c* have been assigned numbers starting with 300.

Acknowledgments

I would like to take this opportunity to thank all the people who contributed to the work described in this thesis. A few people should be mentioned here specifically. First and foremost is my supervisor Gary Brayer who encouraged my scientific curiosity as well as my independence as a researcher during the last five years. Connie Leung, Teresa Yang, Teresa Hii and Yuoguang Luo were invaluable for their technical assistance. My fellow grad-students, Gordon Louie, Michael Murphy, Terry Lo and David Burk were a constant source of information and a sounding board for ideas. Specifically, I would like to thank Mike, who not only was the driving force behind the development of some of the graphics software used in this thesis (Preras3d/Raster3d package), but also helped develop some of the ideas expressed in Appendix A. Moreover we became close friends in the process. I also like to express my gratitude to our collaborators Grant Mauk and Michael Smith here at U.B.C., George McLendon and Fred Sherman at the University of Rochester, and Arieh Warshel at U.S.C., who were very generous in providing protein samples and sharing their results. This acknowledgment page would be incomplete without mentioning Alfred Gärtner, Guy Guillemette and Steve Rafferty, without whom the contents of this thesis would be quite different. I would also like to acknowledge the two other members of my supervisory committee Phil Bragg and Ian Clark-Lewis. Finally, I would like to thank my wife Lida, without whose support I might have not entered graduate school, and would surely have not finished it.

Chapter 1

Introduction

1.1 The Biological Role of Eukaryotic Cytochromes *c*

Eukaryotic cytochromes *c* are small (~13,000 Daltons) soluble proteins which are found in the inter-membrane space of mitochondria. They form part of the respiratory chain and primarily shuttle electrons between the membrane bound cytochrome *c* reductase and cytochrome *c* oxidase complexes (see Figure 1.1). In order to accomplish this task cytochrome *c* possesses a covalently attached heme as a prosthetic group (see Figure 1.2). Positioned in the center of the heme group is an iron atom. By varying the valence state of the heme iron atom between 2+ and 3+, the reduced and oxidized form respectively, cytochrome *c* can donate and accept electrons and thus function as an electron transport protein.

$$\underset{\text{oxidized}}{\text{Cytochrome } c[\text{Fe(III)}] + e^-} \rightleftharpoons \underset{\text{reduced}}{\text{Cytochrome } c[\text{Fe(II)}]}$$

Cytochrome *c* reductase and cytochrome *c* oxidase are not the only redox partners for cytochrome *c* (Figure 1.1). In yeast, cytochrome *c* peroxidase (CCP) and cytochrome b_2 have been identified as accepting electrons from or donating electrons to cytochrome *c* (Altschul *et al.*, 1940; Abrams *et al.*, 1942; Bach *et al.*, 1942a,b). Cytochrome *c* also functions as an electron acceptor from cytochrome b_5 and sulfite oxidase in animal systems (McLeod *et al.*, 1961; Cohen & Fridovich, 1971; Ito, 1980a,b; Lederer *et al.*, 1983). The cytochrome *c*/CCP and cytochrome *c*/cytochrome b_5 redox couples deserve special mention since they have served as model systems for studying biological electron transfer (McLendon & Miller, 1985; Cokic & Erman, 1987; Liang *et al.*, 1987; Mauk *et al.*, 1991). Studies of these two complexes have not only focussed on the thermodynamics and kinetics of biological electron transfer, but have also

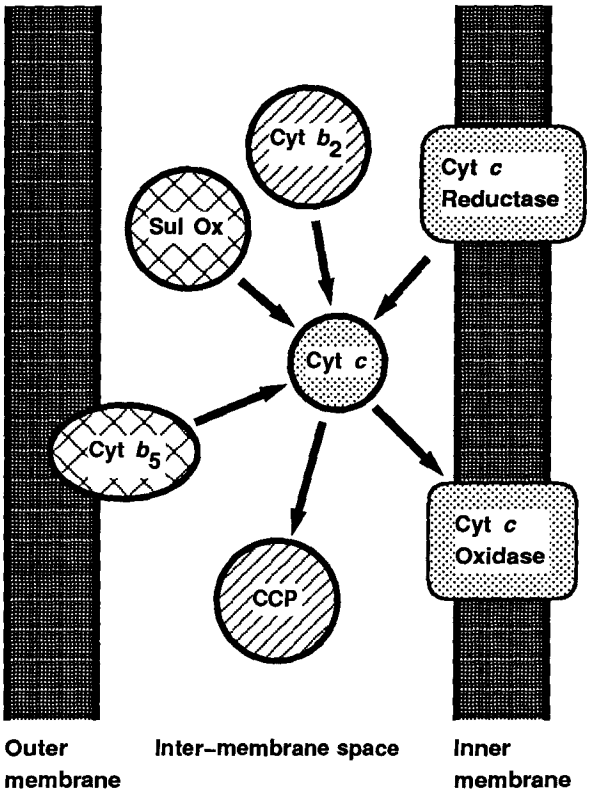

Figure 1.1: A schematic representation of the mitochondrial inter-membrane space showing the location of cytochrome *c* and its physiological redox partners. The arrows in the diagram indicate the direction of the flow of electrons. The abbreviations used are: Cyt *c* - cytochrome *c*; Cyt b_2 - (flavo)cytochrome b_2; Cyt b_5 - cytochrome b_5; Sul Ox - sulfite oxidase; CCP - cytochrome *c* peroxidase; Cyt *c* Reductase - cytochrome *c* reductase (complex III); Cyt *c* Oxidase - cytochrome *c* oxidase (complex IV). The interaction between cytochrome *c* and cytochrome b_2 and CCP is only observed in yeast. The interaction between cytochrome *c* and sulfite oxidase and cytochrome b_5 is only observed in animal systems.

explored the structural aspects of complex formation (Salemme, 1976; Poulos & Kraut, 1980; Lum *et al.*, 1987; Wendoloski *et al.*, 1987; Pelletier & Kraut, 1992).

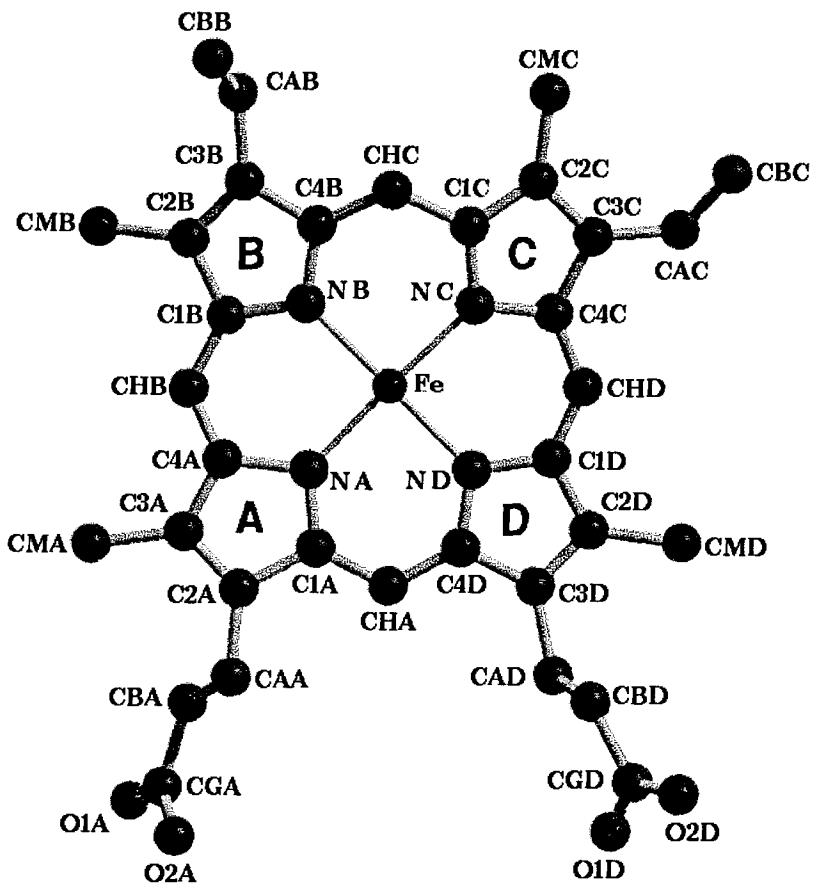

Figure 1.2: A schematic representation of the atomic skeleton of the protoheme IX group of cytochrome *c* and the atom and pyrrole ring labelling convention used herein (Bernstein *et al.*, 1977). In eukaryotic cytochromes *c* the porphyrin ring is covalently attached to the protein through two thioether linkages (Cys14 SG - heme CAB; Cys17 SG - heme CAC). The iron atom is six coordinated with His18 NE2 and Met80 SD providing the remaining ligands. In the view presented, the solvent exposed heme edge observed in cytochromes *c* is located at the right, the histidine ligand is behind the porphyrin ring and the methionine ligand is positioned in front.

1.2 The Structure of Eukaryotic Cytochromes *c*

1.2.1 Amino acid sequences

Cytochromes *c* from eukaryotes have long been the subject of extensive study. As a consequence, a large number of amino acid sequences are presently known. This wealth of information has

been the basis for phylogenetic studies as well as functional studies (Dayhoff *et al.*, 1976; Dickerson & Timkovich, 1975; Amati *et al.*, 1988; Moore & Pettigrew, 1990). Moore and Pettigrew (1990) compiled sequences from 96 eukaryotic species. Examination of this data showed that within the eukaryotes the amino acid sequence is highly conserved, the length of the polypeptide chain varies between 103 and 113 residues and alignment of the sequences from 94 species can be accomplished without creating any gaps (the two remaining sequences are from *Euglina gracilis* and *Tetrahymena pyriformis*). In these 94 sequences, 27 residues are invariant and an additional 16 amino acids are conserved in at least 90 of the sequences (see Table 1.1).

In many cases the reason for the high degree of sequence homology in eukaryotic cytochromes *c* is puzzling. Although several residues are critical for maintaining the fold and function of cytochrome *c*, a large number of invariant residues can be mutated without abolishing biological activity (Hampsey *et al.*, 1986). Furthermore, by mutating the conserved residues 52 and 67, variants of cytochrome *c* have been made which appear not to be affected in their function but which are actually more stable than the wild-type protein (Das *et al.*, 1989; McLendon *et al.*, 1991; Hickey *et al.*, 1991; Luntz *et al.*, 1989). This has not only raised questions regarding the specific role of these conserved residues but also instigated a debate on the evolutionary development of cytochrome *c* (Margoliash, 1990).

1.2.2 Three dimensional structures

The high degree of homology seen in sequence alignments is reflected in the tertiary structures of the cytochromes *c*. To date a total of six structures, two of which are in both oxidation states, have been determined by x-ray crystallographic techniques (see Table 1.2). Analysis of the five high resolution structures available shows that the polypeptide fold is virtually identical (see Figure 1.3). Pair-wise comparison gives a maximum average deviation for main-chain atoms of only 0.57 Å (between horse and yeast iso-2 cytochrome *c*, sequence identity is 58%). A thorough analysis of the subtle differences in conformation between the different eukaryotic

Table 1.1: Sequence alignment for yeast iso-1, yeast iso-2, tuna, horse and rice cytochromes *c*

```
       -9                  1                   10                  20
                           •         •         •         ○   • • • ○         ○         ○   •
Iso-1  - - - - T E F K A G S A K K G A T L F K T R C L Q C H T V E K G G P H K V G
Iso-2  A K E S T G F K P G S A K K G A T L F K T R C Q Q C H T I E E G G P N K V G
Tuna   - - - - - - - - - G D V A K G K K T F V Q K C A Q C H T V E N G G K H K V G
Horse  - - - - - - - - - G D V E K G K K I F V Q K C A Q C H T V E K G G K H K T G
Rice   - A S F S E A P P G N P K A G E K I F K T K C A Q C H T V D K G A G H K Q G

       30                  40                  50                  60
       • ○ •     •     ○ •     •         ○         •     ○ •               •
Iso-1  P N L H G I F G R H S G Q A E G Y S Y T D A N I K K N V L W D E N N M S E
Iso-2  P N L H G I F G R H S G Q V K G Y S Y T D A N I N K N V K W D E D S M S E
Tuna   P N L W G L F G R K T G Q A E G Y S Y T D A N K S K G I V W N N D T L M E
Horse  P N L H G L F G R K T G Q A P G F T Y T D A N K N K G I T W K E E T L M E
Rice   P N L N G L F G R Q S G T T P G Y S Y S T A N K D M A V I W E E N T L Y D

           70                  80                  90                  100
       • •   ○ • • • ○ ○ • • ○ • •     •     •         •     ○ ○     ○ ○
Iso-1  Y L T N P J K Y I P G T K M A F G G L K K E K D R N D L I T Y L K K A C E -
Iso-2  Y L T N P J K Y I P G T K M A F A G L K K E K D R N D L I T Y M T K A A K -
Tuna   Y L E N P K K Y I P G T K M I F A G I K K K G E R Q D L V A Y L K S A T S -
Horse  Y L E N P K K Y I P G T K M I F A G I K K K T E R E D L I A Y L K K A T N E
Rice   Y L L N P J K Y I P G T K M V F P G L K J P Q E R A D L I S Y L K E A T S -
```

The sequences of yeast iso-1 (Smith *et al.*, 1979), yeast iso-2 (Montgomery *et al.*, 1980), tuna (Kreil, 1965), horse (Margoliash *et al.*, 1961) and rice (Mori & Morita, 1980) cytochromes *c* have been aligned so as to maximize the structural homology present. The single letter code is used to identify the amino acids and the residue numbering is based on the sequence of tuna cytochrome *c*. The amidation states of residues 52 and 54 in rice are the opposite of that given by Mori and Morita (1980). The present amide assignment fits the chemical data as well as the original assignment and is more consistent with other sequences (Moore & Pettigrew, 1990). Note the single-letter code J is used to denote ε-N-trimethyl lysines. Those amino acid residues identical in all five protein sequences are enclosed in boxes. Residues which are observed to be invariant in 94 eukaryotic sequences are marked with a bullet (•), those which are conserved in at least 90 sequences are marked with a circle (○).

cytochrome *c* structures can be found in Brayer and Murphy (1993).

Given the high degree of similarity between structures, reduced yeast iso-1-cytochrome *c* will be used here as the archetypal cytochrome *c* for describing structural conformations. This structure has been determined to the highest resolution and forms the basis of the structural

Table 1.2: Eukaryotic cytochrome *c* structures determined by x-ray crystallography

Source	Oxidation state	Resolution (Å)	R-factor (%)	PDB code †	Reference
Yeast iso-1	reduced	1.2	19.2	1YCC	(Louie & Brayer, 1990)
Yeast iso-1 ‡	oxidized	1.9	19.7	2YCC	(Berghuis & Brayer, 1992)
Yeast iso-2	reduced	1.9	19.0	1YEA	(Murphy *et al.*, 1992)
Tuna	reduced	1.5	15.9	5CYT	(Takano & Dickerson, 1981a)
Tuna	oxidized	1.8	20.8	3CYT	(Takano & Dickerson, 1981b)
Horse	oxidized	1.9	17	-	(Bushnell *et al.*, 1990)
Rice	oxidized	1.5	19	1CCR	(Ochi *et al.*, 1983)
Bonito §	reduced	2.3	-	1CYC	(Tanaka *et al.*, 1975)

† Brookhaven protein data bank (PDB) codes (Bernstein *et al.*, 1977) are given to specify the coordinates used for comparison studies. The horse cytochrome *c* coordinates were obtained as the result of other studies ongoing in this laboratory.
‡ The oxidized yeast iso-1 structure is included in this table for completeness but will not be discussed until Chapter 3.
§ The coordinates available for the bonito structure have not undergone crystallographic refinement. Therefore bonito cytochrome *c* is not used in further comparisons.

work described in this thesis.

Yeast iso-1-cytochrome *c* is a globular protein roughly 35 × 35 × 30 Å in size, consisting of five α-helixes and a number of β-loops and turns which wrap the polypeptide chain around the protoheme IX group (see Figures 1.4 and 1.5, and Table 1.3). The porphyrin ring is covalently attached to the polypeptide chain through two thioether linkages with cysteines 14 and 17. The two residues His18 and Met80 form the fifth and sixth ligands to the heme iron atom.

The heme group is almost completely buried inside the protein matrix leaving less than 10% of its surface exposed to solvent (see Figure 1.6). The two heme propionates, each of which probably carries a negative charge (Moore, 1983) are also buried inside the protein. The propionate carboxyl groups form extensive hydrogen bond interactions with proton donor groups of nearby polypeptide chain and with two internal water molecules. Examination of the other cytochrome *c* structures and the available amino acid sequences suggests that most of

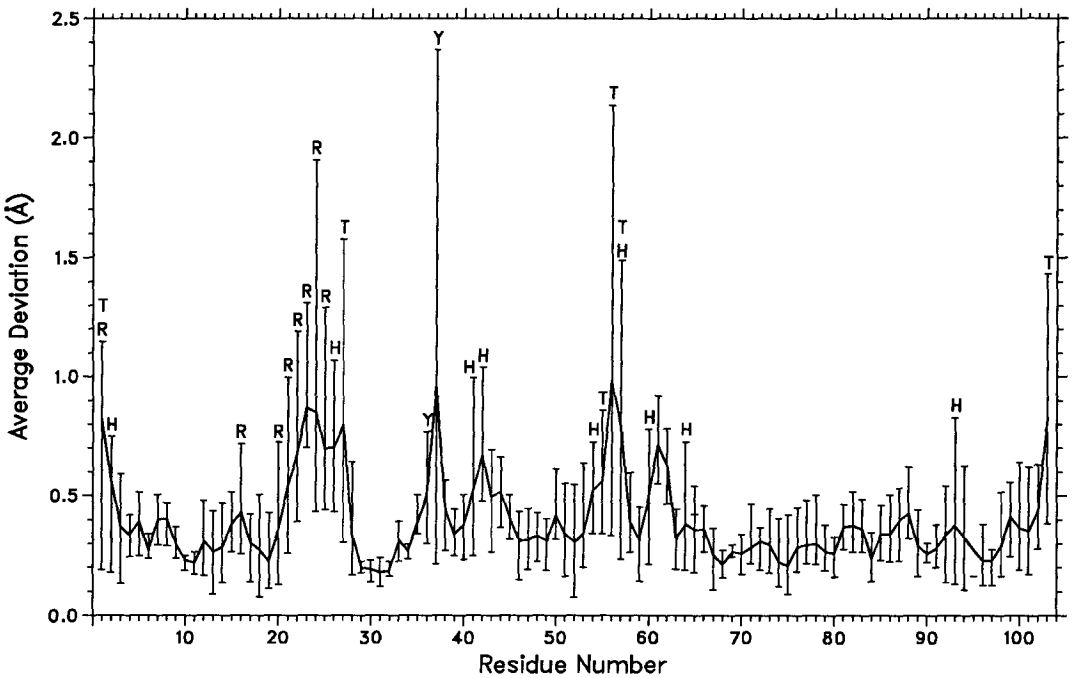

Figure 1.3: A plot of the overall average deviations of main-chain atoms of yeast iso-2, tuna, horse and rice cytochromes *c* from those of yeast iso-1-cytochrome *c* along the course of the polypeptide chain. Only residues 1 to 103, which are common to all five cytochromes *c*, are represented. The vertical bars represent the range of individual pairwise average deviations of main-chain atoms between yeast iso-1 and the other cytochromes *c*. Structures with pairwise deviations both larger than 0.7 Å and 30% greater than the overall average deviations of all structures are labelled at the particular residues involved (H = horse, R = rice, T = tuna and Y = yeast iso-2). *(Figure reproduced with permission of M.E.P. Murphy; Brayer & Murphy (1993))*

these hydrogen bonds are conserved in all eukaryotic cytochromes *c*.

Located around the exposed heme edge is a "ring" of positively charged residues which are functionally conserved in most eukaryotic cytochromes *c* (see Figure 1.6). Chemical modification studies have implicated these residues as being involved in complexation with complementary negatively charged residues on the surfaces of redox partners (for a review see Margoliash & Bosshard, 1983). This has led to the proposal that the region around the exposed heme edge is the binding site of cytochrome *c* (Dickerson & Timkovich, 1975). Model-building studies of cytochrome *c* with cytochrome b_5 as well as the recent structure of the complex of cytochrome *c* with CCP reaffirm this hypothesis (Salemme, 1976; Pelletier & Kraut, 1992).

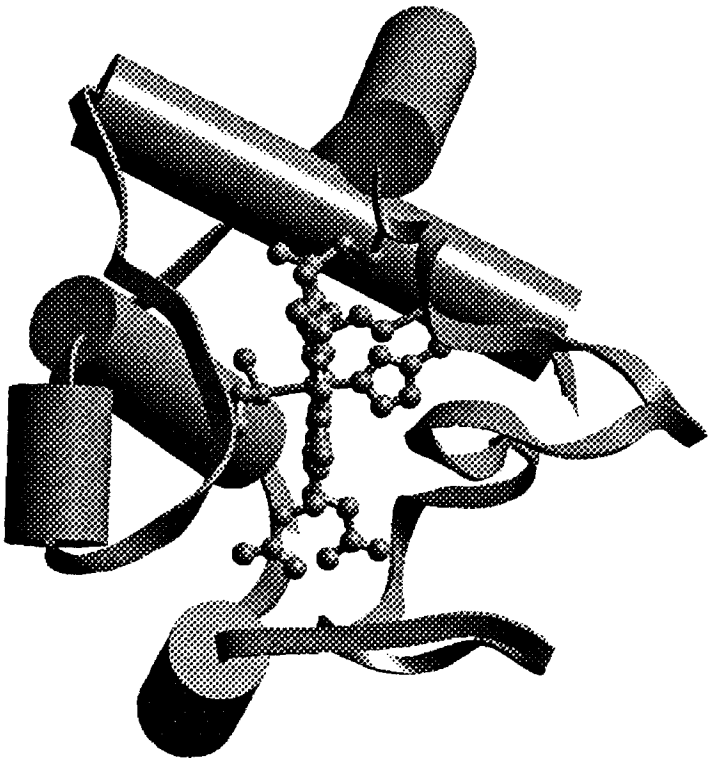

Figure 1.4: A schematic representation of the structure of yeast iso-1-cytochrome *c* in the standard orientation used herein, i.e. heme group is viewed edge on, propionates point downwards and the His18 and Met80 ligands are located right and left of the porphyrin ring, respectively. To highlight the cytochrome *c* fold, α-helixes are represented as cylinders and the remaining peptide chain as a ribbon. Also shown are the two cysteine residues Cys14 and Cys17 which are covalently attached to the heme group.

Further comparisons between the five high resolution cytochrome *c* structures available have identified four conserved water molecules (Brayer & Murphy, 1993). Two of these are of particular interest since they are located inside the protein matrix. The first is Wat121 which is positioned adjacent to Arg38 and the propionate A carboxyl group. It forms hydrogen bonds to both groups as well as to backbone atoms of residues 39 and 42. The second conserved internal water molecule (Wat166) is located in a cavity on the left side of the cytochrome *c* molecule in the standard orientation (see Figure 1.4). It forms three hydrogen bonds with the side-chains of the highly conserved residues Asn52, Tyr67 and Thr78 (see Table 1.1). Structural studies have implicated this water molecule in the function of cytochrome *c* (Takano & Dickerson,

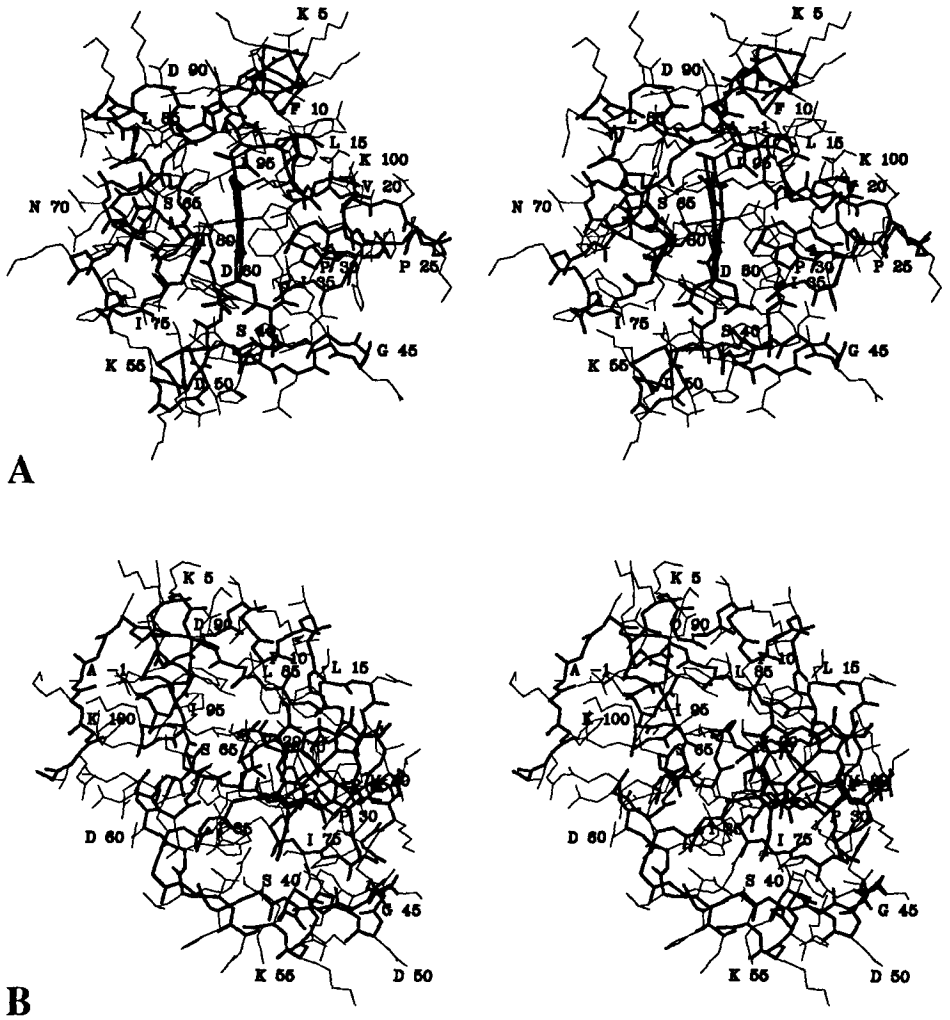

Figure 1.5: A stereo-drawing of the conformations of all main-chain (thick lines) and side-chain (thin lines) atoms in reduced yeast iso-1-cytochrome *c* in (a) the standard view and (b) a view which is rotated 90°. The heme group has also been drawn with thick lines. Also drawn are the two heme ligand bonds to His18 and Met80, and the two covalent thioether linkages to cysteines 14 and 17. For clarity, every 5th α-carbon atom has been labelled with its one-letter amino acid designation and sequence number.

1981b; Bushnell *et al.*, 1990). Other studies suggest, however, that Wat166 is not essential for maintaining the function of cytochrome *c* (McLendon *et al.*, 1991), and possibly serves only to destabilize the protein (Luntz *et al.*, 1989).

Table 1.3: Secondary structural elements present in yeast iso-1-cytochrome *c*

Secondary structure element	Residues involved	
α-Helix	2–14	
β-Turn (type II)	21–24	
γ-Turn	27–29	
β-Turn (type II)	32–35	
β-Turn (type II)	35–38	
β-Turn (type II)	43–46	(distorted†)
α-Helix	49–55	(residue 55 is distorted)
α-Helix	60–70	
α-Helix	70–75	
β-Turn (type II)	75–78	
α-Helix	87–102	

† Mediated through a water molecule.

1.3 Oxidation State Dependent Conformational Differences in Cytochrome *c*

Despite all that is now known about the structure of eukaryotic cytochromes *c*, a fundamental controversy still remains regarding the nature and role of oxidation state conformational changes. On the one hand, a number of techniques give strong indications that the protein undergoes a dramatic alteration in its conformation upon oxidation. For example, studies suggest that the oxidized state has a significantly increased adiabatic compressibility (Eden *et al.*, 1982). In addition, the radius of gyration as observed by small-angle X-ray scattering is larger for the oxidized protein (Trewhella *et al.*, 1988). The oxidized state of cytochrome *c* is also more susceptible to proteolytic digestion (Nozaki *et al.*, 1958) and thermal and chemical denaturation (Butt & Keilin, 1962; Dickerson & Timkovich, 1975). An increase in hydrogen exchange rates for the oxidized state of cytochrome *c* has been measured as well (Ulmer & Kagi, 1968; Liu *et al.*, 1989). Second derivative amide I infrared spectroscopy also suggests that there are differences in the secondary structure between the oxidation states of cytochrome *c* (Dong *et al.*, 1992). And finally, a differential chemical modification technique (Bosshard & Zurrer,

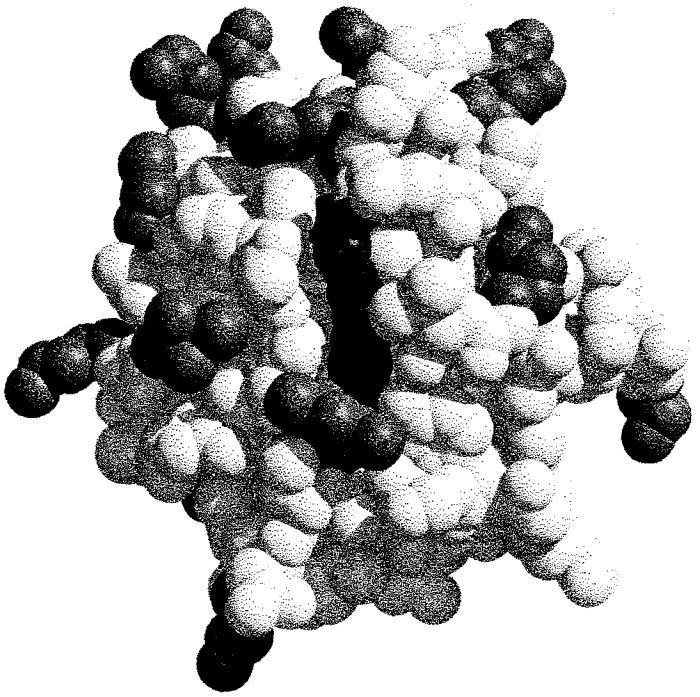

Figure 1.6: A space-filled representation of the structure of yeast iso-1-cytochrome *c* in the standard orientation. Heme group atoms are drawn in black with the solvent exposed edge facing the viewer. Dark grey atoms are associated with positively charged lysine or arginine residues that encircle the exposed heme edge and are believed to be involved in complexation with redox partners.

1980) and one-dimensional NMR studies (Moore, 1983; Williams *et al.*, 1985b) provide further indications that the reduced and oxidized forms of cytochrome *c* have distinctly different conformations.

On the other hand, examination of the three-dimensional structures of eukaryotic cytochromes *c* in differing oxidation states as revealed by x-ray crystallography does not show large differences between the two forms (Takano & Dickerson, 1981a,b; Bushnell *et al.*, 1990). For the reduced and oxidized states of tuna cytochrome *c*, only a few small changes in conformation and a rearrangement of hydrogen bonds near the site of a conserved internal water molecule

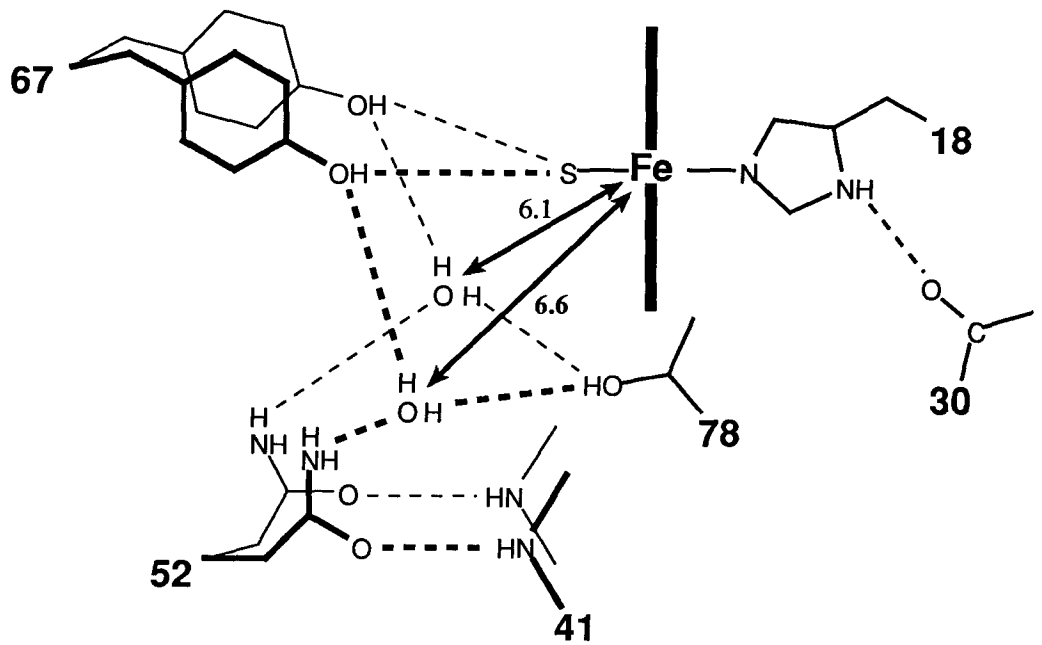

Figure 1.7: A schematic representation of the oxidation state dependent changes in tuna cytochrome *c* as described by Takano and Dickerson (1980,1981b). The reduced state of the protein is shown in thick lines, the oxidized form is displayed in thin lines, and hydrogen bonds are represented by dashed lines. Distances from the internal water molecule (equivalent to Wat166 in yeast iso-1-cytochrome *c*) to the heme iron atom are also indicated on the figure. The distance for the oxidized state is an average for the two molecules in the crystallographic asymmetric unit.

(equivalent to Wat166 in yeast iso-1-cytochrome *c*) were noted (see Figure 1.7). Further, no loss of hydrogen bonds upon oxidation was observed, nor was a change in the main or side-chain flexibility between the oxidation states observed (Takano & Dickerson, 1980, 1981b). Essentially the same results are obtained when comparing cytochrome *c* structures from different species (Bushnell *et al.*, 1990). In this context it should be mentioned that there has been some confusion about the exact assignments of hydrogen bonds around the internal water site. Originally it was considered that residues 41 and 52 hydrogen bond to each other, but based on geometric constraints this cannot be the case. Actually, both Gly41 NH and Asn52 ND2 form hydrogen bonds to the nearby pyrrole A propionate group. Therefore, Figure 1.8 which is based on structures of five eukaryotic cytochromes *c* represents a more correct description of

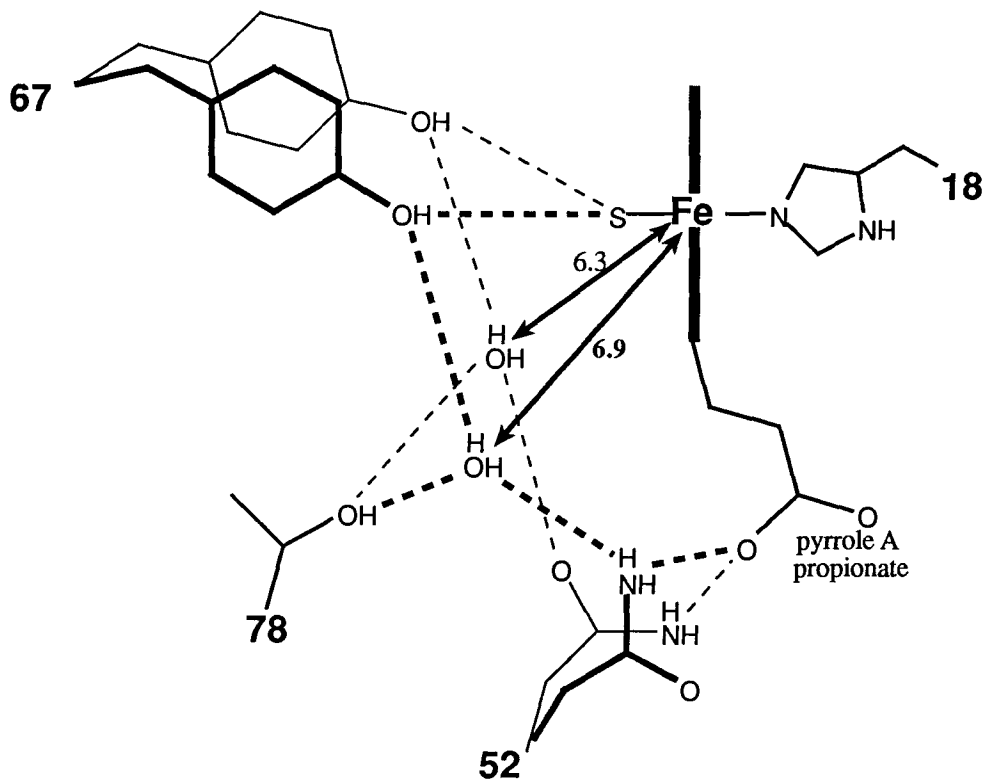

Figure 1.8: A schematic representation of the oxidation state dependent changes found in all five eukaryotic cytochromes *c* whose structures are known as described by Bushnell *et al.* (1990). As in Figure 1.7, the reduced state of the protein is plotted in thick lines, the oxidized form is drawn in thin lines, and hydrogen bonds are represented by dashed lines. Distances from the internal water molecule to the heme iron atom are average values based on the available high resolution structures. The discrepancies with the description of Takano and Dickerson (see Figure 1.7) are due to a more careful analysis of the hydrogen bonds made by Asn 52.

the oxidation state dependent conformational differences than Figure 1.7.

To rationalize the dramatically different experimental results obtained by x-ray crystallography and other biophysical techniques it has been suggested that crystal contacts and high salt concentrations might mask the conformational differences between oxidation states (Williams *et al.*, 1985b; Trewhella *et al.*, 1988; Liu *et al.*, 1989; Dong *et al.*, 1992). Recent two-dimensional NMR experiments, however, disagree with this view and conclude that the solution structure is very similar to that seen by x-ray crystallography (Wand *et al.*, 1986; Feng *et al.*, 1990; Feng & Englander, 1990; Gao *et al.*, 1991). Thus the question of the nature of oxidation state

dependent conformational differences remains to be resolved.

1.4 The Midpoint Reduction Potential of Cytochromes

1.4.1 Introduction

The driving force for biological electron transfer is the difference in the midpoint reduction potentials of the reacting proteins. Among the eukaryotic cytochromes *c* the value of this property appears to be extremely conserved. Although the precise value of reduction potentials is difficult to compare due to their sensitivity to experimental conditions, it appears that the variation in E_m for eukaryotic cytochromes *c* is on the order of only 10–20 mV (Pettigrew & Moore, 1987; McLendon *et al.*, 1991). For example, at 25°C, pH 7.0 and μ=0.01 M, the midpoint reduction potentials for such diverse species as yeast iso-1, tuna and turkey cytochromes *c* are 261, 265 and 260 respectively (Margalit & Schejter, 1973).

The small range of midpoint reduction potentials observed suggests that maintenance of this value is critical for the functioning of cytochrome *c*. It is to be expected that if the reduction potential is raised or lowered by a certain amount the driving force will be too small or large for controlled and efficient electron transfer to occur between cytochrome *c* and its redox partners. However, *in vivo* studies on a yeast mutant cytochrome *c* which has an E_m that is 50 mV lower but does not inhibit growth has raised questions regarding this assumption (McLendon *et al.*, 1991).

1.4.2 Theories for the control of midpoint reduction potential

Because of the importance of the midpoint reduction potential to biological electron transfer, it is not surprising that much research has been done to understand the factors that determine these values. A large number of these studies have focused on the many types of cytochromes that have thus far been isolated. The reduction midpoint potentials of these span a range of nearly 800 mV (Churg & Warshel, 1986). Marchon *et al.* (1982) showed that a variation of ~150 mV can be attributed to different axial ligands to the heme iron atom. Substitutions to

the porphyrin ring such as the thioether linkages in *c*-type cytochromes can cause an alteration in the midpoint potential of an additional ~150 mV (Margoliash & Schejter, 1966; Mashiko *et al.*, 1981; Marchon *et al.*, 1982). However, a variation of nearly 500 mV can still be observed when both the heme ligands and porphyrin ring substitutions are kept constant (Cusanovich *et al.*, 1988). An explanation for this latter observation must be sought in the differences imposed on the heme by the surrounding protein matrix.

Polarity of the heme environment

Kassner showed that the midpoint reduction potential of iron porphyrin ring groups is highly dependent on the dielectric constant of the solution the experiment is carried out in. He proposed that the reduction potential of cytochromes may, therefore, be controlled by the polarity of the heme environment (Kassner, 1972; Kassner, 1973). An examination of seven heme containing proteins by Stellwagen (1978), however, revealed that while the midpoint potential varies over a range of 300 mV for these proteins the polarity of the heme environment is nearly constant. It was observed that the amount of heme solvent exposure did correlate reasonably well with the variation in midpoint potential. Stellwagen concluded that heme solvent exposure rather than the polarity of the heme environment determines the value of the midpoint reduction potential. It should be mentioned that in this study the heme groups of the seven proteins had different axial ligands and varying substituents to the porphyrin ring suggesting caution should be used in the interpretation of these results.

Axial ligand effects

Based on differences in the chemical shift of the methionine ligand in a number of His/Met ligated cytochromes *c*, Moore and Williams (1977) proposed that the midpoint potential can be altered by regulating the iron-sulfur ligand bond length. In this view the conformation of the protein constrains the position of the methionine ligand so as to result in a specific midpoint reduction potential. For example, shortening of the iron-sulfur ligand bond by ~0.1 Å should

result in a drop of 400 mV for the reduction potential. However, it is unlikely that this is the mechanism used by nature to modulate midpoint reduction potential. EXAFS studies have shown that iron-sulfur bond lengths in a variety of cytochromes *c* prove to be independent of the midpoint potential (Korzun *et al.*, 1982).

Another possible mechanism for influencing the midpoint potential is variation of the orientation of the ligand histidine imidazole plane. Korzun *et al.* (1982) remark that upon reduction the added electron will occupy the non-bonding d_{xz} or d_{yz} orbitals of the iron. Steric interference of these orbitals with the histidine ligand imidazole plane will alter the energy levels of the orbitals and thus affect the midpoint potential of the protein. In tuna cytochrome *c* the orientation of the imidazole ring is dependent on oxidation state (Takano & Dickerson, 1981b).

Electrostatic effects

In its simplest form oxidation of cytochrome *c* is essentially a change in the charge of the heme iron atom, and it is expected that electrostatic factors will play an important role in determining reduction potential. The relationship between electrostatic properties and midpoint reduction potential is, however, not a simple one. It is clear from a comparison of related cytochromes *c* that the overall net charge of the protein does not correlate with the midpoint reduction potential (Moore *et al.*, 1986; Cusanovich *et al.*, 1988). Nonetheless, it has been shown that by chemically modifying lysine residues the reduction potential of horse cytochrome *c* and Euglina cytochrome *c*-552 can be altered in a computationally predictable manner (Schejter *et al.*, 1982).

Moore (1983) proposed that midpoint potential variations among homologous cytochromes *c* are caused by differences in the heme propionate charge. One electrostatic interaction that has specifically been implicated in affecting midpoint potential is that between the propionate A carboxyl group and Arg38. This salt-bridge is the only one to the heme propionates and therefore could be important for regulating the midpoint reduction potential. Mutation of Arg38 to an alanine does in fact result in a drop of 50 mV in the reduction potential (Cutler

et al., 1989).

Conformational stabilization

The midpoint reduction potential reflects the relative stability of the two oxidation states. Stabilization of either state by rearrangements in protein conformation will, therefore, influence the midpoint potential. In the past this factor has been considered to play a small role in establishing the reduction potential of eukaryotic cytochromes *c* (Moore *et al.*, 1986). Based on the oxidation state differences seen in tuna cytochrome *c*, computational methods estimate the conformational stabilization to be ~40 mV (Churg & Warshel, 1983; Warshel, 1983). This might be an underestimation since as discussed above there is conflicting data regarding the amount and nature of conformational adjustments upon oxidation. Based on the pH dependence of the reduction potential it has been estimated that the conformational stabilization may account for 70–110 mV (Rogers & Moore, 1988). Measurements of the reorganization energy in cytochrome *c*/cytochrome b_5 kinetic experiments arrive at similar values for the conformational stabilization (~80 mV) (McLendon & Miller, 1985).

1.5 Cytochrome *c* Mutant Studies

1.5.1 Introduction

Site-directed mutagenesis techniques have had a profound impact on the study of eukaryotic cytochromes *c*. Thanks to the ability to alter specific residues, theories regarding the properties of cytochrome *c* and the role certain residues play in establishing these properties can be tested. A good example of this is the study on the role of the invariantly conserved residue Phe82.

Phenylalanine 82 in yeast iso-1-cytochrome *c* was the first residue to be targeted for structure-function studies using site-directed mutagenesis (Pielak *et al.*, 1985). The reason for investigating this residue was that it was implicated in being directly involved in the electron transfer process with CCP (Poulos & Kraut, 1980) and cytochrome b_5 (Wendoloski *et al.*, 1987). Due to its location adjacent to the porphyrin ring, Phe 82 was also considered to be important in

maintaining the midpoint reduction potential (Kassner, 1972, 1973). Studies of several mutants made at the 82 position confirm that this residue is important for both biological electron transfer (Liang *et al.*, 1987, 1988) and regulation of midpoint potential (Rafferty *et al.*, 1990). Structural studies have also shown that the lowered midpoint potentials observed for the serine and glycine mutants can be explained in terms of a decreased hydrophobic environment and an increased heme exposure (Louie *et al.*, 1988b; Louie & Brayer, 1989).

1.5.2 Water-switch mutants

Often mutations at certain residues raise questions rather than provide answers. Recently attention has focused on a group of mutants of highly conserved residues which possess some unusual properties. The location of these mutation sites adjacent to a conserved internal water molecule that moves upon changes in oxidation state has prompted the name "water-switch" mutants for this family of variant cytochromes *c* (see Figure 1.9).

The two most studied mutations of this class of variants are N52I and Y67F. The N52I variant was discovered as a second site revertant to two nonfunctional mutants (Das *et al.*, 1989). Subsequent studies concentrated on this mutant's midpoint reduction potential (Burrows *et al.*, 1991; Langen *et al.*, 1992), electron transfer kinetics (Whitford *et al.*, 1991), *in vivo* growth rates (McLendon *et al.*, 1991), and thermodynamic stability (Hickey *et al.*, 1991). The replacement of Tyr67 to a phenylalanine was made by semi-synthesis techniques more than ten years ago (Koul *et al.*, 1979), but its unusual physiological properties were not recognized until recently. Most studies on Y67F have described the stability of this protein and its dramatically lower midpoint reduction potential (Luntz *et al.*, 1989; Wallace *et al.*, 1989; Rafferty, 1992). But this variant also exhibits altered pH dependent behaviour and electron transfer kinetics (Barker *et al.*, 1991; J.G. Guillemette, private communication).

Table 1.4 gives a comprehensive list of the properties of water-switch mutants and associated references. A number of unique results have been obtained. For example, replacement of an internal hydrophilic amino acid for a hydrophobic one should result in an increase of the

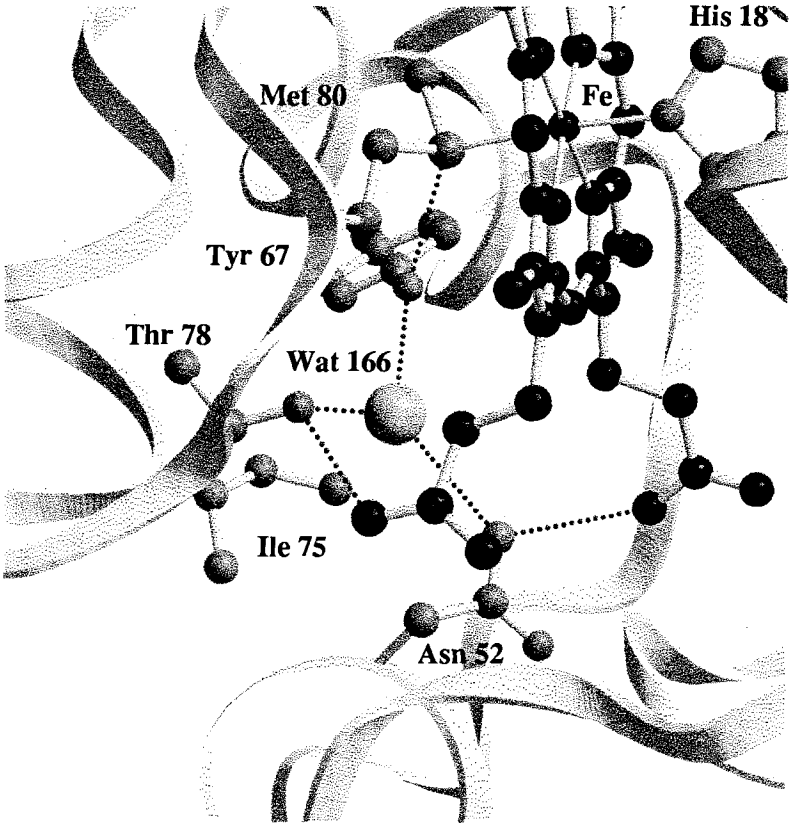

Figure 1.9: A drawing showing the region about an internally bound water molecule (Wat166) in reduced yeast iso-1-cytochrome *c* (Louie & Brayer, 1990). Mutation of residues at positions 52, 67, 75 and 78 results in variants with several unusual properties. In this drawing the course of the polypeptide chain is shown with a ribbon, heme-ligand interactions are indicated by thin white bonds and hydrogen bonds are shown by thin black dotted lines.

midpoint reduction potential of cytochrome *c* (Kassner, 1972,1973). The opposite is, however, observed for all such water-switch mutants, the most striking of which is the N52I-Y67F double mutant which has a 56 mV lower midpoint reduction potential (Table 1.5). In spite of the dramatically lower midpoint potentials of several of the variants, most of them do not appear to be affected in terms of biological activity. Also the stability of four of the water-switch mutants as measured by denaturation studies and alkaline isomerization is increased. In fact, the isoleucine replacement at position 52 is one of the most stabilizing single site mutations made to date (Pace, 1990).

Table 1.4: Properties of water-switch mutants†

Mutant	Properties	Reference
N52A	- E_m is 30 mV lower	(Burrows *et al.*, 1991; Rafferty, 1992)
	- 4° increase in T_m	(Hickey *et al.*, 1991)
	- similar reorganization energy	(Whitford *et al.*, 1991)
	- does not interfere in complexation	(Whitford *et al.*, 1991)
	- similar pK_a	(Rafferty, 1992)
	- lower reorganization energy	(Rafferty, 1992)
N52I	- E_m is 50 mV lower	(Burrows *et al.*, 1991; McLendon *et al.*, 1991; Langen *et al.*, 1992)
	- 13–17° increase in T_m	(Das *et al.*, 1989; Hickey *et al.*, 1991)
	- similar reorganization energy	(Whitford *et al.*, 1991)
	- lower reorganization energy	J.G. Guillemette, private communication
	- does not interfere in complexation	(Whitford *et al.*, 1991)
	- Growth rates unaffected or even better	(McLendon *et al.*, 1991)
	- pK_a of alkaline isomerization increases 1.5 pH units	J.G. Guillemette, private communication
Y67F	- E_m is 40–55 mV lower	(Luntz *et al.*, 1989; Wallace *et al.*, 1989; Frauenhoff & Scott, 1992; Rafferty, 1992)
	- more resistant to extreme pH	(Luntz *et al.*, 1989; Wallace *et al.*, 1989)
	- stronger Fe-S bond	(Luntz *et al.*, 1989)
	- complicated pH curve	(Barker *et al.*, 1991)
	- 56% biological activity ‡	(Koul *et al.*, 1979)
	- biological activity unaffected	(Margoliash, 1990)
	- lower reorganization energy	(Rafferty, 1992)
	- pK_a of alkaline isomerization increases 1.5 pH units	J.G. Guillemette, private communication
N52I-Y67F	- E_m is 55 mV lower	J.G. Guillemette, private communication
	- lower reorganization energy	J.G. Guillemette, private communication
	- pK_a of alkaline isomerization increases 2.5 pH units	J.G. Guillemette, private communication
I75M	- E_m is 45 mV lower	(Rafferty, 1992)
	- similar reorganization energy	(Rafferty, 1992)
	- similar pK_a of alkaline isomerization	(Rafferty, 1992)

† Most of the variants discussed in the references are of yeast iso-1-cytochrome *c* created by site-directed mutagenesis. However, Luntz *et al.*, (1989) studied a mutant of rat cytochrome *c* and Koul *et al.*, (1979), Wallace *et al.*, (1989) and Frauenhoff & Scott (1992) studied Y67F mutants made by semi-synthesis techniques from horse cytochrome *c*.

‡ This observation is not supported by more recent studies (Margoliash, 1990; McLendon *et al.*, 1991).

Table 1.5: Thermodynamic properties for the midpoint reduction potential of wild-type and mutant yeast iso-1-cytochromes *c*

	E_m† (mV)	$\Delta G°$ (kcal/mole)	$\Delta S°_{rc}$ (e.u.)	$\Delta H°$ (kcal/mole)
Wild-type	290	-6.7	-9.1	-14.0
N52A	257 (-33)	-5.9 (+0.8)	-8.0 (+1.1)	-12.9 (+1.1)
N52I	232 (-58)	-5.3 (+1.4)	-8.2 (+0.9)	-12.4 (+1.6)
Y67F	234 (-56)	-5.4 (+1.3)	-7.5 (+1.6)	-12.3 (+1.7)
N52I-Y67F	234 (-56)	-5.4 (+1.3)	-8.5 (+0.6)	-12.6 (+1.4)
I75M	245 (-45)	-5.7 (+1.0)	-11.7 (-2.6)	-13.8 (+0.2)
T78G	249 (-41)	-5.7 (+1.0)	-6.3 (+2.8)	-12.3 (+1.7)

This data is from S.P. Rafferty and J.G. Guillemette (University of British Columbia). Values given in brackets indicate the difference between the mutant and wild-type thermodynamic parameters.
†25°C, pH 6.0, μ=0.1 M, and SHE reference

Several fundamental questions are posed by these results. Why are residues surrounding this internal water molecule so highly conserved while they do not appear to be essential for biological activity? What is the reason for retaining the midpoint potential for eukaryotic cytochromes *c* at such a precise value? How can single (or double) site mutations have such a large effect on the midpoint reduction potential of cytochrome *c* and the stability of this protein?

1.6 Thesis Objectives

It is clear from the above overview that, although much is known about cytochrome *c*, many questions still remain. The nature of the conformational changes that occur upon change in oxidation state and the function of these changes in the electron transfer event are particularly unclear. The main objective of the work in this thesis is to identify the structural differences occurring between the two oxidation states in eukaryotic cytochromes *c* and evaluate the role they play in biological electron transfer.

Identification of the conformational differences between oxidation states has been accomplished by solving the three-dimensional structure of the oxidized form of yeast iso-1-cytochrome *c* to high resolution using x-ray crystallographic techniques, and comparing this structure to that of the reduced form (Louie & Brayer, 1990). Further assessment of these structural differences has been done by studying mutants of yeast iso-1-cytochrome *c* with amino acids altered at one of the focal points of oxidation state dependent conformational changes, namely the internal water molecule Wat166. In total, the structures of five mutants have been solved (N52A, N52I, Y67F, N52I-Y67F and I75M). Of these, three have been studied in both oxidation states. These results provide insight into the nature of a number of functional properties of cytochrome *c* and their relationship to the structural framework on which they must operate.

Chapter 2

General Overview of Experimental Methods

2.1 Crystallization

2.1.1 The hanging drop method

A wide variety of techniques have been developed for growing protein crystals (Giegé & Mikol, 1989; McPherson, 1990). For the growth of wild-type and mutant yeast iso-1-cytochrome *c* crystals the method of vapor diffusion in hanging drops has proven very successful (Sherwood & Brayer, 1985; Brayer & Murphy, 1993).

In a typical setup 5 μl of protein solution is suspended from a siliconized cover slip over 1 ml of mother-liquor (Figure 2.10). The mother-liquor used is 0.1 M sodium phosphate buffer in the pH range of 5–8 which is 80–95% saturated with ammonium sulfate. For the purpose of keeping the protein in a specific oxidation state 10–70 mM of a reducing or oxidizing agent is added. As reducing agents DTT and sodium dithionite have been used, whereas sodium nitrite

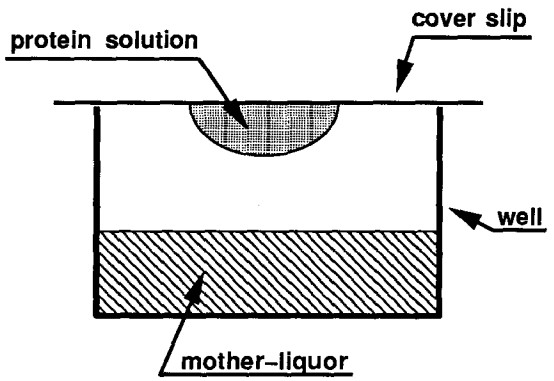

Figure 2.10: Experimental setup for vapor diffusion in hanging drops used in crystallizing yeast iso-1-cytochrome *c* and mutants thereof.

and potassium ferricyanide have been added to the mother-liquor to maintain an oxidizing environment. The solution in the hanging droplet is identical to that of the reservoir with the exception that it is less saturated with ammonium sulfate (70–85%) and there is protein present (~20 mg/ml).

The ammonium sulfate concentration in the hanging drop is chosen such that cytochrome *c* is just soluble. The concentration of ammonium sulfate in the mother-liquor is made slightly higher. Over time the protein solution and mother-liquor equilibrate through vapor diffusion. The result is that the protein solution will gradually attain the same ammonium sulfate concentration as that of the mother-liquor and the protein droplet will become supersaturated. If conditions are favorable nucleation sites form from which protein crystals grow.

2.1.2 The hair-seeding technique

Unfortunately spontaneous crystal formation has rarely been observed for wild-type or mutant forms of yeast iso-1-cytochrome *c*. This has been attributed to a lack of nucleation sites present in the protein droplet (Louie, 1990). To overcome this deficiency a hair-seeding technique has been used (Leung *et al.*, 1989). In this method, miniscule remnants of crushed cytochrome *c* crystals are inserted into the protein droplet using a hair. These remnants form the nucleation sites for crystal growth. A major advantage of using crushed yeast iso-1-cytochrome *c* crystals as seeds, is that the resulting crystals grown are found to be isomorphous with those of the wild-type protein. This property not only greatly simplifies the structure solution process for mutant proteins, but also allows for careful comparisons to be made between structures without the added complexity of differing crystal contacts. Thus, all the crystals discussed in this thesis have the same space group as wild-type yeast iso-1-cytochrome *c*, namely $P4_32_12$, with comparable cell dimensions $a = b = 36.46$–36.56 Å and $c = 137.02$–139.12 Å (cell dimensions for reduced wild-type yeast iso-1-cytochrome *c* are $a = b = 36.46$ Å and $c = 136.86$ Å).

2.1.3 Changing oxidation states of crystals

Except for the oxidized wild-type yeast iso-1-cytochrome *c* structure, all crystals used for determining cytochrome *c* structures in the oxidized form were initially grown in a reducing environment. To obtain crystals containing cytochrome *c* in the oxidized state, reduced crystals were transferred to mother-liquor which contained potassium ferricyanide instead of a reducing agent. The transfer was done gradually, i.e. crystals were washed with a series of solutions containing mother-liquor with decreasing levels of reducing agent and increasing levels of the oxidizing agent. During this process crystals were observed to change color from red to deep brown-red, signifying the oxidation of the heme iron atom from Fe^{2+} to Fe^{3+}. After the transfer was completed crystals were soaked in mother-liquor containing only potassium ferricyanide for at least one hour before mounting.

2.2 Data Collection and Data Processing

2.2.1 Methods of data collection

For the structures described in this thesis two methods of data collection have been employed. Six structures were solved with data sets obtained using an Enraf-Nonius CAD4-F11 diffractometer. The remaining three data sets were collected using the Rigaku R-Axis II imaging plate area detector. The reason that two data collection methods were used is historical. During most of the period in which the work described was done the laboratory was equipped with only the CAD4-F11 diffractometer for collecting diffraction data. Recently the laboratory acquired an area detector and thus the latest data sets were all measured using this new technology.

The approach taken by the two data collection methods is substantially different. With the Enraf-Nonius diffractometer each reflection is measured individually, whereas the Rigaku area detector measures many reflections simultaneously in a manner similar to the photographic rotation method (Arndt & Wonacott, 1977). The result is that a complete data set can be collected much more rapidly using the area detector. For example, collection of a 2.0 Å data

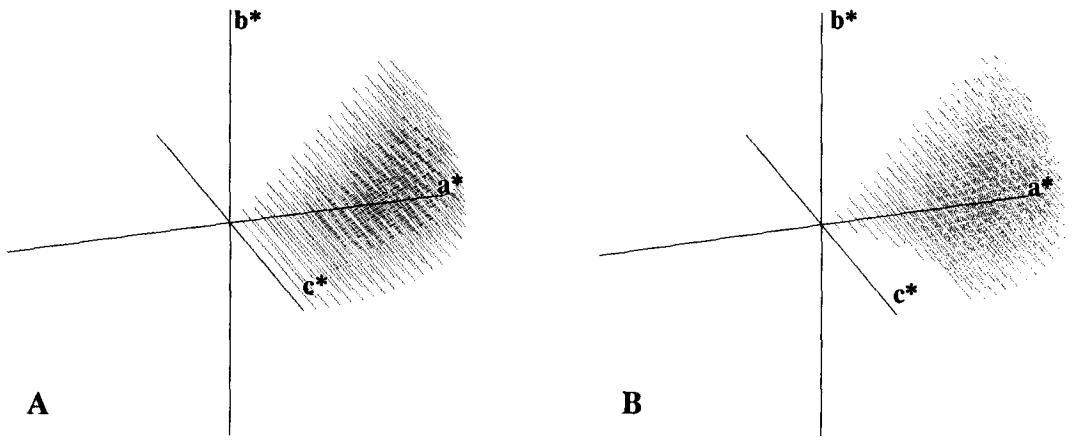

Figure 2.11: Graphical representation of the completeness of the measured diffraction data obtained from two mutant cytochrome *c* crystals using two different data collection tools (see also Figure 2.12). Displayed in (a) is a view of reciprocal space showing the 100% complete 2.0 Å N52A data set obtained from an Enraf-Nonius CAD4-F11 diffractometer. Each point in this figure represents a measured reflection. As shown, it is necessary to only collect $\frac{1}{16}$ of reciprocal space since this is the unique part for the tetragonal space group $P4_32_12$. Shown in (b) is an identical representation for a 2.0 Å data set from an N52I-Y67F crystal collected on the R-Axis II imaging plate area detector under unfavorable geometric conditions. While the diffractometer data set is 100% complete, the cusp region around the c* axis is absent in the data set obtained with the area detector.

set takes about two weeks on a diffractometer versus one day on the Rigaku area detector. However, due to geometrical reasons, it is not always possible to collect 100% of the available data with an area detector, while this is not a problem with a diffractometer (Figure 2.11). Nonetheless, by using multiple scans on the area detector at alternative settings, almost all the available data is usually accessible. A further advantage of area detector data is the quality of the data collected (as measured by counting statistics) which is far superior to that obtained from a diffractometer (see Figure 2.12). This allows for the collection of data sets to higher resolution and from much smaller crystals than is possible with a diffractometer.

2.2.2 Processing of diffraction intensity data

In order to convert measured intensities into structure factors, a number corrections have to be applied to the data set. First, they must be corrected for background scatter which is mainly

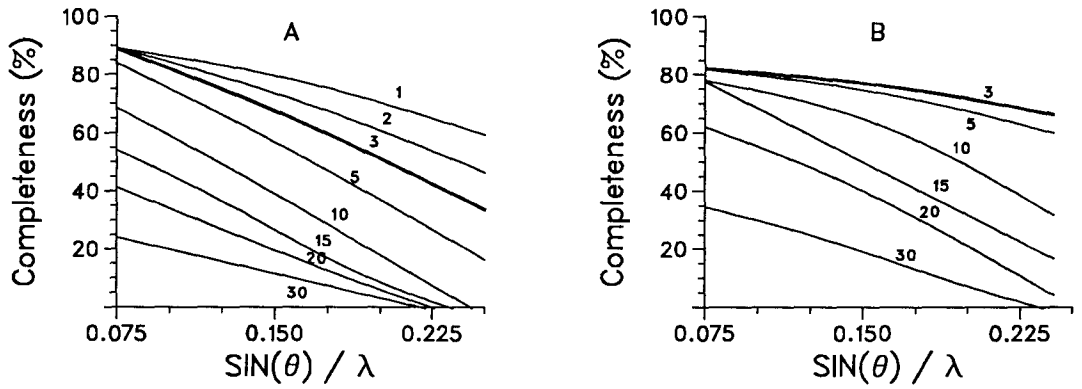

Figure 2.12: A plot of the completeness of the diffraction data obtained from two mutant cytochrome *c* crystals using two different data collection tools as a function of $\frac{F}{\sigma_F}$ (see also Figure 2.11). Displayed in (a) is the 2.0 Å N52A data set obtained from an Enraf-Nonius CAD4-F11 diffractometer. Shown in (b) is a 2.0 Å data set from an N52I-Y67F crystal collected on an R-Axis II imaging plate area detector under unfavorable geometric conditions. As can be seen, the number of reflections above 3 sigma (thick lines) is larger in the area detector data set compared to the one obtained with the diffractometer, despite the fact that the data set is less complete overall. This is the result of the large number of multiple reflections measured by this method and the improved counting statistics obtained as a result.

caused by non-crystalline material in the beam path. A second correction is for absorption. This arises since crystals are generally not spherical in shape resulting in unequal path-lengths through the crystal for different reflections. As a consequence, the amount of absorption of x-rays by the crystal will also vary for different reflections. A further correction is for decay occurring due to crystal degradation during x-ray exposure. The final two corrections to be applied are those for Lorentz and polarization effects. The Lorentz correction is required because reflection intensity is dependent on the geometry of the data collection method. The polarization correction adjusts for the variation of reflection intensity with the amount of refracted x-ray beam polarization. This latter correction factor is dependent on the source of the x-ray beam as well as on the geometry of the diffraction experiment. Once these corrections have been made reflection intensities can be converted into structure factors by taking the square root of the intensity.

Diffractometer data processing

With a diffraction intensity data set obtained using a diffractometer, background correction is relatively easy. The measurement of each reflection is divided into three parts: the initial and final $\frac{1}{6}$ of the scan are taken as the background radiation; the intermediary $\frac{2}{3}$ of the scan represents the intensity of the reflection. The measured background for each reflection can thus be used to correct for the background radiation of that reflection. To improve the background correction, background averaging can be used. In this context background averaging means that an estimate of the background radiation for each reflection is not only based on the background measurements of that reflection but also on the backgrounds of reflections which are close in reciprocal space. For the processing of yeast cytochrome *c* diffraction data it was found that using reflections within a radius of 0.0005–0.0010 Å^{-1} (~6–18 neighbouring reflections) gave good results.

The absorption correction is applied using the empirical method of North *et al.*, (1968). In this method a reflection which is aligned along the ϕ-axis of data collection is used. The measured intensity of this "ϕ-independent" reflection varies with rotation of the ϕ-axis due to absorption and this variation can be used to estimate the amount of correction required for a general reflection. Measurements conducted on several mutant cytochrome *c* crystals show that absorption is not only dependent on ϕ, but also on the resolution of the reflection (Figure 2.13). By measuring several ϕ-independent reflections at different θ values an estimate can be made of the resolution component of the absorption correction.

Correction for intensity decay is performed by measuring a number of standard reflections repeatedly over the course of the full exposure time. Generally the standard reflections are grouped according to resolution so that the decay correction can be applied on this basis as well since high resolution reflections are much more sensitive to crystal decay than low resolution reflections (Figure 2.14). For a detailed description of the applied corrections and their implementation in the diffractometer data processing program ICP, which was used herein for structure determinations, see Murphy (1993).

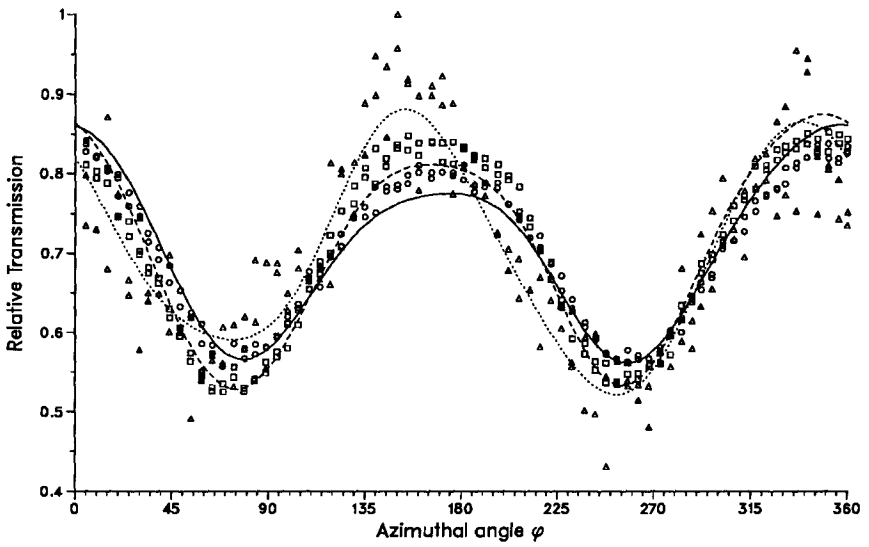

Figure 2.13: A plot of a typical absorption curve for a cytochrome *c* crystal (N52A mutant yeast iso-1-cytochrome *c*). Shown are the relative transmissions as a function of the azimuthal angle for three ϕ independent reflections with different resolutions (○ 8.7 Å, □ 3.9 Å and △ 2.2 Å). Also shown are fitted curves for the three resolutions (solid line 8.7 Å, dashed line 3.9 Å and dotted line 2.2 Å) which were used to correct for absorption by the method of North *et al.*, (1968).

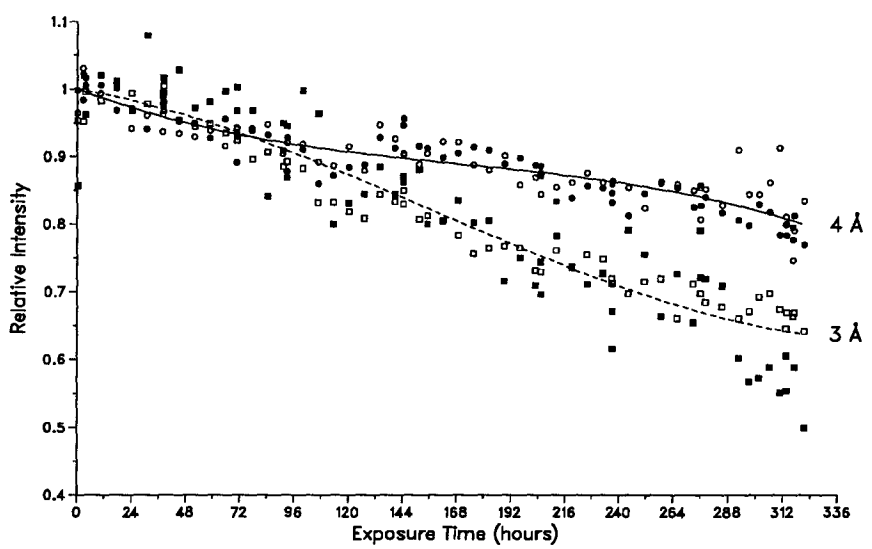

Figure 2.14: A plot of the decay profile for four reflections which were periodically measured during diffractometer data collection for a representative cytochrome *c* crystal (I75M mutant yeast iso-1-cytochrome *c*). For decay correction these four reflections were divided in two groups depending on their resolution (group 1: ○ 4.3 Å and ● 3.5 Å; group 2: □ 3.4 Å and ■ 2.8 Å) to which curves were fitted (solid line - group 1, dashed line - group 2). The fitted curves were then subsequently used for a resolution dependent decay correction of all general reflections measured.

Area detector data processing

Since the method of data collection on the Rigaku R-Axis II is essentially identical to the screen-less oscillation or rotation method (Arndt & Wonacott, 1977), processing of collected data resembles the procedure described by Rossmann *et al.*, (1979). This means that for background corrections the measured intensity directly surrounding a peak is used for estimating the background radiation for that reflection. The correction for decay is performed by scaling between data collection frames. Two frames are scaled with respect to each other by minimizing the function:

$$\varphi = \sum_{hkl} (I^a_{hkl} - Ke^{-B(\frac{\sin\theta}{\lambda})^2} I^b_{hkl})^2 \tag{2.1}$$

In this function I^a_{hkl} and I^b_{hkl} are identical reflections present on the two frames a and b. Since the scale factor is of the form $Ke^{-B(\frac{\sin\theta}{\lambda})^2}$ a resolution dependent decay correction is applied. There is no explicit absorption correction, but absorption effects are corrected for by the above mentioned inter-frame scaling and by merging of symmetry related reflections.

An indicator of the quality of the merged diffraction data is the agreement between reflections which are measured more than once (i.e. duplicates and symmetry mates). This agreement is generally expressed in the form of the merging R-factor:

$$R_{merge} = \frac{\sum_{hkl} \sum_{i=0}^{n} |I_{i_{hkl}} - \overline{I}_{hkl}|}{\sum_{hkl} \sum_{i=0}^{n} I_{i_{hkl}}} \tag{2.2}$$

In this expression I_i is one of n identical reflections and $\overline{I}$ is the average intensity for these n reflections. For the three data sets collected on the Rigaku R-Axis II area detector and discussed in this thesis, values of 6.3–8.4% were obtained. In contrast, merging R-factors of 13.4% and higher were obtained for diffractometer data. This difference in merging R-factors is, however, not completely caused by a difference in quality of the diffraction data. For data sets collected on a diffractometer the merging R-factor is intrinsically higher due to the limited number of repeat measurements.

2.2.3 Scaling of data sets

The structure factor data obtained from a diffraction experiment is on an arbitrary scale and must therefore be scaled such that it can be compared to calculated structure factors. There are several methods available to obtain the necessary absolute scale factor. For the structures discussed in this thesis a method analogous to that described by Wilson (1942) was used for determining an initial absolute scale factor. The value of this scale factor was subsequently improved upon during structure refinement. For a detailed description of the methods used to determine the absolute scale factor see Appendix A.

2.3 Structure Refinement

An immediate advantage in growing crystals of mutant cytochromes *c* isomorphous to those of the reduced wild-type structure is that it facilitates the process of structure solution since difference Fourier techniques can be used to build an initial structural model.

2.3.1 Construction of the starting model

The first step in constructing a starting model for refinement is the examination of a difference Fourier map. The coefficients for this map are $(F^{wt}_{hkl} - F^{mut}_{hkl})$, where F^{wt} is the observed structure factor of the wild-type yeast iso-1-cytochrome *c* structure (reduced state) and F^{mut} is the structure factor from the mutant cytochrome *c* protein that is being studied. The phases used are those of the refined wild-type structure. This Fourier map shows positive electron density peaks where there are atoms in the wild-type protein, which are absent in the to be refined structure, and *vice versa* for the negative peaks (see Figure 2.15a).

However, such maps are generally insufficient for building a complete starting model and therefore, for the structures discussed in this thesis a more conservative approach was taken. Starting with the wild-type yeast iso-1-cytochrome *c* structure, atoms which were close to either negative or positive peaks in the difference Fourier map were removed from this structure. In practice this means that mutated residues were represented by alanine residues and water

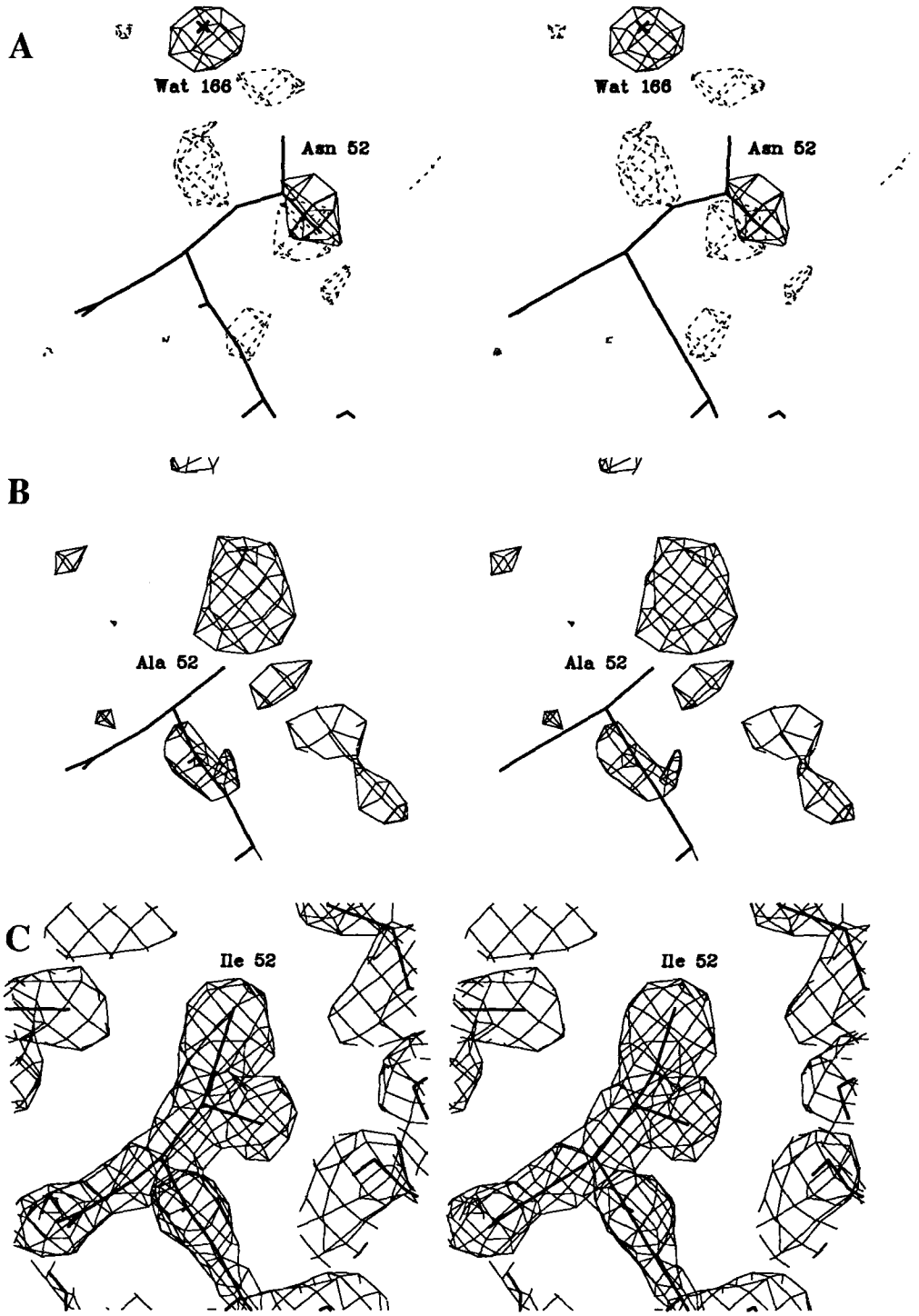

Figure 2.15: Stereo-diagrams of electron density maps in the region of the mutation site of the N52I mutant yeast iso-1-cytochrome *c* (oxidized) at different stages of structural refinement. In (a) is shown the initial difference electron density map used for constructing a starting model. Positive peaks are represented by the cages drawn with solid lines, whereas the dashed line cages indicate negative peaks. A drawing of the $(F_o - F_c)$ difference electron density map after 11 cycles of refinement prior to fitting the missing atoms of the mutated residue is shown in (b). The $(2F_o - F_c)$ difference map of the final refined structure is displayed in (c).

molecules near mutation sites were deleted from the model.

2.3.2 Reciprocal-space refinement

The refinement of the cytochrome *c* structures examined herein was carried out with the restrained parameter least-squares reciprocal-space refinement program PROLSQ (Hendrickson & Konnert, 1980, 1981; Hendrickson, 1985). A discussion of the theory behind this structure refinement approach is given in Appendix B. Refinements were terminated when overall shifts were small (overall r.m.s. shift smaller than 0.05 Å) and when no further improvements could be made by inspection of difference electron density maps (Figure 2.16). The total number of refinement cycles necessary for each structure was highly dependent on the amount of diffraction data available. In general, cytochrome *c* structures for which more than 4000 unique reflections had been collected could be refined in less than 60 cycles. When fewer reflections were available, refinement required at least 70 cycles to reach convergence. The structures obtained all had good agreement between the observed and calculated structure factors (R-factor range 17–22%) and good stereochemistry (see Table 2.6).

2.3.3 Manual interventions

Inspection of $(F_o - F_c)$ and $(2F_o - F_c)$ difference electron density maps calculated over the course of refinement was used to identify missing atoms excluded from the original starting model. This applies not only to mutated residues (see Figure 2.15b) but also to solvent molecules. New solvent molecules were only included in the refinement model if they had reasonable hydrogen bond partners and were consistently observed in $(2F_o - Fc)$ difference electron density maps. In latter stages of refinement $(3F_o - 2F_c)$ difference maps were also examined and proved to be very helpful in repositioning the side-chains of surface residues.

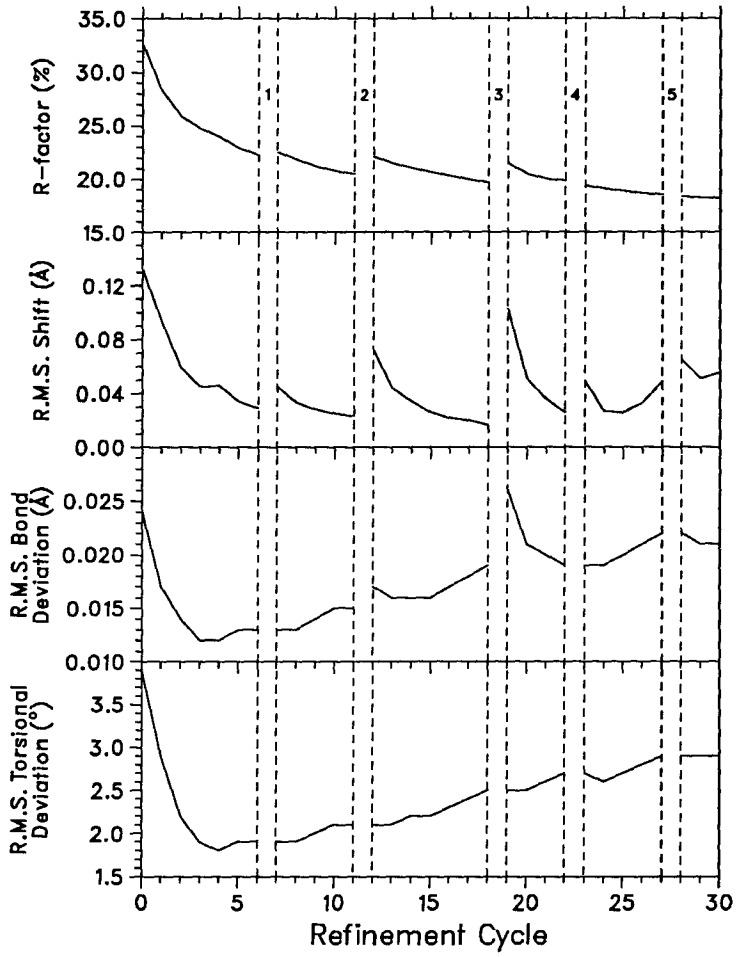

Figure 2.16: A plot of the variation of the R-factor, overall r.m.s. shifts, and r.m.s. bond length and torsional angle deviations from ideal values, during the course of refinement for the oxidized yeast iso-1-cytochrome *c* N52I mutant. Five manual interventions were carried out during refinement as indicated by the numbered vertical breaks. In the first intervention water molecules with thermal factors over 40 $Å^2$ were deleted and the absolute scale of the observed structure factors was adjusted. In the second intervention the Ile52 side-chain was fitted (see Figure 2.15b) and further adjustments to the solvent structure and the absolute scale were made. The remaining three manual interventions were used to further adjust solvent structure and position solvent exposed side-chain groups with high degrees of mobility.

2.3.4 Accuracy of Structures

Several methods have been devised to estimate the errors in the atomic coordinates of macromolecules determined using x-ray diffraction methods. For the structures described in this thesis

Table 2.6: Typical stereochemistry for refined structures of yeast iso-1-cytochrome *c*

Stereochemical refinement parameters	r.m.s. deviation from ideal values	Typical refinement restraint weighting values
Bond distances (Å)		
1-2 bond distance	0.017–0.024	0.020
1-3 bond distance	0.044–0.050	0.030
1-4 bond distance	0.052–0.067	0.050
special distances†	0.048–0.101	0.060
Planar restraints (Å)	0.016–0.018	0.020
Chiral volume ($Å^3$)	0.197–0.286	0.150
Non-bonded contacts‡ (Å)		
single-torsion	0.218–0.238	0.250
multi-torsion	0.179–0.242	0.250
possible hydrogen bonds	0.212–0.262	0.250
Torsion angles (°)		
planar (0°or 180°)	2.3–2.7	2.5
staggered (±60°,180°)	22.3–26.7	20.0
orthonormal (±90°)	18.3–24.6	15.0

† The special distances define the bonds between the heme iron and the His18 and Met80 ligands.
‡ The r.m.s. deviations from ideality for this class of restraint incorporates a reduction of 0.2 Å from the radius of each atom involved in a contact.

two methods are used: the Luzzati plot (Luzzati, 1952) and Cruickshank formula (Cruickshank, 1949, 1954, 1985). Both methods give similar estimates for the accuracy of the cytochrome *c* structures described herein with overall r.m.s. coordinate errors in the range of 0.15–0.25 Å. A thorough discussion of coordinate error determination is given in Appendix C.

Chapter 3

Oxidation State-dependent Conformational Changes in Cytochrome *c*

3.1 Experimental Procedures

3.1.1 Crystallization

Yeast iso-1-cytochrome *c* was isolated from *Saccharomyces cerevisiae* as previously described (Pielak *et al.*, 1985, 1986). Crystals of the oxidized form of this protein were grown using the hanging drop method and employing a hair seeding technique (Leung *et al.*, 1989; see also Section 2.1). Crystal growth was observed under similar conditions to those previously reported for the reduced form of yeast iso-1-cytochrome *c* (Sherwood & Brayer, 1985; Louie *et al.*, 1988a). The buffer used was 0.1 M sodium phosphate, pH 6.2 which contained 30 mM sodium nitrate instead of a reducing agent such as DTT. As precipitant a 90% saturated ammonium sulfate solution was used. In this manner small crystals ($0.2 \times 0.1 \times 0.05$ mm) were grown in 2 to 3 days. To obtain larger crystals, macro seeding was performed. After 7 to 10 days, crystals suitable for high resolution x-ray diffraction ($0.5 \times 0.4 \times 0.2$ mm) were obtained. In order to ensure complete oxidation of the protein, crystals were transferred into a solution containing 20 mM potassium ferricyanide prior to mounting for diffraction studies. Spectroscopic analyses of crystalline material confirmed that the iso-1-cytochrome *c* present was in the oxidized form. Precession camera photography and further diffractometry measurements showed crystals of oxidized iso-1-cytochrome *c* were isomorphous with those grown for the reduced form of this protein. The space group is $P4_32_12$ with cell dimensions of $a = b = 36.47$ Å and $c = 137.24$ Å (cell dimensions of reduced iso-1-cytochrome *c* are $a = b = 36.46$ Å and $c = 136.86$ Å).

3.1.2 Data collection and data processing

Diffraction data was collected to 1.9 Å resolution from a single crystal on an Enraf-Nonius CAD4-F11 diffractometer, with a 36.8 cm crystal to counter distance and a helium purged path for the diffracted beam. The radiation used was nickel filtered and generated from a copper target X-ray tube operating at 26 mA and 40 kV. A total of 8632 intensities were measured using continuous ω scans of 0.6° in width at a scan speed of 0.55° min^{-1}, with backgrounds taken as the terminal one-sixth of the total scan width on either side of each reflection. Three standard reflections were measured every 2.8 h of X-ray exposure time to monitor crystal decay and slippage. The ambient temperature during data collection was maintained at 15°C.

To obtain structure factors, diffraction intensities were first corrected for backgrounds. In this process the measured backgrounds for each reflection, as well as those of neighboring reflections, were used to estimate the correction factor. Intensities were further corrected for absorption using an empirical curve (North *et al.*, 1968) obtained by measuring a phi-independent reflection in 72 consecutive 5° steps. The correction for crystal decay was performed using a polynomial fit to the decay profiles of the three monitor reflections. Diffraction intensities were then merged, corrected for Lorentz and polarization effects and converted into structure factor amplitudes (7925 in total; see also Section 2.2). These were put on an absolute scale using the Wilson plot statistical method (Wilson, 1942; see also Appendix A).

3.1.3 Refinement and analyses

Initially, a reduced minus oxidized iso-1-cytochrome *c* difference electron density map was calculated. The reduced structure used to phase difference map coefficients had previously been refined to 1.7 Å resolution with an R-factor of 19.7% (Louie & Brayer, 1990). Since this map indicated that no large positional shifts were present between the two oxidation states, restrained parameter refinement (Hendrickson & Konnert, 1981; see also Appendix B) of the oxidized protein was initiated using the reduced protein structure as a starting model. Included were 65 water molecules from the reduced structure with thermal factors less than 30 $Å^2$. Also

Table 3.7: Final stereochemistry of oxidized yeast iso-1-cytochrome *c* at 1.9 Å resolution

Stereochemical refinement parameters	r.m.s. deviation from ideal values	Refinement restraint weighting values
Bond distances (Å)		
1-2 bond distance	0.022	0.020
1-3 bond distance	0.046	0.030
1-4 bond distance	0.057	0.050
Planar restraints (Å)	0.018	0.020
Chiral volume ($Å^3$)	0.231	0.150
Non-bonded contacts† (Å)		
single-torsion	0.226	0.250
multi-torsion	0.227	0.250
possible hydrogen bonds	0.254	0.250
Torsion angles (°)		
planar (0°or 180°)	2.8	2.5
staggered (±60°,180°)	26.7	19.0

† The r.m.s. deviations from ideality for this class of restraint incorporates a reduction of 0.2 Å from the radius of each atom involved in a contact.

included were four internal water molecules and a sulfate anion. Refinement utilized 3929 structure factors which had a $F/\sigma(F)$ ratio $\geq$ 2.0 and were within the resolution range of 6.0–1.9 Å. During the course of refinement three manual interventions were carried out based on $F_o - F_c$ difference and fragments maps, and $2F_o - F_c$ and $3F_o - 2F_c$ difference maps. In the first, four more water molecules were removed from the model, several water molecules were repositioned, and some surface lysine and arginine side-chains were placed in better fitting density. During the second manual intervention special effort was put into the positioning of the Asn52 side-chain which had ill-defined density. In the last manual intervention only the terminal side-chain group of a number of asparagine and glutamine residues were adjusted to optimize electron density fits and hydrogen bonding interactions.

After a total of 117 cycles of refinement, convergence was reached to give a final crystallographic R-value of 19.7%. As documented in Table 3.7, the final structure determined for oxidized yeast iso-1-cytochrome *c* has good stereochemistry. From a plot of R-factor versus

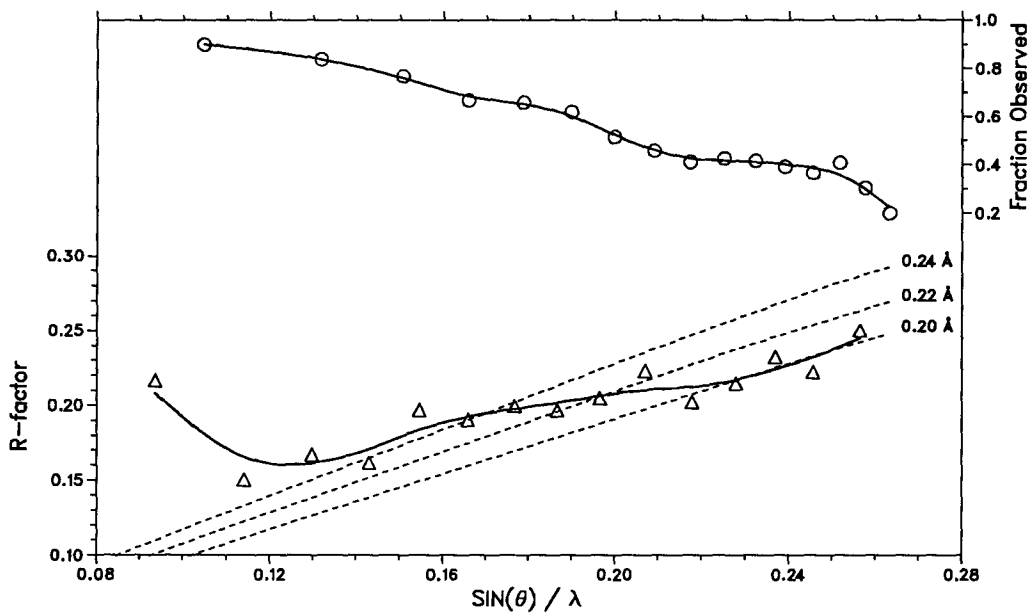

Figure 3.17: Plots of the dependence on resolution of both the R-factor agreement between calculated and observed structure factors (△ – △), and the fraction of data used (○ – ○, axis at top right), for the final refined model of oxidized yeast iso-1-cytochrome *c*. For this analysis reciprocal space was divided into shells according to $sin(\theta)/\lambda$ with each containing at least 260 reflections. For the purpose of assessing the accuracy of the atomic coordinates, curves representing the theoretical dependence of R-factor on resolution assuming various levels of r.m.s. error in the atomic positions of the model (Luzzati, 1952) are also drawn (dashed lines). This analysis suggests an overall r.m.s. coordinate error of ~0.22 Å.

resolution (Luzzati, 1952) one can estimate the r.m.s. error in the coordinates of the structure to be about 0.22 Å (see Figure 3.17). A separate estimate of the atomic coordinate error can be obtained from an unrestrained refinement of the final model (Read *et al.*, 1983). Four cycles of unrestrained refinement caused the R-factor to drop to 14.5%. The resulting model had an r.m.s. deviation of 0.21 Å from the final structure. A third method for estimating coordinate errors is that of Cruickshank (1949,1954). Using this method an overall r.m.s. error for all protein atoms of 0.23 Å is obtained. One has to realize that these coordinate error estimates are average values, and those parts of the structure that are well defined will have smaller coordinate errors, whereas the uncertainty in the positions of mobile surface side-chains is undoubtedly much higher.

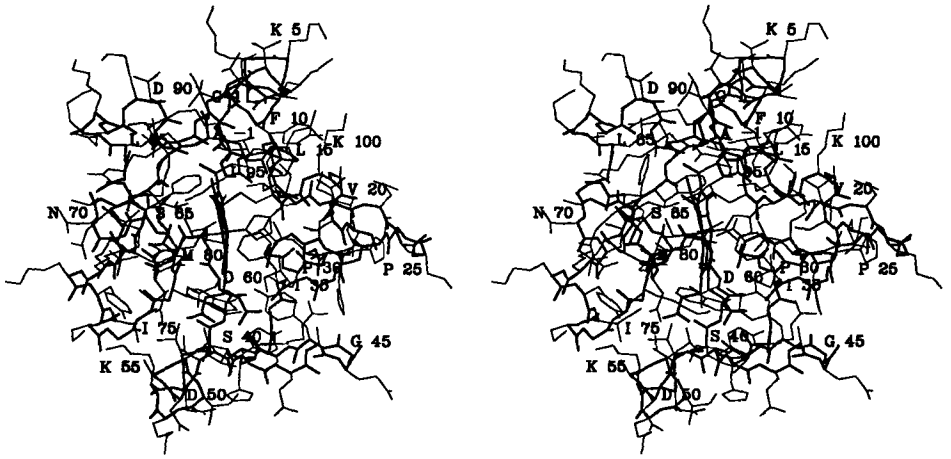

Figure 3.18: A stereo-drawing of the conformations of all main-chain (thick lines) and side-chain (thin lines) atoms in yeast iso-1-cytochrome *c* in the oxidized state. The heme group has also been drawn with thick lines. Also drawn are the two heme ligand bonds to His18 and Met80, and the two covalent thioether linkages to cysteines 14 and 17. For clarity, every 5th α-carbon atom has been labelled with its one-letter amino acid designation and sequence number.

3.2 Results

3.2.1 Polypeptide chain conformation

For the comparison with the reduced state of this protein the oxidized structure was superimposed on the basis of all main-chain and heme atoms on the 1.2 Å reduced iso-1-cytochrome *c* structure using a least-squares algorithm. A stereo-drawing of the structure of oxidized yeast iso-1-cytochrome *c* is presented in Figure 3.18. A similar illustration for the reduced protein is shown in Figure 1.5 as reproduced from the data of Louie & Brayer (1990). A plot of the observed average positional deviations between corresponding main-chain atoms in oxidized and reduced yeast iso-1-cytochrome *c* is shown in Figure 3.19. The conformational changes observed are small and it is notable that none of the hydrogen bond interactions between main-chain atoms are lost upon change in oxidation state. The overall average value for main-chain atom differences in all residues is 0.31 Å. Note that Thr(-5) and Glu(-4) are substantially disordered in both structures of yeast iso-1-cytochrome *c* and the differences observed are therefore likely

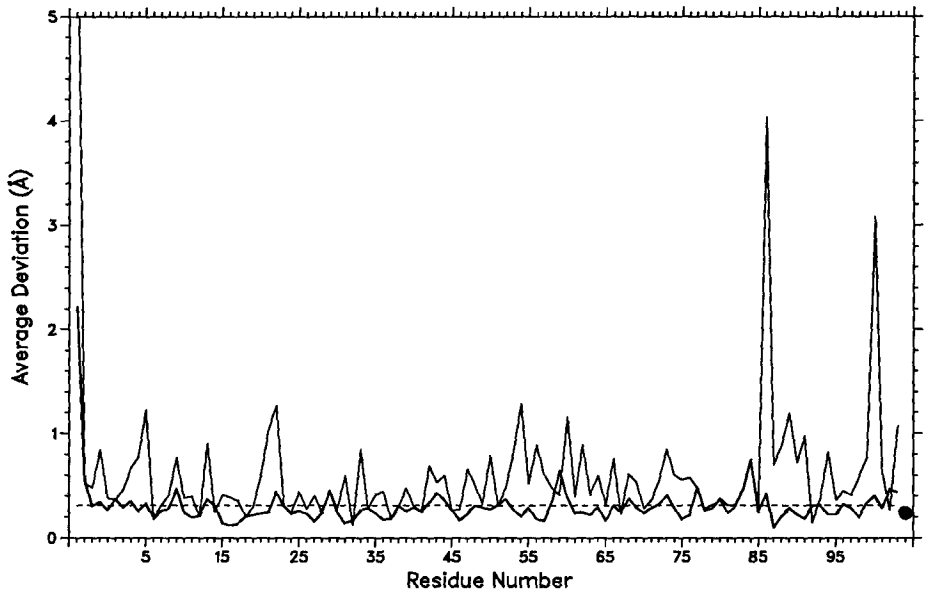

Figure 3.19: A comparison of the average positional deviations between main-chain (thick line) and side-chain (thin line) atoms of the oxidized and reduced forms of yeast iso-1-cytochrome *c*. The horizontal dashed line represents the overall value of 0.31 Å observed for the average deviation between all main-chain atoms. The filled dark circle at position 104 represents the overall average deviation for all heme atoms of 0.24 Å.

due to differential fits to the same poor electron density rather than a reflection of oxidation state. This disorder is apparent from the large thermal B values (~50–55 $Å^2$) assigned to these residues in both structures.

Two conformational changes do appear to result from differing oxidation states. The largest involves Gly84 (average deviation of 0.75 Å) and results in the formation of a new hydrogen bond in the oxidized protein (length 2.6 Å) between the side-chain of Arg13 (NH1 moves ~1.0 Å) and the carboxyl oxygen atom of Gly84 (moves ~1.1 Å; see Figure 3.20). In the reduced protein both of these groups are found associated with surface solvent molecules (Louie & Brayer, 1990). As Figure 3.21 illustrates, a second conformational change in the main-chain of Trp59 (average deviation 0.64 Å) is probably the result of a lengthening in the oxidized protein of the hydrogen bond between the side-chain of this residue and the nearby heme propionate group (from 3.1

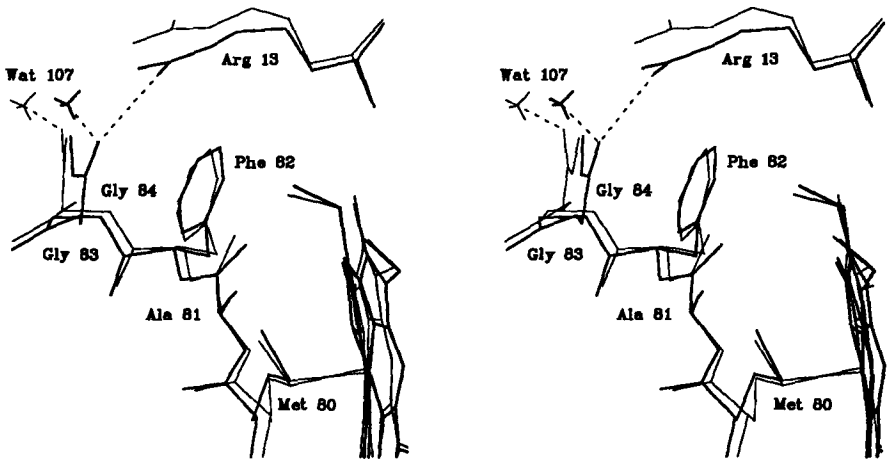

Figure 3.20: A composite stereo-drawing showing the course of the polypeptide chain about Arg13 and Gly84 in the oxidized (thick lines) and reduced (thin lines) states of yeast iso-1-cytochrome *c*. Upon oxidation these two residues move closer together to form a new hydrogen bond (dashed line). Wat 107 also undergoes a comparable shift in position and is hydrogen bonded to the carboxyl group of Gly84 in both oxidation states. Also drawn are the relative positions of the heme moiety and the side-chain of Phe82.

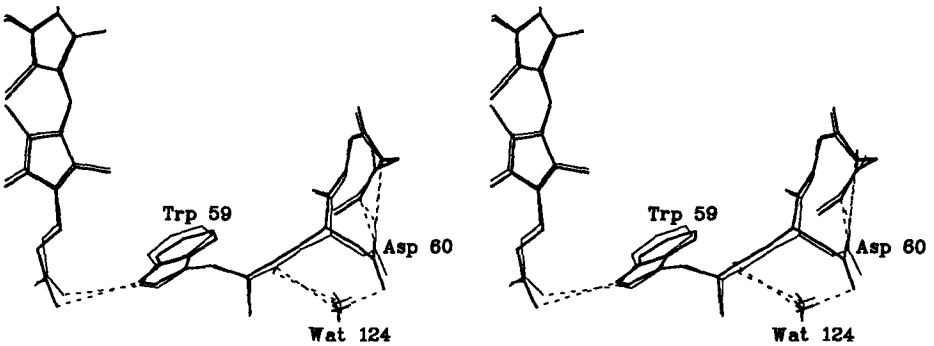

Figure 3.21: A composite stereo-drawing showing a portion of the heme group and the adjacent polypeptide chain about Trp59 in the oxidized (thick lines) and reduced (thin lines) states of yeast iso-1-cytochrome *c*. In the oxidized protein the hydrogen bond (dashed lines) between the side-chain of Trp59 and the nearby heme propionate lengthens leading to a shift in the main-chain atoms of this amino acid. Corresponding shifts are observed for Wat124 and the side-chain of Asp60.

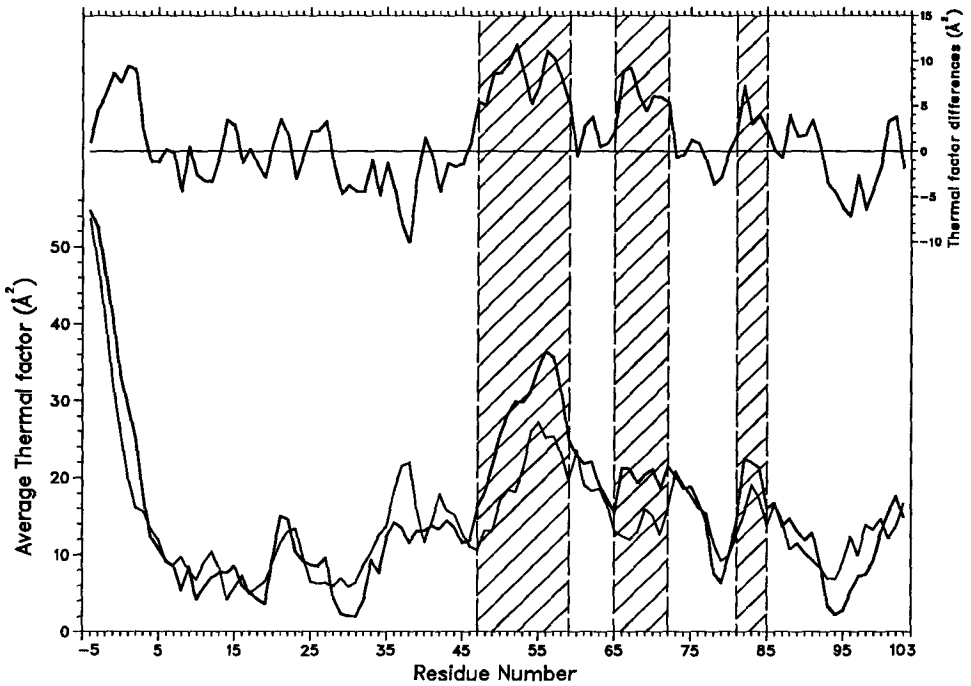

Figure 3.22: Plots of the average thermal factors of main-chain atoms along the polypeptide chain in oxidized (thick line) and reduced (thin line) yeast iso-1-cytochrome *c*. The upper plotted line (axis designation at top right) shows the observed differences between the mean main-chain temperature factors in the two oxidation states. Three regions of polypeptide chain with significantly higher thermal factors in the oxidized protein (residues 47–59 65–72 and 81–85) have been highlighted with cross hatched boxes.

to 3.4 Å). This is reflected in an increase in the average thermal parameters for the side-chain of Trp59 from 15 to 21 Å^2. Associated with this movement is the reorientation of the side-chain of Asp60. Other large side-chain displacements in Figure 3.19 are associated with poorly defined surface residues and would appear to result from positional uncertainty rather than as a consequence of change in oxidation state.

Between oxidation states, comparable values are observed for the overall average thermal factor for all atoms in the polypeptide chain of yeast iso-1-cytochrome *c* (16.4 and 16.5 Å^2 for reduced and oxidized, respectively). However, as evident from Figure 3.22, if comparisons are made of thermal factors along the course of the polypeptide chain, substantive differences are present. These differences are further illustrated in the matrix representation of Figure 3.23. In

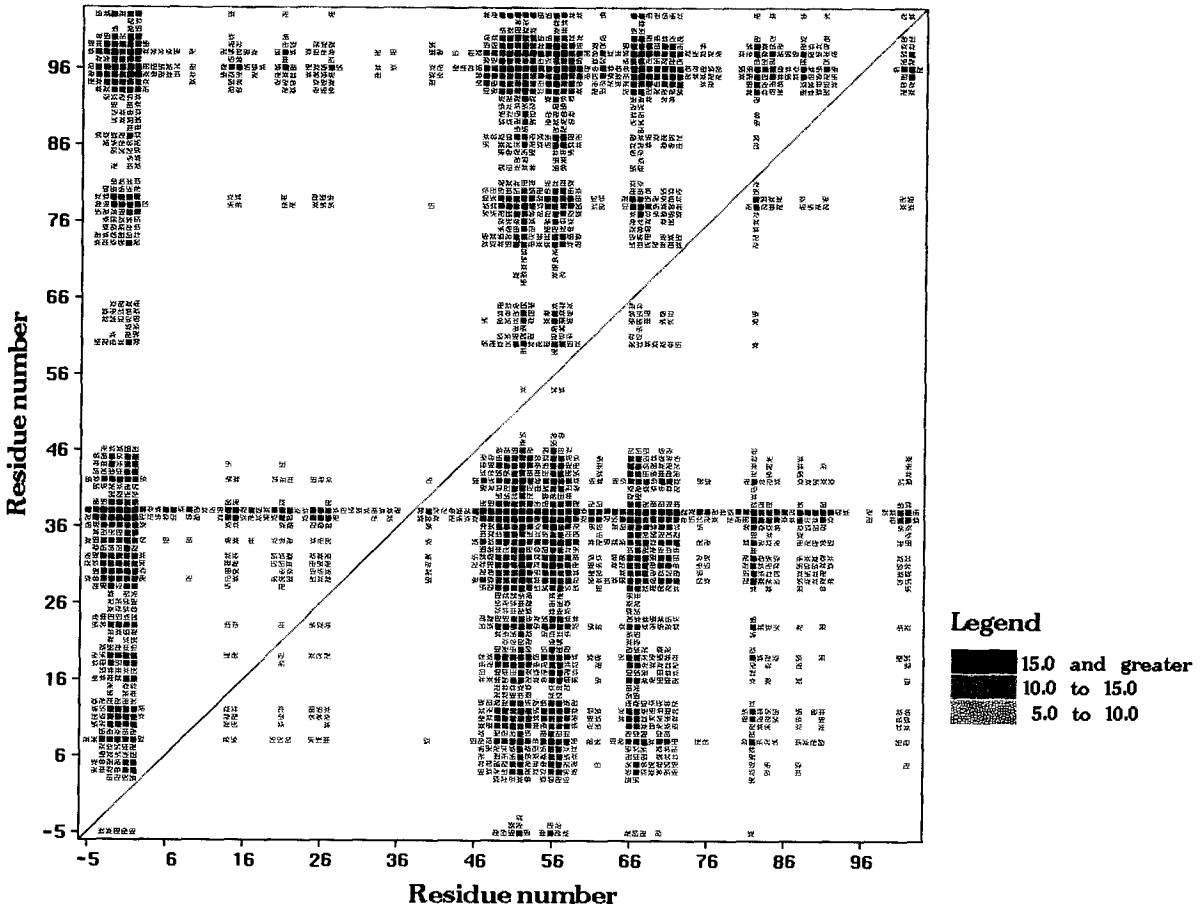

Figure 3.23: A matrix representation of the differences in average main-chain thermal factors between the oxidized and reduced states of yeast iso-1-cytochrome c. Each matrix point $P_{x,y}$ represents an amino acid pairing (x, y) and was calculated using the equation: $P_{x,y} = (B_x - B_y)_{oxidized} - (B_x - B_y)_{reduced}$ where B is the average main-chain thermal factor of a given amino acid. Positive matrix values are displayed as squares of different levels of blackness according to the scale on the right. As this matrix has inverse symmetry across the diagonal line drawn, negative values are redundant and are omitted for clarity. The advantage of this approach is that displayed values are not affected by differences in overall thermal factor between the two structures. Within the matrix, amino acids which have significantly higher average main-chain thermal factors in the oxidized structure produce vertical streaks. These include residues 47–59, 65–72 and 81–85, with maximal differences observed for Asn52, Tyr67 and Phe82. As discussed in the text, the N-terminal region is not considered in this analysis due to the presence of positional disorder in both oxidation states. Amino acids producing horizontal streaks indicate the presence of significantly larger thermal factors for their main-chain atoms in the reduced state of yeast iso-1-cytochrome c. Residues of this type are fewer in number and are centered about Arg38, Thr96 and Leu98.

terms of lower thermal factors in the oxidized protein, the most significant main-chain difference involves Arg38, a residue having a water mediated interaction with the adjacent pyrrole ring A propionate group. Also affected are the two immediately adjacent residues Gly37 and His39. The side-chain atoms of Arg38 exhibit an even larger $\sim$14 Å^2 lowering of average thermal B values in the oxidized state.

In total, four regions of polypeptide chain have significantly higher thermal factors in the oxidized form of yeast iso-1-cytochrome *c*. The significance of the first, which involves the N-terminus, is questionable since this region is disordered and poorly resolved in electron density maps. The three other segments affected (residues 47–59, 65–72 and 81–85) are particularly interesting since they appear to be related to oxidation state dependent conformational changes in internal water structure and the heme group. As Figures 3.22 and 3.23 illustrate, those portions of polypeptide chain are sharply delineated, with maximal increases in thermal factors in the oxidized state being observed for Asn52, Tyr67 and Phe82. Figure 3.24 shows that these three segments are located to the Met80 ligand side of the heme group.

3.2.2 Heme structure

High resolution studies of reduced yeast iso-1-cytochrome *c* have shown that the heme group is substantially distorted from planarity into a saddle shape (Louie & Brayer, 1990). In the oxidized protein the type of distortion observed is similar, but considerably more pronounced, suggesting the degree of heme planarity is dependent on oxidation state (Table 3.8). None of the heme iron coordinate bonds is significantly different between oxidation states, although the largest deviation observed is a lengthening of the Met80 ligand bond. Overall, the average displacement of side-chain atoms in Met80 is 0.36 Å between redox states, with the largest shifts observed for the CB (0.55 Å) and CE (0.47 Å) carbon atoms. These displacements result in a small 7° change in the torsion angle around the Met80 - heme iron ligand bond. It is notable that a substantial increase in thermal factors for the side-chain of Met80 is also observed. In reduced yeast iso-1-cytochrome *c* the average side-chain thermal factor is 5 Å^2

Table 3.8: Heme conformation and ligand geometry in the two oxidation states of yeast iso-1-cytochrome *c*†

I. Angular deviations between both the plane normals of individual pyrrole rings and the heme coordinate bonds and both the pyrrole nitrogen plane and the porphyrin ring plane (°) ‡

	Pyrrole ring	Pyrrole N plane	Porphyrin ring plane
a)	A	12.6 (9.4)	8.3 (6.7)
	B	14.1 (11.1)	12.5 (11.9)
	C	9.6 (8.8)	13.8 (9.8)
	D	12.8 (8.1)	10.0 (6.0)
b)	Heme coordinate bonds		
	Fe - His18 NE2	7.2 (2.1)	6.5 (3.3)
	Fe - Met80 SD	3.3 (4.9)	5.4 (7.5)

II. Heme iron coordinate bond distances (Å)

His18 NE2	2.01 (1.99)
Met80 SD	2.43 (2.35)
Heme NA	1.97 (1.97)
Heme NB	1.98 (2.00)
Heme NC	2.01 (1.99)
Heme ND	2.05 (2.00)

III. Heme propionate hydrogen bond interactions (distances in Å) §

Heme atom	Hydrogen bond partners and distances
O1A	Tyr48 OH 2.83 (2.83), Wat121 2.85 (2.81), Wat168 2.87 (2.85)
O2A	Gly41 N 2.60 (3.21), Asn52 ND2 3.54 (3.34), Trp59 NE1 3.43 (3.09), Wat121 3.34 (4.01)
O1B	Thr49 OG1 2.79 (2.64), Thr78 OG1 3.07 (2.90), Lys79 N 2.67 (3.17)
O2B	Thr49 N 2.75 (2.94)

† Values for reduced yeast iso-1-cytochrome *c* are shown in parentheses.

‡ Each pyrrole ring is defined by nine atoms, which include the five ring atoms plus the first carbon atom bonded to each ring carbon. The porphyrin ring plane is defined by the five atoms in each of the four pyrrole rings, the four bridging methine carbons, the first carbon atom of each of the eight side-chains and the heme iron (33 atoms in total). The pyrrole nitrogen plane is defined by only the 4 pyrrole nitrogens (see Figure 1.2 for the heme atom labeling convention used).

§ Distances are provided for the heme O2A to Wat121 (reduced state) and Asn52 ND2 (oxidized state) interactions for the sake of comparison, even though these are too long to represent hydrogen bonds.

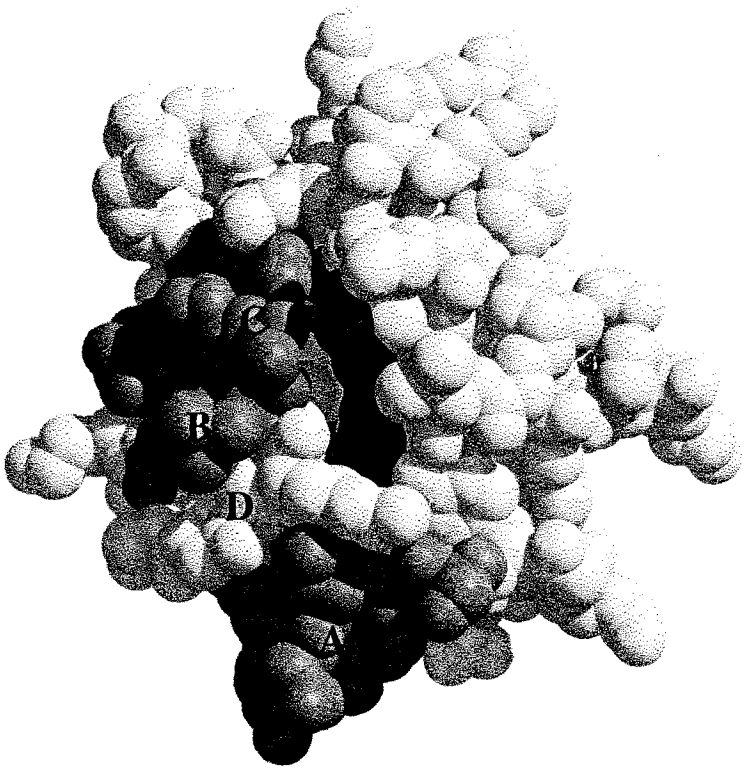

Figure 3.24: A space-filling representation of yeast iso-1-cytochrome *c* showing the location of the three segments of polypeptide chain (drawn in gray spheres) which have main-chain thermal factors which are significantly larger in the oxidized protein. A general indication of the location of the individual affected segments is provided by the letters A–C, corresponding to residues 47–59, 65–72 and 81–85, respectively. Heme atoms are shown with black spheres, while white spheres indicate regions of polypeptide chain with comparable or lower thermal factors in oxidized yeast iso-1-cytochrome *c*. Although contiguously linked through interactions in the protein interior, the three segments having higher thermal factors are subdivided at the protein surface by the polypeptide strand composed of residues 73–80 (labelled D).

whereas it increases to 12 $Å^2$ in the oxidized protein. Coupled with the observed shifts in bond orientation, these thermal factors suggest a weakening of the Met80 SD - heme iron ligand bond in the oxidized state. In contrast, no significant change in ligand bond distance or thermal parameters is observed for His18 between oxidation states. However, the precise orientation of the His18 side-chain may be a function of oxidation state in that a 46.7° angle is found between the imidazole ring plane and a vector drawn through the NA and NC pyrrole

Table 3.9: Heme solvent accessibility in the two oxidation states of yeast iso-1-cytochrome *c*

		Oxidized	Reduced
1.	Solvent accessible heme atoms and surface area exposed (Å^2)		
	CHD	0.0	2.6
	CMC	9.5	9.4
	CAC	5.1	3.3
	CBC	19.2	17.4
	CMD	11.7	9.8
2.	Total heme exposure (Å^2)	45.5	42.5
3.	Total heme surface (Å^2)	503.3	495.7
4.	Heme surface area exposed (%)	9.0	8.6

Computations were done using the method of Connolly (1983) and the results represent the accessible molecular surfaces of the atoms listed (see Figure 1.2 for heme atom nomenclature) The probe sphere used had a radius of 1.4 Å.

nitrogen atoms in reduced yeast iso-1-cytochrome *c*, whereas this value is 55.8° in the oxidized protein.

An analysis of the heme solvent accessibility in the two oxidation states is documented in Table 3.9. These results show that total heme accessibility is comparable, although slightly higher in the oxidized protein.

Another aspect of heme structure differing between oxidation states is the positioning of the pyrrole ring A propionate group (see Figure 3.25). This group's alternative placement in oxidized yeast iso-1-cytochrome *c* can be ascribed to changes in three torsion angles. The first involving the C2A–CAA bond which rotates ~20°, the second involving the CAA–CBA bond which rotates ~30°, the third involving the CBA–CGA bond which is rotated ~45°. The net result is that the O2A oxygen atom of the propionate carboxyl group moves ~0.6 Å in a direction towards Gly41, with the O1A oxygen atom showing a more modest 0.2 Å displacement.

As evident from Table 3.8 and Figure 3.25, not all groups that hydrogen bond to this propionate can accommodate the observed positional shifts in the oxidized state. In particular,

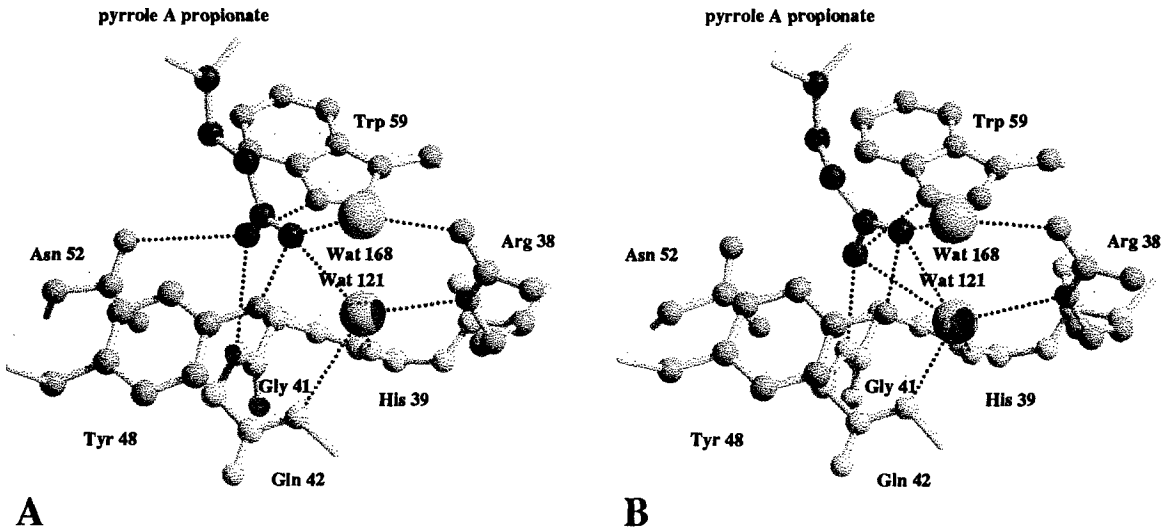

Figure 3.25: Drawings of the region about the pyrrole A propionate group in (a) reduced and (b) oxidized yeast iso-1-cytochrome *c*, illustrating the positional shifts and altered hydrogen bonding patterns observed. The pyrrole ring A propionate group is highlighted with dark shaded balls. Hydrogen bonds are indicated by thin dashed lines. The two internally bound water molecules Wat121 and Wat168, which mediate the interaction of Arg38 with this heme propionate, are shown with larger spheres.

much weaker interactions are made to Asn52 and Trp59, and both of these side-chains have increased thermal factors. Also affected is the position of the internally bound water molecule Wat121, which is conserved in all eukaryotic cytochrome *c* structures determined to date (Bushnell *et al.*, 1990). In reduced yeast iso-1-cytochrome *c*, Wat121 forms a hydrogen bond to the O1A oxygen of the pyrrole ring A propionate group. As can be seen in Figure 3.25, a ~0.5 Å shift of Wat121 in the oxidized protein allows the formation of bifurcated hydrogen bonds to both the propionate O1A and O2A carboxyl oxygen atoms. Notably, the side-chain of Arg38, whose interaction with the propionate group is mediated by Wat121 and Wat168 shows no significant positional displacement in the oxidized protein.

3.2.3 Internal water structure

The internally bound water molecule, Wat166 undergoes a large shift in position in response to oxidation state. In oxidized yeast iso-1-cytochrome *c*, Wat166 moves 1.7 Å almost directly

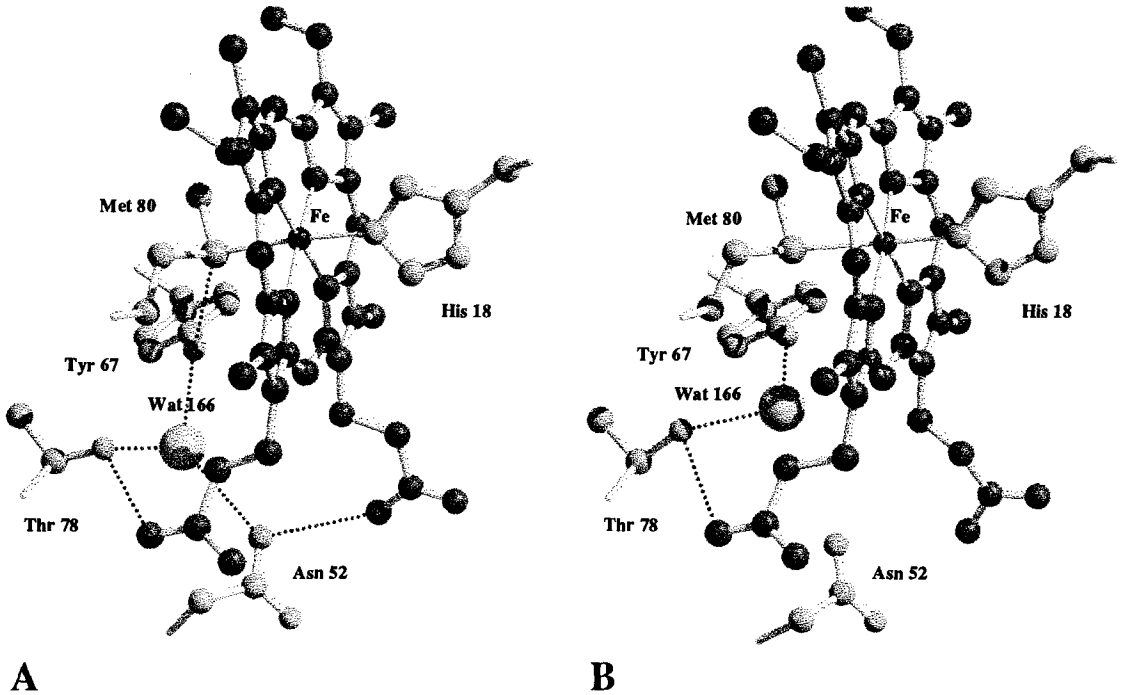

Figure 3.26: Heme and polypeptide chain structure about the internally bound water molecule Wat 166 in (a) reduced and (b) oxidized yeast iso-1-cytochrome *c*. Heme ligand interactions are indicated by thin white bonds, whereas hydrogen bonds are shown by dotted lines.

towards the heme iron atom, closing the distance between these groups from 6.6 Å to 5.0 Å (Figure 3.26). Most affected by this movement is Asn52, to which Wat166 no longer forms a hydrogen bond. As mentioned before, the hydrogen bond between Asn52 and the pyrrole ring A propionate group is also lost (Table 3.8 and Figure 3.25). Despite the large shift in Wat166, hydrogen bonds to both Tyr67 and Thr78 are retained and the distances between the hydrogen bonding atoms involved are similar in both oxidation states.

As illustrated in Figure 3.27a, in reduced iso-1-cytochrome *c* Wat166 is very tightly constrained into a small spherical cavity having a volume of ~10 Å^3. To accommodate Wat166 movement, the densely packed protein matrix must undergo small concerted conformational adjustments, the largest of these involving the side-chain of Asn52 (average side-chain displacement is 0.47 Å). As shown in Figure 3.27b, the resultant cavity surrounding Wat166 in the

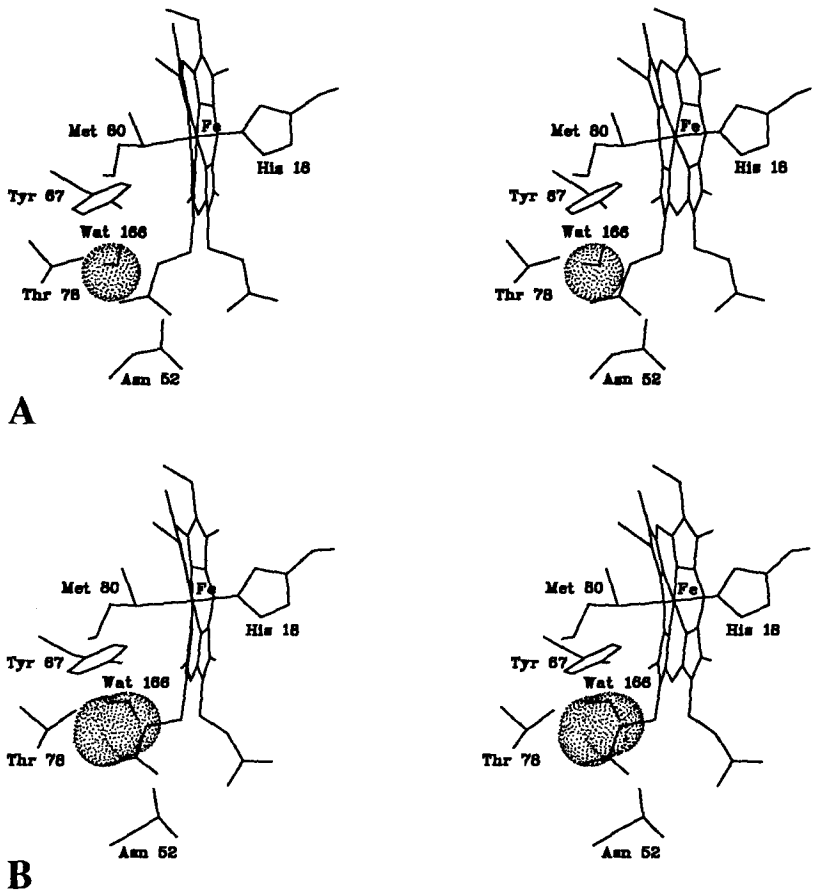

Figure 3.27: Stereo-drawings of the internal cavity occupied by Wat166 in (a) the reduced and (b) the oxidized form of yeast iso-1-cytochrome *c*. Also drawn are the nearby residues to which Wat166 is hydrogen bonded and the adjacent heme group and heme ligands. In the reduced protein the volume of the cavity is ~10 Å^3, which expands to ~25 Å^3 in oxidized yeast iso-1-cytochrome *c*.

oxidized protein is significantly expanded and is now ~25 Å^3 in volume and highly asymmetrical, with Wat166 occupying only that portion of the available volume closest to the heme iron atom. This leaves considerable free volume available to amino acids adjacent to Wat166 and they are therefore conformationally less constrained.

The looser packing evident about Wat166 in oxidized yeast iso-1-cytochrome *c* could provide an explanation for the observed increased thermal factors of the nearby polypeptide segments

47–59 and 65–72. For the 47–59 segment this trend is undoubtedly accentuated by the loss of hydrogen bonds from Wat166 and the pyrrole A propionate to the Asn52 side-chain, as well as the lengthening of the hydrogen bond from the same propionate to the side-chain of Trp59. As Figures 3.22 and 3.23 demonstrate, the largest average main-chain thermal factor increase in the 47–59 segment is localized at Asn52, with the side-chain of this residue also experiencing a similar increase (average increase +16 $Å^2$). Another potentially important aspect of the loss of the Wat166 to Asn52 ND2 hydrogen bond, is the removal of a structural link between the 47–59 and 65–72 polypeptide segments. Despite retention of the Wat166 to Tyr67 OH hydrogen bond, the main-chain of residues 65–72 show increased thermal factors in the oxidized state. This increase is maximal at Tyr67 and is mirrored in a substantive overall mean increase in its overall side-chain thermal factor from 13 $Å^2$ to 27 $Å^2$ on going from the reduced to oxidized states. The primary factor involved in the mobility changes in the 65–72 region again appears to be the increased conformational volume accessible to these residues.

Although the hydrogen atoms of Wat166 cannot be resolved experimentally, the nature of the surrounding hydrogen bonding network makes it possible to speculate on the orientation of the dipole moment of this internal water molecule in both the reduced protein where the heme iron atom is formally uncharged, and in the oxidized state where the iron atom has a positive charge. As illustrated in Figure 3.26a, the complex of hydrogen bond interactions involving Wat166, Asn52, Tyr67, Thr78 and Met80 in reduced yeast iso-1-cytochrome *c*, restrictively defines the orientation of Wat166 and thus its dipole moment. Based on Tyr67 OH and Thr78 OG1 being able to function both as hydrogen bond donor and acceptor groups, three potential orientations for Wat166 can be delineated. Two of these orientations cannot be realized without breaking hydrogen bonds between protein groups and therefore seem unlikely. The third alternative orientation maximizes hydrogen bond interactions (see Figure 3.28a). In this Wat166 orientation the positive end of the dipole of this group is pointing in the direction of the heme iron atom with a $\sim 50°$ declination. A nearly identical dipole moment orientation for Wat166 can be modelled in reduced yeast iso-2 and tuna cytochromes *c*, the only other two

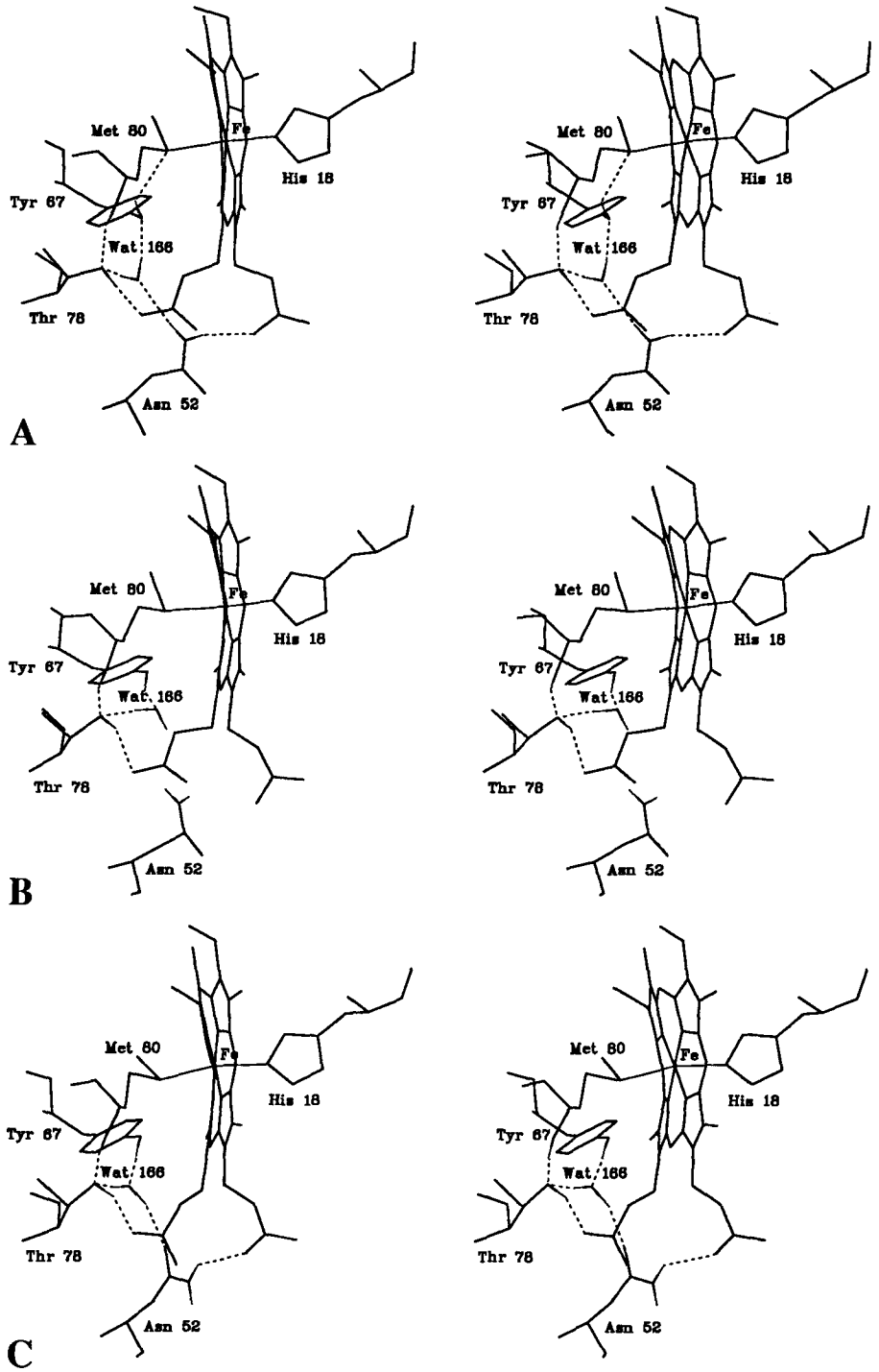

Figure 3.28: Stereo-drawings showing the immediate vicinity about Wat166 in (a) reduced and (b) oxidized yeast iso-1-cytochrome *c*, as well as in (c) oxidized horse cytochrome *c*. Also shown is the hypothetical placement of hydrogen atoms for the purpose of discerning the dipole moment of Wat166 in the presence and absence of a positive charge at the heme iron atom.

reduced cytochromes *c* whose structures are known (Murphy *et al.*, 1992; Takano & Dickerson, 1981a).

The placement of the dipole moment of Wat166 in the oxidized state is less clearly definable since the interaction formed to the side-chain of Asn52 is not present, allowing for several orientational possibilities. However, given that the heme iron atom in the oxidized state is positively charged, it is reasonable to expect that the resultant electrostatic field will be an important factor in the orientation of the dipole of Wat166. Based on a simple model, where propionate oxygen atoms each have a charge of $-\frac{1}{2}$, the heme iron atom a charge of $+1$ and a water molecule has a dipole moment of 1.85 Debye, the following equation can be used to calculate the potential energies for the various possible orientations for Wat166:

$$U = -1|\vec{p}||\vec{E}|\cos\theta \tag{3.3}$$

where U is the potential energy of the dipole, $\vec{p}$ is the dipole moment, $\vec{E}$ is the intensity of the electric field caused by the surrounding charges and θ is the angle between $\vec{p}$ and $\vec{E}$. The electric field strength can be evaluated using Coulomb's law.

From this analysis two energetically favorable orientations can be derived. In one, which is illustrated in Figure 3.28b, the negative end of the dipole of Wat166 is pointing almost directly at the heme iron atom, but requires the Tyr67 OH to Met80 SD hydrogen bond to be broken. In the second possible orientation, the hydrogen bond between the side-chains of Tyr67 and Met80 is retained, but as a consequence the dipole moment is less favorably oriented. The difference in potential energy of these two water dipole orientations is ~55 KJ/mole, with the first being more favorable. Clearly which dipole orientation will dominate depends on the strength of the Tyr67 OH to Met80 SD hydrogen bond which is broken in one orientation and not in the other. Suggested values for the energy of a hydrogen bond range from -15 to -25 KJ/mole (Schultz & Shirmer, 1979; Creighton, 1984). Since the energy gain in optimally orienting the dipole in the electric field is more than twice the energy lost by breaking a hydrogen bond, it is reasonable to expect the dipole orientation shown in Figure 3.28b to predominate in the oxidized yeast iso-1-cytochrome *c* structure.

The loss of the Tyr67 OH to Met80 SD hydrogen bond in the oxidized state could be expected to be an additional stabilizing feature in two ways. Firstly, it would tend to make the Met80 heme ligand less electron withdrawing, a situation favoring stabilization of the positively charged heme iron atom. Secondly, the new hydrogen bond interaction from Wat166 to the side-chain of Tyr67 assists in properly orienting the dipole moment of Wat166 adjacent to the heme group in order to maximally stabilize the positive charge resident there. That the hydrogen bond to the side-chain of Met80 is lost, is supported by the observed increase in thermal factors for both the side-chains of Tyr67 and Met80 in the oxidized state.

Further confirmation of the dipole orientation of Wat166 in oxidized yeast iso-1-cytochrome *c* can be derived from the available three-dimensional structures of oxidized horse (Bushnell *et al.*, 1990), tuna (Takano & Dickerson, 1981b) and rice (Ochi *et al.*, 1983) cytochromes *c*. The presence and position of Wat166 is conserved in all these proteins (Bushnell *et al.*, 1990). In addition, the determination of the direction of the dipole moment of Wat166 has greater certainty, since for these proteins the hydrogen bond to Asn52 is retained in the oxidized state. In each of these structures, the most likely dipole orientation for Wat166 is similar to that proposed herein for oxidized yeast iso-1-cytochrome *c*. This can be seen in Figure 3.28c, which shows the dipole orientation of Wat166 in horse cytochrome *c*. Collectively, these results clearly suggest a major role for Wat166 in stabilizing the oxidized as well as the reduced forms of cytochrome *c* through the differential orientation of dipole moment, shift in distance to the heme iron atom and alterations in the surrounding hydrogen bond network.

3.3 Discussion

3.3.1 Focal points for oxidation state dependent structural alterations

A compilation of significant structural differences observed on going from the reduced to oxidized states of yeast iso-1-cytochrome *c* is presented in Table 3.10. It is clear from this data that oxidation state dependent conformational differences are for the most part expressed in terms of changes in thermal parameters of specific regions, adjustments to heme structure, movement of

Table 3.10: Structural changes observed on going from the reduced to the oxidized state in yeast iso-1-cytochrome *c*

1. Positional displacements of polypeptide chain (see Figure 3.19)
   (a) Movement of Arg13 and Gly84 to form a hydrogen bond (see Figure 3.20)
   (b) Lengthening of the interaction between Trp59 and the heme pyrrole ring A propionate group (see Figure 3.21)
2. Thermal factor parameters of main-chain atoms (see Figures 3.22 and 3.23)
   (a) Lower values observed for residues 37–39, focussed at Arg38; side-chain of Arg38 also has reduced values
   (b) Higher values found for three polypeptide chain segments
      i. residues 47–59, focussed at Asn52
      ii. residues 65–72, focussed at Tyr67
      iii. residues 81–85, focussed at Phe82

   All the side-chains of Asn52, Tyr67 and Phe82 show higher mobility
3. Heme structure and ligands
   (a) Increased distortion of heme planarity (see Table 3.8)
   (b) Readjustment of the pyrrole A propionate group with a realignment of hydrogen bonding interactions (see Figure 3.25)
   (c) Higher thermal parameters for the Met80 side-chain.
4. Internal water structure
   (a) Large displacement of Wat166 towards the heme iron atom coupled with a change in hydrogen bonding interactions (see Figure 3.26)
   (b) Wat166 movement is facilitated by shifts in the protein matrix to enlarge the available internal cavity space (see Figure 3.27)
   (c) Reorientation of the dipole of Wat166 to favor stabilization of the charged heme iron atom (see Figure 3.28)
5. Hydrogen bond interactions (see Figures 3.20, 3.21, 3.25 and 3.26)

(a)	Stronger :	Gly41 N -	Heme O2A
(b)	Weaker:	Trp59 NE1 -	Heme O2A
(c)	Lost:	Asn52 ND2 -	Heme O2A
		Asn52 ND2 -	Wat 166
		Tyr67 OH -	Met80 SD
(d)	New:	Arg13 NH1 -	Gly84 O
		Wat121 -	Heme O2A

internal water molecules and the realignment of hydrogen bonding patterns, rather than in terms of explicit shifts in the polypeptide chain which are found to be minimal. In total there appear to be three points about which the observed conformational differences are focussed. These involve the pyrrole ring A propionate group, the internal water molecule Wat166 and the Met80 heme ligand. All three of these features are structurally linked together through a hydrogen bonding network. They are furthermore localized for the most part on the Met80 ligand side of the heme group, adjacent to the solvent exposed edge of this latter group (see Figure 3.24). As a whole, the amino acids involved also form a large portion of the protein surface implicated in being central to the complexation interactions formed with electron transfer partners (Salemme, 1976; Margoliash & Bosshard, 1983; Pelletier & Kraut, 1992).

It is instructive to examine each focal point for structural changes in oxidized yeast iso-1-cytochrome *c* in light of the conformational shifts observed and their probable role in stabilizing this state of the protein.

Heme pyrrole ring A propionate group

Upon oxidation there is an apparent change in the nature of the relationship between the heme iron atom and the pyrrole ring A propionate, leading to a conformational adjustment in this latter group (see Figure 3.25). As a consequence a realignment of the hydrogen bonds to this group takes place (Table 3.8). Arg38 whose interaction with this propionate group is mediated by two internal water molecules (Wat121 and Wat168) is also affected by a change in oxidation state. In the oxidized state Arg38 becomes positionally more stable as can be deduced from the observed decrease in thermal factors for this residue.

Precisely what triggers conformational changes about the pyrrole A propionate group and how these serve to stabilize the oxidized state of the protein is uncertain. Nonetheless there are two conformational states in this region corresponding to the different oxidation states of the heme iron atom. It is likely that the intricate network of hydrogen bonds present in this region facilitates these changes, perhaps by oxidation state dependent differential delocalization

of the negative charge of the ionized heme propionate group (Barlow & Thornton, 1983; Singh *et al.*, 1987). Alternatively, direct through space interactions of the positively charged heme iron atom and the negatively charged pyrrole A propionate may be the primary factor involved and the observed conformational changes serve to enhance this interaction (Moore, 1983). Also possible are electrostatic interactions between the positively charged side-chain of Arg38 and the heme iron atom, as well as the nearby negatively charged propionate group. Replacement of this arginine by other amino acids is known to substantially lower the reduction potential of cytochrome *c* (Proudfoot & Wallace, 1987; Cutler *et al.*, 1987) indicating the ability of this amino acid to directly affect the functional properties of the heme. Finally it is conceivable that the observed increased distortion of the heme from planarity in the oxidized state leads to a need for conformational adjustments that are focussed in the area of the pyrrole ring A propionate group. In this view, the adjustments to the surrounding protein matrix may be simply designed to accommodate the alternative positions of this propionate group. It is notable that structurally similar changes have been observed in horse (Feng *et al.*, 1990) and yeast iso-1-cytochrome *c* (Gao *et al.*, 1991) using two-dimensional NMR, and in yeast iso-1-cytochrome *c* mutants having lowered reduction potentials and apparently hybrid oxidation state conformations using x-ray crystallographic techniques (Louie *et al.*, 1988b; Louie & Brayer, 1989).

The internal water molecule Wat166

A second focal point for oxidation state dependent structural changes is Wat166. This internally buried water molecule is directly linked to shifts about the pyrrole ring A propionate group through interactions with the side-chain of Asn52 (see Figure 3.26a). In the oxidized state, the Wat166 to Asn52 hydrogen bond is lost and this water molecule moves closer to the heme iron atom. These changes are accommodated by the enlargement of the cavity containing Wat166 due to small concerted movements of the surrounding protein matrix. The result of these events in conjunction with the changes around the pyrrole A propionate group, is an increased

flexibility for the two protein segments consisting of residues 47–59 and 65–72. It should be noted that the Asn52 residue would appear to play a central role in mediating these flexibility changes through breakage of its hydrogen bond to both the propionate group and the internal water molecule in the oxidized state.

The apparent, more open and flexible nature of the structure of oxidized cytochrome *c* has been documented in many other studies. For example, reduced cytochrome *c* has a greater resistance to ligand displacement by exogenous substituents, to proteolytic digestion and to unfolding by both chemical denaturants and by heat (see reviews by Margoliash & Schecter, 1966; Salemme, 1977). It also has a lower viscosity in solution (Fisher *et al.*, 1973), a lower compressibility (Eden *et al.*, 1982) and a smaller radius of gyration as indicated by small-angle X-ray scattering (Trewhella *et al.*, 1988). It is notable that elimination of the Wat166 cavity by replacement of Asn52 with an isoleucine residue leads to significantly increased stabilization of the oxidized form of yeast iso-1-cytochrome *c* (Das *et al.*, 1989; Hickey *et al.*, 1991).

The trigger for Wat166 movement in the oxidized state of cytochrome *c* and the subsequent cascade of related conformational changes, appears to be an electric field induced response to the oxidation state of the heme. The shift of Wat166 could serve several functions. It results in direct stabilization of the positively charged heme iron by the dipole orientation of Wat166 (see Figure 3.28). Through the side-chain of Asn52, Wat166 is also physically linked to conformational changes about the pyrrole A propionate group and may help mediate these adjustments. Wat166 movement also appears to be a primary factor in the increased mobility of nearby polypeptide chain segments which form a part of the putative surface binding site with electron transfer partners and in this way could potentially influence binding in an oxidation state dependent manner (see Figure 3.24). Finally, as discussed below, its positioning may result in modification of the strength of the Met80 SD heme iron atom ligand interaction.

The oxidation state dependent positioning of Wat166 has also been observed in other mitochondrial cytochrome *c* structures in which this internal water molecule is invariably present (Bushnell *et al.*, 1990; see also Section 1.3). As can be seen from Figure 3.29, there are clearly

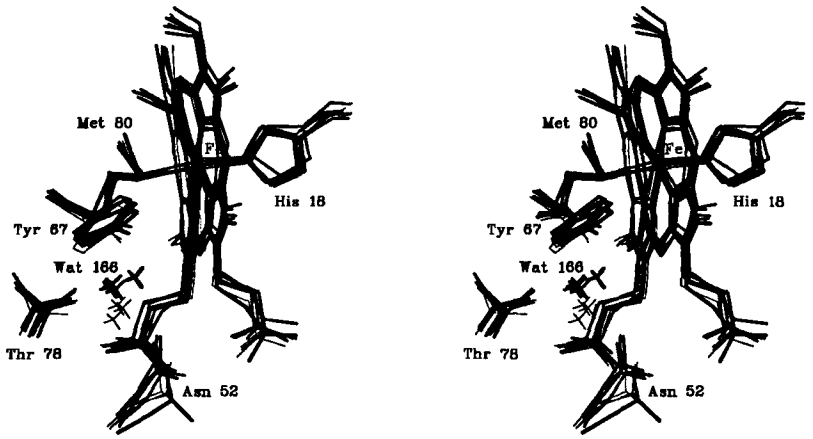

Figure 3.29: A stereo-drawing of the internal water site as found in all of the high resolution structures of eukaryotic cytochromes *c* thus far solved. In thin lines are displayed: tuna, yeast iso-1 and yeast iso-2-cytochrome *c* all in the reduced form The oxidized structures of tuna (both the inner and outer molecule), rice and horse are displayed in medium thick lines. The yeast iso-1-cytochrome *c* in the oxidized state is shown in heavy lines.

two areas over which the equivalent internal water molecules in yeast iso-1 and iso-2, horse, rice, and the tuna cytochromes *c* are distributed. The first area, furthest away from the heme, contains the water molecules of reduced yeast iso-1 and iso-2, and tuna cytochromes *c*. The area closer to the heme holds those water molecules from the oxidized structures of yeast iso-1, tuna, rice and horse cytochromes *c*.

Despite the obvious similarities among the different oxidized structures, there are differences between the oxidized yeast iso-1-cytochrome *c* and the oxidized tuna, horse and rice structures (see Figure 3.30). First, the shift made by the internal water in yeast iso-1-cytochrome *c* is roughly twice that seen in tuna. Furthermore, its distance to the heme iron is smaller than that observed in any of the other oxidized cytochrome *c* structures. The second difference is in the side-chain conformation of Asn52. As mentioned before, in oxidized yeast iso-1-cytochrome *c*, hydrogen bonds from Asn52 ND2 to both the heme and the internal Wat166 are lost. In the case of tuna, rice and horse cytochromes *c*, the Asn52 side-chain amide has flipped over 180° in the oxidized state. In these proteins, the ND2 group is still hydrogen bonded to the heme pyrrole A propionate and the nearby internal water is now hydrogen bonded to the OD1 atom

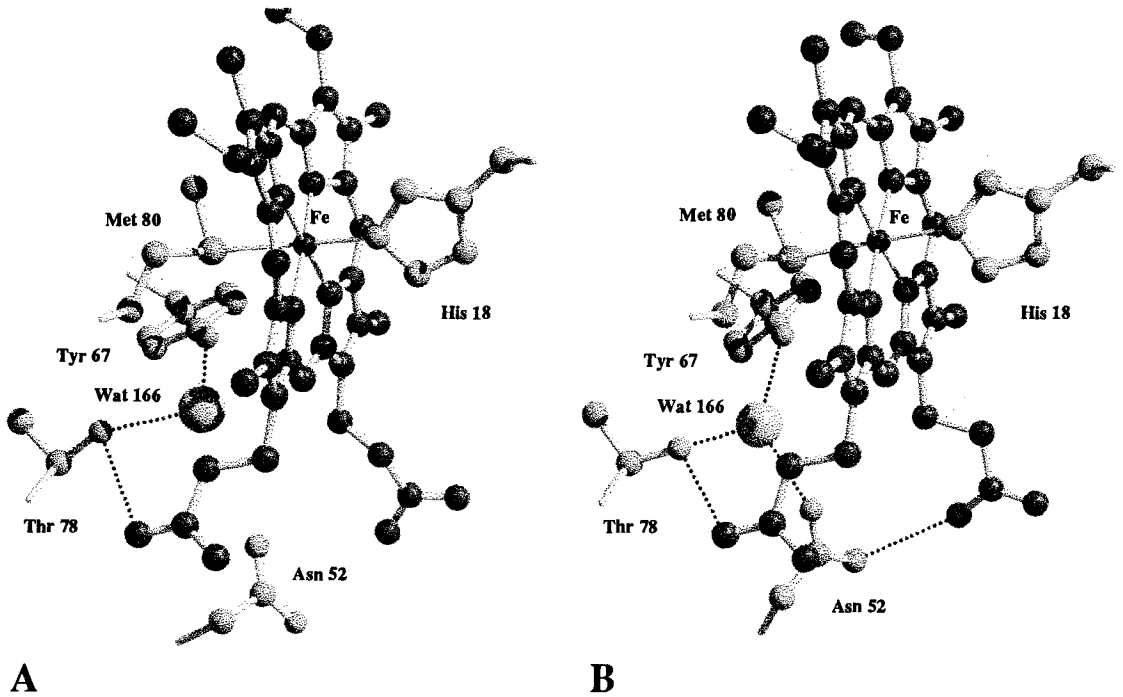

Figure 3.30: Heme and polypeptide chain structure about the internally bound water molecule Wat 166 in (a) oxidized yeast iso-1-cytochrome *c* and (b) oxidized horse cytochrome *c*. Heme-ligand interactions are indicated by thin white bonds, whereas hydrogen bonds are shown by thin black dotted lines. An identical hydrogen bonding network as present in horse cytochrome *c* is also seen in the oxidized structures of rice and tuna.

of Asn52. This rearrangement allows the internal water molecule to move closer to the heme iron without the loss of a hydrogen bond to the side-chain of residue 52. Notably, the retention of a hydrogen bond to Asn52 does not affect the orientation of the dipole moment of the internal water molecule (Figure 3.28c), nor does it interfere with the possible role this water molecule has in modulating the strength of the iron sulfur ligand bond, as discussed below.

The Met80 ligand region

The region about the Met80 heme ligand and Phe82 forms a third focus for oxidation state dependent conformational changes. Coincident with observations showing that oxidized cytochrome *c* has a more open structure and the present study which suggests this is likely localized on the Met80 side of the heme group, are data indicating a substantive weakening of the Met80

ligand bond in the oxidized state (Dickerson & Timkovich, 1975; Williams *et al.*, 1985b). Only in the oxidized state is cytochrome *c* known to undergo a conformational change at alkaline pH in which this bond is lost (Pearce *et al.*, 1989) and a new ligand is formed by some other group whose identity has not yet been established (Gadsby *et al.*, 1987). The presence of a weaker ligand bond in the oxidized state is supported by our observation that the side-chain of Met80 has an increased overall thermal factor in oxidized yeast iso-1-cytochrome *c*. This change in thermal factor values is very localized in that neither the heme group average thermal factor nor that of the His18 ligand is affected. Also apparent is some displacement of side-chain atoms and a small ~7°change in the rotation bond angle between the Met80 side-chain and the heme iron atom. However bond weakening does not appear to cause a lengthening of the Met80 SD heme iron bond (Korzun *et al.*, 1982). This observation is supported by the current study, where although this bond distance shows the largest deviation amongst heme iron coordinate bonds (Table 3.8), this deviation is within experimental error (see Appendix C).

An essential parameter in the regulation of the relatively high reduction potential of cytochrome *c* is the strong withdrawing power that the Met80 ligand exerts on the heme iron atom (Marchon *et al.*, 1982; Mathews, 1985). Our results and those of others suggest that relative to the reduced state, in the oxidized protein the Met80 ligand is weakened and exerts less of an electron withdrawing effect, a situation favoring stabilization of the positively charged heme iron atom. An important factor in modulating Met80 ligand bond strength is likely to be the interaction this group has with the side-chain of Tyr67. In reduced yeast iso-1-cytochrome *c* the Met80 SD sulfur atom is the acceptor in a hydrogen bond to the Tyr67 OH (see Figure 3.28). This hydrogen bond could be expected to increase the electron withdrawing strength of the Met80 ligand and thereby serve to further increase the heme reduction potential. This is clearly illustrated by the ~55 mV drop in reduction potential observed when this hydrogen bond is eliminated by mutation of Tyr67 to a phenylalanine (Luntz *et al.*, 1989; Wallace *et al.*, 1989; see also Section 1.5 and Chapters 4 and 5).

As discussed earlier, in the oxidized state it appears that the Met80 SD to Tyr67 OH

hydrogen bond is broken (see Figure 3.28). This change would facilitate stabilization of oxidized cytochrome *c* through two mechanisms. First, breakage of this bond would diminish the electron withdrawing power of the Met80 ligand. Second, the side-chain of Tyr67 could interact to allow the internal water molecule Wat166 to reorient its dipole moment to maximally stabilize the positive charge resident on the heme iron atom. It should be stressed here that these features are not unique to yeast iso-1-cytochrome *c*. Model-building of the dipole of the internal water molecule in other eukaryotic cytochrome *c* structures provide the same results (see Figure 3.28c). Thus, stabilization of the oxidized state through differential orientation of the dipole moment of the internal water molecule, which not only interacts favorably with the altered electrostatic field, but also regulates the electron withdrawing power of the Met80 sulfur ligand, appears to be a universal feature among eukaryotic cytochromes *c*.

Perturbations about the Met80 ligand also appear to be associated with the greater mobility observed in the adjacent polypeptide chain segment composed of residues 81-85. Maximal increases in thermal factors are localized at Phe82, a residue implicated in the electron transfer mechanism of cytochrome *c* (Liang *et al.*, 1987, 1988; Louie *et al.*, 1988b; Louie & Brayer, 1989; Nocek *et al.*, 1991). The side-chain of this residue is positioned parallel to the plane of the heme, near its solvent exposed edge and is in van der Waals contact with the side-chain of Met80 (see Figure 3.20). It is also positioned in that portion of the cytochrome *c* surface believed to form the contact face in complexes with electron transfer partners (Salemme, 1976; Lum *et al.*, 1987; Pelletier & Kraut, 1992). A small change in the orientation of the phenyl group of Phe82 with respect to the heme plane is observed in the oxidized state of yeast iso-1-cytochrome *c*. It has been postulated that the side-chain of Phe82 may undergo very large motions in complexes with electron transfer partners (Wendoloski *et al.*, 1987), however this suggestion has not been proven experimentally (Burch *et al.*, 1990).

Another conformational change in this region is the formation of a new hydrogen bond between Gly84 O and Arg38 NH1. Arg13 has also been implicated in being an important component in electron transfer complexation interactions (Margoliash & Bosshard, 1983; Satterlee

et al., 1987; Lum *et al.*, 1987). This new hydrogen bond has also been observed in mutant yeast iso-1-cytochrome *c* structures which mimic the oxidized state (Louie *et al.*, 1988b; Louie & Brayer, 1989). Given its apparent importance to the electron transfer process, it is not unexpected that the region about Phe82 is sensitive to oxidation state changes.

3.3.2 Mechanistic implications

Taken together, our results suggest that there is a concerted and cooperative deployment of interconvertible structural changes which facilitate stabilization of the alternative oxidation states of the heme group of cytochrome *c*. Based on these observations and the surface presentation of oxidation state dependent conformational changes, it is possible to propose a trigger mechanism for electron transfer events mediated by cytochrome *c*. As illustrated in Figure 3.24, those residues about the three focal points of conformation readjustments are localized to the Met80 ligand side of the heme, adjacent to the solvent exposed edge of this group. In terms of the outer surface of cytochrome *c*, these changes are centrally positioned in that region of the surface believed to form complementary complexation interactions with redox partners (Margoliash & Bosshard, 1983; Poulos & Kraut, 1980; Pelletier & Kraut, 1992). One rather unusual aspect of the surface representation of those residues conformationally adjusted according to oxidation state is the unperturbed nature of residues 73-80 which bisects this region.

It is notable that the 73-80 segment in cytochrome *c* has the most highly conserved sequence, with almost complete absence of natural variance at the amino acids involved (Hampsey *et al.*, 1986; Louie *et al.*, 1988a). The major secondary structural feature formed by this portion of polypeptide chain is a type II β-turn composed of residues 75-78 (Louie & Brayer, 1990). It is also the side-chain of Thr78 that forms a hydrogen bond to the oxidation state sensitive Wat166.

Given its central location within the putative contact zone with redox partners, the high degree of sequence conservation, its unique response to oxidation state and the linkage through Thr78 to the three focal points of conformational change, it seems possible that the 73-80

segment acts to trigger the alternative structural changes required to stabilize the two oxidation states of cytochrome *c*. This peptide segment is appropriately positioned to carry out this triggering function by virtue of its surface location next to the solvent exposed edge of the heme, a likely route for an electron to travel to and from the heme group. One might think of residues 73-80, or perhaps more specifically the 75-78 β-turn, as a push-button contact trigger operated by protein-protein contacts with redox partners, leading to the generation of the necessary conditions required to facilitate electron transfer.

Precisely how this proposed push-button trigger might function cannot be determined in detail on the basis of our current studies. This arises from only being able to resolve the two endpoints of the oxidation reaction, the reduced and oxidized cytochrome *c* structures, not the intermediary transition state conformation (or possibly conformations). Nonetheless, the current results suggest that the proposed push-button trigger would retain the same spatial positioning in the two end oxidation states, but once pushed by protein-protein contacts in a complex with a redox partner, would act to promote changes in protein conformation that would facilitate electron transfer in either direction. One role of differential mobility in the region about residues 73-80 could be that of mediating these changes by providing the necessary conformational flexibility. This proposal is analogous to the type of switching carried out in a conventional electrical circuit having a push-button contact switch that reversibly alters the direction of current flow.

It must be emphasized, as diagrammatically illustrated in Figure 3.31, that the structural details of the intermediary conformational states through which cytochrome *c* must pass on going between alternative oxidation states are missing from our understanding of this process. Also missing is any information on the potential role played by redox partner interactions beyond the proposed recognition and triggering events. While it seems likely that the intermediary state is some hybrid of the fully oxidized and reduced cytochrome *c* structures, it is entirely possible that redox partner complexation results in transient conformations (of either residues 73-80 or elsewhere) that are not evident in the two end states. Such surface transitions have

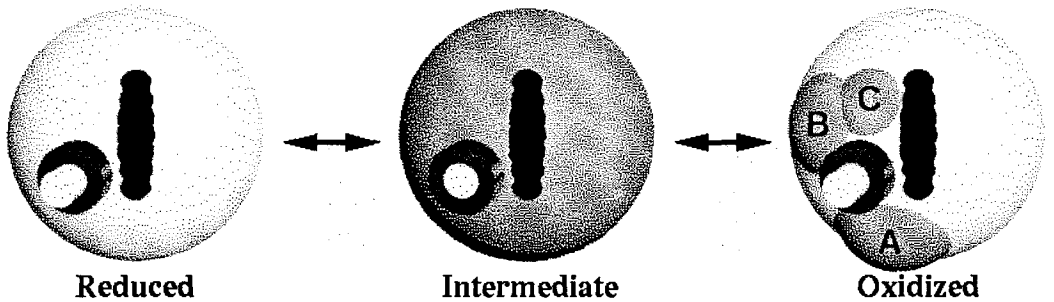

Figure 3.31: A schematic representation of the proposed push-button mechanism for electron transfer events in yeast iso-1-cytochrome *c*. Upon complexation with redox partners residues 73–80, or perhaps more specifically 75–78, alter their conformation so as to facilitate electron transfer and the necessary concomitant structural readjustments. After electron transfer has taken place and the oxidation state dependent conformational changes have occurred, cytochrome *c* disengages from the complexed redox partner and residues 75–78 return to their original conformation. The exposed heme edge is shown in black in this figure, whereas areas labelled A, B, and C correspond to those segments of polypeptide chain which become more mobile in the oxidized state of cytochrome *c* (see Figure 3.24). Residues 73–80 are represented by the push-button trigger.

been suggested from the results of other studies (Koloczek *et al.*, 1987; Weber *et al.*, 1987; Hildebrandt & Stockburger, 1989). Indeed, it could very well be that our minimalist three-state scheme may be too simple and that multiple transition states span the gap between oxidized and reduced cytochrome *c*.

Our current results do suggest a number of approaches to investigating the intermediary state that is proposed to occur during electron transfer. For example, an obvious target for future structure-function-mutagenesis studies is the protein segment composed of residues 73-80. Of particular interest here would be the effect of three classes of replacements. Those involving residues that outwardly face into the surrounding solvent and presumably contact the surfaces of complexed redox proteins (Lys73, Tyr74, Pro76, Gly77, Lys79), those residues that contact Wat166 or line its internal cavity (Ile75, Thr78, Met80) and those residues likely to be instrumental in setting the folded conformations (Pro76, Gly77). In this latter group, Gly77 is a particularly interesting residue, since it is not only at the third position of the type II β-turn,

but it also adopts ϕ,ψ angles not accessible to other amino acids with side-chains.

At this point only limited attention has been directed at the role of the 73-80 segment. The available studies show that mutation of residues in this region leads to destabilization of cytochrome *c* and to low biological activity (Boon *et al.*, 1979; Wood *et al.*, 1988; ten Kortenaar *et al.*, 1985; Wallace *et al.*, 1989). These results are in accord with the mechanistic proposals discussed herein.

Chapter 4

Mutation of Tyrosine-67 in Cytochrome *c* Significantly Alters The Local Heme Environment

4.1 Experimental Procedures

4.1.1 Crystallization

Protein samples for these studies were obtained from J.G. Guillemette (University of British Columbia). Crystals of the Y67F mutant protein in the reduced state were grown in a 0.1 M sodium phosphate buffer (pH 5.8) containing 94% saturated ammonium sulfate and 40 mM DTT, using the hanging drop method and employing a hair seeding technique (Leung *et al.*, 1989). In most instances crystals did not grow large enough using this method so that supplementary macro-seeding techniques were used. To obtain protein in the oxidized state, crystals were soaked for a day in 20 mM ferricyanide. Both the reduced and oxidized mutant protein crystals were isomorphous with those grown for wild-type yeast iso-1-cytochrome *c* (Louie & Brayer, 1990; Chapter 3). The space group was $P4_32_12$ and unit cell dimensions for the reduced and oxidized mutant protein crystals were $a = b = 36.46$ Å, $c = 137.71$ Å and $a = b = 36.47$ Å, $c = 139.08$ Å, respectively (unit cell dimensions for reduced and oxidized wild-type protein crystals are $a = b = 36.46$ Å, $c = 136.86$ Å and $a = b = 36.47$ Å, $c = 137.24$ Å, respectively).

4.1.2 Data collection and data processing

Diffraction data was collected on an Enraf-Nonius CAD4-F11 diffractometer with a 36.8 cm crystal to counter distance and a helium purged path for the diffracted beam. The radiation used was generated from a copper target x-ray tube operating at 26 mA and 40 kV and was nickel filtered. For each reduced and oxidized protein data collection only one crystal was

used. A total of 9612 reflections to 1.9 Å resolution were measured for the Y67F mutant in the reduced state, and for its oxidized counterpart a total of 9211 reflections to 2.2 Å resolution were collected. The lower resolution of the oxidized Y67F dataset was due to the smaller crystal size in this case. The ambient temperature during data collection was maintained at 15°C.

The two data sets were converted to structure factors as previously described for the oxidized form of yeast iso-1-cytochrome *c* (see Section 3.2) except for the decay and absorption corrections. To correct for decay, the six intensity control reflections periodically measured during data collection, were sorted into two groups so that a resolution dependent correction could be applied. A similar approach was taken for the absorption correction in that two phi curves were collected at different theta angles and used to estimate the resolution dependence of the absorption profile (see Section 2.2). After merging of duplicate reflections the reduced Y67F data set had 7882 structure factors and the oxidized data set contained 5274 structure factors.

4.1.3 Refinement and analyses

Before restrained parameter refinement (Hendrickson & Konnert, 1981) was initiated, each structure factor data set was put on an absolute scale using the method described by Wilson (1942). The starting model used for refinement was the 1.2 Å resolution refined structure of reduced yeast iso-1-cytochrome *c* (Louie & Brayer, 1990) with Tyr67 replaced by a phenylalanine and Cys102 by a threonine. In total, 64 well defined water molecules present in the wild-type reduced structure were used in the starting model. These included the internal water molecules Wat110, Wat121, and Wat168, but excluded Wat166 which is located adjacent to the mutation site. Also included was a sulfate ion which is observed in both the reduced and the oxidized forms of the wild-type protein. Refinement of the reduced mutant protein structure utilized 3565 structure factors that had a $F/\sigma(F)$ ratio ≥ 2.0 and were within the resolution range of 6.0 to 1.9 Å; for the oxidized form 3207 structure factors were employed with a $F/\sigma(F)$ ratio ≥ 1.5 and were in the resolution range 6.0 to 2.2 Å.

Table 4.11: Final stereochemistry for reduced and oxidized Y67F yeast iso-1-cytochrome *c*

Stereochemical refinement parameters	r.m.s. deviation from ideal values		weighting† parameters
	Reduced	Oxidized	
Bond distances (Å)			
1-2 bond distance	0.024	0.020	0.020
1-3 bond distance	0.046	0.046	0.030
1-4 bond distance	0.061	0.058	0.050
Planar restraints (Å)	0.017	0.016	0.200
Chiral volume ($Å^3$)	0.223	0.220	0.150
Non-bonded contacts‡(Å)			
single-torsion	0.238	0.225	0.250
multi-torsion	0.223	0.222	0.250
possible hydrogen bonds	0.237	0.251	0.250
Torsion angles (°)			
planar (0°or 180°)	2.3	2.4	2.5
staggered (±60°,180°)	24.8	25.4	19.0

†The weighting parameters used in the final cycles of refinement were identical for the reduced and oxidized yeast iso-1-cytochrome *c* Y67F mutant.
‡The r.m.s. deviations from ideality for this class of restraint incorporates a reduction of 0.2 Å from the radius of each atom involved in a contact.

During the course of refinement several manual interventions were carried out for both the reduced and the oxidized mutant structures. Based on $F_o - F_c$ difference and fragment maps, and $2F_o - F_c$ and $3F_o - 2F_c$ difference maps, a number of side chains of surface residues were repositioned and several water molecules were added and deleted from the model. At the end of refinement the reduced Y67F structure contained 48 water molecules and 1 sulfate ion, while the oxidized structure had 30 water molecules plus the sulfate ion. Each refinement was concluded when positional shifts became small (r.m.s. shifts $\leq$ 0.03 Å), indicating that convergence had been reached.

The final mutant protein structures display excellent stereochemistry (see Table 4.11) and final crystallographic R-factors were 20.1% and 19.6% for the reduced and oxidized structures, respectively. By plotting R-factor versus resolution an estimate can be made of the r.m.s.

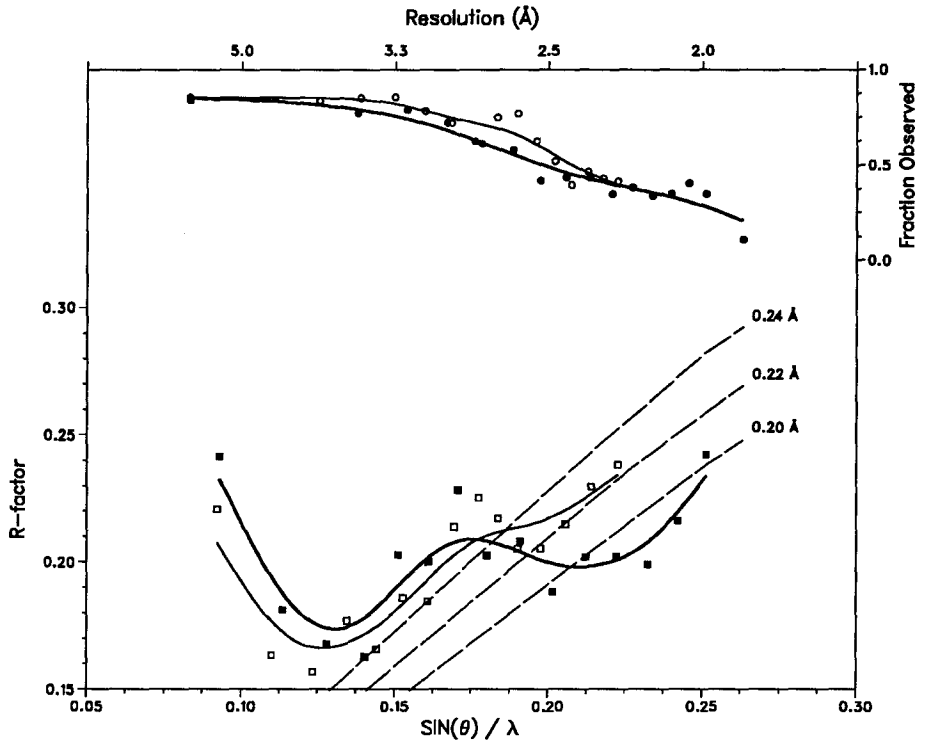

Figure 4.32: Plots of the dependence on resolution of both the R-factor agreement between calculated and observed structure factors (□ – □), and the fraction of data used (○ – ○, axis at top right), for the final refined structures of reduced (filled symbols) and oxidized (open symbols) Y67F yeast iso-1-cytochrome *c*. For this analysis reciprocal space was divided into shells according to $sin(\theta)/\lambda$, with each containing at least 260 reflections. For the purpose of assessing the accuracy of the atomic coordinates of the mutant structures, curves representing the theoretical dependence of R-factor on resolution assuming various levels of r.m.s. error in the atomic positions of the model (Luzzati, 1952) are also drawn (dashed lines). This analysis suggest an overall r.m.s. coordinate error in the range of 0.22 – 0.24 Å for both structures.

coordinate error in each mutant structure (Luzzati, 1952; see also Appendix C). Inspection of Figure 4.32 shows this error to be in the range of 0.22–0.24 Å for both the reduced and oxidized structures. A separate estimate of coordinate errors was made using the method developed by Cruickshank (1949, 1954, 1985). Evaluating Cruickshank's formulas for all protein atoms indicates the overall r.m.s. coordinate errors are 0.21 Å and 0.22 Å for the reduced and oxidized Y67F mutant proteins. These values are in good agreement with those determined by Luzzati's method.

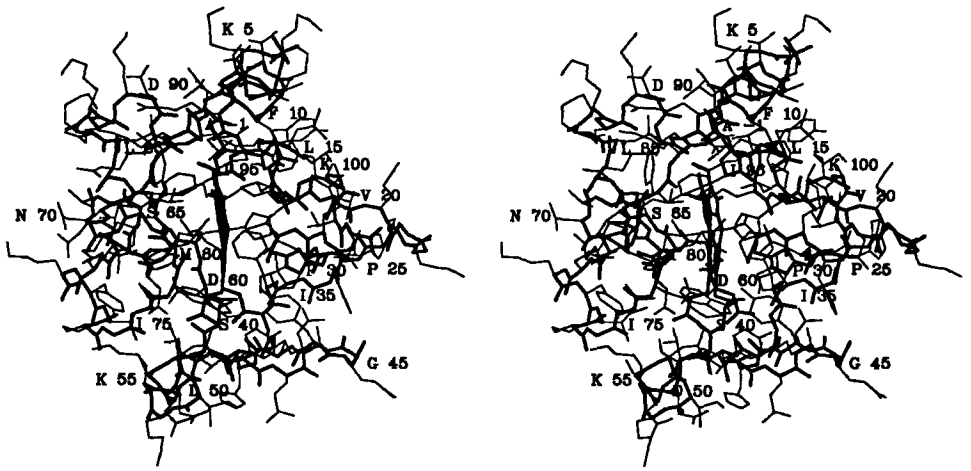

Figure 4.33: A stereo-drawing of the conformations of all main-chain (thick lines) and side-chain (thin lines) atoms in the reduced yeast iso-1-cytochrome *c* Y67F mutant protein. The heme group has been drawn with thick lines. Also drawn are the two heme ligand bonds to His18 and Met80, and the two covalent thioether linkages to cysteines 14 and 17. For clarity, every 5th α-carbon atom has been labelled with its one-letter amino acid designation and sequence number.

4.2 Results

4.2.1 Polypeptide chain conformation

A stereo-drawing of the structure of reduced yeast iso-1-cytochrome *c* in which Tyr67 has been replaced by a phenylalanine residue is shown in Figure 4.33. Note the amino acid numbering used is based on the alignment in Table 1.1. The overall fold of the oxidized form of this protein is identical to its reduced counterpart. The related reduced and oxidized structures of wild-type yeast iso-1-cytochrome *c* have been previously described and comparable figures displaying these structures are shown in Figures 1.5 and 3.18.

As illustrated in Figure 4.34b, a comparison of the reduced forms of the Y67F mutant and wild-type yeast iso-1-cytochrome *c* reveals little perturbation in the placement of main-chain atoms (overall average deviation is 0.29 Å), with the exception of the first two N-terminal

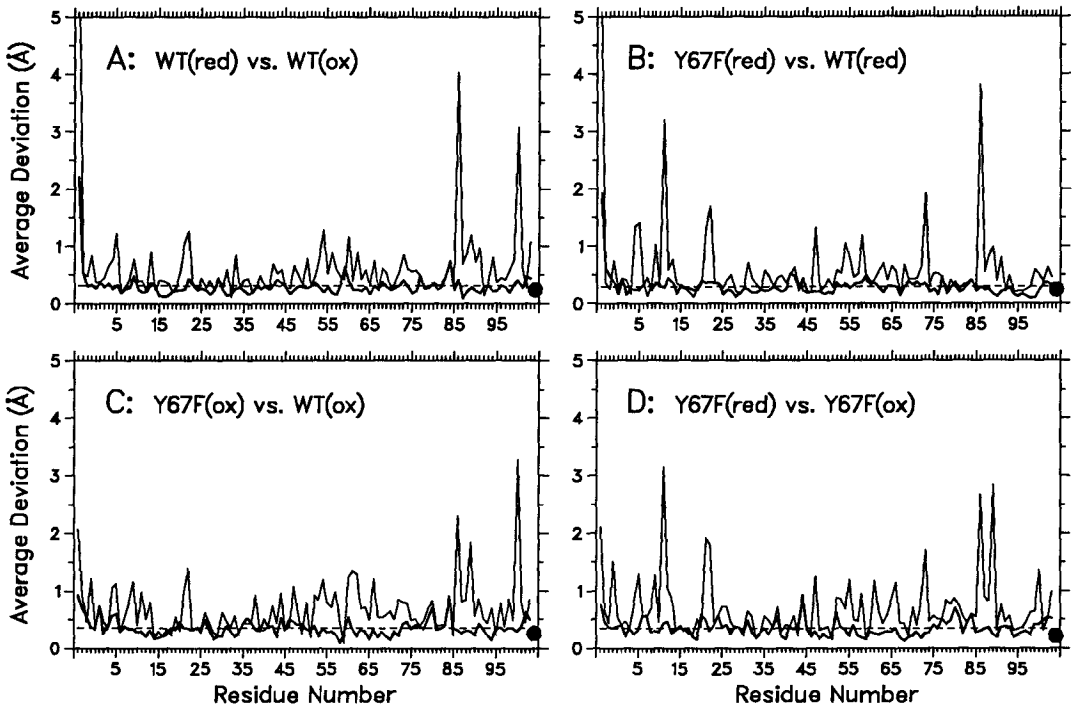

Figure 4.34: A comparison of the average positional deviations between main-chain (thick lines) and side-chain (thin lines) atoms of the reduced and oxidized forms of both the yeast iso-1-cytochrome *c* wild-type and Y67F proteins. In (a) wild-type reduced and oxidized are compared, (b) shows the deviations between reduced Y67F mutant and reduced wild-type, (c) is the comparison between the the oxidized mutant and wild-type structures, and (d) shows the average deviations between the two oxidation states of the Y67F mutant. The horizontal dashed lines represent the average deviations between all main-chain atoms. The filled dark circles at position 104 represent the overall average deviation for heme atoms.

residues. However, these two amino acids (Thr(-5) and Glu(-4)) have been found to be substantially disordered in all yeast iso-1-cytochrome *c* structures resolved thus far and therefore the differences observed are undoubtedly due to differential fits to poor electron density in this region rather than the result of the introduced mutation. Between the two reduced proteins, the next two largest deviations occur at Gly41 (0.5 Å) and Trp59 (0.5 Å). These displacements, however, are less than two times the overall average deviation between all main chain atoms and thus they do not appear to be of marked significance. It is notable that all of the main-chain to main-chain hydrogen bonds present in the wild-type protein are also observed in the Y67F mutant. A detailed listing of these hydrogen bonds can be found in Table 8 of Louie & Brayer

(1990).

Comparison of the oxidized wild-type and Y67F yeast mutant iso-1-cytochromes *c* (see Figure 4.34c) indicates an overall average deviation of 0.35 Å for main-chain atoms. Two amino acids have displacements greater than two times this value. The first is Gly84 which moves ~1 Å in the oxidized state of the wild-type protein to form a hydrogen bond between its carbonyl group and the side-chain of Arg13 (see Chapter 3). The formation of this hydrogen bond does not occur in the oxidized state of the Y67F mutant. The second largest positional deviation (0.7 Å) involves the main-chain atoms of Met80, one of the two residues forming an axial heme ligand. This residue is immediately adjacent to the mutation site and the observed shift may be related to the incorporation of an additional water molecule in this region of the Y67F protein. As with the reduced wild-type and Y67F mutant yeast iso-1-cytochromes *c* a comparable complement of main-chain to main-chain hydrogen bonds is retained in the two oxidized proteins.

A detailed analysis of the conformational changes between the reduced and oxidized forms of the Y67F mutant protein (see Figure 4.34d), reveals an average main-chain displacement of 0.34 Å. One major difference is observed. In the oxidized Y67F protein, Glu44 shifts 0.9 Å such that its carbonyl oxygen can form a new hydrogen bond with the side chain of His26. The next largest residue shift involves Met80, although this is only slightly greater than two times the overall average observed main-chain displacements. One factor in the shift of Met80 may be the loss of the hydrogen bond from Met80 N to Thr78 OG1 in the reduced Y67F protein (see below). This hydrogen bond, which is present in the oxidized protein, appears to be lost in the reduced form in favor of a new interaction with an additional water molecule bound in this region.

With respect to side-chain conformations between the reduced and oxidized Y67F mutant structures, many side-chains at solvent exposed locations appear to have altered conformations. As Figure 4.34d shows, these include virtually all lysine (-2, 5, 11, 22, 55, 86, 89, 100) and several glutamic acid residues (21, 61, 66), as well as Ser47 and Leu9. However, these apparent

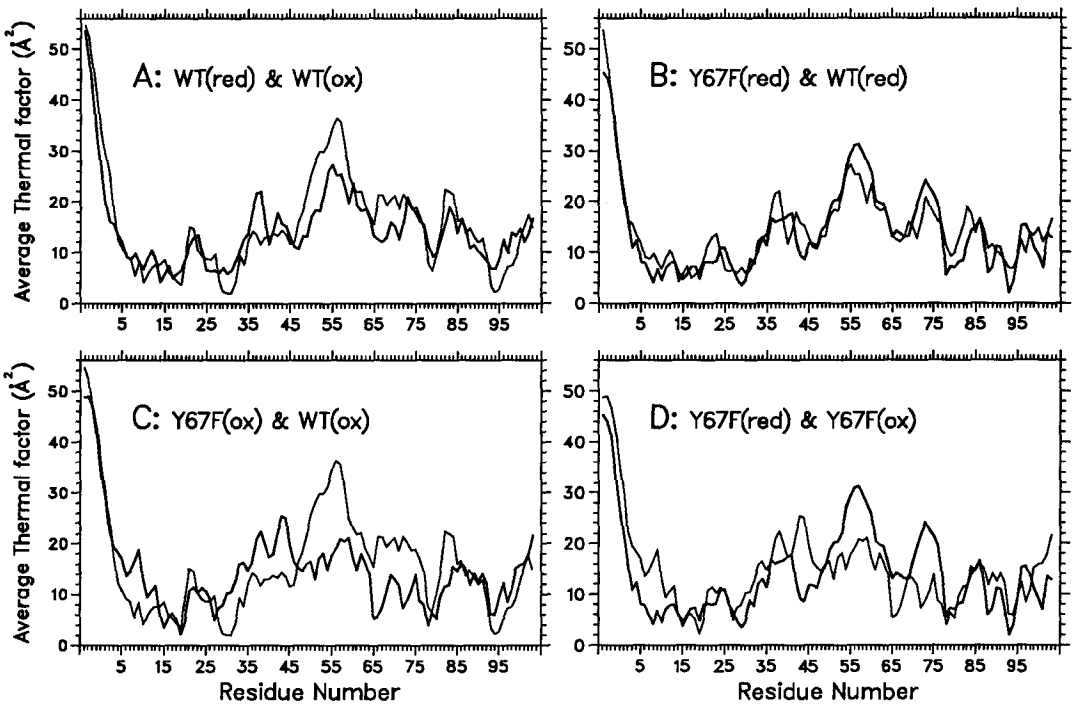

Figure 4.35: Plots of the average thermal factors of main-chain atoms along the polypeptide chain in the reduced and oxidized structures of Y67F and wild-type yeast iso-1-cytochrome *c*. The frames show (a) the wild-type reduced (thick line) and oxidized (thin line) mean thermal factors, (b) the values for reduced Y67F (thick line) and wild-type (thin line), (c) the oxidized structures of Y67F (thick line) and wild-type (thin line), and (d) the values for the reduced (thick line) and oxidized (thin line) Y67F mutant structures.

differences likely result from the mobile nature of these side-chains rather than being related to oxidation state. A detailed analysis of conformation changes in main-chain and side-chain atoms between reduced and oxidized wild-type iso-1-cytochrome *c* (Figure 4.34a), can be found in Chapter 3.

An analysis of main-chain thermal factors along the course of the polypeptide chain in the reduced and oxidized Y67F mutant, as well as comparisons to the related wild-type structures are shown in Figure 4.35. As shown in Figure 4.35a, a number of differences are observed between thermal factor profiles in the two oxidation states of the wild-type iso-1-cytochrome *c*. These have been proposed to play a role in promoting electron transfer and stabilizing the protein in its alternative oxidation states (see Section 3.3). Between the reduced Y67F and

wild-type proteins (see Figure 4.35b), thermal factor variability along the polypeptide chain is quite similar. The only notable difference involves a small increase in mobility for residues 56–59 in the mutant protein. This change is maximal at Trp59, a residue which has an average main-chain thermal factor of 20 $Å^2$ in the reduced wild-type protein and 28 $Å^2$ in the mutant.

As Figure 4.35c shows, the oxidized Y67F structure has a very different thermal factor profile when compared to the oxidized wild-type protein. Thus, the nature of changes observed for thermal factors between oxidation states in the wild-type protein and those found in the Y67F protein are substantially different (see Figures 4.35a, d). In particular, the oxidized Y67F mutant appears to lack the increased polypeptide chain mobility changes observed in the oxidized wild-type protein. As Figure 4.35d shows, two portions of polypeptide chain have marked lower average thermal factor values. These include residues 51–61 and 70–77. Although there is a general trend to lower thermal factor values in the oxidized state of the Y67F protein, two small regions composed of residues 37–39 and 42–46, do exhibit greater mobility. Thus while the mutation of Tyr67 to a phenylalanine appears to have minimal impact on polypeptide chain positioning, it strongly affects the profile of average thermal factors observed, especially in the oxidized state.

4.2.2 Heme structure

The heme groups of the reduced and oxidized Y67F mutant are distorted from planarity into a saddle shape (see Table 4.12). This distortion is more pronounced for both mutant forms in comparison to that of reduced wild-type protein, but comparable to that seen in the case of the oxidized wild-type protein. Thus the increase in heme planar distortion observed upon oxidation of wild-type iso-1-cytochrome *c* (Section 3.2) is not seen between the two oxidation states of the mutant protein. The average angular deviations of pyrrole rings from the pyrrole N plane are 11.5° and 11.6° for reduced and oxidized Y67F respectively, and 9.4° and 12.2° for the reduced and oxidized forms of the wild-type protein.

As Table 4.12 shows heme iron coordinate bond lengths are comparable in both oxidation

Table 4.12: Heme geometry of wild-type and Y67F mutant proteins

	Reduced Wild-type	Reduced Y67F	Oxidized Wild-type	Oxidized Y67F
a) Angular deviations (°) between the pyrrole nitrogen plane normal and the four individual pyrrole ring plane normals and the heme coordinate bonds.				
A	9.4	13.7	12.6	13.4
B	11.1	9.1	14.1	10.7
C	8.8	11.7	9.6	12.0
D	8.1	11.3	12.8	10.2
Fe - His18 NE2	2.1	3.9	7.2	1.9
Fe - Met80 SD	4.9	3.9	3.3	6.3
b) Angular deviations (°) between the porphyrin ring plane normal and the four pyrrole ring normals, the pyrrole nitrogen plane normal and the heme coordinate bonds.				
A	6.7	9.0	8.3	7.4
B	11.9	12.3	12.4	12.4
C	9.8	11.6	13.8	10.8
D	6.0	6.4	10.0	4.4
NNNN	2.6	4.9	4.3	6.0
Fe - His18 NE2	3.3	5.1	6.5	5.8
Fe - Met80 SD	7.5	6.9	5.4	7.6
c) Heme iron ligand bond distances (Å).				
His18 NE2	1.99	1.96	2.01	2.05
Met80 SD	2.35	2.36	2.43	2.44
Heme NA	1.97	2.03	1.97	2.01
Heme NB	2.00	1.98	1.98	2.01
Heme NC	1.99	1.97	2.02	2.01
Heme ND	2.00	2.03	2.05	2.02

Each pyrrole ring is defined by nine atoms, which include the five ring atoms plus the first carbon atom bonded to each ring carbon. The porphyrin ring plane is defined by the five atoms in each of the four pyrrole rings, the four bridging methine carbons, the first carbon atom of each of the eight side chains and the heme iron (33 atoms in total). The pyrrole nitrogen plane is defined by only the 4 pyrrole nitrogens. Heme atom nomenclature used in this table and elsewhere follows that of the Protein Data Bank as illustrated in Figure 1.2.

states, and in both wild-type and mutant proteins. In contrast further analysis shows the angle between the pyrrole bonds to the heme and the imidazole ring plane of the heme ligand His18 appears to be dependent upon oxidation state. This effect was previously noted in the wild-type protein (Section 3.2). In reduced Y67F protein, the His18 ring plane and a vector through the NA and NC heme atoms have a relative angle of 44.8° (in reduced wild-type this value is 46.7°). In the oxidation form the plane of the His18 side-chain is rotated ~10° towards the A and C pyrrole rings of the heme, making an angle of 54.6°. A similar rotation is observed in the oxidized wild-type protein where this angle becomes 55.8°.

A comparison of the Met80 ligand in the reduced wild-type and Y67F mutant proteins shows the structural differences present are minimal, even though this residue is adjacent to the mutation site and has substantially altered interactions with nearby residues in the mutant protein. Particularly notable is the loss of the hydrogen bond from the Met80 SD sulfur atom to the hydroxyl group of Tyr67 present in the wild-type protein and its substitution for a hydrogen bond to Wat300 in the Y67F mutant. Also the hydrogen bond between Met80 N and Thr78 OG1 is absent in the mutant structure. In contrast a comparison of the oxidized wild-type and Y67F mutant proteins shows that more substantial conformational differences are present. The average positional shift of Met80 side-chain atoms is ~0.8 Å (compared with ~0.3 Å between the reduced proteins) with the terminal Met80 CE methyl group showing the largest displacement (~1.8 Å). In particular there is a ~40° rotation of this methyl group about the ligand bond. As will be discussed herein, this conformation may be the result of the disordering of water molecules in the nearby internal protein cavity. In the oxidized wild-type protein a well ordered water molecule (Wat166) is located adjacent to Met80 and believed to stabilize this conformation of the protein (Chapter 3).

An analysis of the solvent accessibility of the heme group in the two oxidation states of both the wild-type and mutant iso-1-cytochromes *c* is presented in Table 4.13. This data shows that heme solvent exposure is not dependent upon the oxidation state of either protein. However, the average heme exposure to solvent in the two states of the mutant protein (~7.7%) is somewhat

Table 4.13: Heme solvent accessibility of wild-type and Y67F mutant proteins

	Reduced Wild-type	Reduced Y67F	Oxidized Wild-type	Oxidized Y67F
1. Solvent accessible heme atoms and surface area exposed (Å^2)				
CHD	2.6	0.0	0.0	0.0
CMC	9.4	9.9	9.5	9.0
CAC	3.3	0.0	5.1	2.9
CBC	17.4	22.1	19.2	15.8
CMD	9.8	7.5	11.7	9.4
2. Total heme exposure (Å^2)	42.5	39.5	45.5	37.1
3. Total heme surface (Å^2)	495.7	500.0	503.3	499.9
4. Heme surface area exposed (%)	8.6	7.9	9.0	7.4

Computations were done using the method of Connolly (1983) and the results represent the accessible molecular surfaces of the atoms listed. The probe sphere used had a radius of 1.4 Å.

smaller than in the wild-type protein (~8.8%).

Another factor which is clearly dependent on whether the protein is in the reduced or oxidized state is the positioning of the pyrrole A propionate group (Figure 4.36). The conformation of this propionate in the Y67F reduced structure is essentially identical to that seen in the wild-type reduced form. In oxidized wild-type protein, the pyrrole A propionate exhibits an alternate conformation where the three bonds C2A–CAA, CAA–CBA, and CBA–CGA have rotated by ~20°, ~30°, and ~45°, respectively (Section 3.2). In the Y67F oxidized protein these same bonds have changed by ~20°, ~25°, and ~35°, respectively, resulting in a virtually identical positioning of the propionate compared to that seen in oxidized wild-type protein. As in the wild-type protein, oxidation of the mutant cytochrome *c* also results in the movement of an internal water molecule (Wat121) such that it can form a bifurcated hydrogen bond to both oxygens of the propionate (Section 3.2).

The hydrogen bond network around the pyrrole A propionate in the two mutant structures

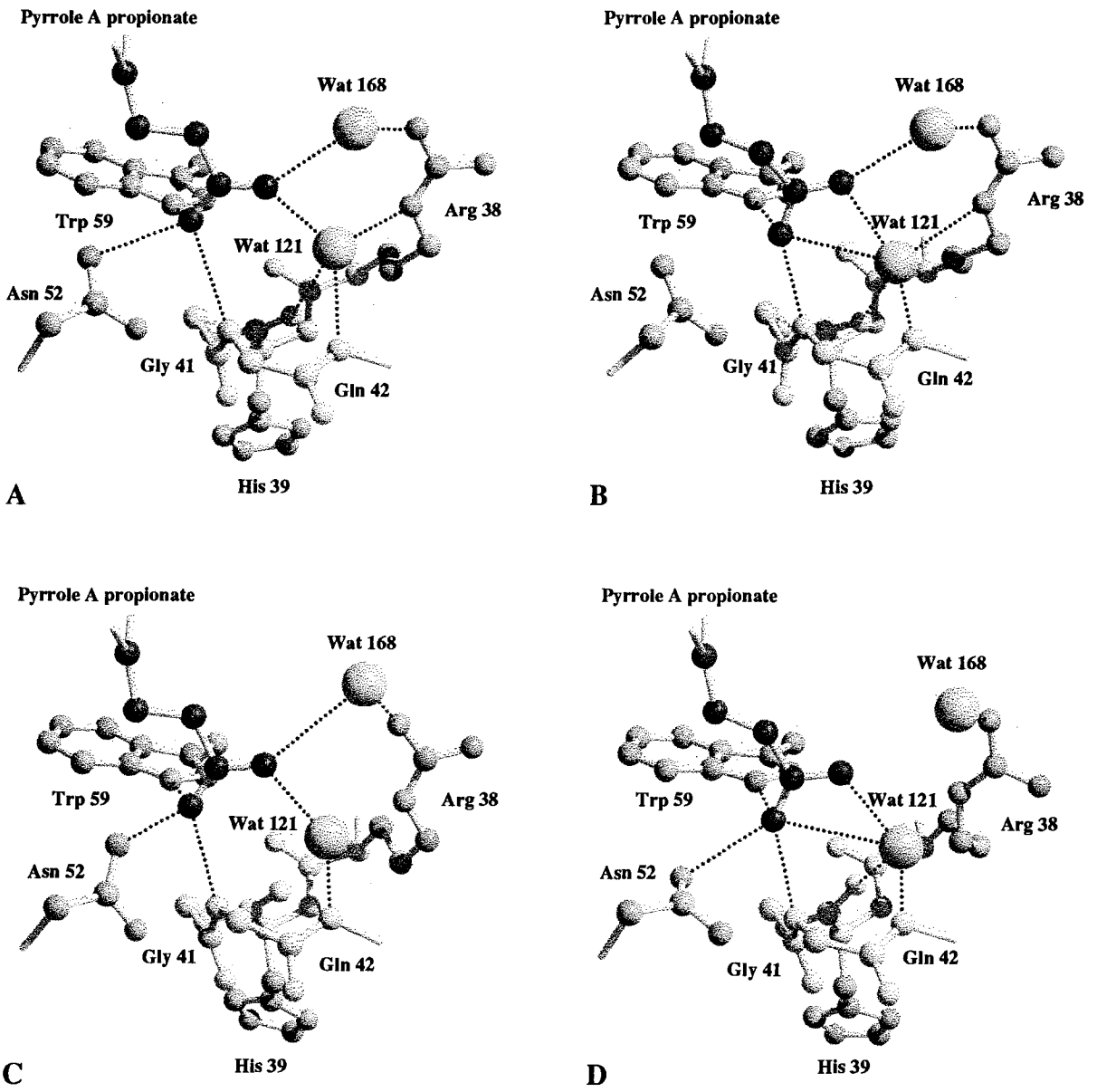

Figure 4.36: Drawings of the pyrrole A propionate group region as observed in (a) reduced and (b) oxidized wild-type yeast iso-1-cytochrome *c*, and in the (c) reduced and (d) oxidized Y67F mutant protein. The pyrrole A propionate group is highlighted with dark shaded balls. The two internally bound water molecules, Wat121 and Wat168, which mediate the interaction of Arg38 with this heme propionate, are shown as larger spheres. Hydrogen bonds are indicated by thin dashed lines.

Table 4.14: Heme propionate hydrogen bond interactions in wild-type and Y67F mutant yeast iso-1-cytochrome *c*

Hydrogen bond partners		Distances (Å)			
		Reduced Wild-type	Reduced Y67F	Oxidized Wild-type	Oxidized Y67F
O1A	Tyr48 OH	2.83	2.97	2.83	3.51
	Wat121	2.81	2.57	2.85	2.69
	Wat168	2.85	3.41	2.87	(3.62)
O2A	Gly41 N	3.21	2.82	2.60	2.88
	Asn52 ND2	3.34	2.67	(3.54)	3.41
	Trp59 NE1	3.09	3.21	3.43	2.89
	Wat121	(4.01)	(3.74)	3.34	3.51
O1D	Thr49 OG1	2.64	2.60	2.79	2.71
	Thr78 OG1	2.90	2.75	3.07	3.08
	Lys79 N	3.17	(3.39)	2.67	2.96
O2D	Thr49 N	2.94	2.77	2.75	2.95

Interactions were accepted as hydrogen bonds only if they met all of the following criteria: a H···A distance < 2.6 Å, a D—H···A angle > 120°, and a C—A···H angle > 90°. Values given in brackets are not considered to be hydrogen bonds by this criteria, but are listed for comparison.

closely resembles that of the reduced and oxidized wild-type structures (see Table 4.14). However, a few differences in this hydrogen bond network are observed (see Figure 4.36). In the reduced Y67F mutant structure Wat121 is shifted by ~1 Å and as a result the hydrogen bond between Wat121 and Arg38 NE, as well as the hydrogen bond between this water molecule and His39 O are absent. Because of the absence of this latter interaction, the His39 CA moves by 0.25 Å and this shift is propagated to the side chain (shift in CB is 0.50 Å), culminating in a 30° rotation of the imidazole ring. In addition to these differences the hydrogen bond between Trp59 and the propionate A carboxyl group is weakened as judged by the increased hydrogen bond length (see Table 4.14) and the increased thermal factors for the tryptophan side-chain (average increase +15 $Å^2$).

In the oxidized mutant structure the interaction of Arg38 with the propionate, which is mediated through Wat121 and Wat168, is severely weakened as indicated by the absence of the Wat121 - Arg38 NE and Wat168 - heme O1A hydrogen bonds, as well as the increase in thermal factors for residues 37-39 (see Figure 4.35). This is in sharp contrast to the wild-type protein where the interaction between Arg38 and the pyrrole A propionate becomes stronger in the oxidized state as deduced from the observed thermal factors (see Chapter 3). In addition, the hydrogen bond interaction between Asn52 ND2 and the propionate is not lost on going from the reduced to the oxidized state in the Y67F mutant as it is in the wild-type protein. This is evident not only from the distance between these two groups, but also from the absence of an increase in thermal factors for the Asn52 side-chain in the mutant protein. In the wild-type protein the Asn52 side-chain becomes substantially more mobile in the oxidized state (average increase +16 $Å^2$).

4.2.3 Mutation site: the reduced state

In the reduced state of wild-type yeast iso-1-cytochrome *c*, the side chain hydroxyl group of Tyr67 is involved in an extensive hydrogen bonding network (Chapter 3). As Figure 4.37a shows, this involves the nearby side-chains of Asn52, Thr78 and Met80, as well as an internally bound water molecule Wat166. This water molecule is a conserved structural feature found in all eukaryotic cytochromes *c* for which three-dimensional structures have been determined (Bushnell *et al.*, 1990; Brayer & Murphy, 1993). The exchange of tyrosine for phenylalanine at position 67 leads to the creation of an internal cavity due to the absence of the tyrosyl hydroxyl group. As Figure 4.37c illustrates, this extra internal volume is filled by an additional water molecule (Wat300) positioned in a location comparable to that of the original hydroxyl group. However, since Wat300 is not covalently bound to Phe67 it is further displaced from this side-chain than a tyrosyl hydroxyl group. In order to accommodate Wat300, a shift in position of Wat166 is required and this group moves downwards (~1 Å) towards the side-chain of Asn52. Thus an important consequence of the tyrosine to phenylalanine mutation at position 67 is to

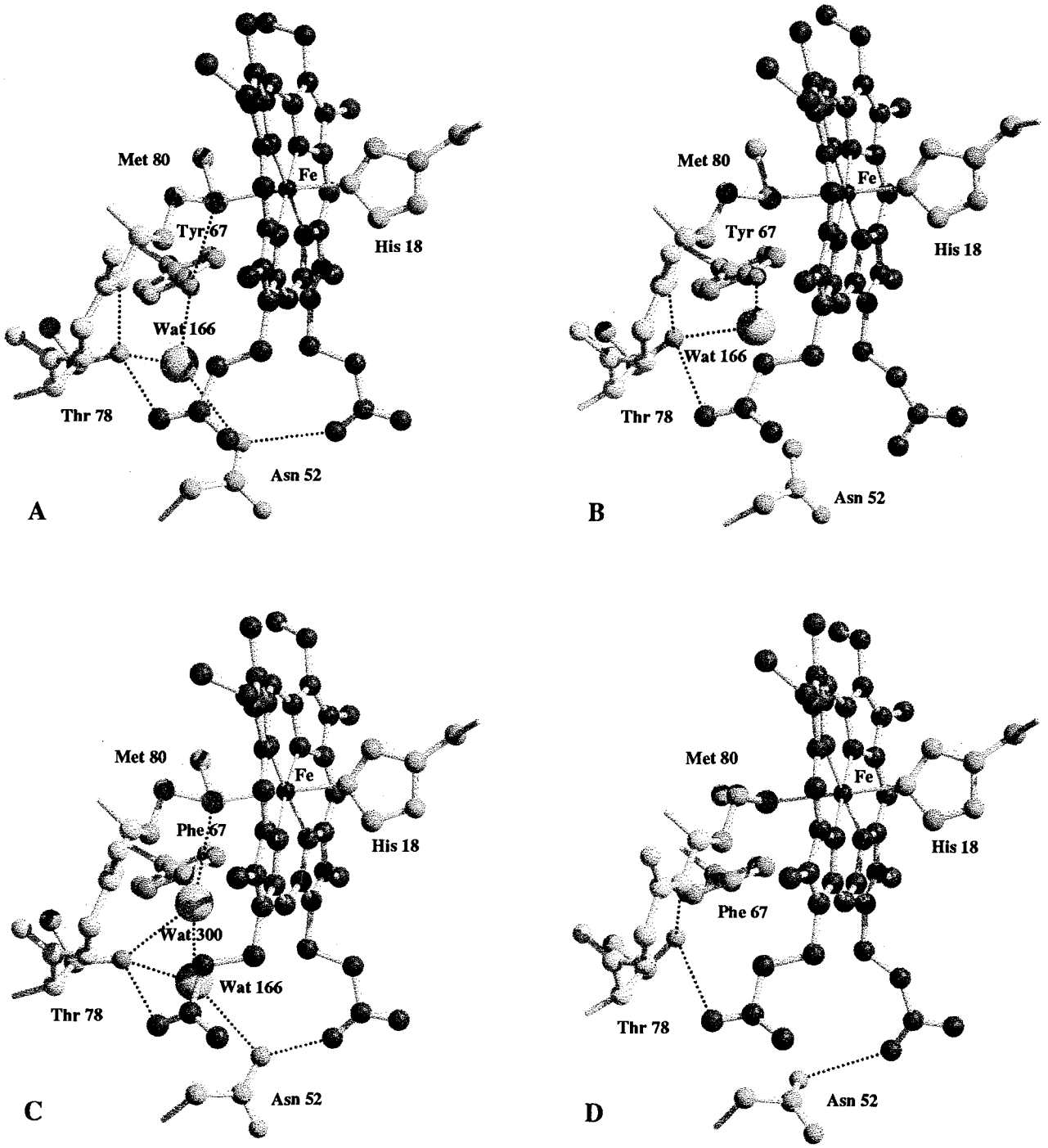

Figure 4.37: Drawings showing the region about residue 67 in the (a) reduced and (b) oxidized wild-type and, (c) reduced and (d) oxidized Y67F proteins. In each case the heme group is shown with dark shaded balls and internal water molecules are represented by larger spheres. Heme ligand interactions are indicated by thin white bonds whereas hydrogen bonds are shown by thin dashed lines.

Table 4.15: Hydrogen bond interactions at the mutation site in the wild-type and Y67F proteins

Hydrogen bond partners		Distances (Å)			
		Reduced Wild-type	Reduced Y67F	Oxidized Wild-type	Oxidized Y67F
Asn52 ND2	Heme O2A	3.34	2.67	(3.53)	3.41
Asn52 ND2	Wat166	3.14	3.15	(4.26)	-
Tyr67 OH/Wat300†	Thr78 OG1	(4.17)	2.94	(4.42)	-
Tyr67 OH/Wat300†	Met80 SD	3.25	3.22	(3.12)	-
Tyr67 OH/Wat300†	Wat166	2.62	2.87	2.63	-
Thr78 OG1	Heme O2D	2.90	2.75	3.07	3.08
Thr78 OG1	Wat166†	2.72	3.25	2.70	-
Met80 N	Thr78 OG1	3.32	(3.69)	2.73	2.81

Interactions were accepted as hydrogen bonds only if they met all of the following criteria: a H···A distance < 2.6 Å, a D—H···A angle > 120°, and a C—A···H angle > 90°. Distances given in brackets are not considered to be hydrogen bonds by this criteria, but are listed for the sake of comparison.

† In the reduced Y67F mutant protein, the hydroxyl group of Tyr67 is replaced by a new water molecule (Wat300) positioned in approximately the same location. In the oxidized Y67F structure, neither Wat300 or Wat166 can be definitively located and both are likely undergoing large dynamic motions in the enlarged internal protein cavity available.

increase the number of water molecules in the nearby internal cavity from one to two.

Surprisingly, the presence of Wat300 leads to relatively little change in the orientation of nearby side-chains. The ring atoms of the side-chain of Phe67 show no significant shifts in position (average deviation 0.1 Å), average thermal factors (average decrease ~-1 $Å^2$) or in overall ring orientation (angular deviation ~1°), when compared to its tyrosyl counterpart in the wild-type protein. The largest shift occurs in the side-chain of Asn52 (average deviation 0.6 Å), which moves to accommodate the displaced Wat166. Movement of Asn52 is accomplished for the most part by a 15° rotation in the torsional angle between its CB and CG atoms.

Insertion of Wat300 does result in a substantial reorganization of the hydrogen bond network in this region of the protein (Table 4.15, Figures 4.37a,c). The most prominent change is the loss of the Tyr67 OH to Met80 SD hydrogen bond, which is replaced by a new interaction

involving Wat300. This latter group is in turn hydrogen bonded to Wat166 and the side-chain of Thr78. Note that in the wild-type protein the hydroxyl group of Tyr67 is too far removed from the OG1 of Thr78 to form a similar interaction. The positioning of the side-chain of Thr78 is largely unaffected by these changes, although a loss of the hydrogen bond between this side-chain and the main-chain amide group of Met80 is observed.

4.2.4 Mutation site: the oxidized state

Neither Wat300 nor Wat166 can be resolved in the Y67F oxidized protein structure, presumably because they are undergoing greater dynamic motion than in the reduced mutant protein where the thermal factor of these water molecules are 30 $Å^2$ and 27 $Å^2$, respectively. This increased mobility likely results from two factors, both of which involve the side chain of Asn52. As Figure 4.37d shows, in the oxidized state this side-chain shifts further away from the internal cavity (average deviation 0.9 Å) and is unlikely to be in a position to form a hydrogen bond to an internal water molecule. Furthermore, the shift of this side chain is primarily responsible for opening up the size of the internal cavity containing Wat300 and Wat166 from a volume of 35 $Å^3$ in the reduced mutant protein, to 60 $Å^3$ in the oxidized state. This increased volume would allow considerably more freedom of movement to both water molecules.

Despite the increased size of the cavity, the shifts in side-chain atoms of residues lining the mutation site are small between the reduced and oxidized Y67F proteins, with the exception of Asn52. Further comparisons show side-chain differences between the oxidized wild-type and Y67F cytochromes *c* are also small. Again, the largest deviation between the two oxidized proteins involves Asn52 (average deviation 0.9 Å). Although in each case the general direction of shifts is similar, the two Asn52 side-chains do adopt different local conformations. In the mutant protein the Asn52 ND2 group can still hydrogen bond to the pyrrole A propionate in the oxidized state, while in the wild-type structure this linkage is broken. One further difference between oxidation states of the Y67F mutant protein is the presence of a Thr78 OG1 to Met80 N hydrogen bond in the oxidized state (see Table 4.15 and Figure 4.37c). This hydrogen bond is

observed in both the reduced and oxidized states of wild-type yeast iso-1-cytochrome *c*.

In contrast to the oxidized wild-type protein, the side-chains of Asn52, Phe67 and Met80 all have lower average thermal factors in the oxidized Y67F protein. This is particularly prominent for the side chain of Phe67 for which the average thermal factor is reduced by ~23 $Å^2$. This observation coincides with other results (Figure 4.35) indicating that the polypeptide chain of the oxidized Y67F mutant protein does not exhibit the regional increases in mobility observed for the oxidized wild-type protein.

4.3 Discussion

4.3.1 Conformational effects of the Y67F mutation

Shown in Table 4.16 is an overview of the structural differences observed between the reduced and oxidized Y67F yeast iso-1-cytochrome *c* mutant structures and their wild-type counterparts. The conformational differences observed are localized for the most part in two regions, the mutation site and the pyrrole A propionate region (see Figures 4.36 and 4.37). These two areas have previously been shown to be sensitive to the oxidation state of cytochrome *c* (Chapter 3).

Mutation site

It has been proposed that by mutating tyrosine 67 to a phenylalanine, and thereby removing the hydroxyl group from this residue, the heme environment would become more hydrophobic (Koul *et al.*, 1979; ten Kortenaar *et al.*, 1985; Wallace *et al.*, 1989). It was also suggested that this mutation would affect the internal water molecule Wat166 which is located adjacent to the mutation site and is important in mediating oxidation state dependent conformational changes (Chapter 3). Indeed, Luntz *et al.* (1989) proposed that the Y67F mutation would exclude Wat166 from the structure of cytochrome *c*. This proposal was based on NMR experiments and stability studies, and referred to work performed by Rashin *et al.* (1986) which indicated that a water molecule buried inside a protein would be energetically unfavorable unless it could form at least three hydrogen bonds. In this line of reasoning, the loss of a hydrogen bond

Table 4.16: Structural differences observed in Y67F mutant structures when compared to their wild-type yeast iso-1-cytochrome *c* counterparts

Reduced Y67F Mutant	Oxidized Y67F Mutant
A. *Positional displacements of polypeptide chain (see Figure 4.34)*	
1. No significant differences in main-chain placement observed	1. Movement of Gly84 seen in oxidized wild-type is not observed
	2. Small shift in Met80 main-chain atoms
B. *Thermal factor parameters of main-chain atoms (see Figure 4.35)*	
1. Small increase in flexibility for residues 56–59, focussed at Trp59	1. Regional increases in flexibility seen in wild-type are not observed
	2. Increase of thermal factors observed for residues 37–39 and 42–46
	3. Substantially lower thermal factors for the side-chains of residues lining the internal water cavity
C. *Heme structure (see Figure 4.36 and Tables 4.12 and 4.14)*	
1. Heme plane more distorted	1. Pyrrole A propionate to Arg38 interaction is lost
2. Pyrrole A propionate hydrogen bond network weakened	
D. *Mutation site (see Figure 4.37 and Table 4.15).*	
1. Additional water molecule (Wat300) present in internal cavity	1. Internal cavity size increased
2. Realignment of hydrogen bonds	2. Large dynamic motion of Wat166 and Wat300
	3. Movement of Asn52 results in interaction with pyrrole A propionate
E. *Hydrogen bond interactions (see Figures 4.36 and 4.37 and Tables 4.14 and 4.15).*	
Weaker: Trp59 NE1 - Heme O2A	Stronger: Trp59 NE1 - Heme O2A
Lost: Tyr67OH - Met80 SD Met80 N - Thr78 OG1 Arg38 NE - Wat121 Wat121 - His39 O	Lost: Arg13 NH1 - Gly84 O Arg38 NE - Wat121 Wat168 - Heme O1A Wat166, Wat300†
New: Wat300 - Thr78 OG1 Wat300 - Met80 SD Wat300 - Wat166	New: Asn52 ND2 - Heme O2A

† The positions of Wat166 and Wat300 cannot be defined in the oxidized Y67F mutant structure and their hydrogen bonding interactions are correspondingly unknown.

to Wat166 in the Y67F mutant protein would be expected to result in the loss of this water molecule.

Our results show that Wat166 is retained in the Y67F mutant cytochrome *c*. Further, the hydrophobicity of this region is not increased and may even be decreased due to the presence of an additional water molecule (Wat300) which is found to occupy a position very similar to the hydroxyl group of Tyr67 in the wild-type protein (see Figure 4.37). In the reduced state these two water molecules are part of an extensive hydrogen bond network in this region (see Table 4.15) and each forms three hydrogen bonds, meeting the minimal requirements suggested by Rashin *et al.* (1986). Therefore this work has clearly shown the interpretation presented by Luntz *et al.* (1989) to be incorrect.

In the oxidized wild-type structure Wat166 reorients its dipole so as to stabilize the positive charge residing on the heme iron atom (see Figure 3.28). In this process several hydrogen bonds involving Asn52 ND2 and Tyr67 OH are broken or realigned. Furthermore, three segments of polypeptide chain located on the Met80 side of the heme group become more mobile (Chapter 3). This increase in flexibility and the resulting decrease in stability of the oxidized state is a well documented property of eukaryotic cytochromes *c* (see Section 1.3). In the reduced Y67F mutant protein the normally resident hydrogen bond network that mediates oxidation state dependent shifts is substantially modified. In addition, the ordered hydrogen bond interactions observed in the oxidized state of the wild-type protein are absent in the Y67F variant. One important consequence of these changes is the absence of regional increases in thermal factors for specific segments of polypeptide chain in the oxidized state of the Y67F mutant (see Figure 4.35). In fact, the segments involved (residues 47–59, 65–72 and 81–85), as well as the side-chains in the vicinity of Wat166 that show large thermal factor increases in the oxidized form of the wild-type protein retain comparable thermal factors in both oxidation states. An additional factor that may contribute to this effect is that in the oxidized state of the mutant protein a hydrogen bond between Asn52 and the pyrrole A propionate is maintained which could serve to prevent the increased flexibility in the region of residues 47–59 observed in the

oxidized wild-type structure (see Figure 4.35). These data correlate well with observations made by other investigators that the oxidized state of the Y67F variant cytochrome *c* is more stable than wild-type (Luntz *et al.*, 1989; Wallace *et al.*, 1989; Margoliash, 1990; Frauenhoff & Scott, 1992). This work suggests that the hydrogen bond network about Wat166 in wild-type cytochrome *c* is critical for modulating the flexibility of nearby polypeptide chain segments between oxidation states, a feature that disappears when this hydrogen bond network is disrupted.

Our results show the two internal water molecules, Wat166 and Wat300, to have greater dynamic motion in the oxidized state of the Y67F mutant protein. Two factors that may explain this behaviour are the increased size of the internal cavity available in the oxidized state due to the shift of Asn52, and the loss of the hydrogen bond between Wat166 and Asn52 ND2. Another potential factor is related to the positive charge residing on the heme iron atom in the oxidized form of the protein. In this state the dipole moments of Wat166 and Wat300 will be strongly influenced to realign in response to the electrostatic field created by the positively charged iron atom. Model-building studies suggest that in this process, several hydrogen bonds observed in the reduced state of the protein will be broken. Thus in the oxidized state neither Wat166 nor Wat300 are likely to form discrete hydrogen bonds with protein groups, a factor undoubtedly contributing to the observed increase in mobility of these water molecules.

Pyrrole A propionate region

Comparison of the reduced forms of the Y67F mutant and wild-type yeast iso-1-cytochromes *c* shows that a change in the nature of the hydrogen bond network around the pyrrole A propionate group has occurred (see Figure 4.36 and Table 4.14). The observed differences involve the invariant residues Trp59 and Arg38, and the internal water molecule Wat121, which is present in all high resolution structures of eukaryotic cytochromes *c* determined thus far (Brayer & Murphy, 1993). As in the reduced state, the hydrogen bond network of the Y67F protein is also perturbed in the oxidized state relative to the wild-type structure. These changes involve not only those observed between the reduced proteins, but in addition Asn52, and a second

internal water molecule Wat168. Surprisingly, despite these observed differences the pyrrole A propionate group of the Y67F protein undergoes an oxidation state dependent conformational change which is virtually identical to that seen in wild-type yeast iso-1-cytochrome *c* (Figure 4.36; Chapter 3).

Our studies suggest that one consequence of the Y67F mutation is that the charge distribution on the pyrrole A propionate carboxyl group is altered, thereby affecting its hydrogen bonding characteristics. A possible source for this perturbation is that the introduced mutation results in modification of the delocalized π-electron system on the porphyrin ring. That delocalized electrons in the porphyrin ring can have an effect on pyrrole A propionate function has been proposed by Moore (1983). Analysis of the electronic spectra of the Y67F protein clearly suggests that the conjugated π-electron system has been disturbed (J.G. Guillemette, private communication). UV-visible spectra of horse cytochrome *c* in which the same mutation has been introduced suggests a similar effect (Wallace *et al.*, 1989). Further evidence for this relationship can been seen in a structural study of reduced F82S yeast iso-1-cytochrome *c* (Louie *et al.*, 1988b). In this mutant protein the π-electron system of the heme is affected by introduction of a solvent channel adjacent to the heme plane. The resultant conformational changes about the pyrrole A propionate are similar to those observed upon change in oxidation state in the Y67F mutant and wild-type yeast iso-1-cytochromes *c*.

4.3.2 Effect of the Y67F mutation on midpoint reduction potential

As shown in Table 1.5 mutation of Tyr67 to a phenylalanine in yeast iso-1-cytochrome *c* causes a drop of ~56 mV in midpoint reduction potential. This same mutation shows a comparable lowering of midpoint reduction potential in both horse and rat cytochromes *c* (Wallace *et al.*, 1989; Luntz *et al.*, 1989; Frauenhoff & Scott, 1992). Some confusion as to the source of this effect is evident in the literature. Conventional theories with regard to the factors that control reduction potential would predict that the Y67F mutant should have a higher midpoint reduction potential since it would increase the hydrophobicity of the heme pocket (Kassner,

1972, 1973; see also Section 1.4). To reconcile this apparent paradox between theory and experiment, Wallace *et al.* (1989) proposed that this mutation must cause a dramatic change in the electronic structure of the heme group. It should be noted that an increase in heme solvent exposure is not responsible for the lower reduction potential of the Y67F mutant as has been suggested by Stellwagen (1978). As Table 4.13 shows, the wild-type and mutant proteins have comparable heme solvent exposures.

Our studies provide an alternative explanation for the lower midpoint reduction potential observed. As discussed in Chapter 3, the hydrogen bond between the hydroxyl group of Tyr67 and the sulfur atom of the Met80 ligand appears to be present or absent depending on the oxidation state of the protein (see Figures 3.28 and 4.37). By modulating this hydrogen bond, stabilization of both oxidation states can be achieved. According to this view, the hydrogen bond between Tyr67 OH and Met80 SD is also an important determinant in setting the value of the midpoint reduction potential by virtue of increasing the electron withdrawing power of the Met80 ligand. The electron withdrawing strength of the Met80 ligand is recognized as a central determinant in setting the reduction potential of cytochrome *c* (Marchon *et al.*, 1982; Moore & Pettigrew, 1990). Thus, weakening or removal of the Tyr67 OH to Met80 SD hydrogen bond would be expected to lead to a decrease in the midpoint reduction potential. This is precisely the effect observed in the Y67F mutant protein.

One factor that might affect the observed value for the reduction potential in the Y67F mutant protein is the presence of Wat300 which fills the space formerly occupied by the hydroxyl group of Tyr67. However, this effect is likely to be limited by the greater mobility and less ideal positioning of Wat300, as well as the absence of a conjugated π-electron system from the phenyl group of residue 67 which would serve to further stabilize a hydrogen bond to the Met80 SD sulfur atom as is found in the wild-type protein. This conclusion is supported by mutations introduced into cytochrome *c* which eliminate both the interaction between Tyr67 and Met80 and all water molecules from the nearby internal cavity (Chapter 5). In these mutant proteins a value for the midpoint reduction potential comparable to that of the Y67F variant is observed.

4.3.3 Role of tyrosine 67 in eukaryotic cytochromes *c*

Our structural studies suggest a dual role for tyrosine 67 in cytochrome *c*. One role is to serve as part of an oxidation state dependent hydrogen bond network that stabilizes the alternative oxidation states of cytochrome *c* by optimally positioning and orienting the dipole moment of Wat166. As part of this process Tyr67 participates in modifying the local flexibility of a nearby polypeptide chain segment (residues 65–72) depending on the oxidation state of the protein. A second role for Tyr67 is in setting the value of the midpoint reduction potential through a hydrogen bond interaction with the sulfur atom of the Met80 ligand. In the Y67F mutant protein both these structural features are affected, resulting in a substantially lower midpoint reduction potential and a structurally more rigid and stable form of the oxidized protein.

Some questions remain as to how important these factors are to biological electron transfer activity. On one hand, the high sequence conservation of Tyr67 would suggest it is essential for biological function. In a compilation of 96 eukaryotic cytochrome *c* sequences, only the protein from *Euglina gracilis* does not have a tyrosine at position 67 (Moore & Pettigrew, 1990). In this organism residue 67 is a phenylalanine instead. On the other hand, the yeast strains used in this study and those which possess the Y67F rat cytochrome *c* mutant do not exhibit diminished growth rates. In fact, it has been reported that these actually grow at a faster rate than the normal wild-type strain (Luntz *et al.*, 1989; Margoliash, 1990).

These observations are rather surprising, especially in the light of the conserved presence of Tyr67 and the properties of the protein with which it is associated. For example, midpoint reduction potential is an extremely conserved property exhibiting limited variation among eukaryotic cytochromes *c* ($\pm$20 mV, with the exception of the *Euglina gracilis* protein; Pettigrew & Moore, 1987). In the absence of other data one would expect the $\sim$56 mV drop in midpoint reduction potential of the Y67F mutant protein to have a clear impact on the biological function of cytochrome *c*. However, as shown by McLendon and co-workers (Komar-Panicucci *et al.*, 1992) a drop in midpoint reduction potential by as much as 120 mV does not prohibit *in vivo* respiration. Similar questions remain as to the importance of Tyr67 modulation of

nearby polypeptide chain flexibility and in assisting the reorientation of Wat166 in response to oxidation state. While our results serve to show some of the roles played by Tyr67, they cannot provide a clear indication of the evolutionary pressures that have lead to the present structure of cytochrome *c*, nor the allowable variance in these features that will maintain sufficiently efficient biological electron transfer.

Chapter 5

Perturbation of a Conserved Internal Water Molecule and its Associated Hydrogen Bond Network in Cytochrome *c* by the Mutations N52I and N52I-Y67F

5.1 Experimental Procedures

5.1.1 Crystallization

Protein samples of the N52I variant of yeast iso-1-cytochrome *c* were provided by F. Sherman (University of Rochester) and used for analysis of the reduced state of the protein. For work with the oxidized protein the additional mutation of C102T was introduced to prevent protein-protein dimerization (Cutler *et al.*, 1987). This protein was provided by J.G. Guillemette (University of British Columbia). For study of the N52I-Y67F protein the C102T mutation was present in both oxidation states. These two latter proteins were also provided by J.G. Guillemette.

Crystals of the reduced N52I and N52I-Y67F mutant proteins were grown in a 0.1 M sodium phosphate buffer (pH 6.0–6.4) containing 86–94% saturated ammonium sulfate, using a combination of the hanging drop method and a hair-seeding technique (Leung *et al.*, 1989). In order to maintain the reduced state of the protein during crystallization either 20 mM DTT or 70 mM sodium dithionite was added to the crystallization buffer. To obtain crystals in which the protein was in the oxidized state, crystals were soaked for one hour in mother-liquor containing 20–40 mM potassium ferricyanide prior to mounting (see Section 2.1). All mutant protein crystals proved to be isomorphous with those grown for wild-type yeast iso-1-cytochrome *c* (see Table 5.17).

Table 5.17: Data collection statistics for the N52I and N52I-Y67F yeast iso-1-cytochromes *c*

	N52I		N52I-Y67F	
	Reduced	Oxidized	Reduced	Oxidized
Cell dimensions (Å)†				
a, *b*	36.46	36.47	36.56	36.52
c	137.82	137.02	137.53	139.12
No. of reflections collected	8449	27230	33491	12907
No. of unique reflections	7910	5790	5700	4771
Merging R-factor (%)‡	13.4	8.4	7.0	6.3
Resolution (Å)	1.9	2.0	2.05	2.0

† The unit cell dimensions for the reduced and oxidized wild-type proteins are $a = b = 36.46$ Å and $c = 136.86$ Å, and $a = b = 36.47$ Å and $c = 137.24$ Å, respectively. The space group is $P4_32_12$ for all proteins.

‡ The merging R-factor is defined as: $R_{merge} = \frac{\sum_{hkl}\sum_{i=0}^{n}|I_{i_{hkl}} - \bar{I}_{hkl}|}{\sum_{hkl}\sum_{i=0}^{n} I_{i_{hkl}}}$. The calculation of the merging R includes all reflections measured more than once (i.e. duplicates and symmetry mates). For the reduced N52I data set which was collected on a diffractometer this value is higher due to the limited number of repeat measurements. The merging R-factor between the reduced and oxidized data sets of the N52I and N52I-Y67F variants was 12.4% in both cases.

5.1.2 Data collection and data processing

Diffraction data for the reduced state of the yeast iso-1-cytochrome *c* N52I mutant was collected from one crystal on an Enraf-Nonius CAD4-F11 diffractometer, with a 36.8 cm crystal to counter distance and a helium purged path for the diffracted beam. The radiation used was nickel filtered and generated from a copper target X-ray tube operating at 26 mA and 40 kV. Data collection and processing methodology for this cytochrome *c* variant were identical to that described for the oxidized structure of wild-type yeast iso-1-cytochrome *c* (see Section 2.2 and 3.1).

Diffraction data for the oxidized form of the N52I mutant protein, and the reduced and oxidized N52I-Y67F mutant proteins, were collected on a Rigaku R-Axis II imaging plate area detector system. The incident radiation consisted of CuK_α X-rays from a rotating anode

generator operating at 80–90 mA and 50–60 kV. Individual data collection frames were exposed for 20–25 minutes. Crystals were oscillated through a ϕ angle of 1.0° for each frame. For each of these three mutant proteins only one crystal was used to collect diffraction data. Processing of area detector diffraction data sets was done as described in Section 2.2. All diffraction data sets were put on an absolute scale using the Wilson plot statistical method (Wilson, 1942). Data collection and processing statistics are shown in Table 5.17.

5.1.3 Refinement and analyses

Inspection of mutant protein minus reduced wild-type yeast iso-1-cytochrome *c* difference electron density maps revealed that structural changes were restricted to those regions directly around mutation sites. Therefore, least-squares restrained parameter refinement (Hendrickson & Konnert, 1981) was initiated using as a starting model the reduced wild-type yeast iso-1-cytochrome *c* structure in which mutated residues were represented by alanines. Included in each model was a selection of well-determined water molecules and a sulfate ion. Excluded from starting models was Wat166 which is located adjacent to the side-chains of residues 52 and 67. Refinement of mutant structures utilized structure factors with a resolution greater than 6.0 Å and that had a $F/\sigma(F)$ ratio ≥ 2.0. During the course of refinement the remaining side-chain atoms for residues 52 and 67 were unambiguously located and added to the refinement models. In addition, based on $F_o - F_c$ difference and fragment maps, and $2F_o - F_c$ and $3F_o - 2F_c$ difference electron density maps, the side-chains of some surface residues were repositioned and water molecules were added or deleted according to how well they could be resolved. Final refinement statistics are given in Table 5.18.

The final refined structures of the reduced and oxidized yeast iso-1-cytochrome *c* N52I and N52I-Y67F mutants display good stereochemistry (see Table 5.18). The coordinate error in these structures was assessed by inspection of Luzzati plots (see Figure 5.38) and by the method of Cruickshank (Luzzati, 1952; Cruickshank, 1985; see also Appendix C). In Luzzati plots the overall r.m.s. coordinate errors for the mutant structures was estimated to be ~0.20–0.22 Å,

Table 5.18: Refinement results and stereochemistry for the final models of the yeast iso-1-cytochrome *c* N52I and N52I-Y67F mutants in both oxidation states

	N52I		N52I-Y67F	
	Reduced	Oxidized	Reduced	Oxidized
I. Refinement results				
R-factor (%)	18.5	18.2	17.9	18.1
No. of reflections in refinement	4142	4898	5406	4538
No. of atoms	956	944	950	937
No. of solvent molecules	59	46	53	40
II. Stereochemistry of final models				
	r.m.s. deviation from ideal values			
Bond distances (Å)				
1-2 bond distance	0.017	0.021	0.022	0.022
1-3 bond distance	0.044	0.046	0.044	0.041
1-4 bond distance	0.062	0.060	0.061	0.058
Planar restraints (Å)	0.017	0.018	0.017	0.017
Chiral volume ($Å^3$)	0.217	0.219	0.215	0.235
Non-bonded contacts† (Å)				
single-torsion	0.226	0.218	0.219	0.213
multi-torsion	0.224	0.179	0.181	0.182
possible hydrogen bonds	0.233	0.230	0.212	0.225
Torsion angles (°)				
planar (0° or 180°)	2.6	2.9	2.7	2.5
staggered (±60°,180°)	23.7	24.0	22.3	21.5

† The r.m.s. deviations from ideality for this class of restraint incorporates a reduction of 0.2 Å from the radius of each atom involved in a contact.

while values of ~0.15 Å were obtained for these structures with Cruickshank's method.

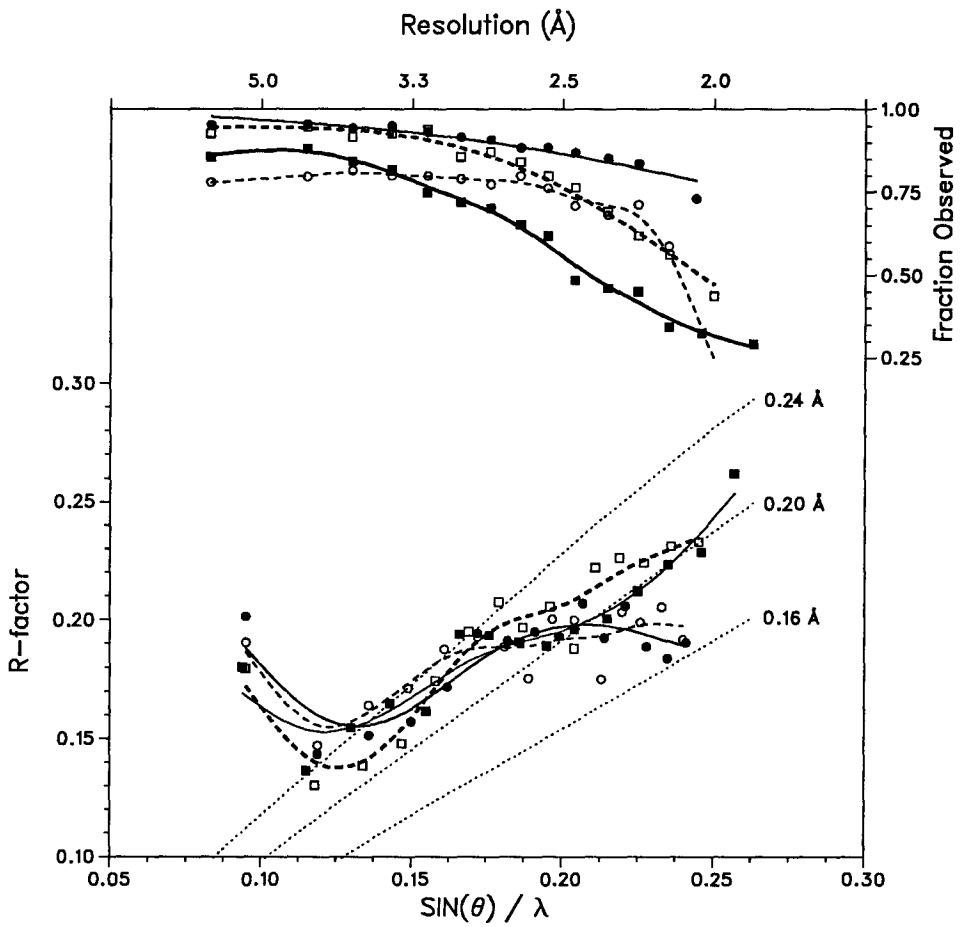

Figure 5.38: Plots of the dependence on resolution of both the R-factor agreement between calculated and observed structure factors (axis lower left), and the fraction of data used (axis at top right), for the final refined structures of reduced (■) and oxidized (□) N52I, and reduced (●) and oxidized (○) N52I-Y67F mutant yeast iso-1-cytochromes *c*. For this analysis reciprocal space was divided into shells according to $sin(\theta)/\lambda$, with each containing at least 230 reflections. For the purpose of assessing the accuracy of the atomic coordinates of the mutant structures, curves representing the theoretical dependence of R-factor on resolution assuming various levels of r.m.s. error in the atomic positions of the model (Luzzati, 1952) are also drawn (dashed lines). This analysis suggests an overall r.m.s. coordinate error for the four mutant structures of between ~0.20 and 0.22 Å.

5.2 Results

5.2.1 Polypeptide chain conformation

Both the reduced and oxidized structures of the wild-type (Louie & Brayer, 1990; Chapter 3) and Y67F mutant (Chapter 4) yeast iso-1-cytochromes *c* have been determined. To allow for a comprehensive analysis of the effects of mutations in the N52I and N52I-Y67F proteins, all the available coordinate sets were superimposed using a least-squares procedure, based on all main-chain and heme atoms. This comparison reveals that the polypeptide chain fold is unaffected by these position 52 and 67 mutations (see Figure 5.39). As indicated in Table 5.19 average deviations for main-chain atoms between the reduced and oxidized N52I and N52I-Y67F structures, and the reduced wild-type protein are of the order of 0.17–0.25 Å, which is slightly less than that observed for the reduced and oxidized Y67F proteins (0.29 and 0.32 Å, respectively). The only region where substantially different conformations of the polypeptide chain are observed is at the N-terminus (see Figure 5.40). However, the first three residues of the

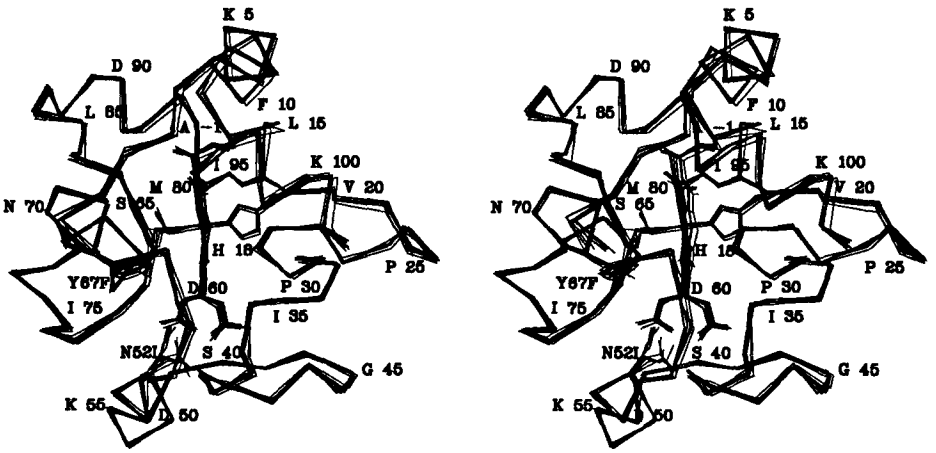

Figure 5.39: A stereo-drawing of the α-carbon backbones of the reduced and oxidized structures of wild-type, N52I, Y67F and N52I-Y67F yeast iso-1-cytochromes *c*. The heme group has also been drawn as well as the two heme ligands, His18 and Met80, and cysteines 14 and 17, which form thioether linkages to the heme porphyrin ring. In addition the side-chain atoms of residues 52 and 67, which are replaced in the mutant proteins, are also displayed. For clarity, every 5th residue has been labelled with its one-letter amino acid designation and sequence number.

Table 5.19: Average deviations of polypeptide chain atoms in the N52I and N52I-Y67F mutants with respect to reduced wild-type yeast iso-1-cytochrome *c*

	N52I		N52I-Y67F	
	Reduced	Oxidized	Reduced	Oxidized
Average deviations (Å)				
main-chain atoms	0.25	0.20	0.17	0.20
buried side-chain atoms †	0.30	0.27	0.24	0.24
exposed side-chain atoms	0.65	0.65	0.52	0.51

† A side-chain was considered buried if less than 20% of its surface area was solvent exposed relative to the unfolded state in any one of the four mutant protein structures (Shrake & Rupley, 1973; Perry *et al.*, 1990).

polypeptide chain (Thr(-5), Glu(-4) and Phe(-3)) are disordered in electron density maps and the differing conformations observed reflect alternate fits to the same poor electron density. In no other instances are shifts of main-chain atoms more than twice the overall average deviation observed. Given this result it is not surprising that all of the main-chain to main-chain hydrogen bonds present in the reduced yeast iso-1-cytochrome *c* wild-type structure are also observed in the four mutant structures (for a complete list of these hydrogen bonds see Table 8 in Louie & Brayer (1990)).

Examination of side-chain conformations reveals that all buried residues have average deviations less than 0.8 Å, except for Leu9 (see Figure 5.40). In the reduced and oxidized N52I and N52I-Y67F structures the side-chain of Leu9 has average deviations of 0.5 Å, 1.0 Å, 0.8 Å and 1.0 Å, respectively. This positional flexibility in Leu9 side-chain placement is also reflected in the multiple conformations observed in the reduced wild-type yeast iso-1-cytochrome *c* structure (Louie, 1990). Overall, average deviations for buried side-chains are comparable to those observed for main-chain atoms (Table 5.19). As evident from Figure 5.40 many solvent exposed side-chains display a variety of different conformations among the various structures, indicative of their high degree of mobility. This is not only reflected in the overall average deviations

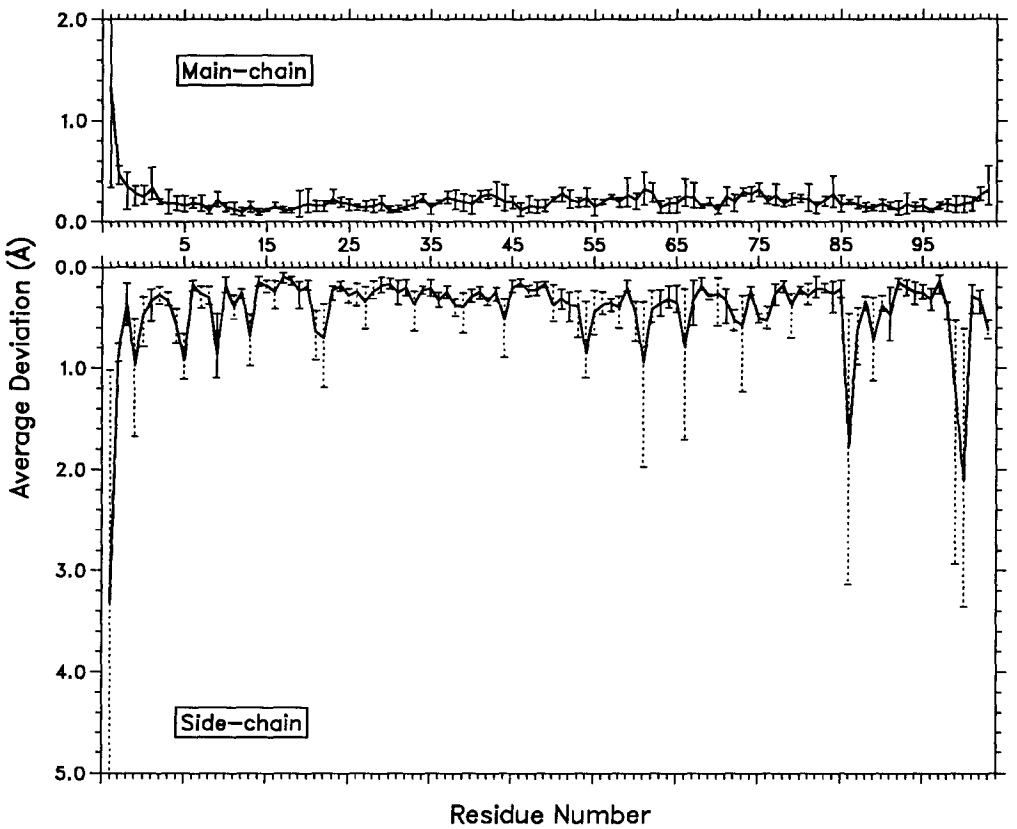

Figure 5.40: A plot of the overall average positional deviations of the reduced and oxidized N52I and N52I-Y67F mutant proteins when compared to the reduced yeast iso-1-cytochrome *c* wild-type structure. In the top frame overall average deviations for main-chain atoms are shown, while in the bottom frame the overall average deviations for side-chain atoms are displayed. The vertical bars represent the range of individual pairwise average deviations. In the bottom frame, dotted vertical bars refer to solvent exposed residues.

observed, which are approximately double that for buried side-chain atoms (see Table 5.19), but also in the higher thermal factors for this group of atoms.

It has been proposed that differences in polypeptide chain flexibility are an essential element in the structural differences between oxidation states of cytochrome *c* (see Section 3.3). Analysis of the thermal factors of main-chain atoms for the reduced and oxidized N52I and N52I-Y67F mutants using a difference matrix method (Section 3.2) provides the following results (see Figure 5.41). For all four mutant proteins the overall thermal factor profile is remarkably similar to that of the reduced wild-type protein. However, two differences appear to be significant. In the

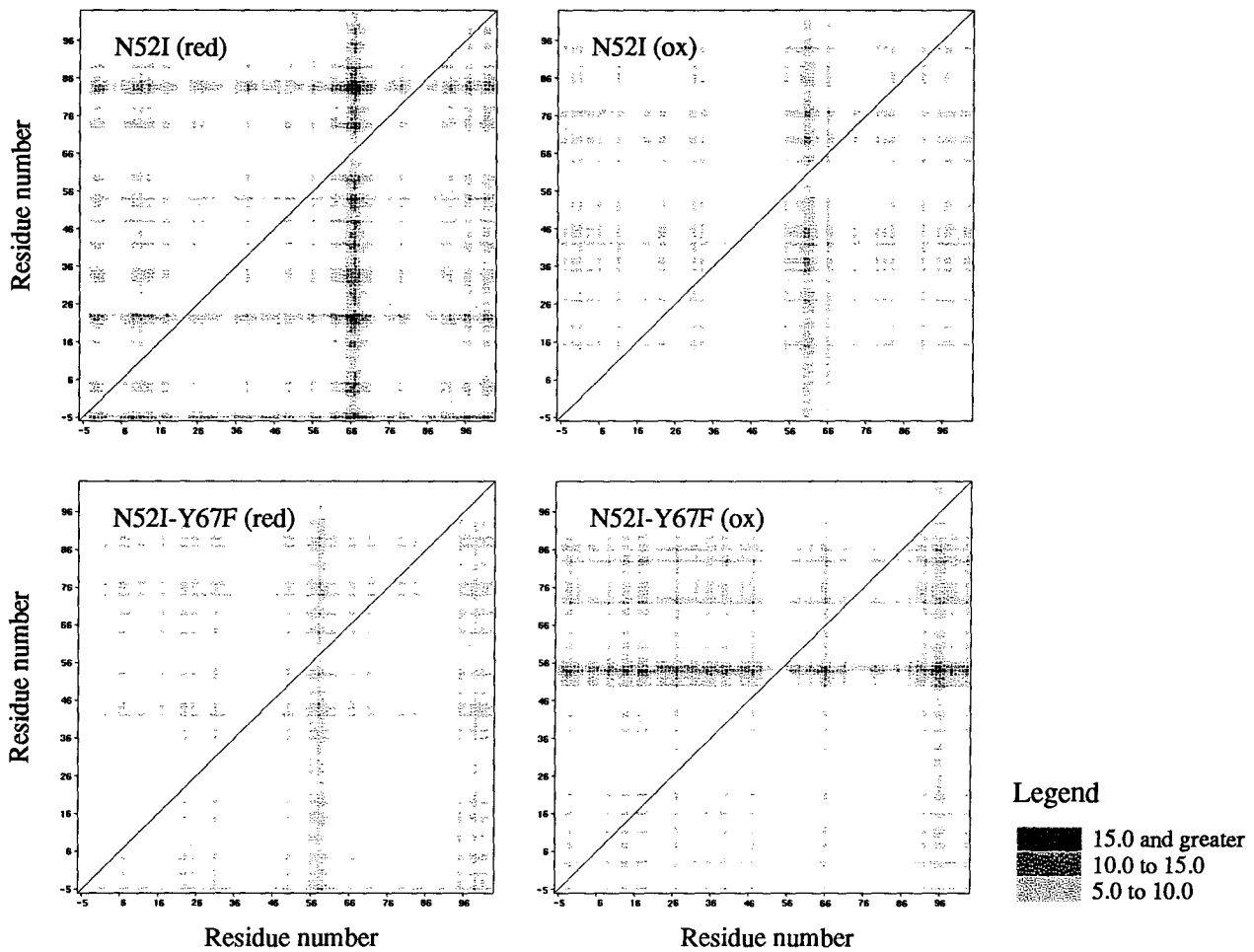

Figure 5.41: A matrix representation of the differences in average main-chain thermal factors between the reduced and oxidized N52I and N52I-Y67F mutant proteins, and the reduced wild-type protein. Each matrix point $P_{x,y}$ represents an amino acid pairing (x, y) and was calculated using the equation: $P_{x,y} = (B_x - B_y)_{\text{mutant}} - (B_x - B_y)_{\text{wild-type}}$ where B is the average main-chain thermal factor of a given amino acid. Positive matrix values are displayed as squares of different levels of blackness according to the scale on the right. As this matrix has inverse symmetry across the diagonal line drawn, negative values are redundant and are omitted for clarity. The advantage of this approach is that displayed values are not affected by differences in overall thermal factor between the two structures. Within the matrix, amino acids which have significantly higher average main-chain thermal factors in the mutant structures produce vertical streaks. These include residues 65–68 in the reduced N52I protein with maximal differences observed for Tyr67. Amino acids producing horizontal streaks indicate the presence of significantly lower thermal factors for their main-chain atoms in the mutant protein. An example of this are residues 54–55 in the oxidized N52I-Y67F variant.

reduced N52I mutant protein a marked increase in main-chain thermal factors is observed for residues 65–68. This effect is focussed at Tyr67 where the average increase is ~+10 Å^2. A corresponding increase in side-chain thermal factors is also observed (average increase ~+10 Å^2). The second difference is seen in the oxidized N52I-Y67F mutant protein. Here residues 54–55, focussed at Lys54, show a significant drop in thermal factors (average decrease ~-10 Å^2). This decrease in flexibility is not only restricted to main-chain atoms but is also observed for the side-chains of these amino acids.

5.2.2 Heme structure

In iso-1-cytochrome *c* the porphyrin ring of the heme group is found to be distorted from planarity into a saddle shape. As evident in Table 5.20, the degree of distortion is oxidation state dependent and most pronounced in the oxidized form. The N52I-Y67F protein shows a similar trend, although overall heme plane distortion is smaller in the oxidized state. However, both the N52I and Y67F mutant proteins in either the reduced or oxidized state show distortions similar to those seen in the oxidized wild-type protein.

Inspection of heme iron coordinate bond lengths between the different reduced and oxidized structures show these to be similar within expected atomic coordinate errors. However, the orientation of the His18 imidazole plane does appear to be a function of oxidation state (see Table 5.20). For the reduced wild-type and mutant cytochromes *c* the average angle between the imidazole plane normal and a line drawn through the NA and NC heme atoms is ~46°. In the oxidized state the His18 side-chain rotates towards the A and C pyrrole rings, resulting in values in the range of 50–55° for this angle. In contrast, the Met80 ligand does not exhibit oxidation state dependent shifts and virtually identical conformations are observed in the wild-type, N52I and N52I-Y67F cytochrome *c* structures (average deviations 0.2–0.3 Å).

Both heme propionates form extensive hydrogen bond interactions with protein groups (see Table 5.21). Between the different reduced and oxidized structures, the hydrogen bonds formed by the pyrrole ring D propionate show particularly high conservation. Nonetheless, this

Table 5.20: Heme conformation and ligand geometry in wild-type, N52I, Y67F and N52I-Y67F mutant yeast iso-1-cytochromes *c*

	Reduced				Oxidized			
Pyrrole	Wild-type	N52I	Y67F	N52I-Y67F	Wild-type	N52I	Y67F	N52I-Y67F
I. Angular deviations between the pyrrole nitrogen plane normal and the four pyrrole ring plane normals (°)								
A	9.4	12.2	13.7	10.7	12.6	13.4	13.4	8.9
B	11.1	15.4	9.1	8.9	14.1	10.3	10.7	11.6
C	8.8	10.2	11.7	9.4	9.6	9.2	12.0	11.6
D	8.1	11.3	11.3	7.5	12.8	10.7	10.2	8.5
Average	9.3	12.7	11.5	9.1	12.3	10.9	11.6	10.2
II. Angular deviations between the porphyrin ring plane normal and the four pyrrole ring plane normals, plus the pyrrole nitrogen plane normal (°)								
A	6.7	8.0	9.0	7.0	8.3	9.2	7.4	5.4
B	11.9	14.3	12.3	10.5	12.4	11.8	12.4	12.8
C	9.8	13.6	11.6	9.5	13.8	10.2	10.8	11.2
D	6.0	8.0	6.4	4.2	10.0	7.2	4.4	5.0
Average	8.6	11.0	9.8	7.8	11.1	9.6	8.8	8.6
NNNN	2.6	4.2	4.9	3.7	4.3	4.3	6.0	3.8
III. Angles between the normal of the His18 imidazole plane and a vector drawn through the NA and NC heme atoms (°)								
	46.7	45.4	44.8	46.9	55.8	49.7	54.6	49.7

Each pyrrole ring is defined by nine atoms, which include the five ring atoms plus the first carbon atom bonded to each ring carbon. The porphyrin ring plane is defined by the five atoms in each of the four pyrrole rings, the four bridging methine carbons, the first carbon atom of each of the eight side-chains and the heme iron (33 atoms in total). The pyrrole nitrogen plane is defined by only the 4 pyrrole nitrogens (see Figure 1.2 for heme atom labeling convention).

propionate group does have more than a single conformation. For example, when comparing the C2D–C3D–CAD–CBD torsion angle (102° in the reduced wild-type protein), a range of values are observed, the largest and smallest coming from the reduced N52I and N52I-Y67F structures (113° and 76°, respectively). However, this torsion angle variability does not affect the position of the two pyrrole D propionate oxygen atoms which in the reduced wild-type and the four mutant structures are all located within 0.3 Å of each other. This is due to smaller

Table 5.21: Heme propionate hydrogen bond interactions in wild-type, N52I, Y67F and N52I-Y67F mutant yeast iso-1-cytochromes *c*

Interaction	Distances (Å)							
	Reduced				Oxidized			
	Wild-type	N52I	Y67F	N52I-Y67F	Wild-type	N52I	Y67F	N52I-Y67F
O1A Tyr48 OH	2.83	2.63	2.97	2.53	2.83	2.67	3.51	2.79
Wat121	2.81	3.47	2.57	2.83	2.85	3.20	2.69	2.45
Wat168	2.85	3.50	3.41	3.00	2.87	3.31	(3.62)	2.91
O2A Gly41 N	3.21	3.07	2.82	3.25	2.60	3.33	2.88	3.34
Asn52 ND2	3.34	-	2.67	-	(3.54)	-	3.41	-
Trp59 NE1	3.09	2.84	3.21	3.05	3.43	2.80	2.89	2.86
Wat121	(4.01)	(3.90)	(3.74)	(3.68)	3.34	(3.89)	3.51	(3.71)
O1D Thr49 OG1	2.64	2.50	2.60	2.68	2.79	2.73	2.71	2.52
Thr78 OG1	2.90	2.75	2.75	3.03	3.07	3.03	3.08	2.95
Lys79 N	3.17	3.03	(3.39)	3.26	2.67	3.10	2.96	3.12
O2D Thr49 N	2.94	2.79	2.77	2.97	2.75	3.17	2.95	2.95

Interactions were accepted as hydrogen bonds only if they met all of the following criteria: a H···A distance < 2.6 Å, a D—H···A angle > 120°, and a C—A···H angle > 90°. Values given in brackets are not considered to be hydrogen bonds by this criteria, but are listed for comparison.

compensatory angular changes in other pyrrole D propionate torsional angles.

In the reduced and oxidized N52I and N52I-Y67F mutant structures the conformation of the pyrrole A propionate is essentially the same as that observed in the reduced wild-type structure (see Figure 5.42). The largest difference is an ~15° rotation of the carboxyl group of the pyrrole A propionate around the CBA–CGA bond in both mutant proteins. This causes the O1A oxygen atom to move towards Tyr48 OH and the O2A oxygen atom to move towards Trp59 NE2. This conformational readjustment appears to occur in response to the absence of the Asn52 side-chain in these mutants. In reduced wild-type and Y67F mutant yeast iso-1-cytochromes *c* Asn52 is positioned adjacent to the heme pyrrole A propionate to which it forms

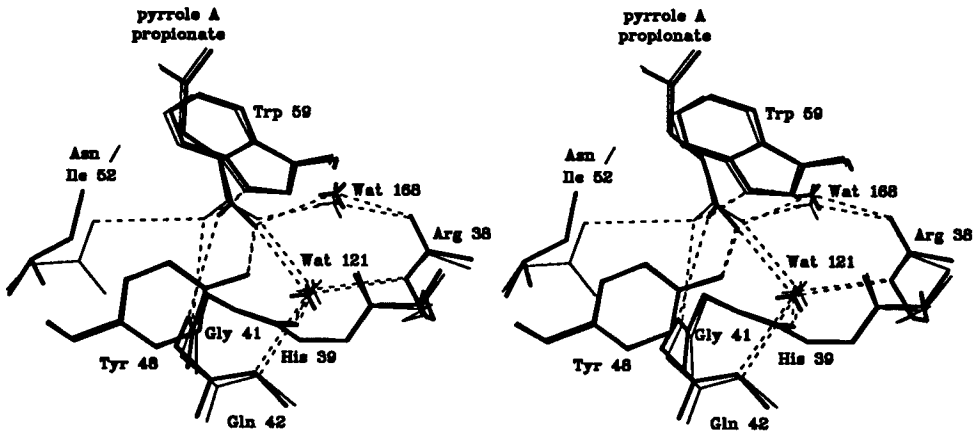

Figure 5.42: A stereo-drawing showing the region around the pyrrole ring A propionate in reduced N52I-Y67F (thick lines) and wild-type (thin lines) yeast iso-1-cytochromes *c*. Mutation of residue 52 from an asparagine to an isoleucine results in removal of a hydrogen bond (dashed lines) to this propionate group and as a consequence it rotates ~15°. Similar rotations of the pyrrole A propionate group are also seen in the oxidized state of the N52I-Y67F mutant, as well as in the reduced and oxidized states of the N52I mutant protein.

a hydrogen bond interaction (see Chapters 3 and 4). Notably, the rearrangement of pyrrole A propionate conformation and hydrogen bond interactions observed in the wild-type and Y67F proteins upon a change in oxidation state (Chapters 3 and 4) does not occur in the N52I and N52I-Y67F variants.

5.2.3 Mutation site region

In reduced wild-type yeast iso-1-cytochrome *c* the side-chains of Asn52, Tyr67 and Thr78 form hydrogen bonds to a structurally conserved internal water molecule (Wat166; see Figure 5.43a). This water molecule shifts and reorients its dipole moment upon oxidation so as to stabilize the positive charge on the heme iron atom (Chapter 3). Previous studies have shown that mutation of Tyr67 to a phenylalanine results in an additional water molecule being introduced into this internal water cavity (Chapter 4; Figure 5.43c). In contrast, our current studies show that substitution of Asn52 for an isoleucine results in the displacement of Wat166 (see Figure 5.43b). In the reduced wild-type protein the size of the cavity occupied by Wat166 is ~10 Å^3. As a consequence of the Ile52 substitution this cavity is absent in both the reduced

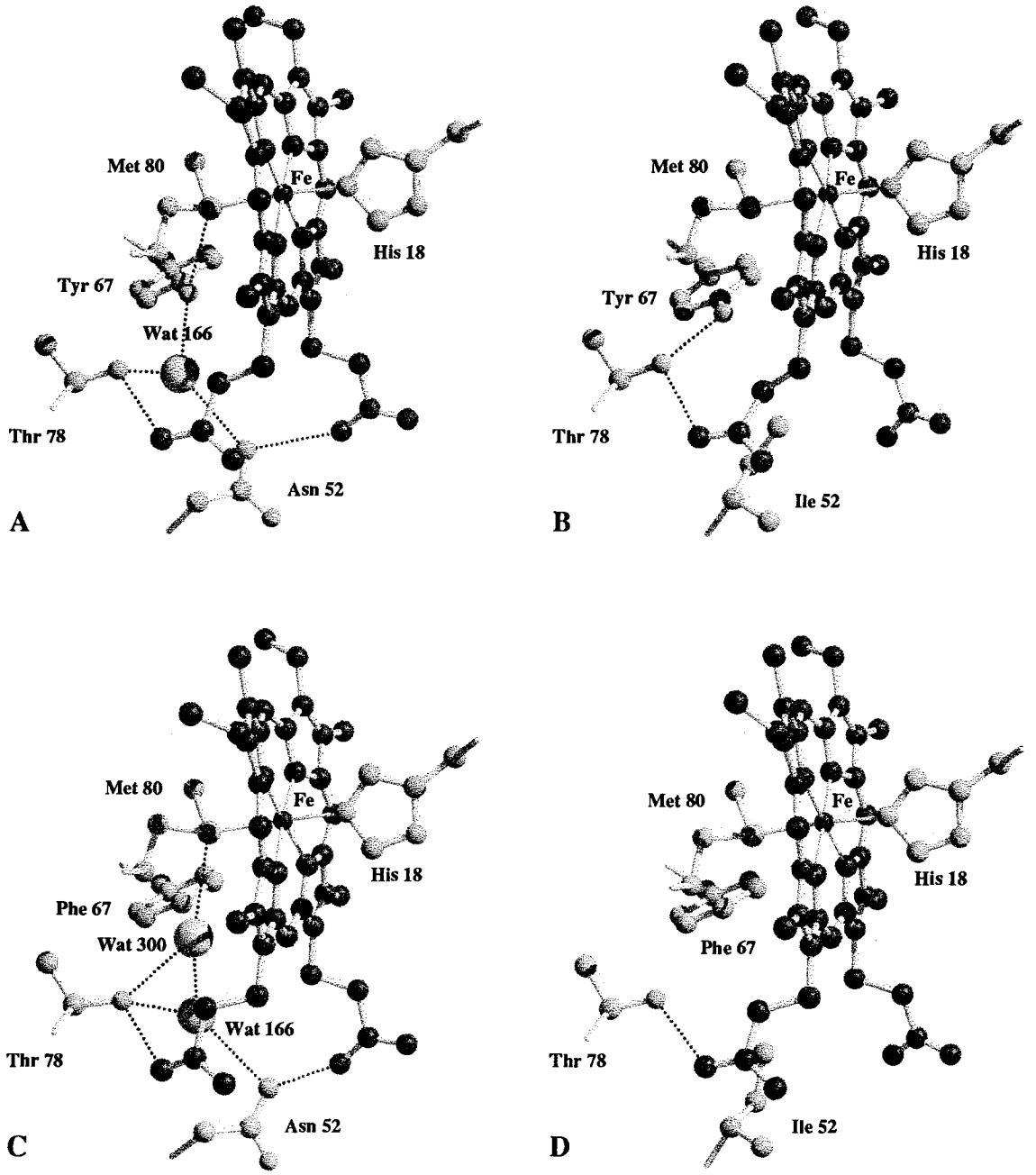

Figure 5.43: Drawings showing the region about residues 52 and 67 in (a) wild-type, (b) N52I, (c) Y67F and (d) N52I-Y67F yeast iso-1-cytochromes *c* in the reduced state. In each case the heme group is shown with dark shaded balls and internal water molecules are represented by larger spheres. Heme ligand interactions are indicated by thin white bonds whereas hydrogen bonds are shown by thin dashed lines.

and oxidized states of the N52I variant thereby excluding a water molecule at this location.

The absence of Wat166 in the N52I mutant protein disrupts the normal hydrogen bond network in this region of the protein. One result of this is that in the reduced N52I protein the side-chain of Tyr67 shifts (average deviation 0.6 Å) towards the side-chain of Thr78 to which its hydroxyl group forms a hydrogen bond (distance 3.4 Å). In the oxidized N52I structure a similar situation is seen. Here the side-chain of Tyr67 has a more modest shift (average deviation 0.3 Å) and according to our normal criteria the interaction between the Tyr67 OH and Thr78 OG1 is too long (3.6 Å) to be considered a hydrogen bond. However, in both oxidation states of the N52I mutant, as a result of the Tyr67 side-chain shift, the hydrogen bond between Tyr67 OH and Met80 SD is absent. The loss of this hydrogen bond is further demonstrated in the reduced N52I protein by an increase in thermal factors for Tyr67.

In the N52I-Y67F reduced and oxidized structures even fewer hydrogen bonds are formed (see Figure 5.43d). In both these structures the size of the internal cavity originally occupied by Wat166 in wild-type yeast iso-1-cytochrome *c* has been reduced to ~8 $Å^3$. This is too small a volume to allow for water binding and as a consequence Wat166 is also excluded from the structure of the N52I-Y67F mutant protein. Like the N52I variant this is the result of the positioning of the side-chain of isoleucine 52. No significant positional shifts or rearrangements of hydrogen bonds are seen to compensate for the substituted amino acids.

5.3 Discussion

5.3.1 Structural effects of mutations

A summary of the differences observed in the reduced and oxidized N52I and N52I-Y67F mutant structures compared to their respective wild-type structures is given in Table 5.22. It is clear from this comparison that the mutations at position 52 and 67 primarily affect the conformation of two regions of the protein, these being the mutation site and the pyrrole A propionate regions. Differences are not only expressed in terms of atom shifts and changes in hydrogen bonding, but also in the altered flexibilities of polypeptide chain segments that make up these two regions of

Table 5.22: Structural differences observed in the N52I and N52I-Y67F mutant structures when compared to wild-type yeast iso-1-cytochrome *c*

N52I	N52I-Y67F
A. *Thermal factors for main-chain atoms (see Figure 5.41)*	
1. Both reduced and oxidized mutants display a similar thermal factor profile as seen in the reduced wild-type protein with the exception of increased flexibility for residues 65–68 (maximal at residue 67) in the reduced protein	1. Both reduced and oxidized mutants display a similar thermal factor profile as seen for reduced wild-type protein, except for decreased flexibility of residues 54–55 (focused at Lys54) in the oxidized variant
B. *Heme structure (see Figure 5.42 and Tables 5.20 and 5.21)*	
1. Distortion of the heme plane in the reduced and oxidized state are similar to that observed in the oxidized wild-type structure	1. Distortion of the heme plane is less pronounced in oxidized state
2. Pyrrole A propionate hydrogen bond network does not display oxidation state dependent conformational differences	2. Pyrrole A propionate hydrogen bond network does not display oxidation state dependent conformational differences
3. Absence of asparagine side-chain at position 52 causes small rotation of propionate A carboxyl group	3. Absence of asparagine side-chain at position 52 causes small rotation of propionate A carboxyl group
C. *Mutation site region (see Figure 5.43).*	
1. Displacement of Wat166 from internal cavity	1. Displacement of Wat166 from internal cavity
2. Loss of Tyr67 OH to Met80 SD hydrogen bond interaction	2. Loss of Tyr67 OH to Met80 SD hydrogen bond interaction
3. New hydrogen bond interaction formed between Tyr67 and Thr78 in the reduced state	

the protein.

Mutation site region

Comparisons of mutation site regions show that the hydrogen bond network about Wat166 is altered in both mutant proteins (see Figure 5.43). In the N52I mutant the hydrogen bond between Tyr67 OH and Met80 SD is severed, Wat166 is displaced, a new hydrogen bond is formed between Tyr67 OH and Thr78 OG1, and the link from Asn52 ND2 to the pyrrole A propionate is lost. In the N52I-Y67F variant only the hydrogen bond between Thr78 OG1 and the heme pyrrole D propionate remains of the original hydrogen bond network in this region.

Surprisingly, the breakdown of the hydrogen bond network about the internal water molecule, Wat166, does not adversely affect the stability of cytochrome *c*, rather the opposite appears to be the case. A dramatic increase of thermodynamic stability for the N52I yeast iso-1-cytochrome *c* mutant has been reported (Das *et al.*, 1989; Hickey *et al.*, 1991). Also, this mutant is less sensitive to alkaline pH. This can be seen in the oxidized wild-type protein where the Met80 SD - heme iron ligand bond is broken at pH 8.5 (Pearce *et al.*, 1989) whereas this alkaline transition occurs at pH 10.0 (J.G. Guillemette, private communication) when Asn52 is mutated to an isoleucine. Resistance to this alkaline transition has been taken as a measure of protein stability (Saigo, 1986). The N52I-Y67F mutant protein has not been as extensively studied, but measurements of its alkaline transition indicates this occurs at pH 11.0 (J.G. Guillemette, private communication) suggesting even greater stability than that found for the N52I mutant. These stability studies (all performed with the oxidized protein) correlate well with observations that neither of these mutant proteins display an increase in polypeptide chain flexibility in the oxidized state as is observed in the wild-type protein (see Figure 5.41). Indeed, in the N52I mutant residues 65–68 show a decrease in flexibility when this protein is in the oxidized state, and for the N52I-Y67F variant residues 54–55 appear to be less mobile under these conditions. It is noteworthy that both these polypeptide chain segments display an increase in flexibility in the oxidized wild-type structure (Chapter 3).

A similar situation is observed for the Y67F mutant cytochrome *c*, which is more resistant to high temperature, pH and denaturation by urea than the wild-type protein (Luntz *et al.*,

1989). The Y67F variant also has a higher alkaline transition (pH 10.1; J.G. Guillemette, private communication) that is similar to that of the N52I mutant protein. Furthermore, structural studies of the Y67F mutant protein have shown a breakdown of the hydrogen bond network involving Wat166, particularly in the oxidized state (Chapter 4). This mutant protein also appears to be more rigid in the oxidized state as inferred from its thermal factor profile (see Figure 4.35).

Taken together, these observations suggest that the hydrogen bond network involving Wat166 destabilizes the tertiary structure of cytochrome *c*, since alteration of this network or removal of Wat166 altogether, leads to considerable structural stabilization. It also appears that one function of Wat166 and the interactions it forms is to specifically increase the flexibility of three nearby segments of polypeptide chain in the oxidized state. This feature is abolished in all three of the N52I, N52I-Y67F and Y67F variant proteins, suggesting much of the observed increase in protein stability may be due to this factor.

Pyrrole A propionate region

The region about the pyrrole A propionate has been previously shown to be sensitive to oxidation state (Chapter 3). In wild-type yeast iso-1-cytochrome *c* this group displays two distinctly different conformations depending on the valence state of the heme iron atom (see Figure 3.25). In the Y67F variant nearly identical conformational changes are observed (see Figure 4.36). Other mutant cytochromes *c* are observed to have the pyrrole A propionate conformation found in the oxidized wild-type protein even though they are in the reduced state (Louie *et al.*, 1988b; Louie & Brayer, 1989).

The origin of the driving force behind oxidation state dependent conformational changes about the pyrrole A propionate remains unclear. Several factors such as, differential delocalization of the negative charge on the heme propionate group (Barlow & Thornton, 1983; Singh *et al.*, 1987); electrostatic interactions between this propionate, the heme iron atom and the side-chain of Arg38 (Moore, 1983); and oxidation state dependent distortions of the heme porphyrin

ring (Chapter 3) have all been suggested to affect the pyrrole A propionate conformation.

Neither the N52I nor the N52I-Y67F mutant proteins have oxidation state dependent conformational rearrangements about the pyrrole A propionate. This absence may provide some insight into the factors behind this rearrangement in the wild-type protein. In the first instance, the charge on the heme iron does not appear to be a factor since this conformational change is not elicited between oxidation states in these mutant proteins. Secondly, changes about the pyrrole A propionate are not a consequence of increased heme plane distortion since, for example, both N52I and Y67F display comparable distortions in both oxidation states while they do not exhibit similar pyrrole A propionate conformations (Tables 5.20 and 5.21). These observations clearly suggest that the chemical character of residue 52 is critical in mediating the changes about the pyrrole A propionate group between oxidation states and could be driven by the change in conformation that the Asn52 experiences as the result of the oxidation state dependent movement of Wat166 (see Figure 3.28).

5.3.2 Importance of the hydrogen bond network around Wat166

Control of midpoint reduction potential

Both the N52I and N52I-Y67F variant proteins have midpoint reduction potentials which are ~56 mV lower than that of wild-type yeast iso-1-cytochrome *c* (see Tables 1.4 and 1.5; Burrows *et al.*, 1991; J.G. Guillemette, private communication). In this regard these mutant proteins are similar to the Y67F variant (Chapter 4). In this latter case the observed lower reduction potential can be attributed to the breakage of the hydrogen bond from the hydroxyl group of Tyr67 to the side-chain of Met80 which forms a heme ligand interaction. In the N52I mutant protein the Tyr67 OH is oriented towards Thr78 (see Figure 5.43) strongly suggesting that the absence of a hydrogen bond interaction to Met80 SD is also responsible for the lower midpoint reduction potential of this mutant protein. Langen *et al.* (1992) have proposed that the removal of the Asn52 side-chain and its associated dipole might be a factor in the observed drop in midpoint reduction potential. However, similarities in midpoint reduction potential between

the Y67F and N52I mutant proteins suggest this is unlikely to be a major contributor to this phenomena (see Table 1.5). For the N52I-Y67F protein a similar situation exists. That is, the interaction to the Met80 side-chain has been deleted leading to a drop in midpoint reduction potential comparable to the Y67F and N52I variants.

Beyond demonstrating the effects of the Tyr67 OH - Met80 SD interaction on the midpoint reduction potential of cytochrome *c*, our collective results clearly delineate the role Wat166 has in this area. As has been shown, this internal water molecule appears to play an integral part in helping to stabilize the alternative oxidation states of this protein. In addition, it is clear that Wat166 is essential to the maintenance of the hydrogen bond network in this area. In particular, the N52I mutant protein shows the presence of Wat166 provides for the proper hydrogen bond interactions between Tyr67 and Thr78. Its absence leads to a disruption in the hydrogen bond network of which it is the central feature, the shift of Tyr67 to interact with Thr78 and a subsequent drop in midpoint reduction potential. Thus these studies also show that Wat166 is a critical determinant in maintaining the value of the midpoint reduction potential of cytochrome *c*.

Modulation of flexible regions

Besides regulation of midpoint reduction potential another possible role of Wat 166 and the hydrogen bond network that it forms an integral part of, is in the modulation of flexible regions in cytochrome *c* between oxidation states. In wild-type yeast iso-1-cytochrome *c* Wat166 shifts and reorients upon heme iron oxidation so that its dipole moment is positioned to stabilize the positive charge residing on the heme iron atom. During this process a rearrangement of the hydrogen bond network about Wat166 takes place and several hydrogen bonds are broken, resulting in residues 47–59, 65–72 and 81–85 exhibiting increased mobility (see Chapter 3). The more open and flexible nature of the oxidized form of cytochrome *c* is a property well documented using a variety of techniques ranging from NMR (Williams *et al.*, 1985a) to small angle X-ray scattering (Trewhella *et al.*, 1988) and infrared spectroscopy (Dong *et al.*, 1992;

see also Section 1.3).

The N52I and N52I-Y67F mutations result in proteins which do not exhibit flexible regions in the oxidized state. The same is also observed for the Y67F variant (Chapter 4). As mentioned above, examination of the three dimensional structures of these three mutants reveals that the hydrogen bond network involving Wat166 is substantially altered, particularly when these proteins are in the oxidized state. These observations strongly suggest that Wat166 is a central component in the mechanism which mediates the oxidation state dependent mobility of selected polypeptide chain segments.

The biological role of oxidation state dependent polypeptide chain flexibility in cytochrome *c* remains unclear, but several theories have been suggested. For example, it has been proposed that this factor is important for complexation and dissociation with redox partners (Bosshard & Zurrer, 1980; Rackovsky & Goldstein, 1984; Zhang *et al.*, 1990; Dong *et al.*, 1992). This is plausible since the determination of the oxidized structure of wild-type yeast iso-1-cytochrome *c* has shown that mobility differences are localized in the area implicated to be the binding site with redox partners (see Chapter 3). These flexibility differences might also be important for the kinetics of biological electron transfer, either through changing the effective heme solvent accessible surface area (Schlauder & Kassner, 1979; Zheng *et al.*, 1990), or through modulation of the reorganization energy (Churg & Warshel, 1983; Warshel, 1983; Marcus & Sutin, 1985; Williams *et al.*, 1985b). Another possibility is that the more open and flexible nature of the oxidized state is a means to regulate the thermodynamic properties of the midpoint reduction potential (Watt & Sturtevan, 1969).

To summarize, our studies suggest that the function of the internal water molecule, Wat166, and its associated hydrogen bond network is three fold. First, the presence of Wat166 provides for a convenient mechanism to modify the hydrogen bond network involving several key residues near the Met80 ligand, depending on the oxidation state of the heme. This appears to be particularly important for modulating the hydrogen bond between Tyr67 OH and Met80 SD. Associated with this is the fact that the presence of Wat166 is necessary to maintain the spatial and

hydrogen bonding relationships between residues in this area and in this way is also an important element in setting the value of the midpoint reduction potential of cytochrome *c*. Finally, Wat166 also appears to mediate oxidation state dependent mobility differences of polypeptide chain segments which might play a role in redox partner recognition (Chapter 3).

Chapter 6

Importance of the Dipole Orientation of a Conserved Internal Water Molecule in the Biological Function of Cytochrome *c* as Revealed by the Mutations N52A and I75M

6.1 Experimental Procedures

6.1.1 Crystallization

Protein samples for these studies were obtained from J.G. Guillemette (University of British Columbia). Crystals of the N52A mutant of yeast iso-1-cytochrome *c* were initially grown in a 0.1 M sodium phosphate buffer (pH 6.4) containing 90% saturated ammonium sulfate and 60 mM DTT (to maintain the reduced state of the protein), using a combination of the hanging drop method and a hair-seeding technique (Leung *et al.*, 1989; see also Section 2.1). However, these crystals were of insufficient size for high resolution x-ray diffraction analyses and additional macro-seeding had to be employed to increase their dimensions. The I75M mutant protein was crystallized in a 0.1 M sodium phosphate buffer (pH 5.3) containing 95% saturated ammonium sulfate and 70 mM sodium dithionite using similar methods. Prior to mounting, mutant crystals were transferred to fresh mother-liquor, with the exception that for the I75M protein 40 mM DTT was used instead of sodium dithionite. Mutant protein crystals proved to be isomorphous to those of wild-type yeast iso-1-cytochrome *c*, having the same space group ($P4_32_12$) and similar unit cell dimensions, $a = b = 36.52$ Å, $c = 137.89$ Å, and $a = b = 36.55$ Å, $c = 138.59$ Å for the N52A and I75M crystals, respectively. Wild-type yeast iso-1-cytochrome *c* crystals have unit cell dimensions of $a = b = 36.46$ Å, $c = 136.86$ Å.

6.1.2 Data collection and data processing

Diffraction data for each mutant protein was collected from one crystal on an Enraf-Nonius CAD4-F11 diffractometer. The radiation used for the diffraction experiment was nickel filtered and generated from a copper target x-ray tube operating at 26 mA and 40 kV. Reflections were measured using continuous Ω scans. The ambient temperature during data collection was maintained at 15°C. To monitor slippage and decay, six reflections were measured periodically. In this manner, a total of 10779 reflections were collected to 2.0 Å resolution for the N52A protein, and 10308 reflections to 1.9 Å resolution for the I75M protein. Diffraction intensities were corrected for background, absorption, decay, Lorentz and polarization effects as previously described (see Section 2.2 and 3.1). The resultant structure factors were put on an absolute scale using a Wilson (1942) plot.

6.1.3 Refinement and analyses

Mutant minus wild-type difference Fourier maps revealed that in the N52A and I75M proteins structural changes were restricted to the direct vicinity of each mutation site. The starting models for least-squares restrained parameter refinement (Hendrickson & Konnert, 1981) for both mutant proteins was the wild-type yeast iso-1-cytochrome *c* structure (Louie & Brayer, 1990) in which mutated residues were replaced by alanines. A subset of well determined solvent molecules was also included in the starting models, with the notable exception of the internal water molecule, Wat166, which is located adjacent to the side-chains of both residues 52 and 75. Several manual interventions were carried out during the course of refinement based on inspection of $F_o - F_c$, $2F_o - F_c$ and $3F_o - 2F_c$ difference electron density maps. These involved adjustments to solvent structure and disordered side-chains in both structures, as well as the addition of the remaining atoms to the side-chain of Met75 in the I75M structure.

As shown in Table 6.23 the final refined models for the N52A and I75M yeast iso-1-cytochrome *c* mutants have good stereochemistry. The final standard crystallographic R-factors for these structures are 18.5% and 22.1%, respectively. The higher value obtained for the

Table 6.23: Final stereochemistry for N52I and I75M yeast iso-1-cytochromes *c*

Stereochemical refinement parameters	r.m.s. deviation from ideal values	
	N52A	I75M
Bond distances (Å)		
1-2 bond distance	0.020	0.024
1-3 bond distance	0.044	0.050
1-4 bond distance	0.052	0.064
Planar restraints (Å)	0.016	0.018
Chiral volume ($Å^3$)	0.196	0.286
Non-bonded contacts†(Å)		
single-torsion	0.219	0.236
multi-torsion	0.207	0.242
possible hydrogen bonds	0.245	0.262
Torsion angles (°)		
planar (0°or 180°)	2.6	2.6
staggered (±60°,180°)	25.9	26.3

†The r.m.s. deviations from ideality for this class of restraint incorporates a reduction of 0.2 Å from the radius of each atom involved in a contact.

I75M mutant reflects the smaller crystals available for this protein and the lower quality of the resultant diffraction data (Figure 6.44). The accuracy of the two structures was assessed by inspection of a Luzzati plot (Luzzati, 1952; Appendix C; see Figure 6.44) and suggests an r.m.s. coordinate error of ~0.20 Å for the N52A structure and ~0.24 Å for the I75M structure. A separate estimate for the overall coordinate error can be obtained using the method of Cruickshank (1949, 1954, 1985; see also Appendix C) and gave an r.m.s. coordinate error of 0.18 Å and 0.25 Å for the N52A and I75M mutant structures, respectively. This is in good agreement with Luzzati's method.

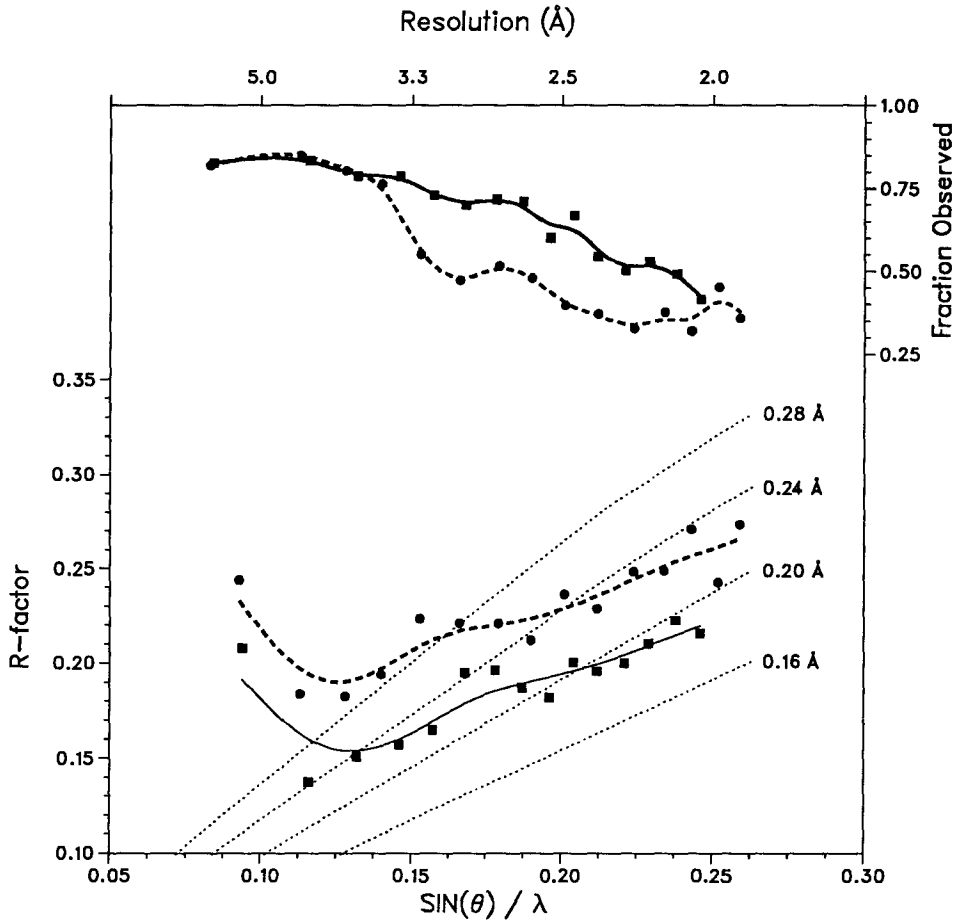

Figure 6.44: Plots of the dependence on resolution of both the R-factor agreement between calculated and observed structure factors (axis at bottom left) and the fraction of data used (axis at top right), for the final refined structures of the N52A (■ and solid lines) and I75M (● and dashed lines) mutant yeast iso-1-cytochromes *c*. For this analysis reciprocal space was divided into shells according to $sin(\theta)/\lambda$, with each containing at least 280 reflections. To assess the accuracy of the atomic coordinates of the mutant structures, curves representing the theoretical dependence of R-factor on resolution assuming various levels of r.m.s. error in the atomic positions of the model (Luzzati, 1952) are also drawn (dashed lines). This analysis suggests an overall r.m.s. coordinate error of ~0.20 Å for the N52A structure and ~0.24 Å for the I75M coordinates.

6.2 Results

6.2.1 Polypeptide chain conformation

As can be seen in Figure 6.45 the mutations N52A and I75M in yeast iso-1-cytochrome *c* do not dramatically affect the overall polypeptide chain fold, even though these residues are buried within the protein and completely inaccessible to solvent. Overall average deviations for all main-chain atoms when compared to wild-type yeast iso-1-cytochrome *c* are 0.24 Å and 0.31 Å for the N52A and I75M structures, respectively (Figure 6.46). These values are comparable to those observed for other yeast iso-1-cytochrome *c* variants (Chapters 4 and 5). However, two regions do display shifts greater than two times the overall average deviation. The first region is at the N-terminal end of the polypeptide chain. As previously discussed these residues are disordered in the wild-type protein and all variant structures determined thus far. Therefore, the differing conformations observed likely represent only alternate fits to the same poorly defined electron density (Chapters 3, 4 and 5). A second region of conformational shifts is focussed at Gly41. In the N52A structure the main-chain atoms of Gly41 and Gln42 have

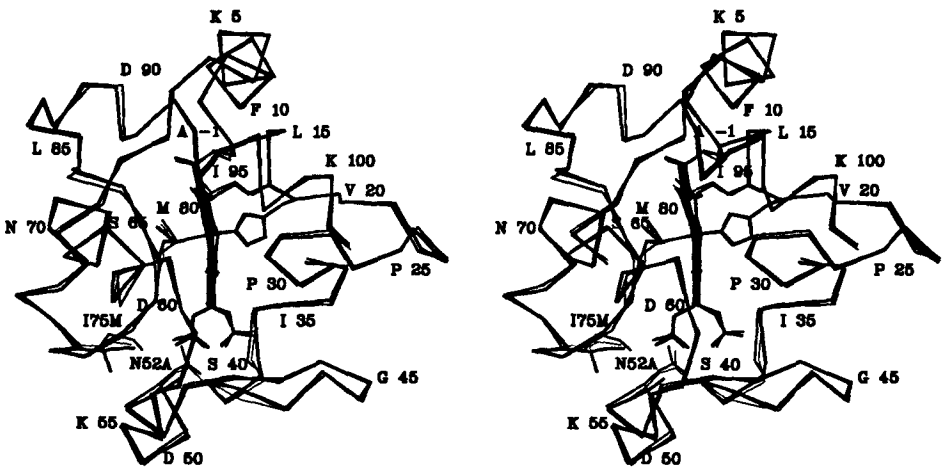

Figure 6.45: A stereo-drawing of the α-carbon backbones of the wild-type (thin lines), N52A (thick lines) and I75M (medium lines) yeast iso-1-cytochromes *c*. The individual heme groups have also been drawn along with the two heme ligands, His18 and Met80, and cysteines 14 and 17, which form thioether linkages to the heme porphyrin ring. In addition, the side-chain atoms of residues 52 and 75 are drawn. Every 5th residue of the wild-type protein has been labelled with its one-letter amino acid designation and sequence number.

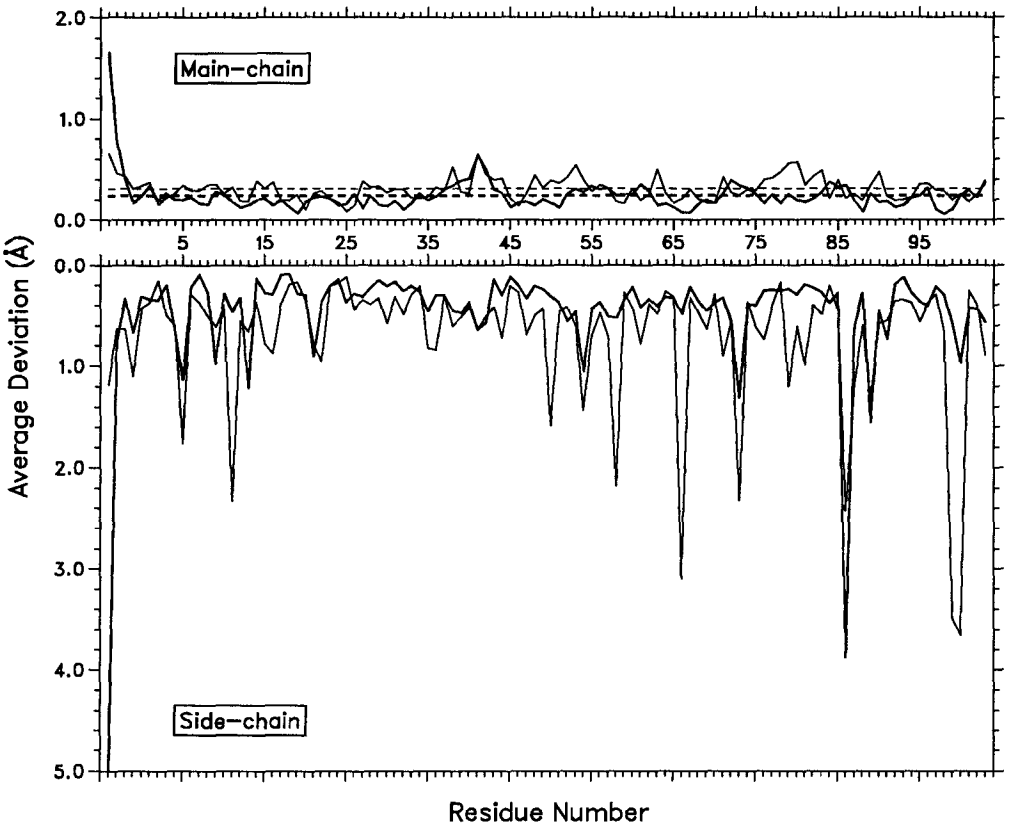

Figure 6.46: A plot of the average positional deviations between the N52A (thick lines) and I75M (thin lines) variants and the wild-type structure of yeast iso-1-cytochrome *c*. The top frame shows the observed deviations for main-chain atoms, while the bottom frame displays the deviations observed for side-chain atoms. The dashed lines in the top frame represent the overall average deviations of all main-chain atoms for the two mutants which are 0.24 Å and 0.31 Å for the N52A and I75M structures, respectively.

average deviations of 0.6 Å and 0.5 Å, respectively. Interestingly, these same two residues also display nearly identical shifts in the I75M structure.

Examination of side-chain conformations reveal that the 55 residues whose side-chains are buried inside the protein matrix have virtually identical conformations in both the mutant and wild-type structures; average deviations for these side-chain atoms are 0.32 Å and 0.43 Å for the N52A and I75M proteins, respectively. The remaining 53 solvent exposed side-chains are observed to have conformations which vary considerably, reflecting their positioning on the protein surface and their high degree of mobility at this location. The average deviations for

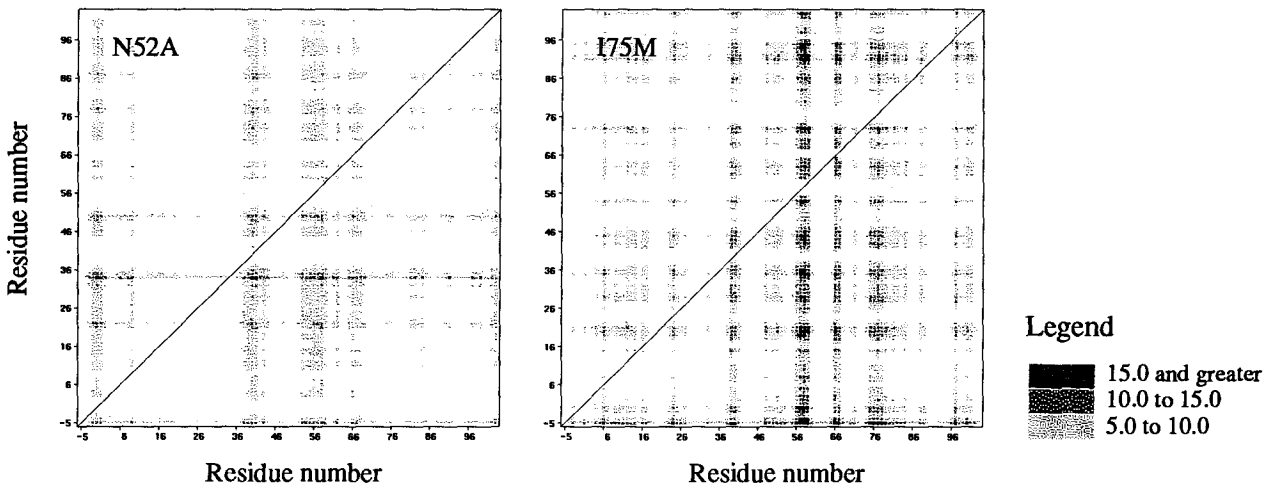

Figure 6.47: A matrix representation of the differences in average main-chain thermal factors between the reduced N52A and I75M mutant proteins, and wild-type yeast iso-1-cytochrome *c*. Each matrix point $P_{x,y}$ represents an amino acid pairing (x, y) and was calculated using the equation: $P_{x,y} = (B_x - B_y)_{\text{mutant}} - (B_x - B_y)_{\text{wild-type}}$ where B is the average main-chain thermal factor of a given amino acid. Positive matrix values are displayed as squares of different levels of blackness according to the scale on the right. As this matrix has inverse symmetry across the diagonal line drawn, negative values are redundant and are omitted for clarity. The advantage of this approach is that displayed values are not affected by differences in overall thermal factor between the two structures. Within the matrix, amino acids which have significantly higher average main-chain thermal factors in the mutant structures produce vertical streaks. These include residues 57–59 and 66–67 in the I75M protein. Amino acids producing horizontal streaks indicate the presence of significantly lower thermal factors for their main-chain atoms in the mutant protein. No residues are observed in either mutant protein to fall in this latter category.

this class of side-chain atoms are 0.64 Å and 0.97 Å for the N52A and I75M mutant proteins, respectively. This high degree of mobility is also reflected in the thermal factors for the solvent exposed side-chains which have average values of 21.4 $Å^2$ and 22.5 $Å^2$ for the N52A and I75M variants, respectively. This can be compared to average thermal factor values of 14.9 $Å^2$ and 16.0 $Å^2$ for the buried side-chain groups in these two mutant proteins.

An analysis of the differences in polypeptide chain flexibility between reduced wild-type yeast iso-1-cytochrome *c* and the N52A and I75M mutant structures is shown in Figure 6.47. Although, overall the mutant and wild-type structures have a similar thermal factor profile a few differences are observed in the I75M mutant structure. These include residues 57–59 which

display an increase of 13 $Å^2$ for main-chain atoms (average thermal factor of 36 $Å^2$ versus 23 $Å^2$ in the wild-type protein). Corresponding increases in side-chain thermal factors are observed as well. A smaller difference involves residues 66-67 which display an increase of 11 $Å^2$ for main-chain atoms (average thermal factor of 23 $Å^2$ versus 12 $Å^2$ in the wild-type protein). Here again, comparable increases in side-chain thermal factors are observed. Comparison of the N52A mutant protein and wild-type yeast iso-1-cytochrome *c* does not indicate the presence of any significant differences in polypeptide chain flexibility.

6.2.2 Heme structure

The heme group of wild-type yeast iso-1-cytochrome *c* is substantially distorted from planarity in both oxidation states (see Tables 6.24 and Table 3.8). The degree of heme distortion is more pronounced for the two reduced N52A and I75M mutant structures compared to the reduced wild-type protein, with average angular deviations of the four pyrrole rings from the pyrrole nitrogen plane being 3–4° higher. This increased distortion is similar to that observed for the N52I and Y67F variants (Chapters 4 and 5), as well as for the oxidized state of the wild-type protein (Chapter 3).

Examination of heme coordinate bond lengths reveal that in the two mutant proteins these bonds are the same within expected atomic coordinate errors. The orientation of the His18 imidazole plane, which is sensitive to the oxidation state of the protein (Chapter 5), is similar in both the reduced N52A mutant and wild-type protein. The angle between the imidazole plane normal and a line drawn through the NA and NC atoms is 48.7° and 46.7°, respectively. However, in the reduced I75M mutant structure this angle is 54.5° and therefore more like that found in the oxidized wild-type protein structure where it is 55.8°. An analysis of the heme solvent accessibility of the two mutant proteins shows this to be comparable to wild-type yeast iso-1-cytochrome *c* (see Table 6.25).

Table 6.24: Heme conformation and ligand geometry in wild-type, N52A and I75M yeast iso-1-cytochromes *c*

Pyrrole	Wild-type	N52A	I75M
I. Angular deviations between the pyrrole nitrogen plane normal and the four pyrrole ring normals (°)			
A	9.4	11.5	16.3
B	11.1	14.0	10.1
C	8.8	8.6	15.7
D	8.1	13.2	12.2
Average	9.3	11.8	13.6
II. Angular deviations between the porphyrin ring plane normal and the four pyrrole ring plane normals, plus the pyrrole nitrogen plane normal (°)			
A	6.7	7.4	10.4
B	11.9	12.8	14.8
C	9.8	12.6	13.6
D	6.0	9.8	6.3
Average	8.6	10.7	11.3
NNNN	2.6	4.3	6.0
III. Average deviations of porphyrin ring atoms from the least squares porphyrin ring plane (Å)			
	0.178	0.200	0.223

Each pyrrole ring is defined by nine atoms, which include the five ring atoms plus the first carbon atom bonded to each ring carbon. The porphyrin ring plane is defined by the five atoms in each of the four pyrrole rings, the four bridging methine carbons, the first carbon atom of each of the eight side-chains and the heme iron (33 atoms in total). The pyrrole nitrogen plane is defined by only the 4 pyrrole nitrogens (see Figure 1.2 for heme atom labeling convention).

6.2.3 The N52A mutation site

In wild-type yeast iso-1-cytochrome *c*, Asn52 is part of two elaborate hydrogen bond networks, one involving Tyr67, Thr78, Met80 and an internal water molecule Wat166, and the second involving the pyrrole A propionate group, Gly41 and Trp59 (see Figures 6.48 and 6.49). Mutation of Asn52 to an alanine residue results in the enlargement of the cavity occupied by Wat166 from ~10 Å^3 to ~70 Å^3. This increase in volume is more than sufficient to allow for a second

Table 6.25: Heme solvent accessibility for wild-type, N52A and I75M yeast iso-1-cytochromes $c^{\dagger}$

	Wild-type	N52A	I75M
1. Solvent accessible heme atoms and surface area exposed (Å^2)			
CHD	2.6	2.4	2.3
CMC	9.4	10.4	9.3
CAC	3.3	4.2	3.8
CBC	17.4	20.6	17.6
CMD	9.8	10.5	8.2
2. Total heme exposure (Å^2)	42.5	48.1	41.2
3. Total heme surface (Å^2)	495.7	504.3	503.0
4. Heme surface area exposed (%)	8.6	9.5	8.2

† Computations were done using the method of Connolly (1983) and the results represent the accessible molecular surfaces of the atoms listed. The probe sphere used had a radius of 1.4 Å.

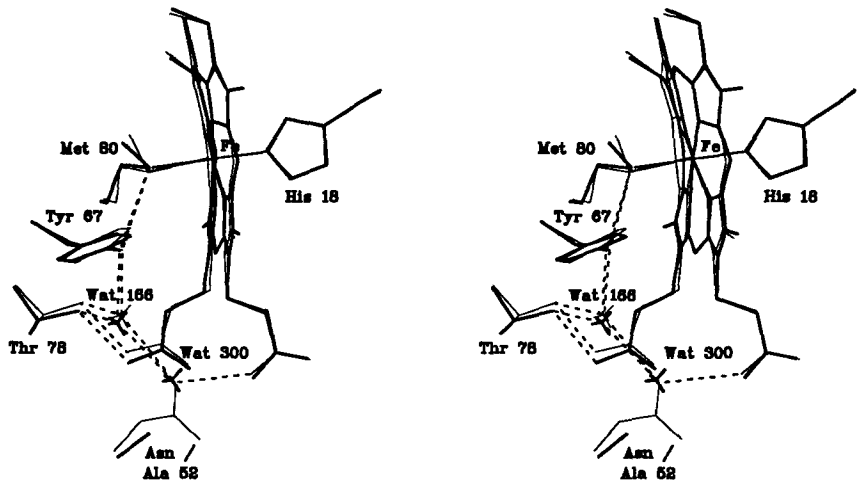

Figure 6.48: Stereo-drawing showing the region about residue 52 in the N52A (thick lines) and wild-type (thin lines) yeast iso-1-cytochromes *c*. This mutation substantially increases the volume of an internal water cavity leading to the inclusion of an additional water molecule (Wat300). The hydrogen bonds (dashed lines) formed by Wat300 are similar to those formed by Asn52 ND2 in the wild-type structure.

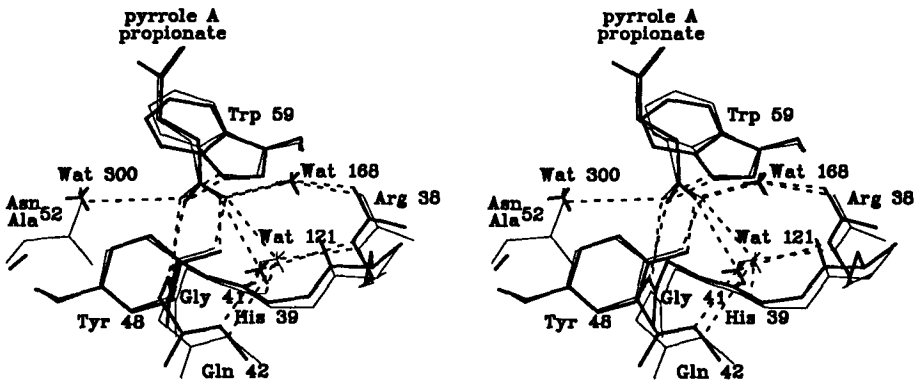

Figure 6.49: A stereo-drawing showing the region around the pyrrole A propionate group in the N52A (thick lines) and wild-type (thin lines) yeast iso-1-cytochromes *c*. Hydrogen bonds are indicated by dashed lines. In the N52A structure Wat300 is located in roughly the same position as the Asn52 ND2 group in wild-type yeast iso-1-cytochrome *c*. However, unlike Asn52 ND2, Wat300 does not form a hydrogen bond to the pyrrole A propionate according to the standard criteria used herein (see Table 6.26).

water molecule to be positioned in this region and an additional water molecule (Wat300) is indeed observed in electron density maps. The location of Wat166 in the N52A structure is virtually identical to that found in wild-type yeast iso-1-cytochrome *c* while the position of Wat300 corresponds roughly to that of the Asn52 ND2 group in the wild-type protein (displacement of 0.3 Å; see Figure 6.48). Both Wat166 and Wat300 have relatively high thermal factors (~50 $Å^2$) indicating a fair degree of mobility in their positioning.

The effect of this mutation on the hydrogen bond network about Wat166 is subtle (see Figure 6.48). Wat300 forms a similar interaction with Wat166 in the N52A mutant structure as did Asn52 ND2 in wild-type yeast iso-1-cytochrome *c*. Other hydrogen bonds formed by Wat166 do not appear to be significantly perturbed by the presence of Wat300. However, unlike the amide group of Asn52, Wat300 could function either as a hydrogen bond donor or acceptor to Wat166. This implies that the dipole orientation of Wat166 is likely to be less clearly defined in the mutant structure compared to that of the wild-type protein (Chapter 3).

The effect of the N52A mutation on the hydrogen bond network about the pyrrole A propionate is illustrated in Figure 6.49. In the wild-type protein Asn52 ND2 forms a hydrogen

Table 6.26: Heme propionate hydrogen bond interactions in wild-type, N52A and I75M yeast iso-1-cytochromes *c*

Hydrogen bond partners		Distances (Å)†		
		Wild-type	N52A	I75M
O1A	Tyr48 OH	2.83	2.89	2.82
	Wat121	2.81	2.97	3.37
	Wat168	2.85	3.24	3.03
O2A	Gly41 N	3.21	2.69	2.63
	Asn52 ND2 (Wat300)‡	3.34	(3.56)	(4.15)
	Trp59 NE1	3.09	2.82	3.04
O1D	Thr49 OG1	2.64	2.44	2.79
	Thr78 OG1	2.90	3.16	2.97
	Lys79 N	3.17	3.14	2.90
O2D	Thr49 N	2.94	2.74	2.55

† Interactions were accepted as hydrogen bonds only if they met all of the following criteria: a H···A distance < 2.6 Å, a D—H···A angle > 120°, and a C—A···H angle > 90°. Values given in brackets are not considered to be hydrogen bonds by this criteria, but are listed for comparison.
‡ In the N52A mutant structure Asn52 ND2 is replaced by Wat300.

bond to the O2A oxygen atom of this propionate. In the N52A variant Wat300 is positioned further from this group (3.56 Å). This factor, coupled with its high thermal factor suggests that the interaction of Wat300 with the O2A oxygen atom of the pyrrole A propionate is weak. (Table 6.26). A consequence of this weaker interaction appears to be a stronger interaction between Gly41 N and the O2A heme propionate oxygen atom. Movement of Gly41 (average deviation 0.6 Å; see Figure 6.46) shortens the hydrogen bond between these two groups by 0.5 Å to an overall length of 2.7 Å. A similar situation is observed in the oxidized state of the wild-type protein where the Asn52 ND2 interaction to the heme is also lost (Chapter 3).

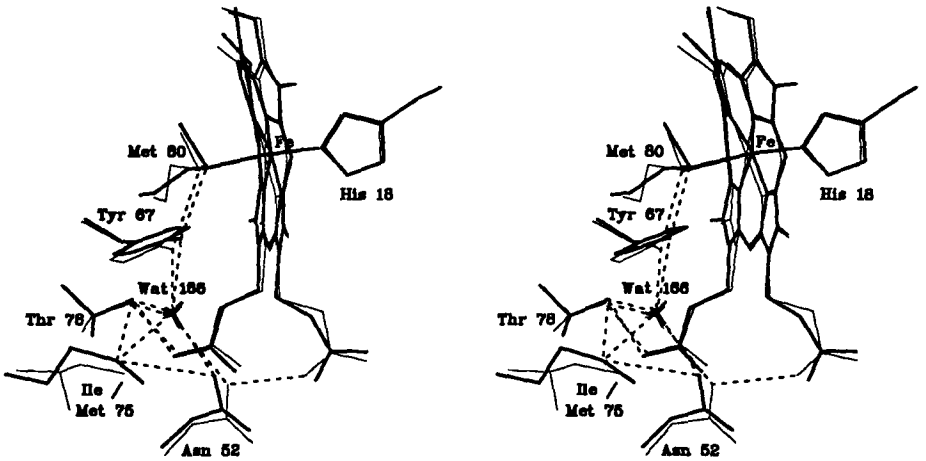

Figure 6.50: Stereo-drawing showing the region about residue 75 in I75M (thick lines) and wild-type (thin lines) yeast iso-1-cytochromes *c*. The presence of the Met75 SD sulfur atom adds another group to the hydrogen bond network (dashed lines) around Wat166.

6.2.4 The I75M mutation site

In wild-type yeast iso-1-cytochrome *c* the side-chain of Ile75 forms part of one side of the internal cavity occupied by Wat166. Substitution of this residue for a methionine effectively adds another group to the hydrogen bond network around Wat166 (see Figure 6.50). This permits new interactions to be made between Met75 SD and Wat166, Asn52 ND2, as well as Thr78 OG1. Thus, one effect of this substitution is that the effective dipole orientation of Wat166 is likely modified.

One group positionally affected by these new interactions is the side-chain of Asn52 which shifts 0.6 Å towards the Met75 sulfur atom. This shift breaks the hydrogen bond link between Asn52 ND2 and the nearby pyrrole A propionate, causing a conformational change in this latter group (see Table 6.26 and Figure 6.51). In comparison to the reduced wild-type protein, the pyrrole A propionate torsion angles C2A–CAA, CAA–CBA and CBA–CGA in the I75M variant are rotated by ~25°, ~20° and ~20°, respectively. A similar alteration is also seen in the oxidized form of wild-type yeast iso-1-cytochrome *c* where these three torsion angles change by ~20°, ~30° and ~45°, respectively (Chapter 3). Another consequence of the loss of the interaction between Asn52 ND2 and the pyrrole A propionate is strengthening of the

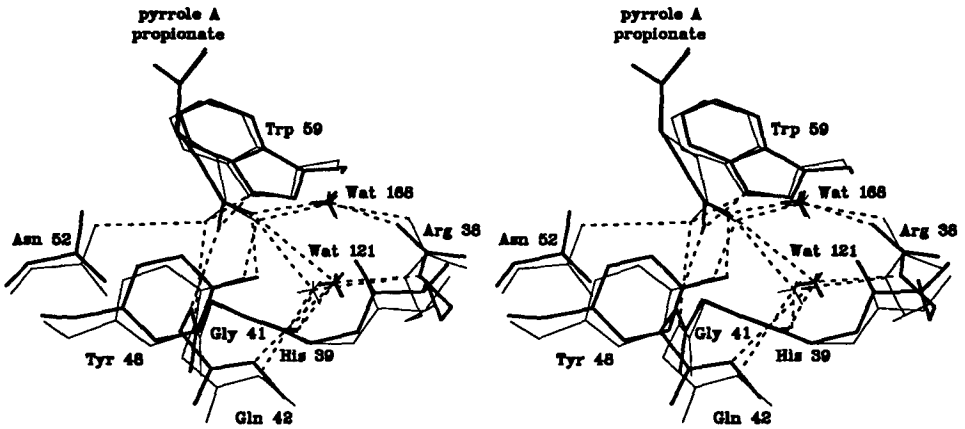

Figure 6.51: A stereo-drawing showing the region around the pyrrole A propionate in I75M (thick lines) and wild-type (thin lines) yeast iso-1-cytochrome *c*. Hydrogen bonds are represented by dashed lines. In the pyrrole A propionate region the I75M mutant structure shows similarities to that of the oxidized form of the wild-type protein in that the hydrogen bond between the propionate group and Asn52 is lost and the propionate group has undergone a conformational change.

interaction between Gly41 N and the O2A heme propionate oxygen atom similar to that in the N52A variant and oxidized wild-type structures (Chapter 3). As part of this process residue 41 moves 0.6 Å towards the propionate group decreasing the distance between them from 3.2 Å to 2.6 Å (Table 6.26).

6.3 Discussion

6.3.1 Plasticity of the pyrrole A propionate region

A summary of the structural differences observed between the N52A and I75M mutant structures and wild-type yeast iso-1-cytochrome *c* is given in Table 6.27. From this overview it can be seen that most of the structural changes present are localized around the pyrrole A propionate group. Previous structural studies of wild-type and mutant forms of yeast iso-1-cytochrome *c* have also shown structural differences in this area. For example, replacement of residues 67 (Chapter 4; Figure 4.36) and 82 (Louie *et al.*, 1988b; Louie & Brayer, 1989) result in such conformational changes, as does going to the oxidized state in the case of the wild-type

Table 6.27: Structural differences observed in the reduced N52A and I75M mutant structures when compared to reduced wild-type yeast iso-1-cytochrome *c*

N52A	I75M
A. *Positional displacements of polypeptide chain (see Figure 6.46)*	
1. Movement of Gly41 towards the pyrrole A propionate	1. Movement of Gly41 towards the pyrrole A propionate
B. *Thermal factor parameters of main-chain atoms (see Figure 6.47)*	
1. No major differences	1. Increased thermal factors for residues 57–59 and 66-67
C. *Heme structure (see Figures 6.49 and 6.51 and Tables 6.24 and 6.26)*	
1. Increased distortion of the heme plane	1. Increased distortion of the heme plane
2. Small readjustments in the pyrrole A propionate hydrogen bond network	2. Rotation of the His18 imidazole plane comparable to that of the oxidized wild-type structure
	3. Pyrrole A propionate conformational changes similar to those of the oxidized wild-type structure
D. *Mutation site (see Figures 6.48 and 6.50).*	
1. Additional water molecule (Wat300) takes the place of original Asn52 ND2	1. Addition of an extra hydrogen bonding group to the Wat166 hydrogen bond network
2. Link lost to pyrrole A propionate	2. Link lost to pyrrole A propionate
E. *Hydrogen bond interactions (see Figures 6.48, 6.49, 6.50 and 6.51 and Table 6.26).*	
Stronger: Gly41 N - Heme O2A	Stronger: Gly41 N - Heme O2A
Lost: Asn52 ND2 - Heme O2A	Weaker: Wat121 - Heme O1A
Asn52 ND2 - Wat166	Lost: Asn52 ND2 - Heme O2A
New: Wat300 - Wat166	New: Met75 SD - Asn52 ND2
	Met75 SD - Thr78 OG1
	Met75 SD - Wat166

protein (see Figure 3.25). It is of interest that mutation of Asn52 to an isoleucine inhibits the conformational changes occurring upon heme oxidation and locks the hydrogen bond network around the pyrrole A propionate in the reduced state (see Figure 5.42 and Table 5.21).

The N52A mutant studied here displays some of the same differences as seen in the oxidized

wild-type structure. The hydrogen bond between the heme O2A oxygen atom and residue 52 is lost and a strengthening of the hydrogen bond between this oxygen atom and Gly41 N is observed (see Figure 6.49 and Table 6.26). Similar conformational differences are observed in the I75M mutant structure, with the additional feature of increased thermal motion for Trp59. As observed in the oxidized wild-type structure, altered torsional angles for the pyrrole A propionate group are also present in the I75M mutant structure (see Figure 6.51).

Apparently the pyrrole A propionate region can adopt a number of stable conformations. Which of these conformations has the lowest energy, and thus predominates, would appear to depend on a number of factors. One factor may be the polarity of the heme environment (Louie *et al.*, 1988; Chapter 4). Mutant structures in the reduced state such as F82S, F82G and to a lesser degree Y67F, which increase the hydrophilicity of the environment around the heme, exhibit pyrrole A propionate conformations which are a hybrid of those found in the reduced and oxidized wild-type structures. Further, the oxidized wild-type structure is also an example of the effect a change in polarity of the heme environment has on the pyrrole A propionate region. On the other hand, mutations which create a more hydrophobic heme environment such as in the N52I and N52I-Y67F proteins result in conformations of the pyrrole A propionate which are very similar to that seen in the wild-type reduced state, even when the heme iron atoms of these proteins are oxidized (Chapter 5). Collectively these studies show pyrrole A propionate conformation could potentially be used as a rough barometer of the polarity of the heme environment.

The structural mechanism linking the polarity of heme environment and pyrrole A propionate conformation could potentially be electrostatic or mechanical. Louie *et al.*, (1988) proposed that a change in the polarity of the heme environment might cause a redistribution of the electrons within the delocalized π-electron system of the heme (see also Section 4.3; Wallace *et al.*, 1989). This in turn would alter the electrostatic interaction between the porphyrin ring and the propionate carboxyl group and result in a conformational change in the latter group as has been suggested by Moore (1983). A more indirect explanation is also possible. Differences

in the polarity of the heme environment might be reflected in different degrees of distortion from planarity of the heme porphyrin ring. The need to accommodate heme plane distortion within the surrounding protein matrix could necessitate a change in pyrrole A propionate conformation. Our studies show that a change in pyrrole A propionate conformation is always accompanied by an increase in the distortion of the heme plane (see Tables 3.8, 4.12, 5.20 and 6.24).

Although polarity of the heme environment appears to be an important factor in determining the exact conformation of the pyrrole A propionate group and associated hydrogen bond network, the N52A and I75M mutant structures presented here show that this is clearly not the only factor. Inspection of these two mutant structures does not reveal an increase in either the polarity of the heme pocket or the solvent exposure of the heme porphyrin ring (Table 6.25). In these cases it must be concluded that the observed differences in pyrrole A propionate conformation are caused by other factors. For example, it is possible that these rearrangements occur as a result of breaking of the hydrogen bond between residue 52 and the heme pyrrole A propionate group in the two mutant structures. Clearly the factors governing pyrrole A propionate conformation are complex and cannot be assigned to a single parameter such as polarity of the heme environment.

6.3.2 Functional role of the internal water molecule Wat166

The effect of the mutations, N52A and I75M, on the positions of atoms involved in the hydrogen bond network about the internal water molecule, Wat166, appears minimal (see Figures 6.48 and 6.50). However, an analysis of the hydrogen bonds present does indicate that some subtle but important changes have occurred. In wild-type yeast iso-1-cytochrome *c* the Asn52 ND2 group functions as a hydrogen bond donor to Wat166. As a result of the mutation of this residue to an alanine the function of the Asn52 ND2 group is taken over by a new water molecule (Wat300). This water molecule can both donate or accept a hydrogen bond from Wat166, effectively both lowering the hydrogen bond donor capacity, and at the same time increasing the hydrogen

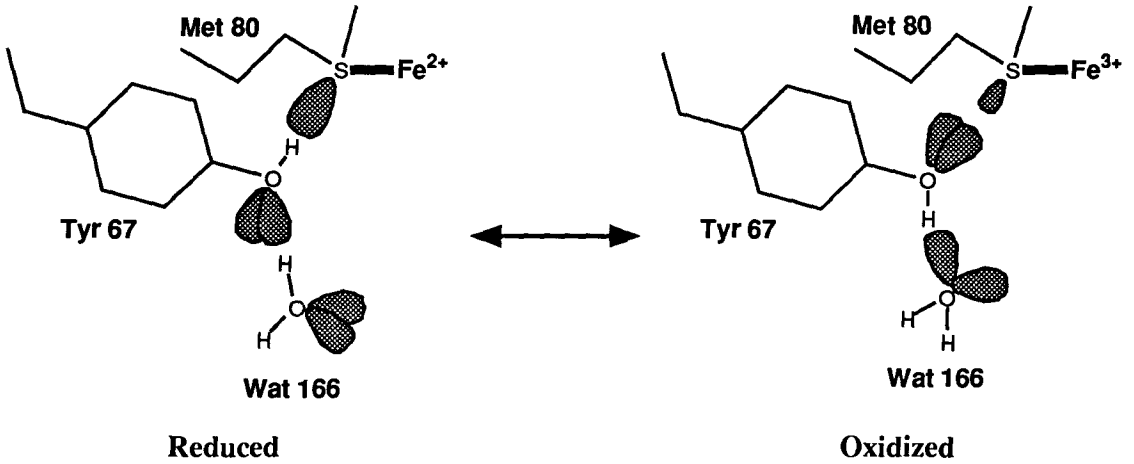

Figure 6.52: A schematic representation of the proposed water-switch mechanism for stabilization of the alternate oxidation states of yeast iso-1-cytochrome *c*.

bond acceptor capacity on the side of Wat166 furthest away from Tyr67 (Figure 6.48). The replacement of Ile75 with a methionine residue produces similar results but only increases the hydrogen bond acceptor capacity on the side of Wat166 that is away from Tyr67 (Figure 6.50). These results suggest that as a result of the N52A and I75M mutations the precise orientation of Wat166 is likely to be affected.

The orientation of Wat166 has been proposed to be crucial for the stabilization of the two oxidation states of cytochrome *c* (Chapter 3). In the reduced state Wat166 is oriented so that the maximum number of hydrogen bonds are formed in this region. In the oxidized form it is proposed that the dipole moment of Wat166 realigns in the electrostatic field generated by the more positively charged heme iron atom. As a consequence of this change in orientation the hydrogen bond between Tyr67 OH and the Met80 SD ligand is broken. The loss of this hydrogen bond is an additional stabilizing feature in the oxidized state of the protein since it lessens the electron withdrawing power of the Met80 ligand (Figure 6.52; see also Figure 3.28).

If this proposed "water-switch" mechanism for the stabilization of the two oxidation states of cytochrome *c* is valid then alteration of the orientation of Wat166 should affect the equilibrium constant for oxidation of yeast iso-1-cytochrome *c*. This is equivalent to modifying the

observed midpoint reduction potential of the protein. The two mutants discussed here test this hypothesis. As has been described above, both the N52A and I75M mutations increase the hydrogen bond acceptor capacity in a region adjacent to Wat166 and furthest away from the Tyr67 hydroxyl group. This should result in Tyr67 OH functioning less as a hydrogen acceptor and more as a hydrogen donor group to Wat166 since the hydrogen atom of Wat166 interacting with the Tyr67 side-chain can also interact with Wat300 in the N52A mutant or with Met75 SD in the I75M mutant. As a consequence, the hydrogen bond between Tyr67 OH and Met80 SD is weakened and the Wat166 hydrogen bond network will tend more towards the conformation seen in the oxidized state of the wild-type protein. Accordingly, the proposed water-switch mechanism suggests the net effect is that the midpoint reduction potential for these mutant proteins should be lower.

As predicted by this mechanism, both mutant proteins do indeed show a drop in midpoint reduction potential. The N52A protein has a midpoint reduction potential 33 mV lower than that of wild-type yeast iso-1-cytochrome *c* and for the I75M variant the midpoint reduction potential is 45 mV lower (Table 1.5; Rafferty, 1992). The magnitude of the drop in midpoint reduction potential observed is also in the range expected. Previous studies have shown that complete removal of the Tyr67 OH - Met80 SD hydrogen bond results in an overall drop of ~56 mV (Chapters 4 and 5), and this therefore represents the maximum limit expected for modifications affecting this hydrogen bond.

These results support the proposed role of Wat166 in stabilizing the two oxidation states of yeast iso-1-cytochrome *c*. However, it is likely that this water-switch mechanism is not only restricted to yeast iso-1-cytochrome *c*. Model-building studies with the five other high resolution eukaryotic cytochrome *c* structures solved to date [reduced and oxidized tuna (Takano & Dickerson, 1981a,b); oxidized rice (Ochi *et al.*, 1983); oxidized horse (Bushnell *et al.*, 1990); and reduced yeast iso-2 (Murphy *et al.*, 1992)] show that the proposed oxidation state dependent orientation of Wat166 is likely a common feature of all these proteins (Chapter 3; see also Figure 3.28). Further, comparison of 94 eukaryotic cytochrome *c* sequences (Moore & Pettigrew,

1990; see also Section 1.2) has shown that the four key residues involved, Asn52, Tyr67, Thr78 and Met80 are invariant except for one case where Thr78 is replaced by an asparagine residue (Amati *et al.*, 1988). It can be concluded that in most eukaryotic cytochromes *c* a comparable hydrogen bond network is present to that in yeast iso-1-cytochrome *c* and that stabilization of the different oxidation states for these cytochromes *c* is similar to the water-switch mechanism described herein.

Summary

The first objective of the work described in this thesis was to identify the structural differences that exist between the two oxidation states of cytochrome *c*. This was accomplished by determining the structure of the oxidized state of yeast iso-1-cytochrome *c* using a crystal form isomorphous to that of the reduced protein whose structure had been previously solved to high resolution (Louie & Brayer, 1990). The use of isomorphous crystalline material and similar structure refinement methodologies for both oxidation states of yeast iso-1-cytochrome *c* allowed for a comparative analysis to be made in the absence of possible systematic errors introduced due to differing structure determination approaches. These results showed that oxidation state differences are expressed for the most part as increased mobility for selected segments of polypeptide chain in the oxidized state, rather than as discrete positional shifts of atoms. This correlates well with a large body of data from a wide variety of techniques which suggest that the oxidized form of cytochrome *c* is more open and less rigid than that of the reduced state (see Section 1.3 for references).

Specifically, three regions of polypeptide chain in yeast iso-1-cytochrome *c* displayed an increase in flexibility in the oxidized state of the protein, and include residues 47–59, 65–72 and 81–85, with maximal increases being observed for Asn52, Tyr67 and Phe82. The side-chains of two of these residues are hydrogen bonded to the internal water molecule, Wat166, which showed a 1.7 Å shift towards the heme iron atom in the oxidized state of the protein. Further study of this water molecule suggested that it might be a major factor in stabilizing both oxidation states through differential orientation of its dipole moment, shift in distance to the heme iron atom and alteration of the surrounding hydrogen bond network. Comparison between oxidation states also revealed some subtle differences in the conformation of the heme pyrrole A propionate and the surrounding hydrogen bond network that had been previously postulated

to occur (Moore, 1983). It also demonstrated that heme planar distortion and the orientation of the imidazole plane of the His18 ligand is dependent on oxidation state (Chapter 3).

Once the structural differences between the two oxidation states of cytochrome *c* were identified, the next objective was to determine the role of these observed differences in the biological function of cytochrome *c*. The method of choice to assess the importance of these differences was to study the structures of yeast iso-1-cytochrome *c* variants with mutations in those regions of the protein that demonstrated oxidation state dependent conformations. Five mutants were selected for this study: N52A, N52I, Y67F, N52I-Y67F and I75M. All of these mutations are located directly adjacent to the internal water molecule, Wat166, which studies of the wild-type protein had suggested was likely to play a central role in stabilizing the alternate oxidation states of cytochrome *c*. A large body of functional data was also available for these mutants (see Section 1.5) which allowed for a careful examination of the impact of structural alterations on functional properties.

In total eight structures were determined to high resolution. The N52I, Y67F and N52I-Y67F yeast iso-1-cytochrome *c* variants were determined in both oxidation states, while the N52A and I75M mutant proteins were resolved in the reduced form of the protein. All these mutations were observed to cause perturbations in the hydrogen bond network about Wat166. In the N52I and N52I-Y67F variants the internal cavity occupied by Wat166 was filled in resulting in the exclusion of this internal water molecule from these structures. The N52A and Y67F mutations affected the hydrogen bond network in a more subtle way in that the Asn52 ND2 and Tyr67 OH groups which form hydrogen bonds to Wat166 were replaced by water molecules. The exchange of Ile75 for a methionine residue resulted in an extra hydrogen bond acceptor group being added to the hydrogen bond network about Wat166. For the three mutants whose structures were determined in both oxidation states (N52I, Y67F and N52I-Y67F), none demonstrated the increased polypeptide chain flexibility observed in the oxidized state of the wild-type protein. Besides these changes, all mutant structures also revealed altered conformations about the pyrrole A propionate group.

Correlation of the structural data provided by the wild-type and mutant structures with functional studies suggests the role of Wat166 is three fold. First, the presence of Wat166 provides a convenient mechanism to modify the hydrogen bond network involving several key residues near the Met80 ligand, depending on the oxidation state of the protein. In the reduced state this water molecule is oriented so that the maximum number of hydrogen bonds are formed in this region. In the oxidized form Wat166 realigns in the electrostatic field generated by the more positively charged heme iron atom. As a consequence of this change in orientation several hydrogen bonds are broken, including the interaction between Tyr67 OH and the Met80 SD ligand. This latter interaction has also been shown to be a factor in controlling the midpoint reduction potential of cytochrome *c* by influencing the electron withdrawing power of the Met80 ligand. In the N52I, Y67F and N52I-Y67F variants where this hydrogen bond is absent, the midpoint reduction potential has decreased by ~56 mV (Chapters 4 and 5). Examination of the N52A and I75M mutant proteins has shown that the strength of the Tyr67 OH to Met80 SD hydrogen bond is strongly correlated to the dipole orientation of Wat166 (Chapter 6). In a second role, related to the first, Wat166 is required to maintain the spatial and hydrogen bond relationships in this vicinity of cytochrome *c*. Absence of this internal water molecule, as demonstrated in the N52I mutant protein, leads to shifts in such critical residues as Tyr67 and subsequent rearrangements in the hydrogen bond network (Chapter 5).

Finally, Wat166 also appears to modulate the oxidation state dependent mobility differences in polypeptide chain segments that have been observed between the reduced and oxidized structures of cytochrome *c*. Although the importance of this phenomena is not fully understood, our studies suggest that it might play a role in the complexation and dissociation of cytochrome *c* with redox partners. An unusual aspect of the outer solvent surface presentation of those segments of polypeptide chain having greater conformational mobility in the oxidized state of the protein is that they are bisected by an unperturbed stretch of polypeptide chain (residues 73–80). This has led to the proposal that this highly conserved segment of polypeptide chain could act as a push-button contact trigger operated by protein-protein contacts with redox

partners. The postulated function of this trigger is to initiate the necessary structural changes required to switch between oxidation states and in this way facilitate electron transfer.

In conclusion, these studies provide a solid foundation for further analyses of structure-function relationships in cytochrome *c* targeted at gaining a deeper understanding of the electron transfer reaction mediated by this protein.

Bibliography

Abrams, R., Altshul, A.M. & Hogness, T.R. (1942). Cytochrome *c* peroxidase II. The peroxidase-hydrogen peroxidase complex. *J. Biol. Chem.*, **142**, 303–316.

Altschul, A.M., Abrams, R. & Hogness, T.R. (1940). Cytochrome *c* peroxidase. *J. Biol. Chem.*, **136**, 777–794.

Amati, B.B., Goldsmidth-Clermont, M., Wallace, C.J.A. & Rochaix, J. (1988). cDNA and deduced amino acid sequences of cytochrome *c* from *Chlamydomonas reinhardtii*: unexpected functional and phylogenetic implications. *J. Mol. Evol*, **28**, 151–160.

Arndt, U.W. & Wonacott, A.J., eds. (1977). *The Rotation Method in Crystallography.* North-Holland Publishing Company, Amsterdam, New York, Oxford.

Bach, S.J., Dixon, M. & Keilin, D. (1942a). A new soluble cytochrome component from yeast. *Nature (London)*, **149**, 21.

Bach, S.J., Dixon, M. & Zerfas, L.G. (1942b). Lactic dehydrogenase of yeast. *Nature (London)*, **149**, 48–49.

Barker, P.D., Mauk, M.R. & Mauk, A.G. (1991). Proton titration curve of yeast iso-1-ferricytochrome *c*. Electrostatic and conformational effects of point mutations. *Biochemistry*, **30**, 2377–2383.

Barlow, D.J. & Thornton, J.M. (1983). Ion-pairs in proteins. *J. Mol. Biol.*, **168**, 867–885.

Berghuis, A.M. & Brayer, G.D. (1992). Oxidation state-dependent conformational changes in cytochrome *c*. *J. Mol. Biol.*, **223**, 959–976.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B, Meyer, E.F., Bruce, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. & Tasmui, M. (1977). The protein databank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.*, **112**, 535–542.

Blundell, T.L. & Johnson, L.N. (1976). *Protein Crystallography.* Academic Press, New York, London, San Francisco.

Boon, P.J., Van Raay, A.J.M., Tesser, G.I. & Nivard, R.J.F. (1979). Semi-synthesis, conformation and cytochrome *c* oxidase activity of eight cytochrome *c* analogues. *FEBS Letters*, **108**, 131–135.

Bosshard, H.R. & Zurrer, M. (1980). The conformation of cytochrome *c* in solution. *J. Biol. Chem.*, **255**, 6694–6699.

Bott, R. (1991). Assessing the accuracy of subtilisin BPN. In *Abstracts of the Annual meeting of the American Crystallographic Association*, vol. 19 of *2*, p. 36, Toledo, Ohio. American Crystallographic Association.

Brayer, G.D. & Murphy, M.E.P. (1993). Structural studies of eukaryotic cytochromes *c*. In *The Cytochrome c Handbook* (Mauk, A.G. & Scott, R.A., eds.). Plenum Press, New York. In press.

Burch, A.M., Rigby, S.E.J., Funk, W.D., MacGillivray, R.T.A., Mauk, M.R., Mauk, A.G. & Moore, G.R. (1990). NMR characterization of surface interactions in the cytochrome b_5-cytochrome *c* complex. *Science*, **247**, 831–833.

Burrows, A.L., Guo, L-H., Hill, A.O., McLendon, G. & Sherman, F. (1991). Direct electrochemistry of proteins; Investigations of yeast cytochrome *c* mutants and their complexes with cytochrome b_5. *Eur. J. Biochem.*, **202**, 543–549.

Bushnell, G.W., Louie, G.V. & Brayer, G.D. (1990). High-resolution three-dimensional structure of horse heart cytochrome *c*. *J. Mol. Biol.*, **214**, 585–595.

Butt, W.D. & Keilin, D. (1962). Absorption spectra and some other properties of cytochrome *c* and of its compounds with with ligands. *Proc. Royal Soc. London*, **B152**, 429–458.

Chambers, J.L. & Stroud, R.M. (1979). The accuracy of refined protein structures: comparison of two independently refined models of bovine trypsin. *Acta Crystallogr. sect. B*, **35**, 1861–1874.

Churg, A.K. & Warshel, A. (1986). Control of the redox potential of cytochrome *c* and microscopic dielectric effects in proteins. *Biochemistry*, **25**, 1675–1681.

Churg, A.K. Weiss, R.M. & Warshel, A. (1983). On the action of cytochrome *c*: correlating geometry changes upon oxidation with activation energies of electron transfer. *J. Phys. Chem.*, **87**, 1683–1694.

Cohen, H.J. & Fridovich, I. (1971). Hepatic sulfite oxidase: purification and properties. *J. Biol. Chem.*, **246**, 359–366.

Cokic, P. & Erman, J.E. (1987). The Effect of complex formation upon the reduction rates of cytochrome *c* and cytochrome *c* peroxidase compound II. *Biochim. Biophys. Acta*, **913**, 257–271.

Connolly, M.L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science*, **221**, 709–713.

Creighton, T.E. (1984). *Proteins: Structure and Molecular Principles*. Freeman and Company, New York.

Cruickshank, D.W.J. (1949). The accuracy of electron-density maps in x-ray analysis with special reference to dibenzyl. *Acta Crystallogr.*, **2**, 65–82.

Cruickshank, D.W.J. (1954). The accuracy of electron-density maps in x-ray analysis: correction. *Acta Crystallogr.*, **7**, 519.

Cruickshank, D.W.J. (1985). Fourier synthesis and structure factors. In *International Tables for X-ray Crystallography* (Kasper, J.S. & Lonsdale, K., eds.), vol. 2, pp. 317–339. D. Reidel Publishing Company, Dordrecht, 3rd edition.

Cusanovich, M.A., Meyer, T.E. & G., Tollin. (1988). *c*-Type cytochromes: oxidation-reduction properties. In *Heme Proteins* (Eichhorn, G.L. & Marzilli, J.G., eds.), vol. 7 of *Advances in Inorganic Biochemistry*, pp. 37–92. Elsevier, New York.

Cutler, R.L., Davies, A.M., Creighton, S., Warshel, A., Moore, G.R., Smith, M. & Mauk, A.G. (1989). Role of arginine-38 in regulation of the cytochrome *c* oxidation- reduction equilibrium. *Biochemistry*, **28**, 3188–3197.

Cutler, R.L., Pielak, G.J., Mauk, A.G. & Smith, M. (1987). Replacement of cysteine-107 of *Saccharomyces cerevisiae* iso-1-cytochrome *c* with threonine: improved stability of the mutant protein. *Protein Eng.*, **1**, 95–99.

Das, G., Hickey, D.R., McLendon, D., McLendon, G. & Sherman, F. (1989). Dramatic thermostabilization of yeast iso-1-cytochrome *c* by an asparagine → isoleucine replacement at position 57. *Proc. Nat. Acad. Sci., U.S.A.*, **86**, 496–499.

Dayhoff, M.O., Schwartz, R.M. & Orcutt, B.C. (1976). In *Atlas of Protein Sequence and Structure* (Dayoff, M.O., ed.), vol. 5, pp. 20–49. National Biomedical Research Foundation, Washington, D.C.

Dickerson, R.E. & Timkovich, R. (1975). Cytochromes *c*. In *The Enzymes* (Boyer, P.D., ed.), vol. 11, part A, pp. 397–547. Academic Press, New York, 3rd edition.

Dong, A., Huang, P. & Caughey, W.S. (1992). Redox-dependent changes in β-extended chain and turn structures of cytochrome *c* in water solution determined by second derivative amide I infrared spectra. *Biochemistry*, **31**, 182–189.

Eden, D., Matthew, J.B., Rosa, J.J. & Richards, F.M. (1982). Increase in apparent compressibility of cytochrome *c* upon oxidation. *Proc. Nat. Acad. Sci., U.S.A.*, **79**, 815–819.

Fauman, E.B. & Stroud, R.M. (1991). An empirical approach to analysis of errors in x-ray protein structures by reference to atomic B factors. In *Abstracts of the Annual meeting of the American Crystallographic Association*, vol. 19 of *2*, p. 34, Toledo, Ohio. American Crystallographic Association.

Feng, Y. & Englander, S.W. (1990). Salt-dependent structure change and ion binding in cytochrome *c* studied by two-dimensional proton NMR. *Biochemistry*, **29**, 3505–3509.

Feng, Y., Roder, H. & Englander, S.W. (1990). Redox-dependent structure change and hyperfine nuclear magnetic resonance shifts in cytochrome *c*. *Biochemistry*, **29**, 3494–3504.

Fisher, W.R., Tanuichi, H. & Anfinsen, C.B. (1973). On the role of heme in the formation of the structure of cytochrome *c*. *J. Biol. Chem.*, **248**, 3188–3195.

Frauenhoff, M.M. & Scott, R.A. (1992). The role of tyrosine 67 in the cytochrome *c* heme crevice structure studied by semisynthesis. *Proteins: Struct. Funt. Genet.*, **14**, 202–212.

Gadsby, P.M.A., Peterson, J., Foote, N., Greenwood, C. & Thomson, A.J. (1987). Identification of the ligand-exchange process in the alkaline transition of horse heart cytochrome *c*. *Biochem. J.*, **246**, 43–54.

Gao, Y., Boyd, J., Pielak, G.J. & Williams, R.J.P. (1991). Comparison of reduced and oxidized yeast iso-1-cytochrome *c* using proton paramagnetic shifts. *Biochemistry*, **30**, 1928–1934.

Giegé, R. & Mikol, V. (1989). Crystallogenesis of proteins. *Trends Biochem. Sci.*, **7**, 277–282.

Hampsey, D.M., Das, G. & Sherman, F. (1986). Amino acid replacements in yeast iso-1-cytochrome *c*, comparison with the phylogenetic series and the tertiary structures of related cytochromes *c*. *J. Biol. Chem.*, **261**, 3259–3271.

Hendrickson, W.A. (1985). Stereochemically restrained refinement of macromolecular structures. In *Methods in Enzymology* (Wyckoff, H.W., Hirs, C.H.W. & Timasheff, S.N., eds.), vol. 115 B, pp. 252–270. Academic Press, Inc., Orlando.

Hendrickson, W.A. & Konnert, J. (1980). Incorporation of stereochemical information into crystallographic refinement. In *Computing in Crystallography* (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds.), pp. 13.01–13.26. Indian academy of sciences, Bangalore.

Hendrickson, W.A. & Konnert, J. (1981). Stereochemically restrained crystallographic least-squares refinement of macromolecular structures. In *Biomolecular Structure, Function, Conformation and Evolution* (Sunivasan, R., ed.), vol. 1, pp. 43–57. Pergamon Press, Oxford.

Hickey, D.R., Berghuis, A.M., Lafond, G., Jaeger, J.A., Cardillo, T.S., McLendon, D., Das, G., Sherman, F., Brayer, G.D. & McLendon, G. (1991). Enhanced thermodynamic stabilities of yeast iso-1-cytochrome *c* with amino acid replacements at position 52 and 102. *J. Biol. Chem.*, **266**, 11686–11694.

Hildebrandt, P. & Stockburger, M. (1989). Cytochrome *c* at charged interfaces. 2. Complexes with negatively charged macromolecular systems studied by resonance Raman spectroscopy. *Biochemistry*, **28**, 6722–6728.

Ito, A. (1980a). Cytochrome b_5-like hemoprotein of outer mitochondrial membrane; OM Cytochrome *b* I: Purification of OM cytochrome *b* from rat liver mitochondria and comparison of its molecular properties with those of cytochrome b_5. *J. Biochem. (Japan)*, **87**, 63–71.

Ito, A. (1980b). Cytochrome b_5-like hemoprotein of outer mitochondrial membrane; OM Cytochrome *b* II. Contribution of OM cytochrome *b* to rotenone insensitive NADH-cyt *c* reductase activity. *J. Biochem. (Japan)*, **87**, 73–80.

Kassner, R.J. (1972). Effects of nonpolar environments on the redox potentials of heme complexes. *Proc. Nat. Acad. Sci., U.S.A.*, **69**, 2263–2267.

Kassner, R.J. (1973). A theoretical model for the effects of local nonpolar heme environments on the redox potentials of cytochromes. *J. Amer. Chem. Soc.*, **95**, 2674–2677.

Koloczek, H., Horie, T., Yonetani, T., Anni, H., Maniara, G. & Vanderkooi, J.M. (1987). Interaction between cytochrome *c* and cytochrome *c* peroxidase: excited-state reactions of zinc- and tin-substituted derivatives. *Biochemistry*, **26**, 3142–3148.

Komar-Panicucci, S., Bixler, J., Bakker, G., Sherman, F. & McLendon, G. (1992). Tuning the redox potential of cytochrome *c* through synergistic site replacements. *J. Amer. Chem. Soc.*, **114**, 5443–5445.

Korzun, Z.R., Moffat, K., Frank, K. & Cusanovich, M.A. (1982). Extended X-ray absorption fine structure studies of cytochrome *c*: structural aspects of oxidation-reduction. *Biochemistry*, **21**, 2253–2258.

Koul, A.K., Folena-Wasserman, G. & Warme, P.K. (1979). Semi-synthetic analogs of cytochrome *c* at positions 67 and 74. *Biochem. Biophys. Res. Commoun.*, **89**, 1253–1259.

Kreil, G. (1965). Die C-terminale Aminosauresequenz des Tunfischcytochroms *c*. *Z. Physiol. Chemie.*, **340**, 86–87.

Langen, R., Brayer, G.D., Berghuis, A.M., McLendon, G., Sherman, F. & Warshel, A. (1992). Effect of the Asn52→Ile mutation on the redox potential of yeast cytochrome *c*: Theory and experiment. *J. Mol. Biol.*, **224**, 589–600.

Lederer, F., Ghrir, R., Guiarad, B., Cortial, S. & Ito, A. (1983). Two homologous cytochromes b_5 in a single cell. *Eur. J. Biochem.*, **132**, 95–102.

Leung, C.J., Nall, B.T. & Brayer, G.D. (1989). Crystallization of yeast iso-2-cytochrome *c* using a novel hair seeding technique. *J. Mol. Biol.*, **206**, 783–785.

Liang, N., Mauk, A.G., Pielak, G.J., Johnson, J.A., Smith, M. & Hoffman, B. (1988). Regulation of interprotein electron transfer by residue 82 of yeast cytochrome *c*. *Science*, **240**, 311–313.

Liang, N., Pielak, G.J., Mauk, A.G., Smith, M. & Hoffman, B.M. (1987). Yeast cytochrome *c* with phenylalanine or tyrosine at position 87 transfers electrons to (zinc cytochrome *c* peroxidase)$^+$ at a rate ten thousand times that of the serine-87 or glycine-87 variants. *Proc. Nat. Acad. Sci., U.S.A.*, **84**, 1249–1252.

Liu, G., Crygon, C.A. & Spiro, T.G. (1989). Ionic strength dependence of cytochrome *c* structure and ultraviolet resonance Raman spectroscopy. *Biochemistry*, **28**, 5046–5050.

Louie, G.V. (1990). *Structural studies of wild-type and variant yeast iso-1-cytochromes c*. PhD thesis, University of British Columbia, Vancouver, Canada.

Louie, G.V. & Brayer, G.D. (1989). A polypeptide chain-refolding event occurs in the Gly82 variant of yeast iso-1-cytochrome *c*. *J. Mol. Biol.*, **210**, 313–322.

Louie, G.V. & Brayer, G.D. (1990). High-resolution refinement of yeast iso-1-cytochrome *c* and comparisons with other eukaryotic cytochromes *c*. *J. Mol. Biol.*, **214**, 527–555.

Louie, G.V., Hutcheon, W.L.B. & Brayer, G.D. (1988a). Yeast iso-1-cytochrome *c*: a 2.8Å resolution three-dimensional structure determination. *J. Mol. Biol.*, **199**, 295–314.

Louie, G.V., Pielak, G.J., Smith, M. & Brayer, G.D. (1988b). The role of Phe-82 in yeast iso-1-cytochrome *c* and remote conformational changes induced by a serine residue at this position. *Biochemistry*, **27**, 7870–7876.

Lum, V.R., Brayer, G.D., Louie, G.V., Smith, M. & Mauk, A.G. (1987). Computer modelling of yeast iso-1-cytochrome *c* - yeast cytochrome *c* peroxidase complexes. *Protein Struct., Folding Design*, **2**, 143–150.

Luntz, T.L., Schejter, A., Garber, E.A.E. & Margoliash, E. (1989). Structural significance of an internal water molecule studied by site-directed mutagenesis of tyrosine-67 in rat cytochrome *c*. *Proc. Nat. Acad. Sci., U.S.A.*, **86**, 3524–3528.

Luzzati, V. (1952). Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallogr.*, **5**, 802–810.

Marchon, J.C., Mashiko, T. & Reed, C.A. (1982). How does nature control cytochrome redox potentials. In *Electron Transport and Oxygen Utilization* (Ho, C., ed.), pp. 67–72. Elsevier, Amsterdam.

Marcus, R.A. & Sutin, N. (1985). Electron transfer in chemistry and biology. *Biochim. Biophys. Acta*, **811**, 265–322.

Margalit, R. & Schejter, A. (1973). Cytochrome *c*: a thermodynamic study of the relationships among oxidation state, ion-binding and structural parameters; 1. The effects of temperature, pH and electrostatic media on the standard redox potential of cytochrome *c*. *Eur. J. Biochem.*, **32**, 492–499.

Margoliash, E. (1990). Was cytochrome *c* properly engineered in the first place. In *Abstracts of the C5 Symposium; Control of charge transfer in cytochrome and chlorophyll complexes*, p. 8, Montreal, Quebec, Canada. International Union for Pure and Applied Biophysics.

Margoliash, E. & Bosshard, H.R. (1983). Guided by electrostatics, a textbook protein comes of age. *Trends Biochem. Sci.*, **8**, 316–320.

Margoliash, E. & Schejter, A. (1966). Cytochrome *c*. *Advan. Protein Chem.*, **21**, 113–286.

Margoliash, E., Smith, E.L., Kreil, G. & Tuppy, H. (1961). Amino Acid Sequence of Horse Heart Cytochrome *c*. *Nature (London)*, **192**, 1125–1127.

Mashiko, T., Reed, C.A., Haller, K.J., Kastner, M. E. & Scheidt, W.R. (1981). Tioether ligation in iron-porphyrin complexes: models for cytochrome *c*. *J. Amer. Chem. Soc.*, **103**, 5758–5767.

Mathews, F.S. (1985). The structure, function and evolution of cytochromes. In *Progress in Biophysics and Molecular Biology* (Noble, D. & Blundell, T.L., eds.), vol. 45, pp. 1–56. Pergamon Press, Oxford.

Mauk, M.R., Barker, P.D. & Mauk, A.G. (1991). Proton linkage of complex formation between cytochrome *c* and cytochrome b_5: electrostatic consequences of protein-protein interactions. *Biochemistry*, **30**, 9873–9881.

McLendon, G., Hickey, D.R., Berghuis, A.M., Sherman, F. & Brayer, G.D. (1991). Effects of reaction free energy in biological electron transfer *In Vitro* and *In Vivo*. In *Electron Transfer in Inorganic, Organic and Biological Systems* (Bolton, J.R., Mataga, N. & McLendon, G., eds.), vol. 228 of *Advances in Chemistry Series*, pp. 179–189. American Chemical Society.

McLendon, G. & Miller, J.R. (1985). The dependence of biological electron transfer rates on exothermicity: the cytochrome *c*/cytochrome b_5 couple. *J. Amer. Chem. Soc.*, **107**, 7811–7816.

McLeod, R.M., Farkas, W., Fridovich, I. & Handler, P. (1961). Purification and properties of hepatic sulfite oxidase. *J. Biol. Chem.*, **236**, 1841–1846.

McPherson, A (1990). Current approaches to macromolecular crystallization. *Eur. J. Biochem.*, **189**, 1–23.

Montgomery, D., Leung, D., Smith, M., Shalit, P. Faye, G. & Hall, B (1980). Isolation and sequence of a gene for iso-2-cytochrome *c* in *Saccharomyces cerevisiae*. *Proc. Nat. Acad. Sci., U.S.A.*, **77**, 541–545.

Moore, G.R. (1983). Control of redox properties of cytochrome *c* by special electrostatic interactions. *FEBS Letters*, **161**, 171–175.

Moore, G.R. & Pettigrew, G.W. (1990). *Cytochrome c. Evolution, Structural and Physiochemical Aspects.* Springer-Verlag, Berlin.

Moore, G.R., Pettigrew, G.W. & Rogers, N.K. (1986). Factors influencing redox potentials of electron transfer proteins. *Proc. Nat. Acad. Sci., U.S.A.*, **83**, 4998–4999.

Moore, G.R. & Williams, R.J.P. (1977). Structural basis for the variation in redox potential of cytochromes. *FEBS Letters*, **79**, 229–232.

Mori, E. & Morita, Y. (1980). Amino Acid Sequence of Cytochrome *c* from Rice. *J. Biochem. (Tokyo)*, **87**, 249–266.

Murphy, M.E.P (1993). *Structural constraints for the folding stability and function of cytochrome c.* PhD thesis, University of British Columbia, Vancouver, Canada.

Murphy, M.E.P., Nall, B.T. & Brayer, G.D. (1992). Structure determination and analysis of yeast iso-2-cytochrome *c* and a composite mutant protein. *J. Mol. Biol.*, **227**, 160–176.

Nocek, J.M., Stemp, E.D.A., Finnegan, M.G., Koshy, T.I., Johnson, M.K., Margoliash, E., Mauk, A.G., Smith, M. & Hoffman, B.M. (1991). Low-temperature, cooperative conformational transition within [Zn-cytochrome *c* peroxidase, cytochrome *c*] complexes: variation with cytochrome *c*. *J. Amer. Chem. Soc.*, **113**, 6822–6831.

North, A.C.T., Phillips, D.C. & Mathews, F.S. (1968). A semi-empirical method of absorption correction. *Acta Crystallogr. sect. A*, **24**, 351–359.

Nozaki, M., Mitzushima, H., Horio, T. & Okunuki, K. (1958). Further studies on proteinase digestion of bakers yeast cytochrome *c*. *J. Biochem.*, **256**, 673–676.

Ochi, H., Hata, Y., Tanaka, N., Kakudo, M., Sakkurai, T., Aihara, S. & Morita, Y. (1983). Structure of rice ferricytochrome *c* at 2.0Å resolution. *J. Mol. Biol.*, **166**, 407–418.

Ohlendorf, D.H., Treharne, A., Weber, P.C., Wendoloski, J.J. & Salemme, F.R. (1991). Accuracy of structures: a comparison of independently refined models of human IL-β. In *Abstracts of the Annual meeting of the American Crystallographic Association*, vol. 19 of *2*, p. 36, Toledo, Ohio. American Crystallographic Association.

Pace, C.N. (1990). Measuring and increasing protein stability. *Trends Biochem. Sci.*, **8**, 93–98.

Pearce, L.L., Gärtner, A.L., Smith, M. & Mauk, A.G. (1989). Mutation-induced perturbation of the cytochrome *c* alkaline transition. *Biochemistry*, **28**, 3152–3156.

Pelletier, H. & Kraut, J. (1992). Crystal structure of a complex between electron transfer partners, cytochrome *c* peroxidase and cytochrome *c*. *Science*, **258**, 1748–1755.

Perry, K.M., Fauman, E.B., Finner-Moore, J.S., Montfort, W.R., Maley, G.F., Meley, F. & Stroud, R.M. (1990). Plastic adaptation toward mutations in proteins: structural comparison of thymidylate synthases. *Proteins: Struct. Funt. Genet.*, **8**, 315–333.

Pettigrew, G.W. & Moore, G.R. (1987). *Cytochrome c. Biological Aspects.* Springer-Verlag, Berlin.

Pielak, G.J., Mauk, A.G. & Smith, M. (1985). Site-directed mutagenesis of cytochrome *c* shows that an invariant Phe is not essential for function. *Nature (London)*, **313**, 152–153.

Pielak, G.J., Oikawa, K., Mauk, A.G., Smith, M. & Kay, C.M. (1986). Elimination of the negative Soret Cotton effect of cytochrome *c* by replacement of the invariant phenylalanine using site-directed mutagenesis. *J. Amer. Chem. Soc.*, **188**, 2724–2727.

Poulos, T.L. & Kraut, J. (1980). A hypothetical model of the cytochrome *c* peroxidase-cytochrome *c* electron transfer complex. *J. Biol. Chem.*, **255**, 10322–10330.

Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T. (1988). *Numerical recipes.* Cambridge University Press, Cambridge.

Proudfoot, A.E.I. & Wallace, C.J.A. (1987). Semisynthesis of cytochrome *c*: the effect of modifying the conserved residues 38 and 39. *Biochem. J.*, **248**, 965–967.

Rackovsky, S. & Goldstein, D.A. (1984). On the redox conformational change in cytochrome *c*. *Proc. Nat. Acad. Sci., U.S.A.*, **81**, 5901–5905.

Rafferty, S.P. (1992). *Functional properties of active site variants of yeast cytochrome c*. PhD thesis, University of British Columbia, Vancouver, Canada.

Rafferty, S.P., Pearce, L.L., Barker, P.D., Guillemette, J.G., Kay, C.M., Smith, M. & Mauk, A.G. (1990). Electrochemical, kinetic, and circular dichroic consequences of mutations at position 82 of yeast iso-1-cytochrome *c*. *Biochemistry*, **29**, 9365–9369.

Rashin, A.A., Iofin, M. & Honig, B. (1986). Internal cavities and buried waters in globular proteins. *Biochemistry*, **25**, 3619–3625.

Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. sect. A*, **42**, 140–149.

Read, R.J. (1990). Structure-factor probabilities for related structures. *Acta Crystallogr. sect. A*, **46**, 900–912.

Read, R.J., Fujinaga, M., Sielecki, A.R. & James, M.N.G. (1983). Structure of the complex of *Streptomyces griseus* protease B and the third domain of the turkey ovomucoid inhibitor at 1.8-Å resolution. *Biochemistry*, **22**, 4420–4433.

Rogers, N.K. & Moore, G.R. (1988). On the energetics of conformational changes and pH dependent redox behaviour of electron transfer proteins. *FEBS Letters*, **228**, 69–73.

Rossmann, M.G., Leslie, A.G.W., Abdel-Meguid, S.S. & Tsukihara, T. (1979). Processing and post-refinement of oscillation camera data. *J. Appl. Cryst.*, **12**, 570–581.

Saigo, S. (1986). Isomerization and denaturation of homologous cytochromes *c*: correlation between local and gross conformational changes. *J. Biochem*, **100**, 157–165.

Salemme, F.R. (1976). A hypothetical structure for an intermolecular electron transfer complex of cytochrome *c* and b_5. *J. Mol. Biol.*, **102**, 563–568.

Salemme, F.R. (1977). Structure and function of cytochrome *c*. *Annu. Rev. Biochem.*, **46**, 299–329.

Satterlee, J.D., Moench, S.J. & Erman, J.E. (1987). A proton NMR study of the non-covalent complex of horse cytochrome *c* and yeast cytochrome *c* peroxidase and its comparison with other interacting protein complexes. *Biochim. Biophys. Acta*, **912**, 87–97.

Schejter, A., Aviram, I. & Goldkorn, T. (1982). The contribution of electrostatic factors to the oxidation-reduction potentials of *c*-type cytochromes. In *Electron Transport and Oxygen Utilization* (Ho, C., ed.), pp. 95–99. Elsevier, Amsterdam.

Schlauder, G.G. & Kassner, R.J. (1979). Comparative solvent perturbation of horse heart cytochrome *c* and *Rhodospirillum rubrum* cytochrome *c*2. *J. Biol. Chem.*, **254**, 4110–4113.

Schultz, G.E. & Shirmer, R.H. (1979). *Principles of Protein Structure* (Cantor, C.R., ed.). Springer-Verlag, New York.

Sherwood, C. & Brayer, G.D. (1985). Crystallization and preliminary diffraction data for iso-1-cytochrome *c* from yeast. *J. Mol. Biol.*, **185**, 209–210.

Shrake, A. & Rupley, J.A. (1973). Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.*, **79**, 351–371.

Singh, J., Thornton, J.M. Snarey, M. & Campbell, S.F. (1987). The geometries of interacting arginine-carboxyls in proteins. *FEBS Letters*, **224**, 161–171.

Smith, M., Leung, D.W., Gillam, S., Astell, C.R., Montgomery, D.L. & Hall, B.D. (1979). Sequence of the gene for iso-1-cytochrome *c* in *Saccharomyces cerevisiae*. *Cell*, **16**, 753–761.

Stellwagen, E. (1978). Haem exposure as the determinate of oxidation-reduction potential of haem proteins. *Nature (London)*, **330**, 86–88.

Takano, T. & Dickerson, R.E. (1980). Redox conformational changes in refined tuna cytochrome *c*. *Proc. Nat. Acad. Sci., U.S.A.*, **77**, 6371–6375.

Takano, T. & Dickerson, R.E. (1981a). Conformation change of cytochrome *c*: I. Ferrocytochrome *c* structure refined at 1.5Å resolution. *J. Mol. Biol.*, **153**, 79–94.

Takano, T. & Dickerson, R.E. (1981b). Conformation change of cytochrome *c*: II. Ferricytochrome *c* refinement at 1.8Å and comparison with the ferricytochrome structure. *J. Mol. Biol.*, **153**, 95–115.

Tanaka, N., Yamane, T., Tsukihara, T., Ashida, T. & Kakudo, M. (1975). The crystal structure of bonito ferrocytochrome *c* at 2.3 Å resolution. *J. Biochem*, **77**, 147–162.

ten Kortenaar, P.B.W., Adams, P.S.H.M. & Tesser, G.I. (1985). Semisynthesis of horse heart cytochrome *c* analogues from two or three fragments. *Proc. Nat. Acad. Sci., U.S.A.*, **82**, 8279–8283.

Trewhella, J., Carlson, V.A.P., Curtis, E.H. & Heidorn, D.B. (1988). Differences in the solution structures of oxidized and reduced cytochrome *c* measured by small-angle X-ray scattering. *Biochemistry*, **27**, 1121–1125.

Ulmer, D.D. & Kagi, J.H.R. (1968). Hydrogen-deuterium exchange of cytochrome *c*: I. Effect of oxidation state. *Biochemistry*, **7**, 2710–2717.

Wallace, C.J.A., Mascagni, P., Chait, B.T., Collawn, J.F., Paterson, Y., Proudfoot, A.E.I. & Kent, S.B.H. (1989). Substitutions engineered by chemical synthesis at three conserved sites in mitochondrial cytochrome *c*. *J. Biol. Chem.*, **264**, 15199–15209.

Wand, A.J., Roder, H. & Englander, S.W. (1986). Two-dimensional ^{1}H NMR studies of cytochrome *c*: hydrogen exchange in the N-terminal helix. *Biochemistry*, **25**, 1107–1114.

Warshel, A. Churg, A.K. (1983). Converting structural changes upon oxidation of cytochrome *c* to electrostatic reorganization energy. *J. Mol. Biol.*, **168**, 693–697.

Watt, G.D. & Sturtevan, J.M. (1969). The enthalpy change accompanying the oxidation of ferrocytochrome *c* in the pH range 6–11 at 25°. *Biochemistry*, **8**, 4567–4571.

Weber, C., Michel, B. & Bosshard, H.R. (1987). Spectroscopic analysis of the cytochrome *c* oxidase-cytochrome *c* complex: circular dichroism and magnetic circular dichroism measurements reveal change of cytochrome *c* heme geometry imposed by complex formation. *Proc. Nat. Acad. Sci., U.S.A.*, **84**, 6687–6691.

Wendoloski, J.J., Matthew, J.B., Weber, P.C. & Salemme, F.R. (1987). Molecular dynamics of a cytochrome *c*-cytochrome b_5 electron transfer complex. *Science*, **238**, 794–796.

Whitford, D., Gao, Y., Pielak, G.J., Williams, R.J.P., McLendon, G.L. & Sherman, F. (1991). The role of the internal hydrogen bond network in first-order protein electron transfer between *Saccharomyces cerevisiae* iso-1-cytochrome *c* and bovine microsomal cytochrome b_5. *Eur. J. Biochem.*, pp. 359–367.

Williams, G., Clayden, N.J. Moore, G.R. & Williams, R.J.P. (1985a). Comparison of the solution and crystal structures of mitochondrial cytochrome *c*. Analysis of paramagnetic shifts in the nuclear magnetic resonance spectrum of Ferricytochrome *c*. *J. Mol. Biol.*, **183**, 447–460.

Williams, G., Moore, G.R. & Williams, R.J.P. (1985b). Biological electron transfer: the structure, dynamics and reactivity of cytochrome *c*. *Comments Inorg. Chem.*, **4**, 55–98.

Wilson, A.J.C. (1942). Determination of absolute from relative X-ray intensity data. *Nature (London)*, **150**, 151–152.

Wood, L.C., Muthukrishnan, K., White, T.B., Ramadas, L. & Nall, B.T. (1988). Construction and Characterization of Mutant Iso-2-cytochrome *c* with replacement of conserved prolines. *Biochemistry*, **27**, 8554–8561.

Zhang, Q.P., Marohn, J. & McLendon, G. (1990). Macromolecular recognition in the cytochrome *c* - cytochrome *c* peroxidase complex involves fast two-dimensional diffusion. *J. Phys. Chem.*, **94**, 8628–8630.

Zheng, C., Wong, C.F. & McCammon, J.A. (1990). Fluctuation of the solvent-accessible surface area of tuna ferrocytochrome *c*. *Biopolymers*, **29**, 1877–1883.

Appendix A

Determination of the Absolute Scale Factor

A.1 Introduction

The structure factors initially obtained after data processing are on a relative scale. In order to perform a meaningful atomic coordinate refinement these structure factors must be multiplied by a scale factor so that they can be compared to calculated structure factors. This scale factor is appropriately called the absolute scale factor. A common approach to determining the absolute scale factor is to first obtain an initial estimate by statistical means which is later improved upon during subsequent structure refinement. In total, three methods have commonly been used for obtaining initial estimates for the absolute scale factor: linear rescale, least-squares rescale and the Wilson plot method.

A.2 Initial Estimate for the Absolute Scale Factor

A.2.1 Linear rescale

Linear rescale requires a structure factor data set from a related isomorphous crystalline form of the protein which is already on an absolute scale. The scale factor K between this structure factor data set and a set of data just collected can be computed as follows:

$$K = \frac{\sum_{hkl} F_{R_{hkl}}}{\sum_{hkl} F_{o_{hkl}}} \tag{A.4}$$

where F_R and F_o are the structure factors from the highly related previously scaled data set and the data set to be scaled, respectively. The sum is taken over all the reflections which are common between the two data sets. This linear rescale method assumes that the total amount of scattering material in the unit cell is identical in the two related crystals. One drawback

to this method is that the scale factor obtained is largely determined by the strong reflections present.

A.2.2 Least-squares rescale

An improvement on the linear rescale approach is to introduce a least-squares analysis where the scale factor is chosen so as to minimize the difference between the two data sets. This translates mathematically into minimizing the following function:

$$\varphi = \sum_{hkl}(KF_{o_{hkl}} - F_{R_{hkl}})^2 \tag{A.5}$$

Solving $\frac{\partial\varphi}{\partial K} = 0$ gives the following expression for the scale factor:

$$K = \frac{\sum_{hkl} F_{R_{hkl}} F_{o_{hkl}}}{\sum_{hkl}(F_{o_{hkl}})^2} \tag{A.6}$$

A drawback of the least-squares rescale method is that it assumes the overall thermal factor, which is related to the fall off in structure factor magnitudes with increasing resolution, is similar for the two data sets. This is not necessarily the case even for highly related data sets. On the other hand, the linear rescale is less sensitive to differences in overall thermal factors since it is biased towards strong reflections which occur mostly at low resolution.

A.2.3 Wilson plot method

An *ab initio* method for estimating absolute scale factor requiring only a knowledge of the chemical composition of the unit cell contents was developed by Wilson (1942). Assuming the atomic contents of the unit cell are randomly distributed, the theoretical mean value of diffraction intensities as a function of resolution is given by:

$$< \overline{I}_T > = \sum_{i=1}^{N}(f_{oi}e^{-B_i(\frac{\sin\theta}{\lambda})^2})^2 \tag{A.7}$$

In this equation N is the total number of atoms in the unit cell, and f_{oi} and B_i are the atomic scattering factor and isotropic thermal factor for atom i. If the individual thermal factors

are replaced by an overall thermal factor, the following expression can be formulated for the absolute scale factor:

$$K^2 < \overline{I} > = < \overline{I}_T > = e^{-2B(\frac{\sin\theta}{\lambda})^2} \sum_{i=1}^{N} f_{oi}^2 \tag{A.8}$$

Note that the scale factor is squared in this equation since intensities instead of structure factors are compared.

Traditionally this equation has been evaluated by rewriting it into the following form:

$$\ln \frac{< \overline{I} >}{\sum_{i=1}^{N} f_{oi}^2} = -\ln 2K - 2B(\frac{\sin\theta}{\lambda})^2 \tag{A.9}$$

A plot of $\ln(< \overline{I} > / \sum_{i=1}^{N} f_{oi}^2)$ versus $(\sin\theta/\lambda)^2$ should give a straight line with a slope of $-2B$ and an intercept of $-\ln 2K$. Figure A.53 gives an example of such a plot. As can be seen, at medium and high resolution (> 4.5 Å) the data does follow a straight line as predicted by Wilson, but at low resolution (< 5 Å) the data points present are scattered. The reason for the anomalous behaviour observed at low resolution is the non-randomness of the unit cell contents at this resolution. In determining the scale factor, resolution cutoffs are therefore first applied, after which a least squares fit to the theoretical line is calculated.

An alternative approach to solving equation A.8 is to apply a non-linear least-squares procedure (Press *et al.*, 1988). Graphically this can be represented by plotting $< \overline{I} > / \sum_{i=1}^{N} f_{oi}^2$ versus $(\sin\theta/\lambda)^2$ and determining the best exponential curve of the form $y = \frac{1}{K^2} e^{-2Bx}$ that fits the data points using a non-linear least-squares algorithm (see Figure A.54). As was the case for the Wilson plot, because of the non-randomness of the unit cell contents at low resolution, resolution cutoffs are first applied after which the least-squares fit to the exponential curve is calculated. Effectively, the major difference between this method and the original Wilson plot is the weighting of the data points in the least-squares procedure. As can be seen from Figures A.53 and A.54, in the Wilson plot method the theoretical curves agree better at high resolution while the non-linear approach forces a better agreement for lower resolution data points. The results obtained with these two fitting methods therefore give an indication of the spread in possible values for the absolute scale factor. Experience from the studies herein has

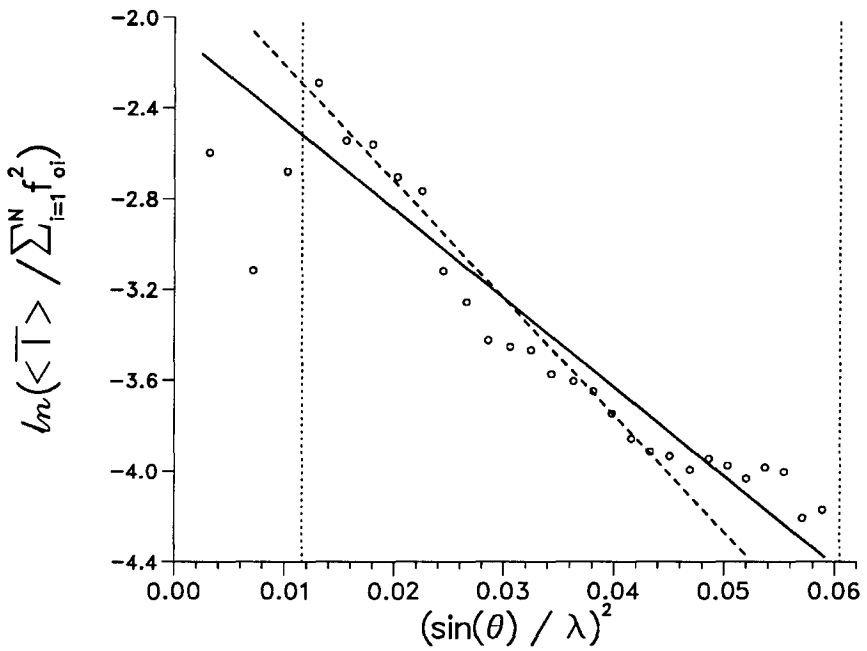

Figure A.53: Wilson plot for determining the absolute scale factor for the reduced N52I-Y67F mutant cytochrome *c* data set. For this analysis reciprocal space was divided into shells according to $(\frac{\sin\theta}{\lambda})^2$ with each containing at least 200 reflections. To determine the absolute scale factor a least-squares fit of the theoretical line (solid) to the data points was calculated using only data points between 4.5–2.0 Å as indicated by the vertical dotted lines. This analysis suggested that the absolute scale factor for this data set was 2.8. For comparison, a theoretical line (dashed) is also shown representing the absolute scale factor determined using a non-linear method (see Figure A.54). This approach suggested a value of 2.3 for the absolute scale factor.

shown that the final absolute scale factor resulting from structure refinement falls between the two values obtained using the linear and non-linear versions of the Wilson plot. Therefore as an initial estimate for the scale factor the average of these values was used.

A.3 Refinement of the Absolute Scale Factor

Beyond the usual scale refinement, a method found useful for estimating the shift in scale factor, and which has been used for the structure determinations presented herein, is the following. In this method the agreement between F_o and F_c is described as a function of both the absolute

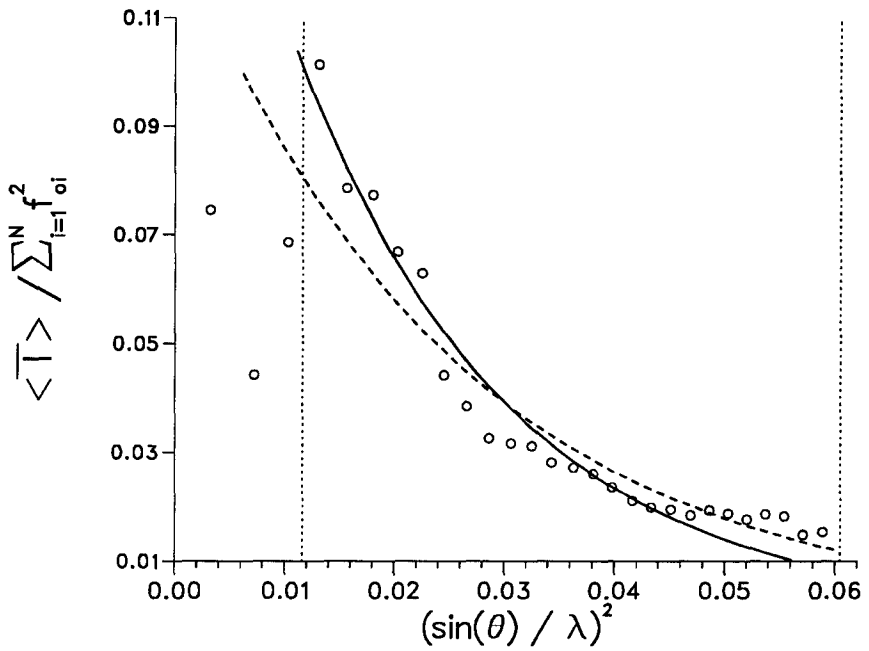

Figure A.54: Absolute scale factor plot for the reduced N52I-Y67F mutant cytochrome *c* data set. For this analysis reciprocal space was divided into shells according to $(\frac{\sin\theta}{\lambda})^2$ with each containing at least 200 reflections. To determine the absolute scale factor a least-squares fit of the theoretical exponential curve (solid line) to the data points was calculated using only data points between 4.5–2.0 Å as indicated by the vertical dotted lines. This analysis suggested a value of 2.3 for the absolute scale factor. Also shown is the theoretical curve (dashed line) derived from the absolute scale factor (2.8) determined using the Wilson plot method (see Figure A.53). The difference in the values obtained with these two methods reflect the accuracy of the absolute scale factor determinations. In practice the average of the two values was used as an initial estimate for the absolute scale factor.

scale factor and the discrepancy in the overall thermal factor between data and model.

$$KF_{o_{hkl}} = F_{c_{hkl}} e^{\Delta B(\frac{\sin\theta}{\lambda})^2} \tag{A.10}$$

Here ΔB represents the discrepancy between the actual overall thermal factor and that of the atomic model. To evaluate this equation one can rewrite it into the form:

$$\ln \frac{F_{o_{hkl}}}{F_{c_{hkl}}} = -\ln K + \Delta B(\frac{\sin\theta}{\lambda})^2 \tag{A.11}$$

A plot of $\ln \frac{F_{o_{hkl}}}{F_{c_{hkl}}}$ versus $(\frac{\sin\theta}{\lambda})^2$ should give a straight line with a slope of ΔB and intercept of $-\ln K$ (see Figure A.55). This is analogous to the Wilson plot method described above with

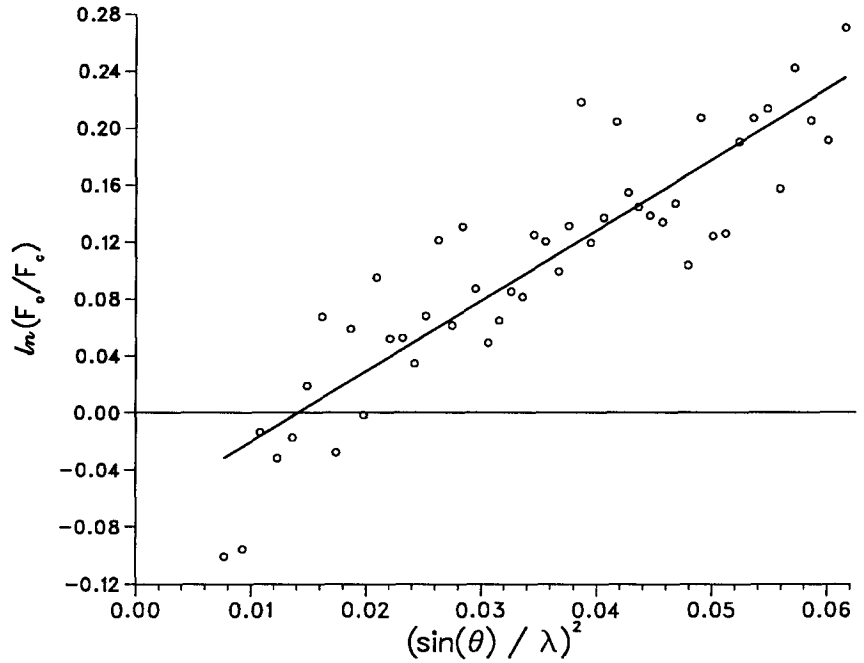

Figure A.55: A comparison of the amplitudes of the observed and calculated structure factors of the oxidized N52I structure after only six cycles of least-squares refinement. For this analysis reciprocal space was divided into shells according to $(\frac{\sin\theta}{\lambda})^2$ with each shell containing at least 100 structure factors. As can be seen, at high resolution (less than 4 Å; $(\frac{\sin\theta}{\lambda})^2 > 0.015$) the amplitudes of F_c's are systematically underestimated due to a high value for the overall thermal factor of the model. By fitting a straight line to the data points an estimate can be obtained for the discrepancy between the actual overall thermal factor and that of the model as well as for the absolute scale factor of the data set. This analysis suggests that the overall thermal factor for the oxidized N52I model after six cycles of refinement is ~3.5 Å^2 too high and that the absolute scale factor can be improved by multiplication of 1.07.

the main difference being the substitution of $\sum_{i=1}^{N} f_{oi}$ for F_c. Alternatively, one can minimize the function:

$$\varphi = \sum_{hkl}(KF_{o_{hkl}} - F_{c_{hkl}}e^{\Delta B(\frac{\sin\theta}{\lambda})^2})^2 \tag{A.12}$$

using non-linear least-squares procedures similar to what has been described above for the modified Wilson plot method. It is only necessary to apply the scale factor correction to the observed structure factors. Subsequent structure refinement will improve the individual atomic thermal factors so that the discrepancy between the actual overall thermal factor and that of the model will disappear.

Appendix B

Theory of Crystallographic Refinement

B.1 Introduction

The foundation of protein structure refinement is that both the structural model and the observed intensities can be expressed in terms of structure factors. Structure factors for the model can be calculated using the following equation:

$$F_{c_{hkl}} = K \sum_{i=1}^{N} f_{oi}\, e^{-B_i(\frac{\sin\theta}{\lambda})^2}\, e^{2\pi i(hx_i+ky_i+lz_i)} \tag{B.13}$$

Here $F_{c_{hkl}}$ is the calculated structure factor (Miller indices h,k,l), K is an overall scale factor and N is the total number of atoms in the unit cell. The atomic scattering factor, isotropic thermal factor and the position of atom i are represented by f_{oi}, B_i and (x_i, y_i, z_i), respectively. The diffraction angle, which is dependent on the Miller indices, is represented by θ, and λ is the wavelength of the x-ray radiation.

To obtain structure factors from the experimental diffraction data the square-root of the measured intensities is taken.

$$|F_{o_{hkl}}| = \sqrt{I_{hkl}} \tag{B.14}$$

Note that while F_c is a complex number only the amplitude of F_o can be measured experimentally. This reflects the fundamental phase problem in crystallography.

A common way of expressing the agreement between the refinement model and the observed structure factors is the crystallographic R-factor:

$$R = \frac{\sum_{hkl} ||F_{o_{hkl}}| - |F_{c_{hkl}}||}{\sum_{hkl} |F_{o_{hkl}}|} \tag{B.15}$$

The R-factor is often expressed as a percentage by multiplication by 100. In the case of a perfectly matching refinement model the R-factor would be 0% while with a randomly incorrect

model this value would be ~59% for acentric data (Blundell & Johnson, 1976). Typical R-factors for small molecule structures are less than 5%, whereas protein structure determinations yield R-factors in the range of 10 to 25%. The higher R-factors of protein structures can be attributed to the intrinsically greater flexibility of these molecules and the higher solvent content of the crystals (from 30–95% of the unit cell volume). Both these properties lead to large thermal vibrations in these molecules, dynamic disorder among possible local conformations and static variations in the structures within different unit cells (Hendrickson & Konnert, 1981).

B.2 Reciprocal-space Refinement

The three-dimensional structures of the proteins described in this thesis were refined using the reciprocal-space refinement method. In this method the object is to minimize the differences between F_o and F_c using a least squares approach. Therefore, refinement is aimed at finding the absolute minimum of the following function:

$$\varphi = \sum_{hkl} \omega_{hkl}(|F_{o_{hkl}}| - K|F_{c_{hkl}}|)^2 \tag{B.16}$$

In this equation K is the overall scale factor from equation B.13 and ω_{hkl} is a weighting factor. This weighting factor can reflect the accuracy of measured structure factor, or can be used to give structure factors different weights depending on their resolution.

The standard way of determining the minimum of the function φ is to take the partial derivatives with respect to all the parameters and set these to zero:

$$\frac{\partial \varphi}{\partial x_i} = 0 \ ; \quad \text{for all parameters } x_i \tag{B.17}$$

Thus for a function φ with N parameters the problem is reduced to solving N equations with N unknowns.

In the case where φ is a linear function a solution can be obtained using a matrix approach since the partial derivatives can also be written as linear equations with the general form of:

$$\frac{\partial \varphi}{\partial x_i} = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ii}x_i + \cdots + a_{iN}x_N - b_i = 0 \tag{B.18}$$

In matrix form the N equations to solve become:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2N} \\ \vdots & & & & & \\ a_{i1} & a_{i2} & \cdots & a_{ii} & \cdots & a_{iN} \\ \vdots & & & & & \\ a_{N1} & a_{N2} & \cdots & a_{Ni} & \cdots & a_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_N \end{pmatrix} \tag{B.19}$$

Or in abbreviated notation:

$$A \times \vec{x} = \vec{b}$$

To solve this, one has to invert matrix A.

$$\vec{x} = A^{-1} \times \vec{b} \tag{B.20}$$

Three problems are encountered when one uses this approach for the refinement of macromolecular structures. These are:

1. the non-linearity of the equations
2. the ratio of the number of observations to the number of parameters to be refined
3. the size of the matrix

B.2.1 The non-linearity of the equations

The least-squares method described above is straight-forward but uses the fundamental assumption that the function φ is linear. The function represented in equation B.16 is not linear. To solve this problem the function φ or its derivatives are approximated by a linear function using a Taylor series:

$$f(z) = \sum_{m=0}^{\infty} \frac{f^{(m)}(a)}{m!}(z-a)^m \tag{B.21}$$

Here $f^{(m)}(a)$ is the m order derivative evaluated at a.

Applied to the structure refinement problem the procedure used is as follows. First simplify equation B.16 by explicitly stating that φ and F_c are a function of the variables $x_1, x_2, \cdots$ (in vector notation $\vec{x}$) and treating the overall scale factor K as one of the variables of F_c.

$$\varphi(\vec{x}) = \sum_{hkl} \omega_{hkl}[F_{o_{hkl}} - F_{c_{hkl}}(\vec{x})]^2 \tag{B.22}$$

Note that the absolute value bars are deleted in this equation and those that follow for readability. Analogous to above, this function is minimal when all derivatives are zero:

$$\frac{\partial \varphi}{\partial x_i} = -2 \sum_{hkl} \omega_{hkl}[F_{o_{hkl}} - F_{c_{hkl}}(\vec{x})] \frac{\partial F_{c_{hkl}}(\vec{x})}{\partial x_i} = 0 \ ; \quad \text{for all parameters } x_i \tag{B.23}$$

At this point the problem can be linearized by expressing $F_c(\vec{x})$ as a Taylor series leaving out all terms after the first derivative.

$$F_c(\vec{x}) = F_c(\vec{x^0}) + \sum_j^N \frac{\partial F_c(\vec{x^0})}{\partial x_j}(x_j - x_j^0) \tag{B.24}$$

Substituting this into equation B.23 and rearranging:

$$\sum_j^N \sum_{hkl} \frac{\partial F_{c_{hkl}}(\vec{x^0})}{\partial x_i} \frac{\partial F_{c_{hkl}}(\vec{x^0})}{\partial x_j}(x_j - x_j^0) = \sum_{hkl} \omega_{hkl}[F_{o_{hkl}} - F_{c_{hkl}}(\vec{x^0})] \frac{\partial F_{c_{hkl}}(\vec{x^0})}{\partial x_i} \tag{B.25}$$

Since there are N of these equations they can be written in the following matrix form.

$$A \times \vec{\delta} = \vec{b} \tag{B.26}$$

In this equation

$$\begin{aligned} a_{ij} &= \sum_{hkl} \frac{\partial F_{c_{hkl}}(\vec{x^0})}{\partial x_i} \frac{\partial F_{c_{hkl}}(\vec{x^0})}{\partial x_j} \\ b_i &= \sum_{hkl} \omega_{hkl}[F_{o_{hkl}} - F_{c_{hkl}}(\vec{x^0})] \frac{\partial F_{c_{hkl}}(\vec{x^0})}{\partial x_i} \\ \delta_i &= x_i - x_i^0 \end{aligned}$$

This equation can be solved by inverting matrix A as discussed previously.

In taking this approach to structure refinement, two factors must be remembered. First, solving for the above linear equations will only provide shifts (δ_i) for the parameters. How

optimal these shifts are will depend on the validity of the Taylor series approximation. In practice, the updated parameters will be used in a new Taylor expansion and the process will be iterated until the shifts are smaller than their estimated errors. Reducing the number of iterations necessary to reach convergence is possible if more terms of the Taylor series are incorporated. However, most crystallographic refinement programs only use first order or at the most second order derivatives because the extra computing time required for calculating additional terms in the Taylor expansion is far more than will be saved in reducing the number of iterations.

Secondly, the above method of minimization does not guarantee that the overall global minimum of function φ will be found. It is more likely that the minimization will converge to a local minimum. To overcome this problem one can evaluate the agreement between F_o and F_c in real-space by inspection of difference electron density maps which can suggest improvements to the structure and thus to the calculated structure factors. After adjustments to the structure have been made reciprocal-space minimization can be continued.

B.2.2 The ratio of the number of observations to the number of parameters to be refined

In the least-squares refinement method the ratio of the number of data points to the number of refinement parameters, also known as the overdeterminancy, is an important indicator of the reliability and feasibility of the refinement. The minimum overdeterminancy required for a least-squares refinement is 1, assuming that the data can be perfectly represented by the parameters. This can never be the case for experimental situations since the diffraction data cannot be perfectly measured and the refinement model cannot be perfectly represented. To compensate for this situation more data points than parameters are necessary. The higher the overdeterminancy the more robust the refinement will be and the more precise the resulting model will become. In the case of small molecule structures the overdeterminancy can be as high as 10 resulting in extremely robust refinements and accurate structures.

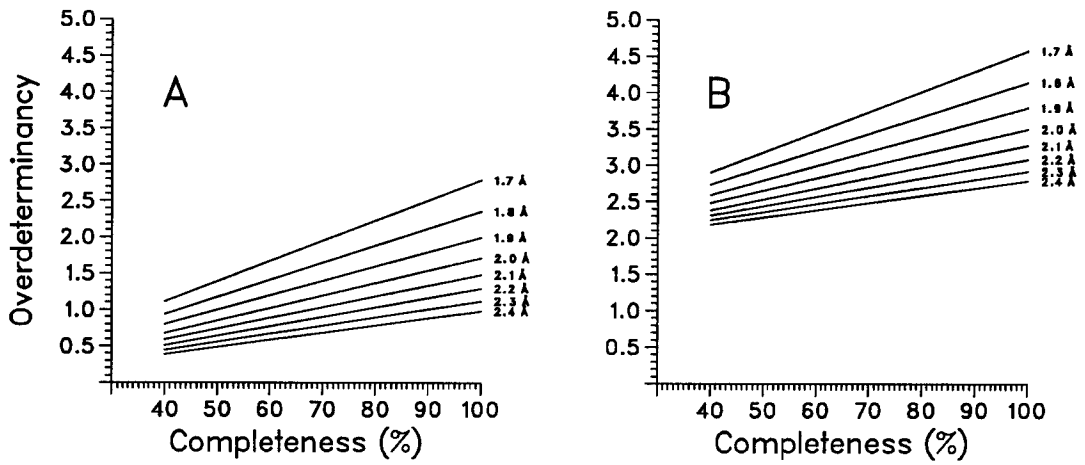

Figure B.56: The overdeterminancy for the refinement of a typical cytochrome *c* structure with respect to the completeness of the diffraction data used in the refinement and the resolution of the data collected. In graph (a) the overdeterminancy is given for unrestrained refinement. In this case only structure factors are considered as observations, and as such the ratio of the number of observations to the number of parameters is generally not greater than 2. When stereochemical information is added as restraints, the observation to parameter ratio can be dramatically improved. In (b) is shown the overdeterminancy for a refinement using PROLSQ (Hendrickson & Konnert, 1981) with standard restraints (~6750 restraints; see Table B.28). This approach would theoretically improve the overdeterminancy by 1.75. However, since the stereochemical restraints are highly correlated and can not be counted as independent observations, the actual improvement of the number of observations to the number of parameters ratio is less than is shown in (b). How much less can, unfortunately, not be determined since this would require inversion of matrix A from equation B.26 (as modified for the restrained parameter refinement approach) which is not feasible as discussed in the text.

Unfortunately, for macromolecular structure determinations the number of observations (diffraction intensities) that can be measured is limited due to resolution dependent falloff in the diffraction data. Figure B.56a shows that for a typical cytochrome *c* structure refinement one must have at least a complete 2.3 Å data set or a 50% complete 1.9 Å data set to obtain an overdeterminancy greater than 1. For an overdeterminancy of 2, which is still insufficient for a stable refinement, one needs to have at least a complete 1.9 Å data set.

Clearly, to make refinement of macromolecular structures feasible, the overdeterminancy ratio must be improved. A method to accomplish this is by increasing the number of observations using an approach known as restrained refinement. Here stereochemical knowledge is included

as data points by treating ideal bond distances, torsional angles, etc. as observations (Hendrickson & Konnert, 1981). Mathematically this is done by minimizing an extended equation B.16 as represented below:

$$\Phi = \varphi_{\text{xray}} + \varphi_{\text{stereochem}_1} + \varphi_{\text{stereochem}_2} + \cdots \tag{B.27}$$

Here φ_{xray} corresponds to the right hand side of equation B.16 and $\varphi_{\text{stereochem}_i}$ describes different stereochemical restraints. An example of a stereochemical restrain is the function that controls the bond lengths:

$$\varphi_{\text{bondlengths}} = \sum_{i}^{\text{No. of bonds}} \omega_{\text{bond}_i}(d_i - d_i^{\text{ideal}})^2 \tag{B.28}$$

In this equation d_i and d_i^{ideal} are the actual and ideal bond lengths for bond i and ω_{bond_i} is a weighting factor. In this method the structural model is allowed to deviate from ideal geometry by adjustments of the various weighting factors and is therefore called restrained refinement. The increase in observations can be quite significant using this approach. For a typical cytochrome *c* refinement, using the least-squares restrained parameter refinement program PROLSQ (Hendrickson & Konnert, 1981), with approximately 3800 variables (950 atoms), a total of about 6600–6900 restraints can be added (depending on the exact conformation of the protein and the number and position of solvent molecules; see Figure B.56 and Table B.28). However since structural restraints are correlated, this does not imply that the overdeterminancy is improved by $\frac{\sim 6750}{3800} \approx 1.75$.

B.2.3 The size of the matrix

The normal matrix A in equation B.26 as modified for the restrained parameter refinement approach can be large for macromolecular refinements. In the case of a typical cytochrome *c* refinement with isotropic temperature factors and fixed occupancies, the number of variables is about 3800. This means that matrix A will contain 3800^2 or more than 14 million elements making the inversion of matrix A computationally impractical. Therefore, an alternative method for the solution of equation B.26 is employed. In the implementation of PROLSQ, the

Table B.28: Typical numbers of stereochemical restraints of different classes used in the refinement of yeast iso-1-cytochrome *c* structures

Stereochemical refinement parameters	Typical number of restraints	Typical refinement restraint weighting values
Bond distances		
1-2 bond distance	925	0.020 Å
1-3 bond distance	1250	0.030 Å
1-4 bond distance	340	0.050 Å
special distances†	14	0.060 Å
Planar restraints	780	0.020 Å
Chiral volume	120	0.150 $Å^3$
Non-bonded contacts		
single-torsion	400	0.250 Å
multi-torsion	90	0.250 Å
possible hydrogen bonds	70	0.250 Å
Torsion angles		
planar (0°or 180°)	110	2.5°
staggered (±60°,180°)	170	20.0°
orthonormal (±90°)	15	15.0°
Isotropic thermal factors		
main-chain bond	500	1.5 $Å^2$
main-chain angle	650	2.5 $Å^2$
side-chain bond	430	2.0 $Å^2$
side-chain angle	620	3.0 $Å^2$

† The special distances define the bonds between the heme iron and the His18 and Met80 ligands. See Hendrickson (1985) for a thorough description of the restraint definitions.

method of conjugate gradients is employed since it efficiently utilizes the properties of matrix A (Hendrickson & Konnert, 1980). For example, matrix A is symmetrical around the diagonal, leading to a reduction of $\sim \frac{1}{2}$ in the number of elements required for the evaluation of matrix A. Additionally, not all elements of the matrix are equally important, with many having values close to zero. It appears that the stability, speed and radius of convergence of refinement are not seriously affected if only those elements of the matrix that are non-zero for stereochemical

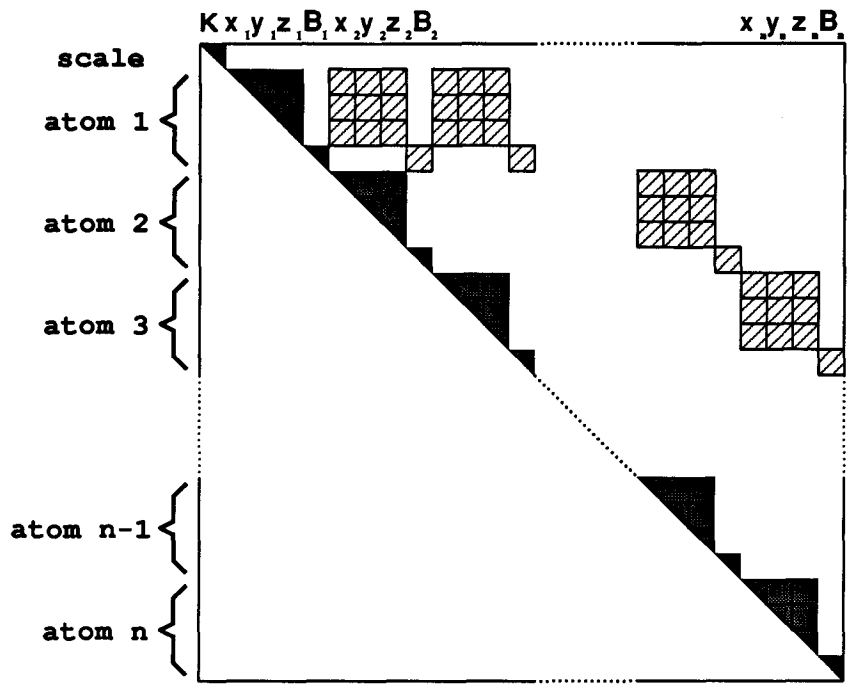

Figure B.57: A diagram of the normal matrix A showing the elements evaluated by the restrained parameter refinement program PROLSQ (Hendrickson & Konnert, 1981). The dark elements are in "self blocks" containing the six elements related to the atomic coordinates plus an additional element for the thermal factor and one element for refinement of the scale factor (K). The hatched blocks are elements that relate parameters that are correlated through stereochemical restraints. Blank elements are assumed to have a value of zero. In the above example the atom pairs 1 & 2, 1 & 3, 2 & (n-1), and 3 & n are correlated through stereochemical restraints.

restraints are retained (see Figure B.57). The result is that generally less than 1% of the elements of the matrix are evaluated. In the case of cytochrome c, approximately 0.2% of the matrix elements are retained in a typical refinement cycle. The conjugate gradient method as implemented in PROLSQ also incorporates knowledge from previous iterations which increases its convergence rate. Theoretically convergence will be reached in N steps (number of parameters), but experience has shown that a minimum can be reached in much fewer (typically 10–30) steps.

Appendix C

Estimating Coordinate Errors in Macromolecular Structures

C.1 Introduction

Any quantitative determination is incomplete without a reliable measure of the associated uncertainty. Without this knowledge interpretation of the obtained results can easily lead to erroneous conclusions. When applied to the study of mutant structures in a protein-engineering environment, without an estimate of the accuracy of the three-dimensional structures involved, differences between structures can be either over or under estimated. Such a situation seriously hampers the interpretation of the relationships between protein structure and expressed biological activity.

Determining the accuracy of a three dimensional atomic structure solved by x-ray diffractions methods is in theory straight-forward. One first determines the errors in the observed structure factors and evaluates how these errors are propagated through subsequent calculations. This will result in estimates for the accuracy of the parameters that define the structure and can be calculated for the case of small molecules as follows:

$$\sigma_i = \sqrt{\frac{b_{ii} \sum_{hkl} \omega_{hkl}(|F_{o_{hkl}}| - |F_{c_{hkl}}|)^2}{M - N}} \tag{C.29}$$

Here σ_i is the standard deviation of parameter i, b_{ii} is the ith diagonal element of the inverse matrix A^{-1}, A being the matrix from equation B.26, ω_{hkl} is the weighting factor used in the least-squares refinement and M and N are the number of structure factors (observations) and parameters, respectively.

From the discussion on refinement presented in Appendix B it can be seen that this procedure is generally not practical in the case of macromolecular crystallography due to the size of

matrix A. As a result, several alternatives for estimating the accuracy of macromolecular structures have been developed. These include those based on empirical rules and those that have a statistical origin (Luzzati, 1952; Cruickshank, 1949,1954). However, the program SHELXL-92 (by G. Sheldrick), which is able to calculate e.s.d.'s for macromolecules using equation C.29, has recently become available.

C.2 Empirical error estimates

Recently a number of studies have been carried out in which the structures of highly related or identical proteins have been compared (Perry *et al.*, 1990; Bott, 1991; Fauman & Stroud, 1991; Ohlendorf *et al.*, 1991). From an analysis of the deviations between these structures, which should represent random errors, empirical rules can be derived. For example, Perry *et al.* (1990) have proposed the following empirical equation:

$$\sigma_{r_i} = \frac{3}{4} SR(0.0015B_i^2 - 0.0203B_i + 0.359) \tag{C.30}$$

In this formula the radial coordinate error of atom i is dependent on its thermal factor B_i, the resolution S, of the structure and the crystallographic R-factor represented as R. This method only provides information regarding the coordinate errors and does not give estimates for the standard deviations of thermal factor parameters. This is a general feature of all the methods for estimating errors in structures.

Clearly there are some drawbacks in the use of empirical rules. First the validity of the database of structures must be considered. If related but not identical structures are used, the differences between these structures not only reflect random errors, but also real differences. Further, real differences between identical or nearly identical structures can exist due to different crystallization conditions or crystal contacts. These situations will result in an overestimation of the real coordinate errors. A second criticism relates to the possibility that not all the variables are present in the equation representing the empirical rule. This will again result in false error estimates in those cases were variables are missing or perhaps a particular set of structures is

dependent on a different mix of variables. For example, in the empirical equation C.30 the atom type is not a variable, while it is to be expected that the position of an iron atom with a certain thermal factor will be more accurate than a carbon atom with the same thermal factor because of the greater scattering power of the iron atom. Thus empirical rules for estimating errors must be used with care particularly in their application to new structures.

C.3 Luzzati error estimates

The method developed by Luzzati (1952) is the one most often used for estimating the coordinate errors in macromolecular structures. In its derivation random errors in the atomic coordinates are expressed in terms of probability distributions. These can be combined into an overall probability distribution which will have a normal Gaussian form. Thus even if errors in different parts of the protein structure do not follow a normal distribution, the overall distribution will still have a normal Gaussian form due to the central limit theorem. By applying the mathematics of the Fourier transform to this normal distribution, a probability distribution for the structure factors can be obtained. Evaluation of this probability distribution gives information on the overall coordinate error in the structure, although the accuracy for individual parts of the structure can not be determined.

Luzzati expressed the probability distribution for the structure factors in the form of R-factors dependent on resolution and assuming that the errors in atomic positions are the only source for the differences between F_o and F_c. Therefore, a Luzzati plot (see Figure C.58) shows the theoretical dependence of R-factor versus resolution (expressed as $\frac{\sin\theta}{\lambda}$) for different overall error estimates of the structure. By inserting the R-factors of different resolution shells for a particular experimentally determined structure an estimate of its overall accuracy can be made.

Read (1986) extended this work by allowing for the fact that the difference between F_o and F_c is caused not only by coordinate errors, but also by an incomplete model. However, in a subsequent paper Read (1990) described a number of flaws in some of the assumptions made by Luzzati for deriving the structure factor probability distributions and how these distributions

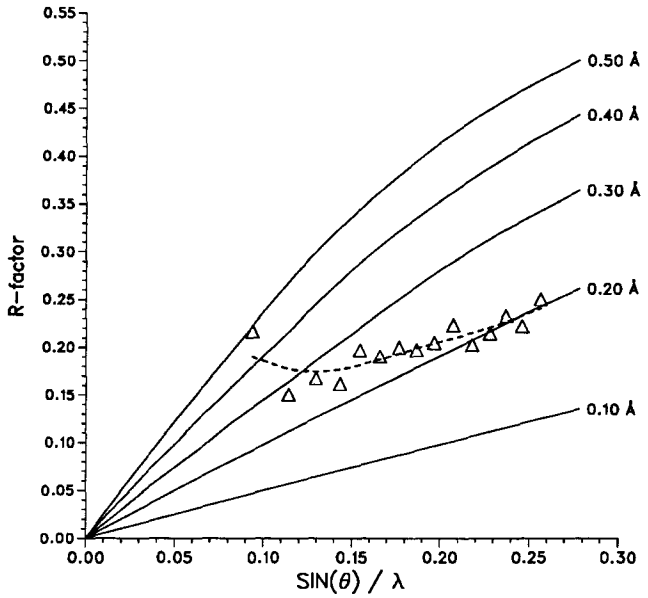

Figure C.58: Plot of the theoretical dependence of the crystallographic R-factor factor as a function of resolution assuming various levels of r.m.s. coordinate errors (Luzzati, 1952). Also shown (dashed line and △ symbols) is the dependence of the R-factor versus resolution for the experimentally determined wild-type yeast iso-1-cytochrome *c* structure in the oxidized state. The analysis suggests that the r.m.s. coordinate error for this structure is ~0.22 Å.

relate to coordinate errors. His main critique is that the Luzzati method does not take into account the fact that a certain portion of the coordinate errors will be absorbed into the thermal factor parameters, thereby skewing the relationship between the structure factor probability distributions and coordinate errors. However, this effect will probably be small at the end of refinement. Nevertheless, the existence of such systematic errors suggests overall coordinate errors obtained with the Luzzati method should be used comparatively (Read, 1990).

C.4 Cruickshank error estimates

Whereas Luzzati's method starts from the assumption that there is a certain error in the coordinates of the model and evaluates how this affects the structure factors, a method developed by Cruickshank (1949,1954) begins by assuming errors are present in the observed structure

factors and determines how this error is reflected in the electron density map. Taking the differences between F_o's and F_c's as an estimate for the errors in the observed structure factors and assuming well resolved spherically symmetric atoms, Cruickshank estimated the standard deviation of a peak position to be (for orthorhombic, tetragonal and cubic cells):

$$\sigma_x = \frac{2\pi\sqrt{\sum_{hkl} h^2(|F_{o_{hkl}}| - |F_{c_{hkl}}|)^2}}{aV\,\frac{\partial^2\rho}{\partial^2 x}} \tag{C.31}$$

Similar equations for σ_y and σ_z can be derived. In this equation a is the length of the crystallographic unit cell a axis, h is one of the Miller indices, and V is the volume of the unit cell. The term $\frac{\partial^2\rho}{\partial^2 x}$ represents the curvature of the electron density at the peak position which can be estimated directly from an F_o map. An alternative method for estimating this curvature is to use the thermal factor of the atom at the peak position (Chambers & Stroud, 1979; Read *et al.*, 1983; Cruickshank, 1985). When this is done the above equation transforms into:

$$\sigma_{x_i} = \frac{a\sqrt{\sum_{hkl} h^2(|F_{o_{hkl}}| - |F_{c_{hkl}}|)^2}}{2\pi\sum_{hkl}\frac{m}{2}h^2 f_{oi} e^{-B_i(\frac{\sin\theta}{\lambda})^2}} \tag{C.32}$$

Here m is 2 for centric reflections, otherwise it is 1. The expression for the radial error in the tetragonal space group $P4_32_12$ is then given by:

$$\sigma_{r_i} = \sqrt{2{\sigma_{x_i}}^2 + {\sigma_{z_i}}^2} \tag{C.33}$$

since $\sigma_{x_i} = \sigma_{y_i}$ ($a = b$ and $F_{hkl} = F_{khl}$) and the angles between the unit cell axes are 90°. Thus the modified Cruickshank equation relates the standard deviation of an atom position to the atom type by virtue of the scattering factor f_{oi} and the thermal factor (see Figure C.59).

Chambers and Stroud (1979) remark that σ_x is equal to the estimated r.m.s. shift in x during a single cycle of real-space difference Fourier refinement. This implies that for cases where the refinement of a model has not yet converged, Cruickshank's method will underestimate the coordinate errors. To counteract this problem, Chambers and Stroud propose that comparisons between refinement models that are separated by more than one cycle of difference Fourier refinement will give more realistic estimates for the coordinate errors. Read *et al.* (1983) adapted

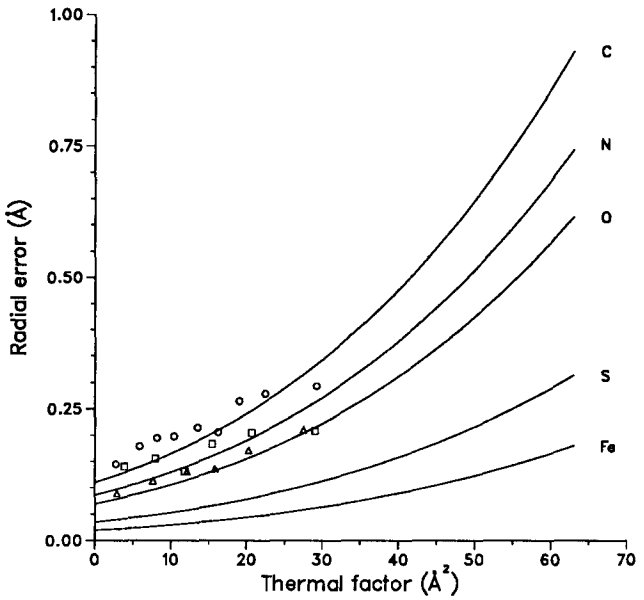

Figure C.59: Plot of the dependence of the radial coordinate errors on the value of the thermal factor for different atom types in oxidized yeast iso-1-cytochrome *c* as estimated by the Cruickshank (1949, 1954) method. Also shown are the r.m.s. shifts (using the numerical values on the vertical axis) after unrestrained refinement as a function of the thermal factor for carbon (○), nitrogen (□) and oxygen (△) atoms.

this idea by suggesting that instead of difference Fourier refinement, shifts from unrestrained least-squares reciprocal-space refinement based on a fully converged structure could be used. When unrestrained refinement is applied to the oxidized form of wild-type cytochrome *c* (see Figure C.59), it is observed that the r.m.s. shifts correspond closely to the radial error estimates derived with equation C.32. On this basis it can be concluded than that this cytochrome *c* structure is well refined and that reliable atomic error estimates can be obtained using the method developed by Cruickshank. However, care must be taken that the assumptions made in deriving equation C.32 are valid for the structure whose coordinate errors are being estimated. Specifically, in the derivation it is assumed that thermal factors correctly reflect the curvature of the electron density at atomic positions. This assumption is generally true, but fails for surface side-chains with high degrees of mobility when thermal factor restraints are applied in the refinement.